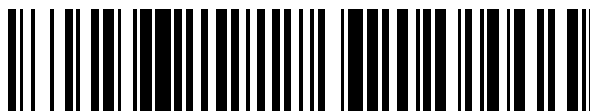


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 701 652**

51 Int. Cl.:

**C12N 15/67** (2006.01)

**A61K 39/12** (2006.01)

**C12N 15/63** (2006.01)

**A61K 39/00** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **15.02.2013 PCT/EP2013/000460**

87 Fecha y número de publicación internacional: **22.08.2013 WO13120628**

96 Fecha de presentación y número de la solicitud europea: **15.02.2013 E 13706423 (4)**

97 Fecha y número de publicación de la concesión europea: **04.07.2018 EP 2814962**

54 Título: **Ácido nucleico que comprende o codifica para un tallo-bucle de histona y una secuencia poli(A) o una señal de poliadenilación para aumentar la expresión de un antígeno patogénico codificado**

30 Prioridad:  
**15.02.2012 WO PCT/EP2012/000673**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**25.02.2019**

73 Titular/es:  
**CUREVAC AG (100.0%)  
Paul-Ehrlich-Str. 15  
72076 Tübingen, DE**

72 Inventor/es:  
**THESS, ANDREAS;  
SCHLAKE, THOMAS y  
PROBST, JOCHEN**

74 Agente/Representante:  
**AZNÁREZ URBIETA, Pablo**

ES 2 701 652 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

**Ácido nucleico que comprende o codifica para un tallo-bucle de histona y una secuencia poli(A) o una señal de poliadenilación para aumentar la expresión de un antígeno patogénico codificado**

5

La presente invención se refiere a una secuencia de ácido nucleico que comprende o codifica para una región codificante, codificando para al menos un péptido o proteína que comprende un antígeno patogénico o un fragmento del mismo, al menos un tallo-bucle de histona y una secuencia poli(A). Además, la presente invención proporciona el uso del ácido nucleico para aumentar la expresión de dicho péptido o proteína codificados *in vitro*. También se describe su uso para la preparación de una composición farmacéutica, en especial una vacuna, por ejemplo para utilizarse en el tratamiento de enfermedades infecciosas. La presente invención describe además un método *in vitro* o *ex vivo* para aumentar la expresión de un péptido o proteína que comprende un antígeno patogénico o un fragmento del mismo utilizando el ácido nucleico que comprende o codifica para un tallo-bucle de histona y una secuencia poli(A).

10

15

El aumento de una inmunidad adaptativa mediante vacunación se dirige a estimular respuestas efectivas contra antígenos específicos presentes en patógenos *in vivo*. Los métodos tradicionales de vacunación utilizando patógenos atenuados vivos o muertos por calor han tenido éxito en la prevención y el tratamiento de enfermedades infecciosas tales como viruela, polio y difteria, pero existen enfermedades graves para las que no está disponible una vacuna efectiva (por ejemplo malaria y VIH), o la vacuna disponible sólo proporciona una protección transitoria o parcial (por ejemplo cólera y gripe). Las estrategias más modernas se dirigen a antígenos seleccionados diana para subconjuntos de células presentadoras de antígenos y que dirigen el sistema inmune hacia las respuestas inmunes de tipo Th1 y/o Th2 asociadas con la protección contra el patógeno específico. Estas estrategias exclusivamente dirigidas también pueden conducir al desarrollo de vacunas terapéuticas capaces de superar algunas de las deficiencias inmunitarias inducidas por patógenos para una evasión inmune (Gamvrellis, A., D. Leong *et al.* (2004), *Immunology and Cell Biology* 82, 506–516.). Una de estas estrategias novedosas es la vacunación genética.

20

25

La terapia génica y la vacunación genética son métodos de medicina molecular que ya se han probado en la terapia y la prevención de enfermedades y, en general, tienen un efecto considerable sobre la práctica médica diaria, en particular respecto al tratamiento de enfermedades como se mencionó anteriormente. Ambos métodos, la terapia génica y la vacunación genética, se basan en la introducción de ácidos nucleicos en las células o tejido del paciente y en el procesamiento posterior de la información codificada por el ácido nucleico introducido en las células o tejido, es decir la expresión (proteína) de los polipéptidos deseados.

30

35

En los procedimientos de terapia génica, se utiliza típicamente ADN incluso aunque también se conoce el ARN de desarrollos recientes. Es de señalar que en todos estos procedimientos de terapia génica el ARNm funciona como mensajero para la información de secuencia de la proteína codificada independientemente de si se utiliza ADN, ARN vital o ARNm.

40

45

En general se considera que el ARN es una molécula inestable: las ARNsas son ubicuas y notoriamente difíciles de inactivar. Además, el ARN también es químicamente más lábil que el ADN. Así, quizá es sorprendente que el “estado predeterminado” de un ARNm de una célula eucariota esté caracterizado por una estabilidad relativa y se requieren señales específicas para acelerar la degradación de los ARNm individuales. La razón principal de este hallazgo parece ser que el ARNm descompuesto dentro de las células es catalizado casi exclusivamente por exonucleasas. Sin embargo, los extremos de los ARNm eucariotas están protegidos contra estas enzimas por estructuras terminales específicas y sus proteínas asociadas: un m7GpppN CAP en el extremo 5' y típicamente una secuencia poli(A) en el extremo 3'. La eliminación de estas dos modificaciones terminales se consideran así limitantes de la velocidad de degradación del ARNm. Aunque se ha caracterizado un elemento estabilizante en el 3' UTR del ARNm de la alfa-globina, las secuencias de ARN que afectan al recambio de los ARNm eucariotas típicamente actúan como un promotor de la degradación en general al acelerar la desadenilación (revisado en Meyer, S., C. Temme, *et al.* (2004), *Crit Rev Biochem Mol Biol* 39(4): 197-216.).

50

55

Como se mencionó anteriormente, los extremos 5' de los ARNm eucariotas típicamente están modificados post-transcripcionalmente para portar una estructura CAP metilada, por ejemplo m7GpppN. Además de las funciones de empalme, estabilización y transporte del ARN, la estructura CAP mejora significativamente el reclutamiento de la subunidad ribosómica 40S en el extremo 5' del ARNm durante el inicio de la traducción. La última función requiere el reconocimiento de la estructura CAP por el complejo factor de iniciación eucariota eIF4F. La secuencia poli(A) adicionalmente estimula la traducción vía el reclutamiento aumentado de la subunidad 40S para los ARNm, un efecto que requiere la intervención de la proteína de unión con poli(A) (PABP). La PABP, a su vez, recientemente ha demostrado interactuar físicamente con eIF4G, que forma parte del complejo eIF4F unido a CAP. Así, se ha postulado un modelo de bucle cerrado del inicio de la traducción

en los ARNm poliadenilados, rematados (Michel, Y. M., D. Poncet, *et al.* (2000), *J Biol Chem* 275(41): 32268-76.).

Prácticamente todos los ARNm eucariotas terminan con esta secuencia poli(A) que se agrega a su extremo 3' mediante el mecanismo de segmentación/poliadenilación ubicua. La presencia de una secuencia poli(A) en el extremo 3' es una de las características más reconocibles de los ARNm eucariotas. Después de la segmentación, la mayoría de los pre-ARNm, con la excepción de los transcritos de histona dependientes de la replicación, adquieren una cola poliadenilada. En este contexto, el procesamiento del extremo 3' es un proceso co-transcripcional nuclear que estimula el transporte de los ARNm desde el núcleo al citoplasma y afecta a la estabilidad y la traducción de los ARNm. La formación de este extremo 3' ocurre en una reacción en dos pasos dirigida por el mecanismo de segmentación/poliadenilación y depende de la presencia de dos elementos de secuencia en los precursores del ARNm (pre-ARNm); un hexanucleótido altamente conservado AAUAAA (señal de poliadenilación) y una secuencia rica en G/U aguas abajo. En un primer paso, los pre-ARNm se segmentan entre estos dos elementos. En un segundo paso estrechamente unido al primer paso, el extremo 3' recién formado se extiende por la adición de una secuencia poli(A) consistente en 200-250 adenilatos, lo cual afecta posteriormente a todos los aspectos del metabolismo del ARNm, incluyendo la exportación, estabilidad y traducción del ARNm (Dominski, Z. y W. F. Marzluff (2007), *Gene* 396(2): 373-90.).

La única excepción conocida a esta regla son los ARNm de histonas dependientes de la replicación que terminan con un tallo-bucle de histona en lugar de una secuencia poli(A). Secuencias tallo-bucle de histona ilustrativas se describen en Lopez *et al.* (Dávila López, M., & Samuelsson, T. (2008), *RNA* (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308.).

Los tallos-bucle de histona en los pre-ARNm típicamente están seguidos por una secuencia rica en purina conocida como elemento aguas debajo de histona (HDE). Estos pre-ARNm se procesan en el núcleo mediante una única segmentación endonucleolítica de aproximadamente 5 nucleótidos aguas abajo del tallo-bucle, catalizada por el U7 RNPsn por apareamiento de bases del U7 ARNsn con el HDE. La secuencia 3'-UTR que comprende la estructura de tallo-bucle de histona y el elemento aguas abajo de histona (HDE) (el sitio de unión del U7 RNPsn) se denominaron en general como la señal de procesamiento 3' de la histona (Véase, por ejemplo, Chodchoy, N., N. B. Pandey, *et al.* (1991). *Mol Cell Biol* 11(1): 497-509).

Debido al requisito de empaquetar un ADN recién sintetizado en la cromatina, la síntesis de histonas es regulada en acuerdo con el ciclo celular. Una síntesis aumentada de proteínas histona durante la fase S se alcanza mediante la activación transcripcional de los genes de histona, así como por la regulación pos-transcripcional de los niveles de ARNm de la histona. Se podía demostrar que el tallo-bucle de histona es esencial para todos los pasos pos-transcripción de la regulación de la expresión de histonas. Para un procesamiento eficiente es necesaria la exportación del ARNm en el citoplasma, la carga sobre los polirribosomas y la regulación de la estabilidad del ARNm.

En el contexto anterior, se identificó una proteína de 32 kDa que está asociada con el tallo-bucle de histona en el extremo 3' de los mensajes de histona tanto en el núcleo como en el citoplasma. El nivel de expresión de esta proteína de unión al tallo-bucle (SLBP) se regula por el ciclo celular y es mayor durante la fase S, cuando aumentan los niveles de ARNm de histona. La SLBP es necesaria para un procesamiento eficiente en el extremo 3' del pre-ARNm de la histona mediante el U7 RNPsn. Después del procesamiento, la SLBP permanece asociada al tallo-bucle en el extremo de los ARNm de la histona madura y estimula su traducción en proteínas histona en el citoplasma (Dominski, Z. and W. F. Marzluff (2007), *Gene* 396(2): 373-90). De manera interesante, el dominio de unión del ARN de la SLBP se conserva en todos los metazoarios y protozoarios (Dávila López, M. y Samuelsson, T. (2008), *RNA* (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308) y se podía demostrar que su unión a la secuencia tallo-bucle de histona depende de la estructura del tallo-bucle y que el sitio de unión mínima contiene al menos 3 nucleótidos 5' y 2 nucleótidos 3' del tallo-bucle (Pandey, N. B., *et al.* (1994), *Molecular and Cellular Biology*, 14(3), 1709-1720 and Williams, A. S., & Marzluff, W. F., (1995), *Nucleic Acids Research*, 23(4), 654-662.).

Incluso aunque los genes de histona se clasifican generalmente como "dependientes de la replicación", dando lugar a la terminación del ARNm en un tallo-bucle de histona, o como de "tipo reemplazamiento", dando lugar a un ARNm que porta una cola poli(A) en su lugar, se han identificado en casos raros ARNm de origen natural que contienen tanto un tallo-bucle de histona y una poli(A) u oligo(A) 3' de los mismos. Sánchez *et al.* examinaron el efecto de las colas oligo(A) de origen natural 3' anexas al tallo-bucle del ARNm de histona durante la ovogénesis de *Xenopus* utilizando luciferasa como proteína reportera y encontraron que la cola oligo(A) es parte activa del mecanismo de represión de la traducción que silencia el ARNm de histona durante la ovogénesis y su eliminación es parte del mecanismo que activa la traducción de los ARNm de histona (Sanchez, R. y W. F. Marzluff (2004), *Mol Cell Biol* 24(6): 2513-25).

Además, se han investigado los requisitos de la regulación de las histonas dependientes de la replicación al nivel del procesamiento de pre-ARNm y la estabilidad del ARNm utilizando estructuras artificiales que codifican para

la proteína marcadora alfa-globina, aprovechando el hecho de que el gen de globina contiene intrones en oposición a los genes de histona, con menos intrones. Para ello, se generaron constructos donde las secuencias codificantes de alfa-globina estaban seguidas de una señal tallo-bucle de histona (el tallo-bucle de histona seguido por el elemento aguas abajo de histona) y una señal de poliadenilación (Whitelaw, E., *et al.* (1986). *Nucleic Acids Research*, 14(17), 7059-7070.; Pandey, N. B., & Marzluff, W. F. (1987). *Molecular and Cellular Biology*, 7(12), 4557-4559.; Pandey, N. B., *et al.* (1990). *Nucleic Acids Research*, 18(11), 3161-3170).

En otra aproximación, Lüscher *et al.* investigaron la regulación dependiente del ciclo celular de un gen H4 de histona recombinante. Se generaron constructos donde la secuencia de codificación H4 estaba seguida de una señal de tallo-bucle de histona y una señal de poliadenilación, las dos señales de procesamiento separadas incidentalmente por una secuencia codificadora de galactoquinasa (Lüscher, B. *et al.*, (1985). *Proc. Natl. Acad. Sci. USA*, 82(13), 4389-4393).

Además, Stauber *et al.* identificaron la secuencia mínima requerida para conferir una regulación del ciclo celular sobre los niveles de ARNm de histona H4. Para estas investigaciones se emplearon constructos comprendiendo una secuencia codificadora para el marcador de selección xantina:guanina fosforibosil transferasa (GPT) que precede a una señal de tallo-bucle de histona seguida por una señal de poliadenilación (Stauber, C. *et al.*, (1986). *EMBO J*, 5(12), 3297-3303).

Examinando el procesamiento del pre-ARNm de histona, Wagner *et al.* identificaron factores necesarios para la segmentación de los pre-ARNm de histona utilizando un constructo reporter situando la EGFP entre una señal de tallo-bucle de histona y una señal de poliadenilación, de forma que la EGFP se expresaba únicamente en caso de interrumpirse el procesamiento del pre-ARNm de histona (Wagner, E. J. *et al.*, (2007). *Mol Cell* 28(4), 692-9).

Cabe señalar que la traducción del ARNm poliadenilado normalmente requiere que la secuencia 3' poli(A) esté próxima al 5' CAP. Esto es mediado mediante la interacción proteína-proteína entre la proteína de unión a poli(A) y el factor de inicio eucariota eIF4G. Con respecto a los ARNm de histona dependientes de la replicación, no se ha descubierto un mecanismo análogo. En este contexto, Gallie *et al.* demuestran que el tallo-bucle de histona es funcionalmente similar a una secuencia poli(A) en que mejora la eficiencia traduccional y es co-dependiente de un 5'-CAP para establecer un nivel de traducción eficiente. Ellos demostraron que el tallo-bucle de histona es suficiente y necesario para aumentar la traducción de un ARNm reporter en células de ovario de hámster chino transfectadas, aunque se debe colocar en el extremo 3' para que funcione óptimamente. Por tanto, de forma similar a la cola poli(A) en otros ARNm, el extremo 3' de estos ARNm de histona parece ser esencial para la traducción *in vivo* y es funcionalmente análogo a una cola poli(A) (Gallie, D. R., Lewis, N. J., & Marzluff, W. F. (1996), *Nucleic Acids Research*, 24(10), 1954-1962).

Además, se podía demostrar que la SLBP está unida al ARNm de la histona citoplásmica y se requiere para su traducción. Incluso aunque la SLBP no interactúe directamente con eIF4G, el dominio requerido para la traducción del ARNm de histona interactúa con la proteína SLIP1, recientemente identificada. En un paso adicional, la SLIP1 interactúa con eIF4G y permite circularizar el ARNm de la histona y apoyar una traducción eficiente del ARNm de histona mediante un mecanismo similar a la traducción de los ARNm poliadenilados.

Como se mencionó anteriormente, los procedimientos de terapia génica normalmente utilizan ADN para transferir la información de codificación en la célula, que luego es transcrita en el ARNm, portando los elementos de origen natural de un ARNm, en particular la estructura 5'-CAP y la secuencia 3' poli(A) para asegurar la expresión de la proteína terapéutica o antigénica codificada.

Sin embargo, en muchos casos, los sistemas de expresión se basan en la introducción de estos ácidos nucleicos en el interior de células o tejido del paciente y la expresión posterior de los polipéptidos deseados codificados por estos ácidos nucleicos no presentan el nivel de expresión deseado o incluso requerido para una terapia eficiente, independientemente de si se utiliza ADN o ARN.

En la técnica anterior, hasta ahora se han realizado diferentes intentos para aumentar el rendimiento de la expresión de una proteína codificada, en particular empleando sistemas mejorados de expresión tanto *in vitro* como *in vivo*. Los métodos para aumentar la expresión descritos en general en la técnica anterior se basan en el uso de vectores de expresión o casetes que contengan promotores específicos y elementos de regulación correspondientes convencionales. Ya que estos vectores de expresión o casetes típicamente se limitan a sistemas celulares particulares, estos sistemas de expresión se deben adaptar para su uso en diferentes sistemas celulares. Estos vectores de expresión o casetes adaptados entonces normalmente se transfectan en células y se tratan típicamente según la línea celular específica.

Por tanto, se da preferencia principalmente a aquellas moléculas de ácido nucleico que son capaces de expresar las proteínas codificadas en una célula diana mediante los sistemas inherentes en la célula, independientemente de los promotores y elementos de regulación que son específicos para los tipos celulares

particulares. En este contexto, puede haber distinción entre los elementos estabilizantes del ARNm y los elementos que aumentan la eficacia de traducción del ARNm.

5 ARNm que están optimizados en su secuencia codificadora y que en general son adecuados para ello se describen en la solicitud WO 02/098443 (CureVac GmbH). Por ejemplo, la WO 02/098443 describe ARNm en general estabilizados y optimizados para la traducción en sus regiones codificantes. La WO 02/098443 describe además un método para determinar modificaciones de secuencia. La WO 02/098443 describe también posibilidades de sustitución de los nucleótidos adenina y uracilo en las secuencias de ARNm para aumentar el contenido guanina/citosina (G/C) de las secuencias. De acuerdo con la WO 02/098443, estas sustituciones y adaptaciones para aumentar el contenido de G/C se pueden utilizar para aplicaciones terapéuticas génicas, aunque también para vacunas genéticas en el tratamiento del cáncer o enfermedades infecciosas. En este contexto, la WO 02/098443 en general menciona secuencias como secuencias base para estas modificaciones, donde los códigos modificados del ARNm para al menos un péptido o polipéptido biológicamente activo, que es traducido en el paciente a tratar, por ejemplo, no son adecuadas en ningún caso o los son con errores. Alternativamente, la WO 02/098443 propone ARNm que codifican para antígenos, por ejemplo antígenos patogénicos o virales, como secuencia base para estas modificaciones.

20 En un procedimiento adicional para aumentar la expresión de una proteína codificada, la solicitud WO 2007/036366 describe el efecto positivo de secuencias poli(A) largas (en particular, más largas de 120 pb) y la combinación de al menos dos regiones no traducidas 3' del gen de beta-globina sobre la estabilidad del ARNm y la actividad traduccional. Además, la WO 2006/024518 A1 describe moléculas de ARN que codifican un antígeno patogénico y que comprenden una secuencia 3'-UTR de alfa- o beta-globina y una secuencia poli(A). Russel J.E. et al. (1996), Blood, (87), 5314-5323, indican elementos dentro de la 3'-UTR de alfa-globina, beta-globina e histonas que pueden estabilizar el ARN.

25 Así, incluso aunque todos estos últimos documentos de la técnica anterior ya tratan de proporcionar herramientas bastantes eficientes para los procedimientos de terapia génica y adicionalmente una estabilidad y actividad traduccional mejoradas del ARNm, siguen existiendo el problema de una estabilidad generalmente inferior de las aplicaciones basadas en ARN frete a las vacunas de ADN y de los procedimientos terapéuticos génicos basados en ADN. Por consiguiente, todavía existe una necesidad en la técnica de proporcionar herramientas mejoradas para los procedimientos de terapia y vacunación genética o como terapias suplementarias de los tratamientos convencionales, como se analizó anteriormente, que permitan una mejor provisión de proteínas codificadas *in vivo*, por ejemplo vía una estabilidad y/o actividad traduccional mejoradas del ARNm, preferentemente para la terapia y la vacunación genética.

Además, a pesar de todo el progreso en la técnica, la expresión eficiente de un péptido o proteína codificados en sistemas libres de células, células u organismos (expresión recombinante) sigue siendo un problema desafiante.

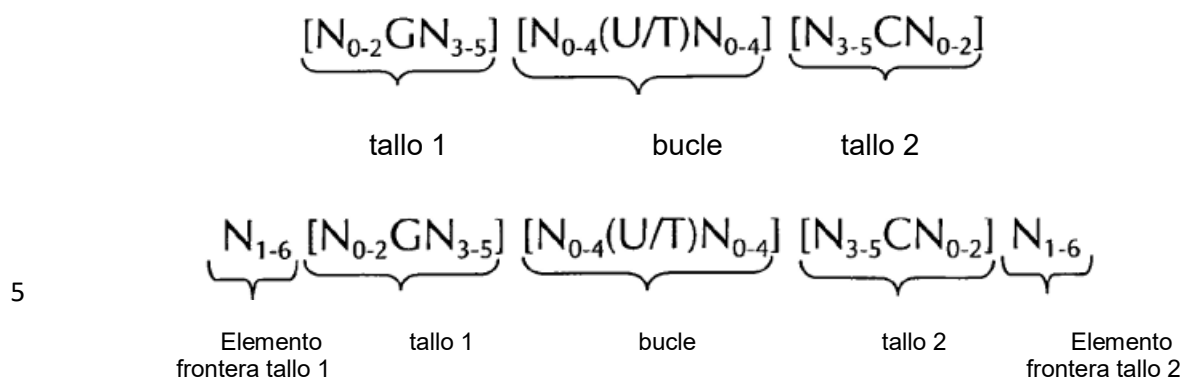
35 Por tanto, el objetivo de la presente invención es proporcionar métodos adicionales y/o alternativos para aumentar la expresión de una proteína codificada, preferentemente vía una estabilización adicional del ARNm y/o un aumento de la eficacia traduccional de este ARNm con respecto a los ácidos nucleicos conocidos de la técnica anterior para el uso en vacunación genética en el tratamiento terapéutico o profiláctico de enfermedades infecciosas.

40 Este objetivo se resuelve según el objeto de las reivindicaciones adjuntas. En particular, el objetivo de la presente invención se resuelve según un primer aspecto mediante una secuencia de ácido nucleico de la invención que comprende o codifica en la dirección 5' → 3':

- 45
- una región codificante, que codifica para al menos un péptido o proteína (a);
  - al menos un tallo-bucle de histona (b), y
  - una secuencia poli(A) (c),
- o
- i)
- una región codificante, que codifica para al menos un péptido o proteína (a);
  - una secuencia poli(A) (c), y
  - 50 - al menos un tallo-bucle de histona (b),

donde el al menos un tallo-bucle de histona de i) o ii) se selecciona de entre las siguientes fórmulas (I) o (II):

fórmula (I) (secuencia tallo-bucle sin elementos frontera de tallo):



donde

10 los elementos frontera de tallo 1 o tallo 2  $N_{1-6}$  son una secuencia consecutiva de 1 a 6, de preferencia de 2 a 6, con mayor preferencia de 2 a 5, incluso con mayor preferencia de 3 a 5, con máxima preferencia de 4 a 5 o 5N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos;

15 Tallo 1  $[N_{0-2}GN_{3-5}]$  es complementario inverso o complementario parcialmente inverso con el elemento de tallo 2 y es una secuencia consecutiva de entre 5 a 7 nucleótidos; donde  $N_{0-2}$  es una secuencia consecutiva de 0 a 2, de preferencia 0 a 1, con mayor preferencia 1N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; donde  $N_{3-5}$  es una secuencia consecutiva de 3 a 5, de preferencia 4 a 5, con mayor preferencia 4 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; y donde G es guanosina o un análogo de la misma y se puede remplazar opcionalmente por una citidina o un análogo de la misma, con la condición de que su nucleótido citidina complementario en el tallo 2 se reemplace por guanosina;

25 Secuencia bucle  $[N_{0-4}(U/T)N_{0-4}]$  está ubicada entre los elementos tallo 1 y tallo 2 y es una secuencia consecutiva de 3 a 5 nucleótidos, con mayor preferencia de 4 nucleótidos, donde cada  $N_{0-4}$  es independientemente una secuencia consecutiva de 0 a 4, de preferencia de 1 a 3, con mayor preferencia 1 a 2 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; y donde U/T representa uridina u opcionalmente timidina;

30 Tallo 2  $[N_{3-5}CN_{0-2}]$  es complementaria inversa o complementaria parcialmente inversa con el elemento tallo 1, y es una secuencia consecutiva de 5 a 7 nucleótidos; donde  $N_{3-5}$  es una secuencia consecutiva de 3 a 5, de preferencia 4 a 5, con mayor preferencia 4 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; donde  $N_{0-2}$  es una secuencia consecutiva de 0 a 2, de preferencia 0 a 1, con mayor preferencia 1 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; y donde C es citidina o un análogo de la misma y se puede remplazar opcionalmente por una guanosina o un análogo de la misma, con la condición de que su nucleótido guanosina complementario en el tallo 1 se reemplace por citidina;

donde

40 el tallo 1 y tallo 2 son capaces de realizar un apareamiento de bases entre sí formando una secuencia complementaria inversa, donde el apareamiento de bases se puede presentar entre el tallo 1 y el tallo 2, o formando una secuencia complementaria inversa parcial, donde el apareamiento de bases incompleto se puede presentar entre el tallo 1 y el tallo 2

donde al menos un tallo-bucle se enlaza a una proteína de unión a tallo-bucle (SLBP); y

45 donde dicho péptido o proteína comprende un antígeno patogénico o un fragmento del mismo con una longitud de al menos seis residuos aminoácidos y tiene la propiedad antigénica específica del péptido o proteína nativa de longitud completa,

donde el antígeno patogénico está asociado a una infección bacteriana, una infección viral o una infección protozoaria y donde el antígeno patogénico se selecciona de un antígeno patogénico del grupo consistente en

- Virus sincitial respiratorio (RSV),
- Virus de la inmunodeficiencia humana (VIH),
- Virus del herpes simplex (HSV),
- Virus de papiloma humano (HPV),
- Virus de parainfluenza humana (HPIV),
- Virus del Dengue,
- Virus de la Hepatitis B (HBV),
- Virus de Influenza,
- Virus de fiebre amarilla,
- Virus de la rabia,
- Plasmodium,
- Cytomegalovirus (CMV),
- Staphylococcus,
- Mycobacterium tuberculosis,
- Chlamydia trachomatis,
- Rotavirus,
- Metapneumovirus Humano (hMPV),
- Virus de fiebre hemorrágica de Crimea-Congo (CCHF),
- Virus de ébola,
- Henipavirus,
- Norovirus,
- Virus Lassa,
- Coronavirus,
- Rinovirus,
- Flavivirus,
- Virus de fiebre Rift Valley y
- Virus de la enfermedad de manos, pies y boca.

Es particularmente preferente que el ácido nucleico de la invención de acuerdo con el primer aspecto de la presente invención se produzca al menos parcialmente mediante síntesis de ADN o ARN, de preferencia como se describe aquí o sea un ácido nucleico aislado.

La presente invención se basa en el hallazgo sorprendente de los inventores de que la combinación de una secuencia poli(A) y al menos un tallo-bucle de histona, incluso aunque ambos representen mecanismos alternativos por naturaleza, actúa sinérgicamente, ya que esta combinación aumenta la expresión proteínica muchas veces por encima del nivel observado con cualquiera de los elementos individuales. El efecto sinérgico de la combinación de poli(A) y al menos un tallo-bucle de histona es visto independientemente del orden de la poli(A) y el tallo-bucle de histona e indeoendientemente de la longitud de la secuencia poli(A).

Por tanto, más en general la presente invención se refiere a un ácido nucleico que comprende o codifica para una región codificante, codificando al menos un péptido o proteína que comprende un antígeno patogénico o un fragmente, al menos un tallo bucle de histona y una secuencia poli(A), preferentemente para aumentar la expresión de dicho péptido o proteína codificado. Alternativamente, se describe (no se reivindica) que puede emplearse cualquier secuencia tallo-bucle apropiada diferente de una secuencia tallo-bucle de histona.

Se prefiere particularmente que la secuencia de ácido nucleico inventiva comprenda o codifique para a) una región codificante, que codifica al menos para un péptido o proteína que comprende un antígeno patogénico o un fragmento tal como se define aquí; b) al menos un tallo-bucle de histona como se define aquí y c) una secuencia poli(A); preferentemente para aumentar el nivel de expresión de dicho péptido o proteína codificado. La proteína codificada según la reivindicación 1 típicamente no es una proteína de histona, en particular ninguna proteína de histona de las familias de histonas H4, H3, H2A y/o H2B o un fragmento, derivado o variante de las mismas que conserve una función de (similar a) histona, esto es que forme un nucleosoma. También, la proteína codificada típicamente no se corresponde con una proteína enlazante de histona de la familia de histonas H1. La molécula de ácido nucleico inventivo típicamente no contiene ninguna de las señales reguladoras (5' y/o en particular 3') de un gen de histonas de ratón, en particular no un gen de histonas de ratón H2A y, además, con mayor preferencia, no el gen de histonas de ratón H2A614. En particular, no contiene un tallo-bucle de histona y/o una señal de procesamiento de tallo-bucle de histona de un gen de histona de ratón, en particular, no un de gen de histona de ratón H2A y con mayor preferencia no de un gen de histona de ratón H2A614.

5 También típicamente el ácido nucleico de la invención no proporciona una proteína reportera (por ejemplo luciferasa, GFP, EGFP,  $\beta$ -galactosidasa, en particular EGFP), galactocinasa (galK) y/o una proteína marcadora o de selección (por ejemplo alfa-globina, galactocinasa y xantina: guanina-fosforribosiltransferasa (GPT)) o una proteína reportera bacteriana, por ejemplo cloranfenicol-acetiltransferasa (CAT) u otras proteínas de resistencia a antibióticos bacterianos, por ejemplo derivadas de *neo* gen bacteriano, en su elemento (a).

10 Típicamente se entiende que una proteína reportera, marcadora o de selección no es una proteína antigénica de acuerdo con la invención. Una proteína reportera, marcadora o de selección o su gen subyacente se utiliza habitualmente como una herramienta de búsqueda en bacterias, cultivos celulares, animales o plantas. Confieren (de preferencia heterológamente) a los organismos que los que expresan una propiedad fácilmente  
15 identificable que se puede medir o que permite la selección. Específicamente, las proteínas marcadoras o de selección tienen una función seleccionable. Típicamente, en esta selección, las proteínas marcadoras o reporteras no son de origen natural en humanos u otros mamíferos, sino que se derivan de otros organismos, en particular de bacterias o plantas. Por consiguiente, las proteínas con una función de selección, marcadora o reportera que se originan de especies inferiores (por ejemplo bacterias) preferentemente son excluidas de entenderse como “proteínas antigénicas” de acuerdo con la presente invención. A este respecto, se debe  
20 entender que una proteína antigénica corresponde a una proteína que desencadena una reacción inmunológica que permite proteger inmunológicamente al sujeto contra una infección mediante un organismo o virus que ejerce una reacción patológica en el sujeto, dando como resultado un estado de enfermedad. En particular, una proteína de selección, marcadora o reportera permite identificar células transformadas mediante ensayos *in vitro*, por ejemplo basados en técnicas de fluorescencia u otras técnicas espectroscópicas y proporcionar resistencia a los antibióticos. Los genes de selección, reporteros o marcadores que otorgan estas propiedades a las células transformadas por tanto típicamente no se entienden como proteínas antigénicas patogénicas de acuerdo con la invención.

25 En cualquier caso, las proteínas reporteras, marcadores o de selección normalmente no ejercen ningún efecto antigénico como resultado de la respuesta inmunológica (del sujeto a tratar) hacia el antígeno patogénico. Si cualquier proteína reportera, marcadora o de selección individual en cualquier caso lo hiciera (además de su función reportera, de selección o marcadora), tal proteína reportera, marcadora o de selección típicamente no se entiende como un “antígeno patogénico” dentro del significado de la presente invención.

30 Por el contrario, un antígeno patogénico (incluyendo sus fragmentos) según la presente invención típicamente no tiene una función de selección, marcadora o reportera. Si cualquier “antígeno patogénico” individual en cualquier caso lo hace (además de su función antigénica), este antígeno patogénico preferentemente no se entiende como una “proteína de selección, marcadora o reportera” dentro del significado de la presente invención.

35 De mayor preferencia se entiende que un antígeno patogénico de acuerdo con la invención es un antígeno patogénico de organismos patogénicos, preferentemente bacterias o virus, que ejercen una función inmunológica. Típicamente estos antígenos no se califican como una proteína de selección, marcadora o reportera.

40 Así, la región codificante (a) que codifica al menos para un péptido o proteína como se define aquí preferentemente es heteróloga para al menos (b) el al menos un tallo-bucle de histona como se define aquí o más ampliamente para cualquier tallo-bucle adecuado. En otras palabras, “heterólogo” en el contexto de la presente invención significa que la al menos una secuencia tallo-bucle no es de origen natural una secuencia (reguladora) (por ejemplo en el 3' UTR) del gen específico, que codifica para la proteína antigénica (patogénica) o el péptido del elemento (a) del ácido nucleico inventivo. Por consiguiente, el tallo-bucle (de histona) del ácido nucleico inventivo preferentemente se deriva del 3' UTR de un gen distinto al que comprende la región  
45 codificante del elemento (a) del ácido nucleico inventivo. Por ejemplo, la región codificante del elemento (a) no codificará una proteína de histona o un fragmento, variante o derivado de la misma (que conserve la función de una proteína de histona) si el ácido nucleico inventivo es heterólogo, pero codificará cualquier otro péptido o secuencia (de la misma u otra especie) como se define aquí que ejerza una función antigénica distinta de la función de (similar)histona, por ejemplo codificará para una proteína antigénica (que ejerza una función  
50 antigénica, por ejemplo desencadenando la reacción del sistema inmune del sujeto, por ejemplo mediante una reacción del anticuerpo) permitiendo así que el ácido nucleico inventivo actúe como una vacuna, por ejemplo en mamíferos, en particular en seres humanos.

De acuerdo con un primer aspecto, el ácido nucleico de la invención según la reivindicación 1 comprende o codifica en la dirección 5'- o 3'-:

- 55
- una región codificante, que codifica para al menos un péptido o proteína que comprende un antígeno patogénico o fragmento del mismo como se define en la reivindicación 1 a);
  - al menos un tallo-bucle de histona como se define en la reivindicación 1, opcionalmente sin un elemento 3'(HDE) aguas debajo de la histona al tallo-bucle de histona b);



- una secuencia poli(A) c).

5 El término "elemento aguas debajo de la histona (HDE)" se refiere a un tramo de polinucleótido rico en purina de aproximadamente 5 hasta 20 nucleótidos 3' de los tallos-bucle de histona de origen natural, lo que representa el sitio de unión para el U7 ARNs implicado en el procesamiento del pre-ARNm de histona en el ARNm de histona maduro. Por ejemplo, en los erizos de mar el HDE es CAAGAAAGA (Dominski, Z. and W. F. Marzluff (2007), Gene 396(2): 373-90).

Además, se prefiere que el ácido nucleico inventivo de acuerdo con el primer aspecto de la presente invención no comprenda un intrón.

10 O, la secuencia del ácido nucleico inventivo de acuerdo con la primera reivindicación según el primer aspecto de la presente invención comprende o codifica para, de 5' a 3':

- una región codificante que codifica para al menos un péptido o proteína que comprende un antígeno patogénico o un fragmento del mismo como se define en la reivindicación 1 a);
- una secuencia poli(A) c); y
- al menos un tallo-bucle de histona como se define en la reivindicación 1 b).

15 La secuencia de ácido nucleico inventivo puede comprender cualquier ácido nucleico adecuado, seleccionado por ejemplo de cualquier ADN (de cadena simple o doble), preferentemente, sin limitarse a, por ejemplo un ADN genómico, un ADN plásmido, moléculas de ADN monohebra, moléculas de ADN de doble hebra, o se puede seleccionar de cualquier PNA (ácido nucleico peptídico) o se puede seleccionar por ejemplo de cualquier ARN (de cadena simple o doble), preferentemente un ARN mensajero (ARNm); etc. La secuencia de ácido nucleico inventivo también puede comprender un ARN viral (ARNv). Sin embargo, la secuencia de ácido nucleico inventivo puede no ser un ARN viral o puede no contener un ARN viral. Más específicamente, la secuencia de ácido nucleico inventivo puede no contener elementos de secuencias virales, por ejemplo intensificadores virales o promotores virales (por ejemplo, sin un promotor viral inactivado o elementos de secuencia, más específicamente no activados por estrategias de reemplazo), u otros elementos de secuencias virales, u otras secuencias de ácido nucleico viral o retroviral. Más específicamente, la secuencia de ácido nucleico inventivo puede no ser un vector retroviral o viral o un vector retroviral o viral modificado.

20 En cualquier caso, la secuencia de ácido nucleico inventivo puede o no contener una secuencia intensificadora y/o promotora, que se pueda modificar o no, o que se pueda activar o no. El intensificador y/o promotor puede ser expresable o no en plantas y/o en eucariotas y/o en procariotas. La secuencia de ácido nucleico inventivo puede contener una secuencia que codifica para un ribosoma (de auto-empalme) o no.

30 En realizaciones específicas, la secuencia de ácido nucleico inventivo puede ser o puede comprender un ARN de auto-replicación (replicón).

Preferentemente la secuencia de ácido nucleico inventivo es un ADN plásmido o un ARN, en particular un ARNm.

35 En realizaciones particulares del primer aspecto de la presente invención, el ácido nucleico inventivo es una secuencia de ácido nucleico comprendida en un ácido nucleico adecuado para la transcripción *in vitro*, en particular en un vector de transcripción *in vitro* adecuado (por ejemplo un plásmido o una secuencia de ácido nucleico lineal que comprende promotores específicos para la transcripción *in vitro*, como los promotores T3, T7 o SP6).

40 En otras realizaciones preferidas particulares del primer aspecto de la presente invención, el ácido nucleico inventivo está comprendido en un ácido nucleico adecuado para la transcripción y/o traducción en un sistema de expresión (por ejemplo en un vector de expresión o plásmido), en particular un sistema de expresión procariota (por ejemplo bacterias similares a *E. coli*) o eucariota (células de mamíferos similares a células CHO, células de levadura o células de insecto u organismos enteros similares a plantas o animales).

45 El término "sistema de expresión" significa un sistema (cultivo celular u organismos enteros) adecuado para la producción de péptidos, proteínas o ARN, en particular ARNm (expresión recombinante).

50 La secuencia de ácido nucleico inventivo de acuerdo con el primer aspecto de la presente invención comprende o codifica para al menos un tallo-bucle de histona como se define en la reivindicación 1. En el contexto de la presente invención, este tallo-bucle de histona típicamente se deriva de genes de histona y comprende un apareamiento de bases intramolecular de dos secuencias complementarias totalmente en vecindad o parcialmente inversas, formando así un tallo-bucle. Un tallo-bucle en general, independientemente de que sea un tallo-bucle de histona o no, se puede presentar en un ADN monohebra o, más comúnmente, en un ARN.

Esta estructura también se conoce como horquilla o asa de horquilla y en general consiste de un tallo y un bucle (terminal) dentro de una secuencia consecutiva, donde el tallo está formado por dos secuencias complementarias totalmente en vecindad o parcialmente inversas separadas por una secuencia corta, como una especie de separador, que constituye el bucle de la estructura tallo-bucle. Las dos secuencias complementarias totalmente en vecindad o parcialmente inversas se pueden definir como, por ejemplo, los elementos de tallo-bucle tallo 1 y tallo 2. El tallo-bucle se forma cuando dos secuencias complementarias en vecindad o parcialmente inversas, por ejemplo los elementos de tallo-bucle tallo 1 y tallo 2, forman pares de bases entre sí, conduciendo a un tramo de secuencia de ácido nucleico de doble hebra que comprende un bucle no apareado en su extremo terminal formado por la secuencia corta ubicada entre los elementos de tallo-bucle tallo 1 y tallo 2 en la secuencia consecutiva. Así, el bucle no apareado típicamente representa una región del ácido nucleico que no es capaz de formar pares de bases con cualquiera de estos elementos del tallo-bucle. La estructura resultante con forma de chupa-chups es un bloque de construcción clave en muchas estructuras secundarias de ARN. La formación de una estructura tallo-bucle depende así de la estabilidad del tallo resultante y de las regiones de bucle, siendo el primer pre-requisito típicamente la presencia de una secuencia que se puede plegar sobre sí misma para formar una doble cadena apareada. La estabilidad de los elementos de tallo-bucle apareado viene determinada por la longitud, el número de mal-apareamientos o protuberancias que contienen (un pequeño número de mal-apareamientos es típicamente tolerable, especialmente en un tramo largo de doble cadena), y la composición de bases de la región apareada. Es concebible una longitud de bucle de 3 a 15 bases, mientras que una longitud de bucle óptima especialmente preferente es de 3-10 bases, con mayor preferencia de 3 a 8, de 3 a 7, de 3 a 6 o incluso con mayor preferencia de 4 a 5 bases, y con la máxima preferencia 4 bases. La secuencia tallo que forma la estructura de doble cadena típicamente tiene una longitud de entre 5 a 10 bases, con mayor preferencia de 5 a 8 bases.

En el contexto de la presente invención, un tallo-bucle de histona según la reivindicación 1 típicamente se deriva de genes de histona (por ejemplo genes de las familias de histonas H1, H2A, H2B, H3, H4) y comprende un apareamiento de bases intramolecular de dos secuencias complementarias totalmente en vecindad o parcialmente inversas, formando así un tallo-bucle. Típicamente, un tallo-bucle 3' UTR de histona es un elemento de ARN implicado en el transporte nucleocitoplásmico de los ARNm de histona y en la regulación de la estabilidad y de la eficiencia de traducción en el citoplasma. Los ARNm de genes de histona en metazoarios carecen de poliadenilación y cola poli-A, en su lugar el procesamiento en el extremo 3' ocurre en un sitio entre este tallo-bucle altamente conservado y una región rica en purina de alrededor de 20 nucleótidos aguas abajo (el elemento aguas abajo de histona o HDE). El tallo-bucle de histona está unido por una proteína de unión al tallo-bucle de 31 kDa (SLBP - también denominada proteína de unión horquilla a histonas o HBP). Estas estructuras de tallo-bucle de histona preferentemente se emplean en la presente invención en combinación con otros elementos de secuencia y estructuras que no son de origen natural (lo que significa que están sin transformar en organismos vivos/células) en los genes de histona, pero que están combinados -de acuerdo con la invención- para proporcionar un ácido nucleico artificial heterólogo. Así, la presente invención se basa en particular en el hallazgo de que una combinación artificial (no natural) de una estructura tallo-bucle de histona con otros elementos de secuencia heterólogos que no se presentan en genes de histona o genes de histona de metazoarios y están aislados de las regiones operacionales y/o reguladoras de la secuencia (lo que influye en la transcripción y/o la traducción) de genes que codifican para proteínas distintas a histonas, proporcionan efectos ventajosos. Por consiguiente, la invención proporciona la combinación de una estructura tallo-bucle de histona con una secuencia poli(A) (3' terminal de una región codificante) que no se presenta en los genes de histonas de metazoarios. La invención también proporciona una combinación de una estructura tallo-bucle de histona con una región codificante que codifica para un antígeno patogénico de un patógeno según la reivindicación 1, que preferentemente no está presente en genes de histona de metazoarios (la región codificante y la secuencia tallo-bucle de histona son heterólogas). Es preferente si estos antígenos patogénicos no se presentan en metazoarios en absoluto, pero se derivan de organismos unicelulares, por ejemplo bacterias, o de virus. En una realización preferida adicional, todos los elementos (a), (b) y (c) del ácido nucleico inventivo son heterólogos entre sí y se combinan artificialmente a partir de tres fuentes diferentes, por ejemplo la región codificante de antígenos proveniente de bacterias o virus como se reivindica, el tallo-bucle de histona proveniente de un gen de histona de metazoario (como se reivindica en la reivindicación 1) y la secuencia poli(A) de por ejemplo un gen de metazoario distinto a un gen de histona.

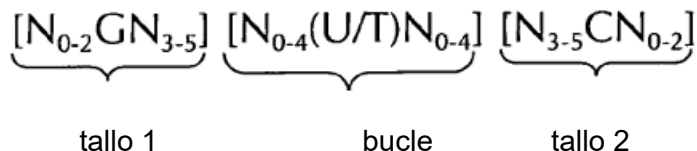
Por tanto, un tallo-bucle de histona es una estructura de tallo-bucle como se describe o se reivindica aquí que exhibe/conserva la propiedad de unión a su pareja de unión natural, la proteína de unión de tallo-bucle (SLBP - también denominada proteína de unión horquilla con histona, o HBP).

De acuerdo con la presente invención, la secuencia de tallo-bucle de histona según el componente (b) de la reivindicación 1 puede no derivarse de una proteína histona de ratón. Más específicamente, la secuencia de tallo-bucle de histona puede no derivarse de un gen de histona de ratón H2A614. También, el ácido nucleico de la invención puede no contener una secuencia tallo-bucle de histona de ratón ni un gen de histona de ratón H2A614. Además, la secuencia de ácido nucleico inventivo puede no contener una señal de procesamiento del tallo-bucle, más específicamente una señal de procesamiento de histona de ratón y, más específicamente, puede no contener una señal de procesamiento de tallo-bucle de ratón H2kA614, incluso si la secuencia de

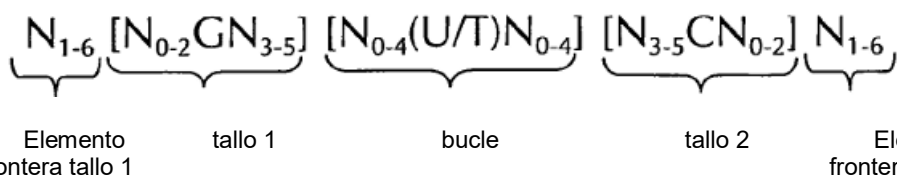
ácido nucleico inventivo puede contener al menos un gen de histona de mamífero. Sin embargo, el al menos un gen de histona de mamífero puede no ser la SEC ID NO: 7 de la WO 01/12824.

La secuencia de ácido nucleico inventivo comprende o codifica para al menos una secuencia tallo-bucle de histona de acuerdo con al menos una de las siguientes fórmulas (I) o (II):

5 fórmula (I) (secuencia tallo-bucle sin elementos frontera de tallo):



fórmula (II) (secuencia tallo-bucle con elementos frontera de tallo):



10

donde

15 los elementos frontera de tallo 1 o tallo 2  $N_{1-6}$  son una secuencia consecutiva de 1 a 6, de preferencia de 2 a 6, con mayor preferencia de 2 a 5, incluso con mayor preferencia de 3 a 5, con máxima preferencia de 4 a 5 o 5N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos;

20 Tallo 1  $[N_{0-2}GN_{3-5}]$  es complementario inverso o complementario parcialmente inverso con el elemento de tallo 2 y es una secuencia consecutiva de entre 5 a 7 nucleótidos; donde  $N_{0-2}$  es una secuencia consecutiva de 0 a 2, de preferencia 0 a 1, con mayor preferencia 1N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; donde  $N_{3-5}$  es una secuencia consecutiva de 3 a 5, de preferencia 4 a 5, con mayor preferencia 4 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; y donde G es guanosina o un análogo de la misma y se puede reemplazar opcionalmente por una citidina o un análogo de la misma, con la condición de que su nucleótido citidina complementario en el tallo 2 se reemplace por guanosina;

30 Secuencia bucle  $[N_{0-4}(U/T)N_{0-4}]$  está ubicada entre los elementos tallo 1 y tallo 2 y es una secuencia consecutiva de 3 a 5 nucleótidos, con mayor preferencia de 4 nucleótidos, donde cada  $N_{0-4}$  es independientemente una secuencia consecutiva de 0 a 4, de preferencia de 1 a 3, con mayor preferencia 1 a 2 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; y donde U/T representa uridina u opcionalmente timidina;

35 Tallo 2  $[N_{3-5}CN_{0-2}]$  es complementaria inversa o complementaria parcialmente inversa con el elemento tallo 1, y es una secuencia consecutiva de 5 a 7 nucleótidos; donde  $N_{3-5}$  es una secuencia consecutiva de 3 a 5, de preferencia 4 a 5, con mayor preferencia 4 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; donde  $N_{0-2}$  es una secuencia consecutiva de 0 a 2, de preferencia 0 a 1, con mayor preferencia 1 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; y donde C es citidina o un análogo de la misma y se puede reemplazar opcionalmente por una guanosina o un análogo de la misma, con la condición de que su nucleótido guanosina complementario en el tallo 1 se reemplace por citidina;

40 donde

el tallo 1 y tallo 2 son capaces de realizar un apareamiento de bases entre sí formando una secuencia complementaria inversa, donde el apareamiento de bases se puede presentar entre el tallo 1 y el tallo 2, por ejemplo un apareamiento de bases Watson-Crick de nucleótidos A y U/T o G y C o un apareamiento de bases no Watson-Crick, por ejemplo un apareamiento de bases por balanceo, un apareamiento de bases Watson-

Crick reverso, un apareamiento de bases Hoogsteen, un apareamiento de bases Hoogsteen reverso o son capaces de un apareamiento de bases uno con otro formando una secuencia parcialmente inversa, donde puese suceder un apareamiento de bases incompleto entre el tallo 1 y el tallo 2, en base a que una o más bases en un tallo no tienen una base complementaria en la secuencia inversa complementaria del otro tallo.

5 En el contexto anterior, un apareamiento de bases de balanceo típicamente es un apareamiento de bases no Watson-Crick entre dos nucleótidos. Los cuatro pares de bases de balanceo principales en el presente contexto que se pueden utilizar son guanosina-uridina, inosina-uridina, inosina-adenosina, inosina-citidina (G-U/T, I-U/T, I-A e I-C) y adenosina-citidina (A-C).

10 Por consiguiente, en el contexto de la presente invención, una base de balanceo es una base que forma un par de bases de balanceo con una base adicional como se ha descrito anteriormente. Así, un apareamiento de bases no Watson-Crick, por ejemplo un apareamiento de bases de balanceo, puede existir en el tallo de la estructura tallo-bucle de histona de acuerdo con la presente invención.

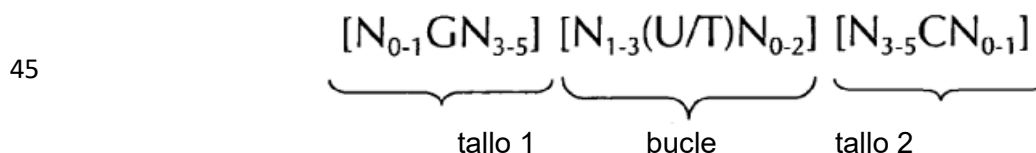
15 En el contexto anterior, una secuencia complementaria parcialmente inversa comprende como máximo 2, de preferencia sólo un mal apareamiento en la estructura tallo de la secuencia tallo-bucle formada por el apareamiento de bases del tallo 1 y el tallo 2. En otras palabras, el tallo 1 y el tallo 2 preferentemente son capaces de formar un apareamiento de bases (completo) entre sí a todo lo largo de la secuencia completa del tallo 1 y el tallo 2 (100% de posibles apareamientos de bases correctos Watson-Crick o no Watson-Crick), formando así una secuencia complementaria inversa, donde cada base tiene su base colgante Watson-Crick o no Watson-Crick como pareja de unión complementaria. Alternativamente, el tallo 1 y el tallo 2 preferentemente son capaces de formar pares de bases parciales entre sí a todo largo de la secuencia del tallo 1 y el tallo 2, donde al menos aproximadamente el 70%, 75%, 80%, 85%, 90%, ó 95% del 100% de los apareamientos de bases correctos posibles Watson-Crick o no Watson-Crick están ocupados con los apareamientos de bases correctos Watson-Crick o no Watson-Crick y como máximo aproximadamente el 30%, 25%, 20%, 15%, 10%, ó 5% de las bases restantes están sin aparear.

25 La al menos una secuencia de tallo-bucle de histona (con los elementos frontera de tallo) de la secuencia de ácido nucleico inventivo como se define aquí comprende una longitud entre aproximadamente 15 a aproximadamente 45 nucleótidos, preferentemente una longitud de aproximadamente 15 a aproximadamente 40 nucleótidos, con preferencia una longitud entre aproximadamente 15 y aproximadamente 35 nucleótidos, incluso con mayor preferencia una longitud de aproximadamente 15 a aproximadamente 30 nucleótidos, con mayor preferencia una longitud entre aproximadamente 20 y aproximadamente 30 y con la máxima preferencia una longitud entre aproximadamente 24 a aproximadamente 28 nucleótidos.

35 La al menos una secuencia de tallo-bucle de histona (sin los elementos frontera de tallo) de la secuencia de ácido nucleico inventivo como se define aquí comprende una longitud entre aproximadamente 10 a aproximadamente 30 nucleótidos, de preferencia una longitud entre aproximadamente 10 a aproximadamente 20 nucleótidos, de preferencia una longitud entre aproximadamente 12 a aproximadamente 20 nucleótidos, con preferencia una longitud entre aproximadamente 14 a aproximadamente 20 nucleótidos, incluso con mayor preferencia una longitud entre aproximadamente 16 y aproximadamente 17 y con la máxima preferencia una longitud de aproximadamente 16 nucleótidos.

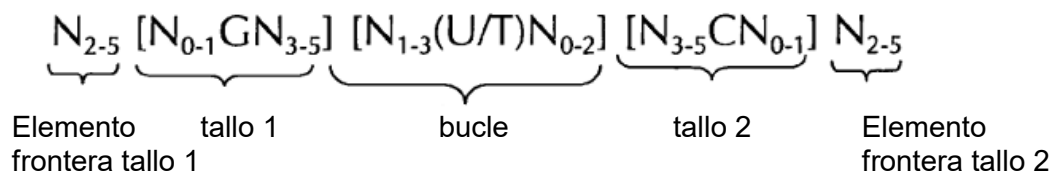
40 De acuerdo con una realización preferida adicional del primer aspecto inventivo, la secuencia de ácido nucleico inventivo según el primer aspecto de la presente invención puede comprender o codificar para al menos una secuencia tallo-bucle de histona de acuerdo con al menos una de las siguientes fórmulas específicas (Ia) o (IIa):

fórmula (Ia) (secuencia tallo-bucle sin elementos frontera de tallo):



fórmula (IIa) (secuencia tallo-bucle con elementos frontera de tallo):

50

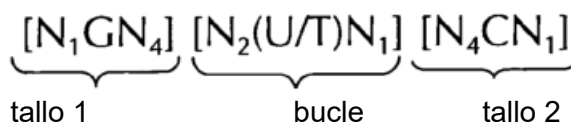


5 donde:

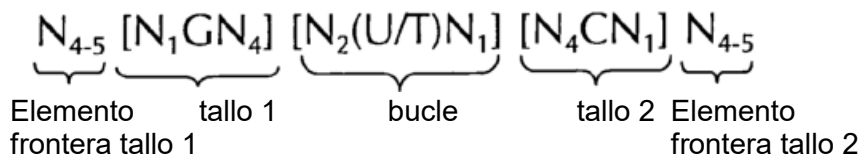
N, C, G, T y U son como se definió anteriormente.

De acuerdo con otra realización preferente de la presente invención, la secuencia de ácido nucleico inventiva puede comprender o codificar para al menos una secuencia tallo-bucle de histona de acuerdo con al menos una de las siguientes fórmulas específicas (Ib) o (IIb):

10 fórmula (Ib) (secuencia de tallo-bucle sin elementos frontera de tallo):



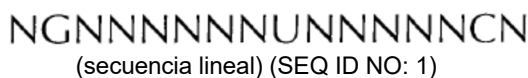
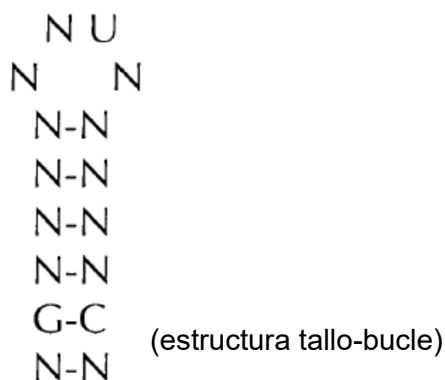
15



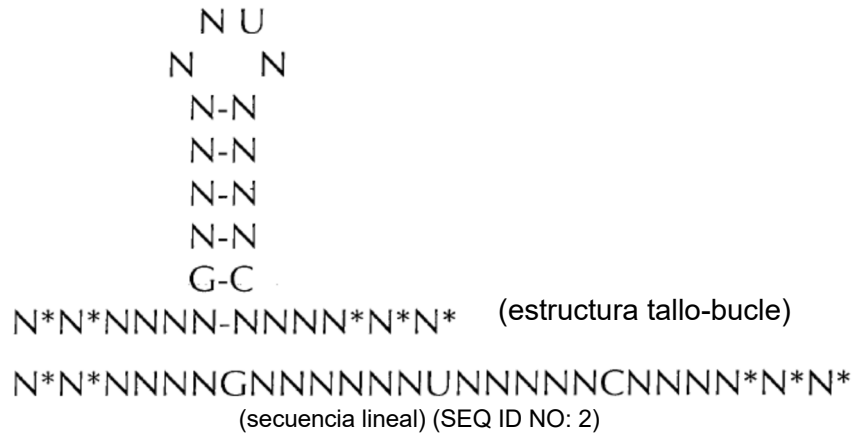
20 De acuerdo con una realización preferente, la secuencia de ácido nucleico inventiva puede comprender o codificar para al menos una secuencia tallo-bucle de histona de acuerdo con al menos una de las siguientes fórmulas específicas (Ic) a (Ih) o (IIc) a (IIh), mostradas alternativamente en su estructura tallo-bucle y como una secuencia lineal que representa las secuencias del tallo-bucle de histona según se generan de acuerdo con

**Ejemplo 1:**

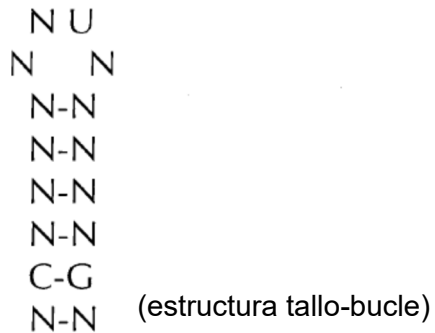
25 Fórmula (Ic): (secuencia consenso tallo-bucle de histona de protozoarios sin elementos frontera de tallo)



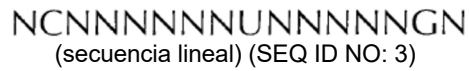
Fórmula (IIc): (secuencia consenso tallo-bucle de histona de metazoarios y protozoarios sin elementos frontera de tallo):



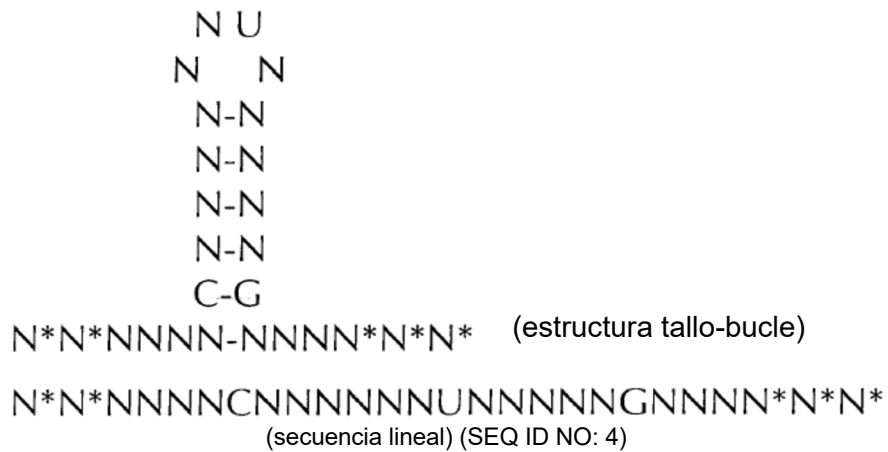
Fórmula (Id): (sin elementos frontera de tallo):



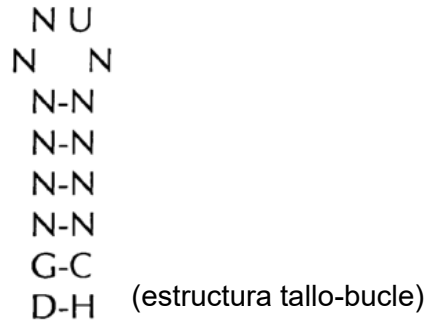
5



Fórmula (IId): (sin elementos frontera de tallo):

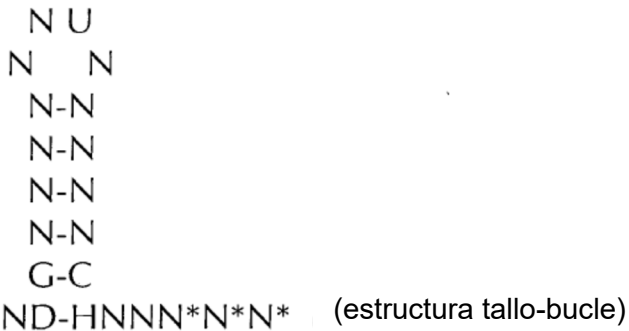


Fórmula (Ie): (secuencia consenso de tallo-bucle de histona de protozoarios sin elementos frontera de tallo):



DGNNNNNNUNNNNNCH  
 (secuencia lineal) (SEQ ID NO: 5)

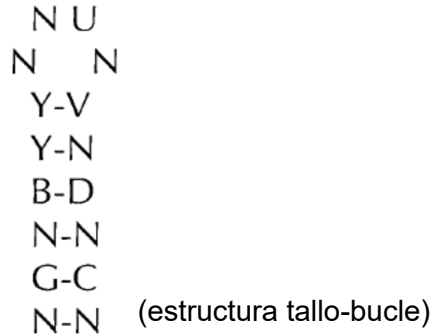
Fórmula (Ile): (secuencia consenso de tallo-bucle de histona de protozoarios con elementos frontera de tallo):



N\*N\*NNNDGNNNNNNUNNNNNCHNNN\*N\*N\*  
 (secuencia lineal) (SEQ ID NO: 6)

5

Fórmula (If): (secuencia consenso de tallo-bucle de histona de metazoarios sin elementos frontera de tallo):



NGNBYNNUNVNDNCN  
 (secuencia lineal) (SEQ ID NO: 7)

Fórmula (If): (secuencia consenso de tallo-bucle de histona de metazoarios con elementos frontera de tallo):

N U  
 N N  
 Y-V  
 Y-N  
 B-D  
 N-N  
 G-C

N\*N\*NNNN-NNNN\*N\*N\* (estructura tallo-bucle)

N\*N\*NNNNGNBYNNUNVNDNCNNNN\*N\*N\*  
 (secuencia lineal) (SEQ ID NO: 8)

Fórmula (Ig): (secuencia consenso de tallo-bucle de histona de vertebrado sin elementos frontera de tallo):

N U  
 D H  
 Y-A  
 Y-B  
 Y-R  
 H-D  
 G-C  
 N-N (estructura tallo-bucle)

NGHYYYDNUHABRDCN  
 (secuencia lineal) (SEQ ID NO: 9)

5

Fórmula (IIg): (secuencia consenso de tallo-bucle de histona de vertebrado con elementos frontera de tallo):

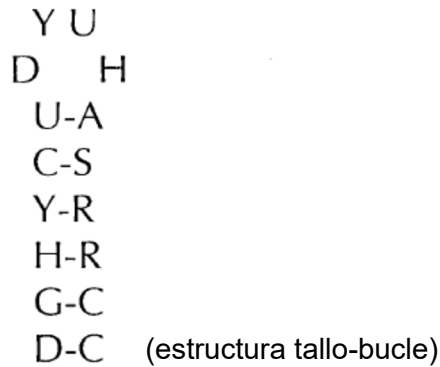
N U  
 D H  
 Y-A  
 Y-B  
 Y-R  
 H-D  
 G-C

N\*N\*HNNN-NNNN\*N\*H\* (estructura tallo-bucle)

N\*N\*HNNNGHYYYDNUHABRDCNNNN\*N\*H\*  
 (secuencia lineal) (SEQ ID NO: 10)

10 Fórmula (Ih): (secuencia consenso de tallo-bucle de histona humana (Homo sapiens) sin elementos frontera de tallo):





DGHYCUDYUHASRRCC  
(secuencia lineal) (SEQ ID NO: 11)

Fórmula (IIh): (secuencia consenso de tallo-bucle de histona humana (Homo sapiens) con elementos frontera de tallo):



N\*H\*AAHD-CVHB\*N\*H\* (estructura tallo-bucle)

5

N\*H\*AAHDGHYCUDYUHASRRCCVHB\*N\*H\*  
(secuencia lineal) (SEQ ID NO: 12)

10 donde en cada una de las fórmulas anteriores (Ic) a (Ih) o (IIc) a (IIh): N, C, G, A, T y U son como se definió anteriormente; cada U se puede remplazar por T; cada G o C (altamente) conservada en los elementos de tallo 1 y tallo 2 se puede reemplazar por su base complementaria de nucleótidos C o G, con la condición de que su nucleótido complementario en el tallo correspondiente se remplace por su nucleótido complementario en paralelo; y/o G, A, T, U, C, R, Y, H, K, S, W, H, B, V, D y N son bases de nucleótidos como se define en la siguiente Tabla:

Abreviatura	Bases de nucleótidos	Comentarios
G	G	Guanina
A	A	Adenina
T	T	Tiamina
U	U	Uracilo
C	C	Citosina
R	G o A	Purina
Y	T/U o C	Pirimidina
M	A o C	Amino
K	G o T/U	Ceto
S	G o C	Fuerte (enlaces 3H)
W	A o T/U	Débil (enlace 2H)
H	A o C o T/U	Sin G
B	G o T/U o C	Sin A
V	G o C o A	Sin T/U
D	G o A o T/U	Sin C
N	G o C o T/U or A	Cualquier base
*	Presente o no	Base puede estar presente o no

La secuencia de tallo-bucle de histona de acuerdo con al menos una de las fórmulas (I) o (Ia) a (Ih) o (II) o (IIa) a (IIh) de la presente invención preferentemente puede seleccionarse de una secuencia tallo-bucle de histona de origen natural, en particular se prefiere en mayor medida que sea a partir de secuencias tallo-bucle de histonas de protozoarios o metazoarios, e incluso con particular preferencia de vertebrados y con total preferencia de secuencias tallo-bucle de histonas de mamífero, en especial provenientes de secuencias tallo-bucle de histonas humanas.

La secuencia del tallo-bucle de histona de acuerdo con al menos una de las fórmulas específicas (I) o (Ia) a (Ih) o (II) o (IIa) a (IIh) de la presente invención preferentemente puede ser una secuencia tallo-bucle de histona que comprende en cada posición de nucleótidos el nucleótido que se presenta con mayor frecuencia o cualquiera del nucleótido que se presenta con mayor frecuencia o el segundo con mayor frecuencia de las secuencias tallo-bucle de histonas de origen natural en metazoarios y protozoarios (Fig. 1), protozoarios (Fig. 2), metazoarios (Fig. 3), vertebrados (Fig. 4) y humanos (Fig. 5) como se muestra en las figuras 1-5. En este contexto, se prefiere particularmente que al menos el 80%, preferentemente al menos el 85%, o con mayor preferencia al menos el 90% de todos los nucleótidos correspondan al nucleótido que se presenta con mayor frecuencia en las secuencias tallo-bucle de histona de origen natural.

En particular, la secuencia de tallo-bucle de histona de acuerdo con al menos una de las fórmulas específicas (I) o (Ia) a (Ih) de la presente invención se puede seleccionar de las siguientes secuencias tallo-bucle de histona (sin elementos frontera de tallo) que representan las secuencias tallo-bucle de histona según se generan de acuerdo con el Ejemplo 1:

- 20 VGYYYYHHHRVRCB (SEC ID NO: 13 de acuerdo con la fórmula (Ic))
- SGYYTTTYTMARRRCS (SEC ID NO: 14 de acuerdo con la fórmula (Ic))
- SGYYCTTTMAGRRCS (SEC ID NO: 15 de acuerdo con la fórmula (Ic))
- DGNBBNNTHTVNNCH (SEC ID NO: 16 de acuerdo con la fórmula (Ie))
- RGNNYHBTDRDNCY (SEC ID NO: 17 de acuerdo con la fórmula (Ie))
- 25 RGNDYHYTHRDHNCY (SEC ID NO: 18 de acuerdo con la fórmula (Ie))
- VGYYYTYHHRVRCB (SEC ID NO: 19 de acuerdo con la fórmula (If))
- SGYYCTTYTMAGRRCS (SEC ID NO: 20 de acuerdo con la fórmula (If))
- SGYYCTTTMAGRRCS (SEC ID NO: 21 de acuerdo con la fórmula (If))
- GGYYCTTYTHAGRRCC (SEC ID NO: 22 de acuerdo con la fórmula (Ig))
- 30 GGCYCTTYTMAGRGCC (SEC ID NO: 23 de acuerdo con la fórmula (Ig))
- GGCTCTTTMAGRGCC (SEC ID NO: 24 de acuerdo con la fórmula (Ig))
- DGHYCTDYTHARRCC (SEC ID NO: 25 de acuerdo con la fórmula (Ih))
- GGCYCTTTTHAGRGCC (SEC ID NO: 26 de acuerdo con la fórmula (Ih))
- GGCYCTTTMAGRGCC (SEC ID NO: 27 de acuerdo con la fórmula (Ih))

35 Además, en este contexto, las siguientes secuencias tallo-bucle de histona (con elementos frontera de tallo) según se generan de acuerdo con el Ejemplo 1 de acuerdo con una de las fórmulas específicas (II) o (IIa) a (IIh) son particularmente preferentes:

- H\*H\*HHVVGYYYYHHHRVRCBVHH\*N\*N\* (SEQ ID NO: 28 según fórmula (IIc))
- M\*H\*MHMSGYYTTYTMARRRCSMCH\*H\*H\* (SEQ ID NO: 29 según fórmula (IIc))
- 40 M\*M\*MMMSGYYCTTTMAGRRCSACH\*M\*H\* (SEQ ID NO: 30 según fórmula (IIc))
- N\*N\*NNNDGNBBNNTHTVNNCHNH\*N\*N\* (SEQ ID NO: 31 según fórmula (IIe))
- N\*N\*HHNRGNNYHBTDRDNCYDHH\*N\*N\* (SEQ ID NO: 32 según fórmula (IIe))
- N\*H\*HHVRGNDYHYTHRDHNCYRHH\*H\*H\* (SEQ ID NO: 33 según fórmula (IIe))
- H\*H\*MHMVGYYTYHHRVRCBVMH\*H\*N\* (SEQ ID NO: 34 según fórmula (IIf))
- 45 M\*M\*MMMSGYYCTTYTMAGRRCSMCH\*H\*H\* (SEQ ID NO: 35 según fórmula (IIf))
- M\*M\*MMMSGYYCTTTMAGRRCSACH\*M\*H\* (SEQ ID NO: 36 según fórmula (IIf))
- H\*H\*MAMGGYYCTTYTHAGRRCCVHN\*N\*M\* (SEQ ID NO: 37 según fórmula (IIg))
- H\*H\*AAMGGCYCTTYTMAGRGCCVCH\*H\*M\* (SEQ ID NO: 38 según fórmula (IIg))
- M\*M\*AAMGGCTCTTTMAGRGCCMCY\*M\*M\* (SEQ ID NO: 39 según fórmula (IIg))
- 50 N\*H\*AAHDGHYCTDYTHARRCCVHB\*N\*H\* (SEQ ID NO: 40 según fórmula (IIh))
- H\*H\*AAMGGCYCTTTTHAGRGCCVMY\*N\*M\* (SEQ ID NO: 41 según fórmula (IIh))
- H\*M\*AAAGGCYCTTTMAGRGCCRM\*H\*M\* (SEQ ID NO: 42 según fórmula (IIh))

La secuencia del ácido nucleico inventivo preferentemente puede comprender o codificar para al menos una secuencia tallo-bucle de histona que muestra al menos aproximadamente el 80%, de preferencia al menos aproximadamente 85%, con mayor preferencia al menos aproximadamente 90%, o incluso con mayor preferencia al menos aproximadamente un 95% de identidad de secuencia con los nucleótidos conservados que no están al 100% en las secuencias del tallo-bucle de histona de acuerdo con al menos una de las fórmulas específicas (I) o (Ia) a (Ih) o (II) o (IIa) a (IIh) o con una secuencia tallo-bucle de histona de origen natural.

La secuencia de tallo-bucle de histona preferentemente puede no contener la secuencia bucle 5'-UUUC-3'. Más específicamente, el tallo-bucle de histona no contiene la secuencia tallo 1 5'-GGCUCU-3' y/o la secuencia tallo 2 5'-AGAGCC-3', respectivamente. O la secuencia de tallo-bucle puede preferentemente no contener la secuencia bucle 5'-CCUGCCC-3' o la secuencia bucle 5'-UGAAU-3'. Más específicamente, el tallo-bucle no contiene la secuencia tallo 1 5'-CCUGAGC-3' o no contiene la secuencia tallo 1 5'-ACCUUUCUCCA-3' y/o la secuencia tallo 2 5'-GCUCAGG-3' o 5'-UGGAGAAAGGU-3', respectivamente. También, las secuencias tallo-bucle de histona preferentemente pueden no derivarse de una región 3'-no traducida del receptor de insulina de mamífero. También, preferentemente, el ácido nucleico inventivo puede no contener señales de procesamiento de tallo-bucle de histona, en particular no aquellas derivadas del gen de histona de ratón H2A614 (H2ka614).

La secuencia de ácido nucleico inventivo de acuerdo con el primer aspecto de la presente invención comprende o codificar para una secuencia poli(A). Preferentemente, esta secuencia poli(A) comprende una secuencia de aproximadamente 30 o con mayor preferencia de aproximadamente 25 a aproximadamente 400 nucleótidos de adenosina, preferentemente una secuencia de aproximadamente 50 a aproximadamente 400 nucleótidos de adenosina, con mayor preferencia una secuencia de aproximadamente 50 a aproximadamente 300 nucleótidos de adenosina, incluso con mayor preferencia una secuencia de aproximadamente 50 a aproximadamente 250 nucleótidos de adenosina, con especial preferencia una secuencia de aproximadamente 60 a aproximadamente 250 nucleótidos de adenosina. En este contexto, el término "aproximadamente" se refiere a una desviación de  $\pm 10\%$  del valor (valores) que se alcanza. Por consiguiente, la secuencia poli(A) contiene al menos 25 o más de 25, con mayor preferencia al menos 30, con mayor preferencia al menos 50 nucleótidos de adenosina. Por tanto, esta secuencia poli(A) no contiene típicamente menos de 20 nucleótidos de adenosina. De manera más particular, no contiene 10 y/o menos de 10 nucleótidos de adenosina.

De preferencia, el ácido nucleico de acuerdo de la presente invención no contiene uno o dos o al menos uno o la totalidad pero uno o la totalidad de los componentes del grupo consistente en: una secuencia que codifica para una ribozima (de preferencia una ribozima de auto-empalme), una secuencia de ácido nucleico viral, una señal de procesamiento de tallo-bucle de histona, en particular una secuencia de procesamiento de tallo-bucle de histona derivada del gen de histona H2A614 de ratón, un gen neo, una secuencia promotora inactivada y una secuencia intensificadora inactivada. Incluso con mayor preferencia, el ácido nucleico de acuerdo con la invención no contiene una ribozima, de preferencia una ribozima de auto-empalme, y una del grupo consistente en: un gen neo, una secuencia promotora inactivada, una secuencia intensificadora inactivada, una señal de procesamiento de tallo-bucle de histona, en particular una secuencia de procesamiento de tallo-bucle de histona derivada del gen de histona H2A614 de ratón. Por consiguiente, en una forma preferente, el ácido nucleico puede no contener una ribozima, de preferencia una ribozima de auto-empalme, ni un gen neo o, alternativamente, ni una ribozima, de preferencia una ribozima de auto-empalme, ni cualquier gen de resistencia (por ejemplo usualmente aplicado para selección). En otra forma preferida, el ácido nucleico de la invención puede no contener una ribozima, de preferencia una ribozima de auto-empalme, ni una señal de procesamiento de tallo-bucle de histona, en particular una secuencia de procesamiento de tallo-bucle de histona derivada del gen de histona H2A614 de ratón.

La secuencia de ácido nucleico inventivo puede comprender opcionalmente una señal de poliadenilación que se define aquí como una señal que transporta la poliadenilación hacia un ARNm (transcrito) mediante factores proteicos específicos (por ejemplo un factor de especificidad para segmentación y poliadenilación (CPSF), un factor de estimulación de la segmentación (CstF), los factores de segmentación I y II (CF I y CF II), la polimerasa poli(A) (PAP)). En este contexto es preferente una señal poliadenilación consenso que comprenda la secuencia consenso NN(U/T)ANA. Preferentemente, la señal de poliadenilación puede comprender una de las siguientes secuencias: AA(U/T)AAA o A(U/T)(U/T)AAA (donde está presente usualmente uridina en el ARN y está presente usualmente timidina en el ADN). La señal de poliadenilación utilizada en el ácido nucleico inventivo no corresponde al U3 ARNsn, U5, la señal de procesamiento de poliadenilación proveniente del gen humano G-CSF o las secuencias de la señal de poliadenilación SV40. En particular, las señales de poliadenilación anteriores no están combinadas con ningún gen de resistencia a antibióticos (o cualquier otro gen reportero, marcador o de selección), en particular no con el gen *neo* de resistencia (neomicina fosfotransferasa) (como el gen de la región codificante de acuerdo con elemento (a) del ácido nucleico inventivo. Y cualquiera de las señales de poliadenilación anteriores (que típicamente no están presentes en el ácido nucleico inventivo) de preferencia no están combinadas con el tallo-bucle de histona o la señal de procesamiento del tallo-bucle de histona proveniente del gen de histona de ratón H2A614 en un ácido nucleico inventivo.

La secuencia del ácido nucleico inventivo comprende una región codificante que codifica para al menos un péptido o proteína, comprendiendo dicho péptido o proteína un antígeno patógeno o un fragmento del mismo con una longitud de residuos aminoácidos y teniendo la propiedad antigénica específica del péptido o proteína nativo de longitud completa, estando el antígeno patógeno asociado con una infección bacteriana, una infección vital o una infección por protozoos y seleccionándose el antígeno patógeno del grupo consistente en

- Antígeno p24 de VIH, proteínas de envoltura de VIH (Gp120, Gp41, Gp160), poliproteína GAG, proteína del factor negativo Nef, trans-activador de transcripción Tat si la enfermedad infecciosa en VIH, preferentemente una infección con el virus de inmunodeficiencia humana;
- 5 • proteína cápside C, proteína premembrana prM, proteína membrana M, proteína de envoltura E (dominio I, dominio II, dominio III), proteína NS1, proteína NS2A, proteína NS2B, proteína NS3, proteína NS4A, proteína 2K, proteína NS4B, proteína NS5, si la enfermedad infecciosa es fiebre del Dengue, preferentemente una infección con el virus del Dengue (DEN-1, DEN-2, DEN-3 y DEN-4)—Flavivirus;
- 10 • antígeno superficial de hepatitis B HBsAg, antígeno núcleo de Hepatitis B HbcAg polimerasa, proteína Hbx, proteína superficial media preS2, proteína superficial L, proteína larga S, proteína viral VP1, proteína viral VP2, proteína viral VP3, proteína viral VP4, si la enfermedad infecciosa es Hepatitis B, preferentemente una infección con el virus de la Hepatitis B (HBV);
- Hemaglutinina (HA), Neuraminidasa (NA), Nucleoproteína (NP), proteína M1, proteína M2, proteína NS1, proteína NS2 (proteína NEP: proteína de exportación nuclear), proteína PA, proteína PB1 (proteína de polimerasa básica 1), proteína PB1-F2 y proteína PB2 (familia Ortomixoviridae, virus de la gripe (flu));
- 15 • nucleoproteína N, proteína estructural grande L, fosfoproteína P, proteína de matriz M, glucoproteína G, si la enfermedad infecciosa es la rabia, preferentemente una infección con virus de la rabia;
- poliproteína genómica, proteína E, proteína M, proteína C de cápside, proteasa NS3, proteína NS1, proteína NS2A, proteína AS2B, proteína NS4A, proteína NS4B, proteína NS5, si la enfermedad infecciosa es fiebre amarilla, preferentemente una infección por el virus de la fiebre amarilla;
- 20 • proteína TRAP anónima relacionada con trombospondina, proteína superficial de SSP2 Sporozoite 2, antígeno de membrana apical 1 AMA1, antígeno de membrana rhoptry RMA1, antígeno de repetición básica ácida ABRA, proteína transversal celular PF, proteína Pvs25, proteína superficial de merozoito 1 MSP-1, proteína superficial de merozoito 2 MSP-2, antígeno superficial de eritrocitos anillo-infectados antígeno de etapa RESALiver 3 LSA-3, proteína Eba-175, antígeno 5 para repetición de serina SERA-5, proteína de circumsporozoitos CS, proteína 3 superficial de merozoito MSP3, proteína superficial 8 de merozoitos MSP8, enolasa PF10, proteína de hepatocitos-eritrocitos de 17 kDa HEP17, proteína con membrana 1 de eritrocitos EMP1, proteína superficial 4/5 de la proteína Kbetamerozoitos MSP 4/5, proteína de choque térmico Hsp90, proteína rica en glutamato GLURP, proteína superficial 4 de merozoitos MSP-4, proteína STARP, precursor antigénico relacionado con la proteína de circumsporozoitos CRA (Plasmodium genus, Malaria);
- 30 • proteína de fusión F, hemaglutinina-neuramidasa HN, glucoproteína G, proteína matriz M, fosfoproteína P, nucleoproteína N, polimerasa L, si la enfermedad infecciosa es la infección por el virus de la parainfluenza humana, de preferencia una infección con los virus de parainfluenza Humanos (HPIV);
- proteína de fusión F, nucleoproteína N, proteína de matriz M, proteína de matriz M2-1, proteína de matriz M2-2, fosfoproteína P, proteína hidrofóbica pequeña SH, glucoproteína G de superficie mayor, polimerasa L, proteína 1 no estructural NS1, proteína 2 no estructural NS2, si la enfermedad infecciosa es una infección por el virus sincitial respiratorio, preferentemente una infección por el virus sincitial respiratorio (RSV);
- 35 • glucoproteína L UL1, uracil-ADN glucosilasa UL2, proteína UL3, proteína UL4, proteína ADN de replicación UL5, proteína portal UL6, proteína para maduración de viriones UL7, ADN helicasa UL8, proteína de unión al origen de replicación UL9, glucoproteína M UL10, proteína UL11, exonucleasa alcalina UL12, proteína quinasa de serina-treonina UL13, proteína tegumentaria UL14, terminasa UL15, proteína tegumentaria UL16, proteína UL17, proteína con cápside VP23 UL18, proteína con cápside principal VP5 UL19, proteína con membrana UL20, proteína tegumentaria UL21, Glucoproteína H (UL22), Timidina quinasa UL23, proteína UL24, proteína UL25, proteína con cápside P40 (UL26, VP24, VP22A), glucoproteína B (UL27), proteína ICP18.5 (UL28), proteína principal de unión a ADN ICP8 (UL29), ADN polimerasa UL30, proteína de matriz nuclear UL31, glucoproteína con envoltura UL32, proteína UL33, proteína con membrana nuclear interna UL34, proteína con cápside VP26 (UL35), proteína tegumentaria larga UL36, proteína de ensamble cápside UL37, proteína VP19C (UL38), ribonucleótido reductasa (subunidad Largo) UL39, ribonucleótido reductasa (subunidad pequeña) UL40, proteína tegumentaria/proteína VHS para interrupción en el huésped de viriones (UL41), factor de procesamiento de ADN polimerasa UL42, proteína con membrana UL43, glucoproteína C (UL44), proteína con membrana UL45, proteína tegumentarias VP11/12 (UL46), proteína tegumentaria VP13/14 (UL47), proteína para maduración de viriones VP16 (UL48, Alfa-TIF), proteína con envoltura UL49, dUTP difosfatasa UL50, proteína tegumentaria UL51, proteína compleja de ADN helicasa/primasa UL52, glucoproteína K (UL53), proteína para regulación transcripcional IE63 (ICP27, UL54), proteína UL55, proteína UL56, proteína de replicación viral ICP22 (IE68, US1), proteína US2, proteína quinasa de serina/treonina US3, glucoproteína G (US4), glucoproteína J (US5), glucoproteína D (US6), glucoproteína I (US7), glucoproteína E (US8), proteína tegumentaria US9, proteína con cápside/tegumentaria US10, proteína Vmw21 (US11), proteína ICP47 (IE12, US12), activador transcriptional principal ICP4 (IE175, RS1), E3 ubiquitina ligasa ICP0 (IE110), proteína 1 latencia-relacionada LRP1, proteína 2 latencia-relacionada LRP2, factor de neurovirulencia RL1 (ICP34.5), transcripción LAT latencia-asociada (Virus de herpes simplex 1 y 2 (HSV-1 y HSV-2), Herpes simplex);
- 40
- 45
- 50
- 55
- 60

- proteína de replicación E1, proteína reguladora E2, proteína E3, proteína E4, proteína E5, proteína E6, proteína E7, proteína E8, proteína con cápside mayor L1, proteína con cápside menor L2 (papilomavirus humano (HPV), infección por papilomavirus humano (HPV));
- 5 • antígeno pp65, proteína de membrana pp15, proteína tegumental cápside-proximal pp150, proteína M45, ADN polimerasa UL54, helicasa UL105, glucoproteína gM, glucoproteína gN, glucoproteína H, glucoproteína B gB, proteína UL83, proteína UL94, proteína UL99 (Citomegalovirus (CMV), infección por Citomegalovirus);
- 10 • proteína IsdB de unión hemo-hierro, colágeno adhesina Cna, factor A de aterronamiento CifA, proteína MecA, proteína de unión A fibronectina FnbA, enterotoxina tipo A EntA, enterotoxina tipo B EntB, enterotoxina tipo C EntC1, enterotoxina tipo C EntC2, enterotoxina tipo D EntD, enterotoxina tipo E EntE, toxina-1 de síndrome de choque tóxico TSST-1, Stafiloquinasa, proteína 2a de unión penicilina PBP2a (MecA), antígeno secretor SssA (Staphylococcus genus, por ejemplo aureus, infección por estafilococos);
- 15 • poliproteína genómica, polimerasa 3D, proteína con cápside viral VP1, proteína con cápside viral VP2, proteína con cápside viral VP3, proteína con cápside viral VP4, proteasa 2A, proteasa 3C (Rhinovirus, infección por Rhinovirus);
- 20 • proteína ESAT-6 (EsxA), antígeno filtrado EsxB de 10 kDa, antígeno secretado 85-B FBPB, proteína A de unión a fibronectina FbpA (Ag85A), serina proteasa PepA, proteína PPE18 de la familia PPE, proteína D de unión a fibronectina FbpD, proteína inmunogénica MPT64, proteína secretada MPT51, catalasa-peroxidasa-peroxinitritasa T KATG, lipoproteína PSTS3 de unión a fosfato periplásmico (PBP-3, Phos-1), hemaglutinina Hbha de unión a heparina hierro-regulada, proteína PPE14 de la familia PPE, proteína PPE68 de la familia PPE, proteína Mtb72F, proteína Apa, proteína inmunogénica MPT63, lipoproteína PSTS1 de unión a fosfato periplásmico (PBP-1), chaperón molecular DnaK, lipoproteína Mpt83 de superficie celular, lipoproteína P23, proteína permeasa pstA del sistema de transporte de fosfato, antígeno de 14 kDa, proteína C de unión a fibronectina FbpC1, alanina deshidrogenasa TB43, glutamina sintetasa 1, proteína ESX-1, proteína CFP10,
- 25 • proteína TB10.4, proteína MPT83, proteína MTB12, proteína MTB8, proteínas similares a Rpf, proteína MTB32, proteína MTB39, cristalina, proteína de choque térmico HSP65, proteína PST-S (usualmente tuberculosis Mycobacterium, Tuberculosis);
- 30 • nucleoproteína N, polimerasa L, proteína de matriz Z, glucoproteína GP (Lassa virus, fiebre de Lassa);
- proteínas spike S, proteínas con envoltura E, proteínas de membrana M, proteínas nucleocápside N (por lo general rinovirus y coronavirus);
- glucoproteína GP, nucleoproteína NP, proteína matriz menor VP24, proteína matriz mayor VP40, activador de transcripción VP30, cofactor de polimerasa VP35, polimerasa L de ARN (Ebolavirus (EBOV), fiebre hemorrágica de Ebola);
- 35 • proteína con envoltura Gc, proteína con envoltura Gn, proteínas nucleocápside (virus de fiebre hemorrágica Crimean-Congo, fiebre hemorrágica Crimean-Congo (CCHF));
- glucoproteína G, proteína matriz M, nucleoproteína N, proteína de fusión F, polimerasa L, proteína W, proteína C, fosfoproteína P, proteína no estructural V (virus Henipavirus (Hendra virus Nipah), infecciones Henipavirus);
- 40 • glucoproteína GP con envoltura, polimerasa L, nucleoproteína N, proteína NSS no estructural (virus de fiebre Rift Valley, fiebre de Rift Valley (RVF)).
- proteína 6 no estructural NS6, proteína 2 no estructural NS2, proteína VP6 con cápside intermedia, proteína VP2 con cápside interna, proteína 3 no estructural NS3, polimerasa L de ARN-ARN-dirigida, proteína VP3, proteína 1 no estructural NS1, proteína 5 no estructural NS5, glucoproteína VP7 con cápside extra, glucoproteína 4 no estructural NS4, proteína VP4 con cápside externa (Rotavirus, infección por Rotavirus);
- 45 • proteína NS1, proteína hidrofóbica pequeña NS2, proteína SH, proteína de fusión F, glucoproteína G, proteína de matriz M, proteína de matriz M2-1, proteína de matriz M2-2, fosfoproteína P, nucleoproteína N, polimerasa L (metapneumovirus Humano (hMPV), infección por metapneumovirus Humano).

50 Más en general, la secuencia de ácido nucleico inventivo puede codificar para una proteína o un péptido que comprende un antígeno patogénico o un fragmento del mismo. Estos antígenos patogénicos se derivan de organismos patogénicos, en particular organismos patogénicos bacterianos, virales o protozoológicos (multicelulares), que evocan una reacción inmunológica en el sujeto, en particular un sujeto mamífero, más particularmente un ser humano. Más específicamente, los antígenos patogénicos preferentemente son antígenos superficiales, por ejemplo proteínas (o fragmentos de proteínas, por ejemplo, la porción exterior de un antígeno superficial) ubicadas en la superficie del virus o el organismo bacteriano o protozoológico.

60 Un gran número de antígenos patogénicos aquí descritos son antígenos peptídicos o proteínicos, que pueden derivarse de un patógeno asociado con una enfermedad infecciosa seleccionada de antígenos derivados de los patógenos *Acinetobacter baumannii*, *Anaplasma* genus, *Anaplasma phagocytophilum*, *Ancylostoma braziliense*, *Ancylostoma duodenale*, *Arcanobacterium haemolyticum*, *Ascaris lumbricoides*, *Aspergillus* genus, *Astroviridae*, *Babesia* genus, *Bacillus anthracis*, *Bacillus cereus*, *Bartonella henselae*, BK virus, *Blastocystis hominis*, *Blastomyces dermatitidis*, *Bordetella pertussis*, *Borrelia burgdorferi*, *Borrelia* genus, *Borrelia* spp, *Brucella* genus, *Brugia malayi*, *Bunyaviridae* family, *Burkholderia* *capacia* y otras especies *Burkholderia*, *Burkholderia mallei*, *Burkholderia pseudomallei*, familia *Caliciviridae*, *Campylobacter* genus, *Candida albicans*,

Candida spp, Chlamydia trachomatis, Chlamydomydia pneumoniae, Chlamydomydia psittaci, CJD prion, Clonorchis sinensis, Clostridium botulinum, Clostridium difficile, Clostridium perfringens, Clostridium perfringens, Clostridium spp, Clostridium tetani, Coccidioides spp, coronaviruses, Corynebacterium diphtheriae, Coxiella burnetii, virus de fiebre hemorrágica de Crimean-Congo, Cryptococcus neoformans, Cryptosporidium genus, Citomegalovirus (CMV), virus del dengue (DEN-1, DEN-2, DEN-3 y DEN-4), Dientamoeba fragilis, Ebolavirus (EBOV), Echinococcus genus, Ehrlichia chaffeensis, Ehrlichia ewingii, Ehrlichia genus, Entamoeba histolytica, Enterococcus genus, Enterovirus genus, Enteroviruses, principalmente un virus Coxsackie A y Enterovirus 71 (EV71), Epidermophyton spp, Epstein-Barr Virus (EBV), Escherichia coli O157:H7, O111 y O104:H4, Fasciola hepatica y Fasciola gigantica, FFI prion, Filarioidea superfamily, Flaviviruses, Francisella tularensis, Fusobacterium genus, Geotrichum candidum, Giardia intestinalis, Gnathostoma spp, GSS prion, Guanarito virus, Haemophilus ducreyi, Haemophilus influenzae, Helicobacter pylori, Henipavirus (Hendra virus Nipah virus), virus de Hepatitis A, virus de Hepatitis B (HBV), virus de Hepatitis C (HCV), virus de Hepatitis D, virus de Hepatitis E, virus de Herpes simplex 1 y 2 (HSV-1 y HSV-2), Histoplasma capsulatum, VIH (virus de inmunodeficiencia humana), Hortaea werneckii, Human bocavirus (HBoV), virus de herpes humana 6 (HHV-6) y virus de herpes humana 7 (HHV-7), Human metapneumovirus (hMPV), virus de papiloma humana (HPV), virus de parainfluenza humana (HPIV), virus encefalitis japonesa, virus JC, virus Junin, Kingella kingae, Klebsiella granulomatis, Kuru prion, virus Lassa, Legionella pneumophila, Leishmania genus, Leptospira genus, Listeria monocytogenes, virus Lymphocytic choriomeningitis (LCMV), virus Machupo, Malassezia spp, virus Marburg, virus Measles, Metagonimus yokagawai, Microsporidia phylum, virus Molluscum contagiosum (MCV), virus Mumps, Mycobacterium leprae y Mycobacterium lepromatosis, tuberculosis Mycobacterium Mycobacterium ulcerans, Mycoplasma pneumoniae, Naegleria fowleri, Necator americanus, Neisseria gonorrhoeae, Neisseria meningitidis, Nocardia asteroides, Nocardia spp, Onchocerca volvulus, Orientia tsutsugamushi, Orthomyxoviridae family (Influenza), Paracoccidioides brasiliensis, Paragonimus spp, Paragonimus westermani, Parvovirus B19, Pasteurella genus, Plasmodium genus, Pneumocystis jirovecii, Poliovirus, virus de la rabia, virus sincitial respiratorio (RSV), Rhinovirus, rinoviruses, Rickettsia akari, Rickettsia genus, Rickettsia prowazekii, Rickettsia rickettsii, Rickettsia typhi, Rift Valley fever virus, Rotavirus, Rubella virus, Sabia virus, Salmonella genus, Sarcoptes scabiei, SARS coronavirus, Schistosoma genus, Shigella genus, Sin Nombre virus, Hantavirus, Sporothrix schenckii, Staphylococcus genus, Staphylococcus genus, Streptococcus agalactiae, Streptococcus pneumoniae, Streptococcus pyogenes, Strongyloides stercoralis, Taenia genus, Taenia solium, virus de encefalitis transmitida por garrapatas (TBEV), Toxocara canis or Toxocara cati, Toxoplasma gondii, Treponema pallidum, Trichinella spiralis, Trichomonas vaginalis, Trichophyton spp, Trichuris trichiura, Trypanosoma brucei, Trypanosoma cruzi, Ureaplasma urealyticum, virus de varicela zoster (VZV), virus de varicela zoster (VZV), Viruela elevada o viruela mínima, vCJD prion, virus de encefalitis equina Venezolana, Vibrio cholerae, virus de Nilo Occidental, virus de encefalitis equina Occidental, Wuchereria bancrofti, virus de fiebre amarilla, Yersinia enterocolitica, Yersinia pestis y Yersinia pseudotuberculosis.

Además, el antígeno patogénico (antígeno derivado de un patógeno asociado con una enfermedad infecciosa) aquí descrito se puede seleccionar de los siguientes antígenos: la proteína de membrana externa OmpA, la biopelícula asociada con la proteína Bap, la proteína de transporte Muck (infecciones por Acinetobacter baumannii, Acinetobacter); la glucoproteína VSG superficial variable, la proteínas MAPP15 microtubular asociada, trans-sialidasa TSA (Trypanosoma brucei, enfermedad del sueño africano (African trypanosomiasis)); Antígeno p24 del VIH, las proteínas con envoltura del VIH (gp120, gp41, gp160), la poliproteína GAG, la proteína del factor negativo Nef, el trans-activador Tat de transcripción (virus de la inmunodeficiencia humana VIH), SIDA (síndrome de inmunodeficiencia adquirida); la proteína de adherencia de galactosa-inhibitable GIAP, antígeno Eh29 de 29 kDa, lectina Gal/GalNAc, proteína CRT, antígeno inmunodominante de 125 kDa, proteína M17, adhesina ADH112, proteína STIRP (Entamoeba histolytica, Amoebiasis); proteínas superficiales mayores 1-5 (MSP1a, MSP1b, MSP2, MSP3, MSP4, MSP5), las proteínas del sistema de secreción tipo IV (VirB2, VirB7, VirB11, VirD4) (Anaplasma genus, anaplasmosis); antígeno protector PA, factor de edema EF, factor letal LF, proteínas de homología con capa S SLH (Bacillus anthracis, ántrax); acranolisina, fosfolipasa D, proteína de unión a colágeno CbpA (infección por Arcanobacterium haemolyticum, Arcanobacterium haemolyticum); proteína nucleocápside NP, precursor de glucoproteína GPC, glucoproteína GP1, glucoproteína GP2 (virus Junin, fiebre hemorrágica argentina); proteínas con capa de proteína-quitina, antígeno superficial de 14 kDa A14, proteína principal de esperma MSP, proteína organizadora de polimerización MSP MPOP, proteína de fibra MSP 2 MFP2, quinasa para activación de polimerización MSP MPAK, proteína similar a ABA-1 ALB, proteína ABA-1, cuticulina CUT-1 (Ascaris lumbricoides, ascariasis); alérgeno de 41 kDa Asp v13, alérgeno Asp f3, la proteína de barra cilíndrica A de superficie conideal principal, proteasa pep1p, proteína anclada a GPI Gel1p, proteína anclada a GPI Crf1p (Aspergillus genus, Aspergillosis); proteína de la familia VP26, proteína VP29 (infección por Astroviridae, Astrovirus); proteína roptri-asociada 1 PAR-1, los antígenos superficiales del merozoitos MSA-1, MSA-2 (a1, a2, b, c), 12D3, 11C5, 21B4, P29, antígeno superficial de eritrocitos variantes VESA1, antígeno de membrana apical 1 AMA-1 (género Babesia, babesiosis); hemolisina, enterotoxina C, pX01-51, glicolato oxidasa, transportador ABC, proteína de unión a penicilina, preteína de la familia transportadora de zinc, pseudouridina sintasa Rsu, proteína para replicación de plásmidos RepX, oligoendopeptidasa F, proteína de membrana profgo, proteína HemK, antígeno flagelar H, antígeno de superficie celular 28.5-kDa (infección por Bacillus cereus, Bacillus cereus); antígeno T largo LT, antígeno T

pequeño, proteína con cápside protein VP1, proteína cápside VP2 (virus BK, infección por el virus BK); proteína de 29 kDa, antígenos similares a caspasa-3, las glucoproteínas (infección por *Blastocystis hominis*, *Blastocystis hominis*); adhesina con superficie de levadura WI-1 (*Blastomyces dermatitidis*, *Blastomycosis*); nucleoproteína N, polimerasa L, proteína matriz Z, glucoproteína GP (Machupo virus, fiebre hemorrágica Boliviana); proteína de superficie externa A OspA, proteína de superficie externa OspB, proteína de superficie externa OspC, proteína de unión a decorina A DbpA, proteína de unión a decorina B DbpB, proteína Fla con núcleo de 41 kDa de filamentos flagelares, proteína de membrana básica A precursora BmpA (Antígeno inmunodominante P39), precursor de lipoproteína 22 kDa de superficie externa (antígeno IPLA7), lipoproteína vlsE de superficie variable (infección por *Borrelia* genus, *Borrelia*); *Botulinum neurotoxins* BoNT/A1, BoNT/A2, BoNT/A3, BoNT/B, BoNT/C, BoNT/D, BoNT/E, BoNT/F, BoNT/G, toxina botulinum recombinante F dominio Hc FHc (*Clostridium botulinum*, Botulism (y botulismo infantil)); nucleocápside, precursor de glucoproteínas (virus Sabia, fiebre hemorrágica Brasileña); superóxido dismutasa de cobre/Zinc SodC, bacterioferritina Bfr, proteína robosómica 50S RplL, proteína que contiene un dominio transmembrana similar a OmpA Omp31, proteína M5 P39 de 39-kDa inmunogénica, proteína de unión a zinc znuA periplásmica transportadora de ABC y zinc, proteína inmunogénica periplásmica Bp26, proteína ribosómica 30S S12 RpsL, gliceraldehido-3-fosfato deshidrogenasa Gap, precursor de la proteína inmunogénica de membrana exterior de 25 kDa Omp25, proteína B de invasión IalB, factor desencadenador Tig, chaperona molecular DnaK, peptidil-prolil cis-transisomerasa putativa SurA, lipoproteína Omp19, proteína de membrana externa MotY Omp16, proteína de membrana externa conservada D15, malato deshidrogenasa Mdh, componente del sistema de secreción Tipo IV (T4SS) VirJ, lipoproteína de función desconocida BAB1\_0187 (*Brucella* genus, *Brucellosis*); miembros de la familia de transportadores ABC (LolC, OppA, y PotF), proteína transmembrana del sistema para liberación de lipoproteína putativa LolC/E, flagelina FlIC, motilidad intracelular Burkholderia A BimA, factor de alargamiento bacteriano-Tu EF-Tu, proteína similar a OmpA de 17 kDa, proteína codificante boaA, proteína codificante boaB (*Burkholderia cepacia* y otras especies de *Burkholderia*, infección por *Burkholderia*); micolil-transferasa Ag85A, proteína de choque térmico Hsp65, proteína TB10.4, antígeno de 19 kDa, proteína PstS3, proteína de choque térmico Hsp70 (*Mycobacterium ulcerans*, *Buruli ulcer*); proteínas cápside virales VP1 y VP2 mayores y menores de norovirus, poliproteína genómica, proteína con cápside de Sapovirus VP1, proteína Vp3, poliproteína genómica (familia de *Caliciviridae*, infección por *Calicivirus* (*Norovirus* y *Sapovirus*)); proteína de membrana externa principal PorA, flagelina FlaA, antígeno superficial CjaA, proteína de unión a fibronectina CadF, proteína transportadora de ABC Peb1A de unión a aspartato/glutamato, proteína FspA1, proteína FspA2 (*Campylobacter* genus, *Campylobacteriosis*); enzima glucolítica enolasa, proteinasas aspartato secretadas SAP1-10, proteína de pared celular enlazada a glucosfosfatidilinositol (GPI), proteína Hyr1, proteína relacionada con el receptor 3 complementario CR3-RP, adhesina Als3p, proteína de choque térmico hsp90 de 90 kDa, proteína de hidrofobicidad de superficie celular CSH (por lo general *Candida albicans* y otras especies de *Candida*, *Candidiasis*); antígeno de 17-kDa, proteína P26, adhesinas autotransportadoras triméricas TAA, adhesina A de *Bartonella BadA*, proteínas Vomps de membrana externa expresada variablemente, proteína Pap3, proteína HbpA, proteasa HtrA envoltura-asociada, proteína OMP89, proteína GroEL, proteína LalB, proteína OMP43, dihirolipoamida succiniltransferasa SucB (*Bartonella henselae*, enfermedad de araña-gato); proteína-2 con superficie amastigote, proteína SSP4 con superficie amastigote-específica, cruzipaina, trans-sialidasa TS, glucoproteína de superficie tripomastigote TSA-1, proteína reguladora complementaria CRP-10, proteína G4, proteína G2, proteína de bastón paraxonemal PAR2, componente4 de bastón paraflagelar Par1, Proteínas superficiales mucina-asociadas MPSP (*Trypanosoma cruzi*, enfermedad de Chagas (*American trypanosomiasis*)); glucoproteínas con envoltura (gB, gC, gE, gH, gI, gK, gL) (virus de varicela zoster (VZV), *Chickenpox*); proteína de membrana externa principal MOMP, proteína de membrana externa probable PMPC, proteína compleja de membrana externa B OmcB, proteínas de choque térmico Hsp60 HSP10, proteína IncA, proteínas provenientes del sistema de secreción tipo III, proteína de cadena pequeña de reductasa y ribonucleótido NrdB, proteína plasmídica Pgp3, proteína externa declamida N CopN, antígeno CT521, antígeno CT425, antígeno CT043, antígeno TC0052, antígeno TC0189, antígeno TC0582, antígeno TC0660, antígeno TC0726, antígeno TC0816, antígeno TC0828 (*Chlamydia trachomatis*, *Chlamydia*); proteína de respuesta a calcio bajo E LCrE, proteína externa de clamidia N CopN, proteína quinasa de serina/treonina PknD, proteína portadora de acilo-S-maloniltransferasa FabD, proteína de unión a ADN de cadena individual de Ssb, proteína de membrana externa principal MOMP, proteína de membrana externa 2 Omp2, familia de proteínas de membrana polimórfica (Pmp1, Pmp2, Pmp3, Pmp4, Pmp5, Pmp6, Pmp7, Pmp8, Pmp9, Pmp10, Pmp11, Pmp12, Pmp13, Pmp14, Pmp15, Pmp16, Pmp17, Pmp18, Pmp19, Pmp20, Pmp21) (infección por *Chlamydia pneumoniae*, *Chlamydia pneumoniae*); toxina B del cólera CTB, pilin corregulado de toxina A TcpA, pilin corregulado de toxina TcpF, proteína para biosíntesis de pilus co-regulado de toxina F TcpF, subunidad A de enterotoxina del cólera, la subunidad B de enterotoxina del cólera, enterotoxina ST estable al calor, hemaglutinina sensible de manosa MSHA, proteína de membrana externa U Porin ompU, proteína Poring B, proteína de membrana polimórfica D (*Vibrio cholerae*, *Cholera*); propionil-CoA carboxilasa PCC, proteína 14-3-3, inhibidora, proteasa de cisteína, transferasas de glutatión, gelsolina, cathepsina L, proteinasa CatL, proteína Tegumental 20.8 kDa TP20.8, proteína tegumental 31.8 kDa TP31.8, fosfatasa de ácido lisofosfatídico LPAP, (*Clonorchis sinensis*, *Clonorchiasis*); proteínas de capa superficial SLP, antígeno de glutamato deshidrogenasa GDH, toxina A, toxina B, cisteína proteasa Cwp84, cisteína proteasa Cwp13, cisteína proteasa Cwp19, proteína de pared celular CwpV, proteína flagelar FlIC, proteína flagelar FlID (infección por *Clostridium difficile*, *Clostridium difficile*); rinovirus: proteínas cápside VP1, VP2, VP3, VP4; coronavirus: proteínas spike S,

proteínas con envoltura E, proteínas de membrana M, proteínas nucleocápside N (por lo general rinovirus y coronavirus, resfriado común (rinofaringitis viral aguda; Acute coryza)); proteína prion Prp (prion CJD, enfermedad de Creutzfeldt-Jakob (CJD)); proteína con envoltura Gc, proteína con envoltura Gn, proteínas nucleocápside (virus de fiebre hemorrágica Crimean-Congo, fiebre hemorrágica Crimean-Congo (CCHF));

5 helicase con ARN de cuadro DEAD virulencia-asociado VAD1, proteína de galactoxilomanano GalXM, glucuronoxilomanano GXM, manoproteína MP (*Cryptococcus neoformans*, Cryptococcosis); proteína ribosómica ácida P2 CpP2, antígenos de mucina Muc1, Muc2, Muc3 Muc4, Muc5, Muc6, Muc7, proteína de adherencia superficial CP20, proteína de adherencia superficial CP23, proteína superficial CP12, proteína superficial CP21, proteína superficial CP40, proteína superficial CP60, proteína superficial CP15, glucopéptidos

10 superficie-asociados gp40, glucopéptidos superficie-asociados gp15, proteína de pared de oocitos AB, profilina PRF, apirasa (*Cryptosporidium* genus, Cryptosporidiosis); proteína-1 de unión a ácido graso y retinol FAR-1, inhibidor de tejidos de metaloproteína TIMP (TMP), cisteína proteinasa ACEY-1, cisteína proteinasa ACCP-1, antígeno superficial Ac-16, proteína 2 secretada ASP-2, metaloproteasa 1 MTP-1, inhibidor de aspartilproteasa API-1, antígeno superficie-asociado SAA-1, factor secretado adulto-específico Xa anticoagulante del inhibidor de proteasa de serina AP, proteasa aspártica similar a catepsina D ARR-1

15 (usualmente *Ancylostoma braziliense*; otros múltiples parásitos, Cutaneous larva migrans (CLM)); proteasas similares a catepsina L, antígeno 53/25-kDa, miembros de la familia de 8kDa, proteína del cisticercos con una actividad similar a tripsina marginal TsAg5, proteína oncosfere TSOL18, proteína oncosfere TSOL45-1A, lactato deshidrogenasa A LDHA, lactato deshidrogenasa B LDHB (*Taenia solium*, Cysticercosis); antígeno pp65, proteína de membrana pp15, proteína tegumental cápside-proximal pp150, proteína M45, ADN polimerasa UL54, helicasa UL105, glucoproteína gM, glucoproteína gN, glucoproteína H, glucoproteína B gB, proteína UL83, proteína UL94, proteína UL99 (Citomegalovirus (CMV), infección por Citomegalovirus); proteína cápside C, proteína premembrana prM, proteína membrana M, proteína con envoltura E (dominio I, dominio II, dominio II), proteína NS1, proteína NS2A, proteína NS2B, proteína NS3, proteína NS4A, proteína 2K, proteína NS4B,

25 proteína NS5 (Virus del dengue (DEN-1, DEN-2, DEN-3 y DEN-4)–Flavivirus, fiebre de Dengue); proteína de 39kDa (*Dientamoeba fragilis*, Dientamoebiasis); precursor de la toxina de difteria Tox, toxina de difteria DT, sortasa pilin-específica SrtA, proteína pilin de eje SpaA, proteína pilin punta SpaC, proteína pilin menor SpaB, proteína superficie-asociado DIP1281 (*Corynebacterium diphtheriae*, Diphtheria); glucoproteína GP, nucleoproteína NP, proteína matriz menor VP24, proteína matriz mayor VP40, activador de transcripción VP30, cofactor de polimerasa VP35, polimerasa L de ARN (Ebolavirus (EBOV), fiebre hemorrágica de Ebola); proteína prion (prion vCJD, enfermedad variante de Creutzfeldt-Jakob (vCJD, nvCJD)); proteína del sistema UvrABC B, proteína Flp1, proteína Flp2, proteína Flp3, proteína TadA, receptor de hemoglobina HgbA, proteína de membrana externa TdhA, proteína CpsRA, regulador CpxR, proteína SapA, antígeno de 18 kDa, proteína de membrana externa NcaA, proteína LspA, proteína LspA1, proteína LspA2, proteína LspB, componente de

35 membrana externa DsrA, lectina DltA, lipoproteína Hlp, proteína de membrana externa principal OMP, proteína de membrana externa OmpA2 (*Haemophilus ducreyi*, Chancroid); aspartilproteasa 1 Pep1, fosfolipasa B PLB, alfa-manosidasa 1 AMN1, glucanosiltransferasa GEL1, ureasa URE, proteína de matriz peroxisomal Pmp1, antígeno rico en prolina Pra, proteína reactiva de linfocitos T humanos TcrP (*Coccidioides immitis* y *Coccidioides posadasii*, Coccidioidomycosis); alérgeno Tri r2, proteína de choque térmico 60 Hsp60, actina fungal Act, antígeno Tri r2, antígeno Tri r4, antígeno Tri t1, proteína IV, glicerol-3-fosfato deshidrogenasa Gpd1, osmodetector HwSho1A, osmodetector HwSho1B, histidina quinasa HwHhk7B, alérgeno Mala s 1, alérgeno Mala s 11, tioredoxina Trx Mala s 13, alérgeno Mala f, alérgeno Mala s (usualmente *Trichophyton* spp, *Epidermophyton* spp., *Malassezia* spp., *Hortaea werneckii*, Dermatophytosis); proteína EG95, proteína EG10, proteína EG18, proteína EgA31, proteína EM18, antígeno EPC1, antígeno B, antígeno 5, proteína P29, proteína

45 14-3-3, 8-kDa proteína, miofilina, proteína de choque térmico 20 HSP20, glucoproteína GP-89, ácido graso proteína de unión FAPB (*Echinococcus* genus, Echinococcosis); proteína 2 de superficial principal MSP2, proteína 4 de superficie principal MSP4, variante MSP SGV1, variante MSP SGV2, proteína de membrana externa OMP, proteína de membrana externa 19 OMP-19, proteína antigénica principal MAP1, proteína antigénica principal MAP1-2, proteína antigénica principal MAP1B, proteína antigénica principal MAP1-3, proteína codificante de Erum2510, proteína GroEL, proteína GroES, proteínas de membrana externa principal de 30-kDa, proteína GE de 100-kDa, proteína GE de 130-kDa, proteína GE de 160-kDa (*Ehrlichia* genus, Ehrlichiosis); antígeno secretado SagA, proteínas similares a sagA, SalA y SalB, colágeno adhesina Scm, proteínas superficiales Fms1 (EbpA(fm), Fms5 (EbpB(fm), Fms9 (EpbC(fm) y Fms10, proteína EbpC(fm), glucoproteína inmunoprotectora G1 de 96 kDa (*Enterococcus* genus, infección por *Enterococcus*); poliproteína de genoma, polimerasa 3D, proteína con cápside viral VP1, proteína con cápside viral VP2, proteína con cápside viral VP3, proteína con cápside viral VP4, proteasa 2A, proteasa 3C (*Enterovirus* genus, infección por *Enterovirus*); proteínas de membrana externa OM, proteína de membrana externa de 60 kDa, antígeno de superficie celular OmpA, antígeno de superficie celular OmpB (sca5), proteína de membrana externa de 134 kDa, proteína de membrana externa de 31 kDa, proteína de membrana externa de 29.5 kDa, proteína de superficie celular SCA4, proteína de superficie celular Adr1 (RP827), proteína de superficie celular Adr2 (RP828), proteína de superficie celular SCA1, proteína de invasión invA, proteína de división celular fts, proteínas de secreción sec 0family, proteínas de virulencia virB, tlyA, tlyC, proteína similar a parvulina Plp, translocasa preproteínica SecA, antígeno de proteína superficial SPA de 120-kDa, antígeno complejo de 138 kD, proteína principal de 100-kD (proteína I), proteína intracitoplásmica D, antígeno de la proteína superficial protectora SPA (*Rickettsia prowazekii*, tifus Epidémico); antígenos nucleares Epstein–Barr (EBNA-1, EBNA-2,

65



EBNA-3A, EBNA-3B, EBNA-3C, EBNA-proteína líder (EBNA-LP)), proteínas de membrana latente (LMP-1, LMP-2A, LMP-2B), anteriormente el antígeno EBV-EA, antígeno de membrana EBV-MA, antígeno cápside viral EBV-VCA, nucleasa alcalina EBV-AN, glucoproteína H, glucoproteína gp350, glucoproteína gp110, glucoproteína gp42, glucoproteína gHgL, glucoproteína gB (virus Epstein-Barr (EBV), Mononucleosis infecciosa del virus Epstein-Barr); proteína con cápside VP2, proteína con cápside VP1, proteína principal NS1 (Parvovirus B19, Erythema infectiosum (quinta enfermedad)); antígeno pp65, glucoproteína 105, proteína con cápside principal, glucoproteína con envoltura H, proteína U51 (herpesvirus humano 6 (HHV-6) y herpesvirus humano 7 (HHV-7), Exanthem subitum); tioredoxin-glutatión reductasa TGR, catepsina L1 y L2, proteína tipo Kunitz KTM, leucina aminopeptidasa LAP, cisteína proteinasa Fas2, proteína-2 similar a saposin SAP-2, tioredoxinperoxidasa TPx, Prx-1, Prx-2, cisteína proteinasa de catepsina 1 CL3, proteasa de catepsina L CL1, fosfoglicerato quinasa PGK, proteína secretora de 27-kDa, proteína de 60 kDa HSP35alpha, glutatión transferasa GST, antígeno tegumental de 28.5 kDa 28.5 kDa TA, proteasa de catepsina B3 CatB3, cistatin stefin-1 Tipo I, catepsina L5, catepsina L1g y catepsina B, proteína de unión con ácido graso FABP, leucina aminopeptidasas LAP (Fasciola hepatica y Fasciola gigantica, Fasciolosis); proteína prion (prion FFI, insomnio familiar fatal (FFI)); proteína similar al homólogo alérgeno de veneno VAL-1, transcripción larval abundante ALT-1, transcripción larval abundante ALT-2, tioredoxin peroxidasa TPX, homólogo alérgeno de avispa VAH, tioredoxin peroxidasa 2 TPX-2, proteína antigénica SXP (péptidos N, N1, N2, y N3), proteína-1 asociada con la activación ASP-1, Tioredoxin TRX, transglutaminasa BmTGA, glutatión-S-transferasas GST, miosina, homólogo del alérgeno de avispa VAH, colagenasa de 175 kDa, gliceraldehido-3-fosfato deshidrogenasa GAPDH, colágeno cuticular Col-4, proteínas ácidas larvales secretadas SLAP, quitinasa CHI-1, proteína de unión a maltosa MBP, enzima glucolítica fructosa-1,6-bisfosfato aldolasa Fba, tropomiosina TMY-1, producto génico nematodo específico OvB20, oncocistatina CPI-2, Cox-2 (superfamilia de Filarioidea, Filariasis); fosfolipasa C PLC, enterotoxina termo-lábil B, componente de la toxina lota Ib, proteína CPE1281, piruvato ferredoxinoxidoreductasa, factor de alargamiento G EF-G, perfringolisina O Pfo, gliceraldehido-3-fosfato deshidrogenasa GapC, Fructosa-bisfosfato aldolasa Alf2, enterotoxina de clostridium perfringens CPE, toxina alfa AT, toxoide alfa ATd, toxoide epsilon ETd, proteína HP, citotoxina larga TpeL, endo-beta-N-acetilglucosaminidasa Naglu, fosfoglucosaminidasa Pgm (Clostridium perfringens, envenenamiento de alimentos por Clostridium perfringens); leucotoxina lktA, adhesión FadA, proteína de membrana externa RadD, proteína de unión a arginina de alto molecular (género Fusobacterium, infección por Fusobacterium); fosfolipasa C PLC, enterotoxina termo-lábil B, componente de toxina lota Ib, proteína CPE1281, piruvato ferredoxinoxidoreductasa, factor de alargamiento G EF-G, perfringolisina O Pfo, gliceraldehido-3-fosfato deshidrogenasa GapC, fructosa-bisfosfato aldolasa Alf2, enterotoxina clostridium perfringens CPE, toxina alfa AT, toxoide alfa ATd, toxoide epsilon ETd, proteína HP, citotoxina larga TpeL, endo-beta-N-acetilglucosaminidasa Naglu, fosfogliceromutasa Pgm (usualmente Clostridium perfringens; otra especie de Clostridium, Gas gangrene (Clostridium myonecrosis)); lipasa A, lipasa B, peroxidasa Dec1 (Geotrichum candidum, Geotrichosis); proteína prion (prion GSS, síndrome de Gerstmann-Sträussler-Scheinker (GSS)); proteína de pares quística CWP1, CWP2, CWP3, proteína superficial variante VSP, VSP1, VSP2, VSP3, VSP4, VSP5, VSP6, antígeno de 56 kDa, piruvato ferredoxinoxidoreductasa PFOR, alcohol deshidrogenasa E ADHE, alfa-giardin, alfa-8-giardin, alfa-1-giardin, beta-giardin, cisteína proteasas, glutatión-S-transferasa GST, arginina deiminasa ADI, fructosa-1,6-bisfosfato aldolasa FBA, antígenos de Giardia trophozoite GTA (GTA1, GTA2), ornitincarboxiltransferasa OCT, proteína similar a asseblin de fibra estriada SALP, proteína similar a uridinfosforilo UPL, alfa-tubulina, beta-tubulina (Giardia intestinalis, Giardiasis); miembros de la familia de transportadores de ABC (LoIC, OppA, y PotF), proteína transmembrana del sistema para liberación de lipoproteínas putativas LolC/E, flagelina FlIC, motilidad intracelular de Burkholderia A BimA, factor de alargamiento bacteriano-Tu EF-Tu, proteína similar a OmpA de 17 kDa, proteína codificante boaA (Burkholderia mallei, Glanders); ciclofilina CyP, proteína de larva en tercera etapa de 24 kDa GS24, productos de excreción-secreción ESPs (40, 80, 120 y 208 kDa) (Gnathostoma spinigerum y Gnathostoma hispidum, Gnathostomiasis); proteínas pilin, subunidad pilin asociada menor pilC, subunidad pilin mayor y variantes pilE, pilS, proteína de variación de fase porA, Porin B PorB, proteína TraD, antígeno de membrana externa Neisserial H.8, antígeno de 70kDa, proteína de membrana externa principal PI, proteínas de membrana externa PIA y PIB, antígeno W, proteína superficial A NspA, proteína de unión a transferrina TbpA, proteína de unión a transferrina TbpB, PBP2, proteína codificante mtrR, proteína codificante ponA, permeasa de membrana FbpBC, sistema de proteína FbpABC, proteínas LbpAB, proteína de membrana externa Opa, transportador de membrana externa FetA, regulador de hierro represado (Neisseria gonorrhoeae, Gonorrhoea); proteína de membrana externa A OmpA, proteína de membrana externa C OmpC, proteína de membrana externa K17 OmpK17 (Klebsiella granulomatis, Granuloma inguinale (Donovanosis)); proteína de unión a fibronectina Sfb, proteína de unión fibronectina/fibrinógeno FBP54, proteína de unión a fibronectina FbaA, proteína M tipo 1 Emm1, proteína tM ipo 6 Emm6, proteína de unión a inmunoglobulin 35 Sib35, Proteína superficial R28 Spr28, superóxido dismutasa SOD, peptidasas C5a ScpA, antígeno I/II AgI/II, adhesina AspA, proteína de unión a macroglobulina alfa2 de G-GRAB relacionada, proteína fibrilar superficial M5 (infección por Streptococcus pyogenes, Grupo A streptococcal); antígeno  $\beta$  de proteína C, proteínas de arginina desiminasa, adhesina BibA, proteína de 105 kDa BPS, antígenos superficiales c, antígenos superficiales R, antígenos superficiales X, proteína resistente a tripsina R1, proteína resistente a tripsina R3, proteína resistente a tripsina R4, proteína inmunogénica superficial Sip, proteína superficial Rib, proteína de repeticiones ricas en Leucina LrrG, proteína de repetición rica en serina Srr-2, antígeno alfa de proteína C Bca, antígeno Beta Bag, antígeno superficial Epsilon, proteína similar a alfa ALP1, proteína similar a alfa ALP5 antígeno delta superficial, proteína

similar a alfa ALP2, proteína similar a alfa ALP3, proteína similar a alfa ALP4, proteína Cbeta Bac (Streptococcus agalactiae, infección por estreptococos Grupo B); proteína de unión a transferrina 2 Tbp2, fosfatasa P4, proteína de membrana externa P6, lipoproteína Pal peptidoglicano-asociada, proteína D, proteína E, proteína de adherencia y penetración Hap, proteína de membrana externa 26 Omp26, proteína de membrana externa P5 (Fimbrin), proteína de membrana externa D15, proteína de membrana externa OmpP2, 5'-nucleotidasa NucA, proteína de membrana externa P1, proteína de membrana externa P2, lipoproteína de membrana externa Pcp, Lipoproteína E, proteína de membrana externa P4, fuculoquinasa FucK, [Cu,Zn]-superóxido dismutasa SodC, proteasa HtrA, proteína O145, alfa-galactosilceramida (Haemophilus influenzae, infección por Haemophilus influenzae); polimerasa 3D, proteína con cápside viral VP1, proteína con cápside viral VP2, proteína con cápside viral VP3, proteína con cápside viral VP4, proteasa 2A, proteasa 3C (Enteroviruses, principalmente virus de Coxsackie A y Enterovirus 71 (EV71), enfermedad de manos, pies y boca (HFMD)); ARN polimerasa L, proteína L, glucoproteína Gn, glucoproteína Gc, nucleoproteína con cápside S, glucoproteína con envoltura G1, nucleoproteína NP, proteína N, poliproteína M (virus sin nombre, Hantavirus, síndrome pulmonar por hantavirus (HPS)); proteína de choque térmico HspA, proteína de choque térmico HspB, citrato sintasa GltA, proteína UreB, proteína de choque térmico Hsp60, proteína activadora de neutrófilos NAP, catalasa KatA, citotoxina vacuolante VacA, ureasa alfa UreA, ureasa beta UreB, proteína Cpn10, proteína groES, proteína de choque térmico Hsp10, proteína MopB, proteína de citotoxicidad asociada de 10 kDa CAG, antígeno de 36 kDa, beta-lactamasa HcpA, Beta-lactamasa HcpB (Helicobacter pylori, infección por Helicobacter pylori); proteínas de membrana integral, proteínas propensas a agregación, O-antígeno, antígenos tóxicos Stx2B, antígeno tóxico Stx1B, fragmento de antígeno de adhesión Int28, proteína EspA, proteína EspB, Intimina, proteína Tir, proteína IntC300, proteína Eae (Escherichia coli O157:H7, O111 y O104:H4, síndrome hemolítico-urémico (HUS)); ARN polimerasa L, proteína L, glucoproteína Gn, glucoproteína Gc, nucleoproteína con cápside S, glucoproteína con envoltura G1, nucleoproteína NP, proteína N, poliproteína M (familia Bunyaviridae, Fiebre hemorrágica con síndrome renal (HFRS)); glucoproteína G, proteína matriz M, nucleoproteína N, proteína de fusión F, polimerasa L, proteína W, proteína C, fosfoproteína P, proteína no estructural V (virus Henipavirus (Hendra virus Nipah), infecciones Henipavirus); poliproteína, glucoproteína Gp2, antígeno superficial de hepatitis A HBsAg, proteína 2A, proteína viral VP1, proteína viral VP2, proteína viral VP3, proteína viral VP4, proteína P1B, proteína P2A, proteína P3AB, proteína P3D (Virus de Hepatitis A, Hepatitis A); antígeno superficial hepatitis B HBsAg, antígeno núcleo de Hepatitis B HbcAg, polimerasa, proteína Hbx, proteína superficial media preS2, proteína superficial L, proteína larga S, proteína viral VP1, proteína viral VP2, proteína viral VP3, proteína viral VP4 (Virus de Hepatitis B (HBV), Hepatitis B); glucoproteína de envoltura E1 gp32 gp35, glucoproteína con envoltura E2 NS1 gp68 gp70, proteína con cápside C, proteína núcleo Core, poliproteína, proteína viral VP1, proteína viral VP2, proteína viral VP3, proteína viral VP4, antígeno G, proteína NS3, proteína NS5A, (Virus de Hepatitis C, Hepatitis C); proteína viral VP1, proteína viral VP2, proteína viral VP3, proteína viral VP4, antígeno delta largo de hepatitis, antígeno delta pequeño de hepatitis (virus de Hepatitis D, Hepatitis D); proteína viral VP1, proteína viral VP2, proteína viral VP3, proteína viral VP4, proteína con cápside E2 (virus de Hepatitis E, Hepatitis E); glucoproteína L UL1, uracil-ADN glucosilasa UL2, proteína UL3, proteína UL4, proteína ADN de replicación UL5, proteína portal UL6, proteína para maduración de viriones UL7, ADN helicasa UL8, proteína de unión al origen de replicación UL9, glucoproteína M UL10, proteína UL11, exonucleasa alcalina UL12, proteína quinasa de serina-treonina UL13, proteína tegumentaria UL14, terminasa UL15, proteína tegumentaria UL16, proteína UL17, proteína con cápside VP23 UL18, proteína con cápside principal VP5 UL19, proteína con membrana UL20, proteína tegumentaria UL21, Glucoproteína H (UL22), Timidina quinasa UL23, proteína UL24, proteína UL25, proteína con cápside P40 (UL26, VP24, VP22A), glucoproteína B (UL27), proteína ICP18.5 (UL28), proteína principal de unión a ADN ICP8 (UL29), ADN polimerasa UL30, proteína de matriz nuclear UL31, glucoproteína con envoltura UL32, proteína UL33, proteína con membrana nuclear interna UL34, proteína con cápside VP26 (UL35), proteína tegumentaria larga UL36, proteína de ensamble cápside UL37, proteína VP19C (UL38), ribonucleótido reductasa (subunidad Largo) UL39, ribonucleótido reductasa (subunidad pequeña) UL40, proteína tegumentaria/proteína VHS para interrupción en el hospedero de viriones (UL41), factor de procesamiento de ADN polimerasa UL42, proteína con membrana UL43, glucoproteína C (UL44), proteína con membrana UL45, proteína tegumentaria VP11/12 (UL46), proteína tegumentaria VP13/14 (UL47), proteína para maduración de viriones VP16 (UL48, Alfa-TIF), proteína con envoltura UL49, dUTP difosfatasa UL50, proteína tegumentaria UL51, proteína compleja de ADN helicasa/primasa UL52, glucoproteína K (UL53), proteína para regulación transcripcional IE63 (ICP27, UL54), proteína UL55, proteína UL56, proteína de replicación viral ICP22 (IE68, US1), proteína US2, proteína quinasa de serina/treonina US3, glucoproteína G (US4), glucoproteína J (US5), glucoproteína D (US6), glucoproteína I (US7), glucoproteína E (US8), proteína tegumentaria US9, proteína con cápside/tegumentaria US10, proteína Vmw21 (US11), proteína ICP47 (IE12, US12), activador transcripcional principal ICP4 (IE175, RS1), E3 ubiquitina ligasa ICP0 (IE110), proteína 1 latencia-relacionada LRP1, proteína 2 latencia-relacionada LRP2, factor de neurovirulencia RL1 (ICP34.5), transcripción LAT latencia-asociada (Virus de herpes simplex 1 y 2 (HSV-1 y HSV-2), Herpes simplex); proteína de choque térmico Hsp60, proteína de superficie celular H1C, dipeptidilpeptidasa tipo IV DppIV, antígeno M, proteína de 70 kDa, proteína similar a histona de 17 kDa (Histoplasma capsulatum, Histoplasmosis); proteína-1 de unión a ácido graso y retinol FAR-1, inhibidor de tejidos de metaloproteína TIMP (TMP), cisteína proteinasa ACEY-1, cisteína proteinasa ACCP-1, antígeno superficial Ac-16, proteína secretada 2 ASP-2, metaloproteasa 1 MTP-1, inhibidor de aspartilproteasa API-1, antígeno superficial-asociado SAA-1, antígeno superficial-asociado SAA-2, factor secretado adulto-específico

Xa, anticoagulante inhibidor de serina proteasa AP, proteasa aspártica similar a cathepsina D ARR-1, glutatión S-transferasa GST, proteasa aspártica APR-1, acetilcolinesterasa AChE (infección por *Ancylostoma duodenale* y *Necator americanus*, Hookworm); proteína NS1, proteína NP1, proteína VP1, proteína VP2, proteína VP3 (infección por bocavirus Humano (HBoV), bocavirus Humano); proteína superficial principal 2 MSP2, proteína superficial principal 4 MSP4, variante MSP SGV1, variante MSP SGV2, proteína de membrana externa OMP, proteína de membrana externa 19 OMP-19, proteína antigénica principal MAP1, proteína antigénica principal MAP1-2, proteína antigénica principal MAP1B, proteína antigénica principal MAP1-3, proteína codificante Erum2510, proteína GroEL, proteína GroES, proteínas de membrana externa principal de 30-kDa, proteína GE de 100-kDa, proteína GE de 130-kDa, proteína GE de 160-kDa (*Ehrlichia ewingii*, Human ewingii ehrlichiosis); proteína principal superficiales 1-5 (MSP1a, MSP1b, MSP2, MSP3, MSP4, MSP5), proteínas del sistema de secreción tipo IV VirB2, VirB7, VirB11, VirD4 (*Anaplasma phagocytofilum*, anaplasmosis granulocítica Humano (HGA)); proteína NS1, proteína hidrofóbica pequeña NS2, proteína SH, proteína de fusión F, glucoproteína G, proteína de matriz M, proteína de matriz M2-1, proteína de matriz M2-2, fosfoproteína P, nucleoproteína N, polimerasa L (metapneumovirus Humano (hMPV), infección por metapneumovirus Humano); proteína superficial principal 2 MSP2, proteína superficial principal 4 MSP4, variante MSP SGV1, variante MSP SGV2, proteína de membrana externa OMP, proteína de membrana externa 19 OMP-19, proteína antigénica principal MAP1, proteína antigénica principal MAP1-2, proteína antigénica principal MAP1B, proteína antigénica principal MAP1-3, proteína codificante Erum2510, proteína GroEL, proteína GroES, proteínas de membrana externa principal de 30-kDa, proteína GE de 100-kDa, proteína GE de 130-kDa, proteína GE de 160-kDa (*Ehrlichia chaffeensis*, Human monocytic ehrlichiosis); proteína de replicación E1, proteína reguladora E2, proteína E3, proteína E4, proteína E5, proteína E6, proteína E7, proteína E8, proteína con cápside mayor L1, proteína con cápside menor L2 (papilomavirus humano (HPV), infección por papilomavirus humano (HPV)); proteína de fusión F, hemaglutinina-neuramidasa HN, glucoproteína G, proteína matriz M, fosfoproteína P, nucleoproteína N, polimerasa L (virus de parainfluenza humana (HPIV), infección por virus de parainfluenza humana); Hemaglutinina (HA), Neuraminidasa (NA), Nucleoproteína (NP), proteína M1, proteína M2, proteína NS1, proteína NS2 (proteína NEP: proteína de exportación nuclear), proteína PA, proteína PB1 (proteína de polimerasa básica 1), proteína PB1-F2 y proteína PB2 (familia Ortomixoviridae, virus de Influenza (gripe)); poliproteína genómica, proteína E, proteína M, proteína con cápside C (virus de encefalitis Japonesa, encefalitis Japonesa); toxina RTX, pili tipo IV, pilus mayor subunidad PIIA, factores de transcripción reguladora PilS y PilR, proteína sigma54, proteínas de membrana externa (*Kingella kingae*, infección por *Kingella kingae*); proteína prion (Kuru prion, Kuru); nucleoproteína N, polimerasa L, proteína de matriz Z, glucoproteína GP (Lassa virus, fiebre de Lassa); lipoproteína peptidoglican-asociada PAL, chaperonina de 60 kDa Cpn60 (groEL, HspB), pilin PilE tipo IV, proteína de membrana externa MIP, proteína de membrana externa mayor MompS, metaloproteínasa de zinc MSP (*Legionella pneumophila*, Legionellosis (enfermedad del Legionario, fiebre Pontiac)); nucleasa P4, proteína WD, reductasa ribonucleótido M2, superficie de membrana glucoproteína Pg46, cisteína proteinasa CP, proteína glucosa-regulada 78 GRP-78, proteína similar al antígeno S en etapa específica A2, ATPasa F1, beta-tubulina, proteína de choque térmico 70 Hsp70, KMP-11, glucoproteína GP63, proteína BT1, nucleósido hidrolasa NH, proteína de superficie celular B1, proteína P1 similar a la proteína ribosómica P1, esteroil 24-c-metiltransferasa SMT, proteína LACK, histona H1, proteína SPB1, antioxidante tiol-específico TSA, antígeno proteínico STI1, señal de peptidasa SP, histona H2B, antígeno superficial PSA-2, cisteína proteinasa b Cpb (género *Leishmania*, Leishmaniasis); proteína con membrana mayor I, antígeno rico en serina de 45 kDa, caperonina de 10 kDa GroES, antígeno HSP kDa, amino-oxononanoato sintasa AONS, recombinasa proteínica A RecA, Acetil-/propionil-coenzima A carboxilasa alfa, alanina racemasa, chaperonina 2 de 60 kDa, proteína EcxB similar ESAT-6 (L-ESAT-6), proteína Lsr2, proteína ML0276, hemaglutinina de unión a heparina HBHA, proteína de choque térmico 65 Hsp65, proteína codificante de mycP1 o ML0041, proteína codificante htrA2 o ML0176, proteína codificante de htrA4 o ML2659, proteína codificante de gcp o ML0379, proteína codificante de clpC o ML0235 (*Mycobacterium leprae* y *Mycobacterium lepromatosis*, Leprosy); proteína de membrana externa LipL32, proteína con membrana LIC10258, proteína con membrana LP30, proteína con membrana LIC12238, proteína similar Ompa Lsa66, proteína superficial LigA, proteína superficial LigB, proteína de membrana externa mayor OmpL1, proteína de membrana externa LipL41, proteína LigAni, proteína superficial LcpA, proteína de adhesión LipL53, proteína de membrana externa Upl32, proteína superficial Lsa63, flagelina FlaB1, lipoproteína de membrana LipL21, proteína con membrana pL40, adhesina superficial leptospiral Lsa27, proteína de membrana externa OmpL36, proteína de membrana externa OmpL37, proteína de membrana externa OmpL47, proteína de membrana externa OmpL54, aciltransferasa LpxA (género *Leptospira*, Leptospirosis); precursor Hly de listeriolisina O (LLO), proteína lap asociada con invasión (P60), proteína reguladora de Listeriolisina PrfA, metaloproteínasa de zinc Mpl, fosfolipasa fosfatidilinositol-específica C PLC (PlcA, PlcB), O-acetiltransferasa Oat, permeasa transportadora de ABC Im.G\_1771, proteína de adhesión LAP, receptor LAP Hsp60, adhesina LapB, hemolisina listeriolisina O LLO, proteína ActA, Internalina A InlA, proteína InlB (*Listeria monocytogenes*, Listeriosis); proteína de superficie externa A OspA, proteína de superficie externa OspB, proteína de superficie externa OspC, proteína de unión a decorina A DbpA, proteína de unión a decorina B DbpB, proteína núcleo Fla de 41 kDa de filamentos flagelares, proteína de membrana básica A BmpA (antígeno inmunodominante P39), precursor de lipoproteína de 22 kDa de superficie externa (antígeno IPLA7), lipoproteína de superficie variable vlsE (usualmente *Borrelia burgdorferi* y otras especie de *Borrelia*, enfermedad de Lyme (Lyme borreliosis)); proteína VAL-1 similar al homólogo alérgico de veneno, transcripción ALT-1 larval abundante, transcriptasa ALT-2 larval abundante, tioredoxina peroxidasa TPX,

homólogo VAH del alérgeno de avispa, tioredoxina peroxidasa 2 TPX-2, proteína antigénica SXP (peptidos N, N1, N2, y N3), proteína-1 asociada con activación ASP-1, tioredoxina TRX, transglutaminasa BmTGA, glutatión-S-transferasas GST, miosina, homólogo VAH del alérgeno de avispa, colagenasa de 175 kDa, gliceraldehído-3-fosfato deshidrogenasa GAPDH, colágeno cuticular Col-4, proteínas ácidas larvales secretadas SLAP, quitinasa CHI-1, proteína de unión a maltosa MBP, fructosa-1,6-bisfosfato aldolasa Fba de enzimas glucolíticas, tropomiosina TMY-1, producto génico nematodo-específico OvB20, oncocistatina CPI-2, proteína Cox-2 (*Wuchereria bancrofti* y *Brugia malayi*, filariasis linfática (*Elephantiasis*)); glucoproteína GP, proteína de matriz Z, polimerasa L, nucleoproteína N (virus de coriomeningitis linfocítica (LCMV), Lymphocytic choriomeningitis); proteína TRAP anónima relacionada con trombospondina, proteína superficial de SSP2 Sporozoite 2, antígeno de membrana apical 1 AMA1, antígeno de membrana rhoptry RMA1, antígeno de repetición básica ácida ABRA, proteína transversal celular PF, proteína Pvs25, proteína superficial de merozoito 1 MSP-1, proteína superficial de merozoito 2 MSP-2, antígeno superficial de eritrocitos anillo-infectados antígeno de etapa RESALiver 3 LSA-3, proteína Eba-175, antígeno 5 para repetición de serina SERA-5, proteína de circumsporozoitos CS, proteína 3 superficial de merozoito MSP3, proteína superficial 8 de merozoitos MSP8, enolasa PF10, proteína de hepatocitos-eritrocitos de 17 kDa HEP17, proteína con membrana 1 de eritrocitos EMP1, proteína superficial 4/5 de la proteína Kbetamerozoitos MSP 4/5, proteína de choque térmico Hsp90, proteína rica en glutamato GLURP, proteína superficial 4 de merozoitos MSP-4, proteína STARP, precursor antigénico relacionado con la proteína de circumsporozoitos CRA (*Plasmodium* genus, Malaria); nucleoproteína N, proteína de VP24 membrana-asociada, nucleoproteína menor VP30, cofactor de polimerasa VP35, polimerasa L, proteína matriz VP40, glucoproteína con envoltura GP (virus Marburg, fiebre hemorrágica Marburg (MHF)); proteína C, proteína matriz M, fosfoproteína P, proteína no estructural V, glucoproteína de hemaglutinina H, polimerasa L, nucleoproteína N, proteína de fusión F (virus de sarampión, Sarampión); miembros de la familia de transportadores de ABC (LoIC, OppA, y PotF), proteína transmembrana LoIC/E del sistema para liberación de lipoproteína putativa, flagelina FliC, motilidad intracelular de *Burkholderia* A BimA, factor de alargamiento bacteriano-Tu EF-Tu, proteína similar a OmpA de 17 kDa, proteína codificante boaA, proteína codificante boaB (*Burkholderia pseudomallei*, Melioidosis (enfermedad de Whitmore)); proteínas pilin, subunidad pilC pilin-asociada menor, subunidad pilin mayor y variantes pilE, pilS, proteína porA de variación de fase, Porin B PorB, proteína TraD, antígeno de membrana externa Neisserial H.8, antígeno de 70kDa, proteína PI de membrana externa mayor, proteínas PIA y PIB de membrana externa, antígeno W, proteína superficial A NspA, proteína de unión a transferrina TbpA, proteína de unión a transferrina TbpB, PBP2, proteína codificante mtrR, proteína codificante ponA, permeaza con membrana FbpBC, sistema proteínico FbpABC, proteínas LbpAB, proteína de membrana externa Opa, transportador de membrana externa FetA, regulador hierro-reprimido MpeR, proteína de unión al factor H fHbp, adhesina NadA, proteína NhbA, represor FarR (*Neisseria meningitidis*, enfermedad Meningococcal); proteína de 66 kDa, proteína de 22 kDa (usualmente *Metagonimus yokagawai*, Metagonimiasis); proteínas de tubo polar (34, 75, y 170 kDa en *Glugea*, 35, 55 y 150kDa en *Encephalitozoon*), proteína quinesina-relacionada, subunidad mayor de ARN polimerasa II, proteína con membrana similar o integral YIPA, proteína 1 anti-identificación, factor de transcripción de choque térmico HSF, proteína quinasa, timidina quinasa, proteína nucleolar similar a NOP-2 (*Microsporidia* phylum, Microsporidiosis); regulador de apoptosis similar a CASP8 y FADD, Glutatión peroxidasa GPX1, ARN helicasa NPH-II NPH2, subunidad catalítica de Poli(A) polimerasa PAPL, proteína con envoltura mayor P43K, subunidad VETFS de 70 kDa con factor de transcripción temprano, subunidad VETFL de 82 kDa con factor de transcripción temprano, metaloendopeptidasa tipo G1, nucleósido trifosfatasa I NPH1, proteína de replicación similar a A28 MC134L, subunidad RPO7 de 7 kDa de ARN polimeasa (virus *Molluscum contagiosum* (MCV), *Molluscum contagiosum* (MC)); proteína matriz M, fosfoproteína P/V, proteína hidrofóbica pequeña SH, nucleoproteína N, proteína V, glucoproteína de fusión F, hemaglutinina-neuraminidasa HN, ARN polimerasa L (virus Mumps, Mumps); proteínas de membrana externa OM, antígeno de superficie celular OmpA, antígeno de superficie celular OmpB (sca5), proteína de superficie celular SCA4, proteína de superficie celular SCA1, proteína intracitoplásmica D, proteína de capa superficial cristalina SLP, antígeno de proteína superficial protectora SPA (*Rickettsia typhi*, Murine typhus (Endemic typhus)); adhesina P1, adhesión P30, proteína p116, proteína P40, proteína citoesquelética HMW1, proteína citoesquelética HMW2, proteína citoesquelética HMW3, proteína codificante MPN152, proteína codificante MPN426, proteína codificante MPN456, proteína codificante MPN-500 (*Mycoplasma pneumoniae*, *Mycoplasma pneumoniae*); NocA, proteína reguladora dependiente de hierro, VapA, VapD, VapF, VapG, proteasa caseinolítica, proteína de 43-kDa punta asociada con filamentos, proteína P24, proteína P61, proteína de 15-kDa, proteína de 56-kDa (usualmente *Nocardia asteroides* y otras especies de *Nocardia*, Nocardiosis); proteína similar al homólogo alérgeno de veneno VAL-1, transcripción larval abundante ALT-1, transcripción larval abundante ALT-2, tioredoxina peroxidasa TPX, homólogo VAH alérgeno de avispa, tioredoxina peroxidasa 2 TPX-2, proteína antigénica SXP (péptidos N, N1, N2, y N3), proteína-1 asociada con la activación ASP-1, Tioredoxina TRX, transglutaminasa BmTGA, glutatión-S-transferasas GST, miosina, homólogo VAH de alérgeno de avispa, collagenasa de 175 kDa, gliceraldehído-3-fosfato deshidrogenasa GAPDH, colágeno cuticular Col-4, proteínas ácidas larvales secretadas SLAP, quitinasa CHI-1, proteína de unión a maltosa MBP, enzima glucolítica fructosa-1,6-bisfosfato aldolasa Fba, tropomiosina TMY-1, producto génico nematodo-específico OvB20, oncocistatina CPI-2, Cox-2 (*Onchocerca volvulus*, Onchocerciasis (River blindness)); glucoproteína secretada de 43 kDa, glucoproteína gp0, glucoproteína gp75, antígeno Pb27, antígeno Pb40, proteína de choque térmico Hsp65, proteína de choque térmico Hsp70, proteína de choque térmico Hsp90, proteína P10, triosefosfato isomerasa TPI, lectina paracoccin de unión a N-acetil-glucosamina,

proteína de 28 kDa Pb28 (Paracoccidioides brasiliensis, Paracoccidioidomycosis (Blastomycosis de sudamérica)); cisteína proteasa similar a cruzipaina de 28-kDa Pw28CCP (usualmente Paragonimus westermani y otras especies de Paragonimus, Paragonimiasis); proteína de membrana externa OmpH, proteína de membrana externa Omp28, proteína PM1539, proteína PM0355, proteína PM1417, proteína de reparación MutL, proteína BcbC, proteína PM0305, formiato deshidrogenasa-N, proteína PM0698, proteína PM1422, girasa ADN, lipoproteína PlpE, proteína adhesiva Cp39, receptor HasR del sistema para hemo-adquisición, proteína capsular de 39 kDa, OMP hierro-regulada IROMP, proteína de membrana externa OmpA87, proteína fimbrial Ptf, proteína de subunidad fimbrial PtfA, proteína de unión a transferrina TbpI, enzima esterasa MesA, toxina de Pasteurella multocida PMT, proteína adhesiva Cp39 (género Pasteurella, Pasteurellosis); hemaglutinina filamentosa FhaB, adenilato ciclase CyaA, precursor PtxD de la subunidad 4 de la toxina pertussis, precursor de pertactina Prn, subunidad 1 de toxinas PtxA, proteína Cpn60, proteína brkA, precursor PtxB de la subunidad 2 de la toxina pertussis, precursor PtxC de la subunidad 3 de la toxina pertussis, precursor PtxE de la subunidad 5 de la toxina pertussis, pertactina Prn, proteína Fim2, proteína Fim3; (Bordetella pertussis, Pertussis (Whooping cough)); antígeno capsular F1, antígeno virulenta-asociado V, proteína efectora secretada LcrV, antígeno V, proteasa Pla de membrana externa, proteína efectora secretada YopD, proteína-tirosina fosfatasa YopH secretada putativa, subunidad mayor YscF de complejo de aguja, proteína quinasa YopO, proteína autotransportadora putativa YapF, membrana interna transportador YbtQ de ABC (Irp7), proteína de unión a azúcar putativa YPO0612, proteína de choque térmico 90 HtpG, proteína de sulfatasa putativa YdeN, proteína LolA portadora de lipoproteína de membrana externa, chaperona de secreción YerA, lipoproteína putativa YPO0420, proteína activadora de hemolisina HpmB, receptor de membrana externa para pesticina/yersiniabactina Psn, proteína YopE efectora secretada, proteína YopF efectora secretada, proteína YopK efectora secretada, proteína de membrana externa YopN, proteína de membrana externa YopM, precursor de coagulasa/fibrinolisis Pla; (Yersinia pestis, peste); proteína PhpA, adhesina superficial PsaA, neumolisina Ply, proteasa Clp dependiente de ATP, lipoato-proteína ligasa LplA, proteína anclada a la superficie de la pared celular psrP, sortasa SrtA, glutamil-ARNt sintetasa GltX, proteína de unión A colina CbpA, proteína de superficie A neumocócica PspA, proteína de superficie C pneumococal PspC, 6-fosfogluconato deshidrogenasa Gnd, proteína de unión a hierro PiaA, Murein hidrolasa LytB, proteon LytC, proteasa A1 (Streptococcus pneumoniae, infección Pneumococcal); proteína superficial mayor B, proteasa similar a quexinaKEX1, proteína A12, antígeno de 55 kDa P55, glucoproteína de superficie mayor Msg (Pneumocystis jirovecii, Pneumocystis pneumonia (PCP)); poliproteína genómica, polimerasa 3D, proteína con cápside viral VP1, proteína con cápside viral VP2, proteína con cápside viral VP3, proteína con cápside viral VP4, proteasa 2A, proteasa 3C (Poliovirus, Poliomyelitis); proteína Nfa1, exendin-3, lipasa secretora, proteasa similar a catepsina B, cisteína proteasa, catepsina, peroxiredoxina, proteína Cry1Ac (usualmente Naegleria fowleri, meningoencefalitis amebica primaria (PAM)); agnoproteína, antígeno T largo, antígeno T pequeño, proteína con cápside mayor VP1, proteína con cápside menor Vp2 (virus JC, leucoencefalopatía multifocal progresiva); proteína con respuesta de calcio bajo E LCrE, proteína externa de clamidia N CopN, serina/teonina-proteína quinasa PknD, proteína acilo-portadora-S-maloniltransferasa FabD, proteína de unión a ADN de cadena individual-Ssb, proteína de membrana externa mayor MOMP, proteína de membrana externa 2 Omp2, familia de proteína con membrana polimórfica (Pmp1, Pmp2, Pmp3, Pmp4, Pmp5, Pmp6, Pmp7, Pmp8, Pmp9, Pmp10, Pmp11, Pmp12, Pmp13, Pmp14, Pmp15, Pmp16, Pmp17, Pmp18, Pmp19, Pmp20, Pmp21) (Chlamydia psittaci, Psittacosis); proteína de membrana externa P1, proteína de choque térmico B HspB, transportador peptídico ABC, proteína de unión con GTP, proteína lcmB, ribonucleasa R, fosfatasa SixA, proteína DsbD, proteína de membrana externa TolC, proteína de unión con DNA PhoB, ATPasa DotB, proteína de choque térmico B HspB, proteína con membrana Com1, proteína de 28 kDa, ADN-3-metiladenina glucosidasa I, proteína de membrana externa OmpH, proteína de membrana externa AdaA, proteína T del sistema de segmentación de glicina (Coxiella burnetii, fiebre Q); nucleoproteína N, proteína estructural grande L, fosfoproteína P, proteína de matriz M, glucoproteína G (Virus de la rabia, Rabias); proteína de fusión F, nucleoproteína N, proteína de matriz M, proteína de matriz M2-1, proteína de matriz M2-2, fosfoproteína P, proteína hidrofóbica pequeña SH, glucoproteína G de superficie mayor, polimerasa L, proteína 1 no estructural NS1, proteína 2 no estructural NS2 (Virus sincitial respiratorio (RSV), infección por virus sincitial respiratorio); poliproteína genómica, polimerasa 3D, proteína con cápside viral VP1, proteína con cápside viral VP2, proteína con cápside viral VP3, proteína con cápside viral VP4, proteasa 2A, proteasa 3C (Rhinovirus, infección por Rhinovirus); proteínas de membrana externa OM, antígeno OmpA de superficie celular, antígeno OmpB de superficie celular (sca5), proteína de superficie celular SCA4, proteína de superficie celular SCA1, proteína PS120, proteína D intracitoplásmica, antígeno SPA de proteína superficial protectora (género Rickettsia, infección por Rickettsial); proteínas OM de membrana externa, antígeno OmpA de superficie celular, antígeno OmpB de superficie celular (sca5), proteína SCA4 de superficie celular, proteína SCA1 de superficie celular, proteína D intracitoplásmica (Rickettsia akari, Rickettsialpox); glucoproteína GP con envoltura, polimerasa L, nucleoproteína N, proteína NSS no estructural (virus de fiebre Rift Valley, fiebre de Rift Valley (RVF)); proteínas OM de membrana externa, antígeno OmpA de superficie celular, antígeno OmpB de superficie celular (sca5), proteína SCA4 de superficie celular, proteína SCA1 de superficie celular, proteína D intracitoplásmica (Rickettsia rickettsii, fiebre maculosa de las Montañas Rocosas (RMSF)); proteína 6 no estructural NS6, proteína 2 no estructural NS2, proteína VP6 con cápside intermedia, proteína VP2 con cápside interna, proteína 3 no estructural NS3, polimerasa L de ARN-ARN-dirigida, proteína VP3, proteína 1 no estructural NS1, proteína 5 no estructural NS5, glucoproteína VP7 con cápside extrna, glucoproteína 4 no estructural NS4, proteína VP4

con cápside externa (Rotavirus, infección por Rotavirus); poliproteína P200, glucoproteína E1, glucoproteína E2, proteína NS2, proteína C con cápside (virus de rubéola, rubéola); chaperonina GroEL (MopA), inositol fosfato fosfatasa SopB, proteína de choque térmico HslU, proteína chaperona DnaJ, proteína TviB, proteína IroN, flagelina FliC, proteína de invasión SipC, glucoproteína gp43, proteína de membrana externa LamB, 5 proteína de membrana externa PagC, proteína de membrana externa TolC, proteína de membrana externa NmpC, proteína de membrana externa FadL, proteína de transporte SadA, transferasa WgaP, proteínas efectoras SifA, SteC, SseL, SseJ y SseF (Salmonella genus, Salmonelosis); proteína 14, proteína NS7b no estructural, proteína NS8a no estructural, proteína 9b, proteína 3a, nucleoproteína N, proteína NS3b no estructural, proteína NS6 no estructural, proteína 7a, proteína NS8b no estructural, proteína con membrana M, 10 proteína EsM con membrana pequeña con envoltura, poliproteína 1a de replicasa, glucoproteína de espira S, replicasa poliproteína 1ab; (SARS coronavirus, SARS (Síndrome respiratorio agudo severo)); serin proteasa, Antígeno 1 ASA1 de sarcoptes atípico, glutatión S-transferasas GST, cisteína proteasa, serina proteasa, apolipoproteína (Sarcoptes scabiei, Scabies); glutatión S-transferasas GST, paramiosina, hemoglobinas SM32, antígeno de huevo mayor, proteína Sm14 de unión con ácido graso de 14 kDa, antígeno P37 de superficie larval mayor, antígeno tegumental de 22.6 kDa, calpaina CANP, trifosfato isomerasa Tim, proteína superficial 15 9B, proteína VP2 con cápside externa, proteína Sm23 con membrana integral de 23 kDa, Cu/Zn-superóxido dismutasa, glucoproteína Gp, miosina (género Schistosoma, Schistosomiasis (Bilharziosis)); chaperonina de 60 kDa, antígeno tipo específico de 56 kDa, piruvato fosfatodiquinasa, 4-hidroxibenzoato octapreniltransferasa (Orientia tsutsugamushi, Scrub typhus); deshidrogenasa GuaB, proteína de invasión Spa32, invasina IpaA, 20 invasina IpaB, invasina IpaC, invasina IpaD, invasina IpaH, invasina IpaJ (Shigella genus, Shigellosis (disentería Bacillary)); proteína P53, homólogo US10 de proteína viriónica, regulador transcripcional IE63, transactivador transcripcional IE62, proteasa P33, proteína de 74 kDa del factor de alfa trans-inducción, desoxiuridina 5'-trifosfato nucleotidohidrolasa, transactivador transcripcional IE4, homólogo UL43 de proteína con membrana, homólogo UL3 de fosfoproteína nuclear, homólogo UL4 de proteína nuclear, proteína de unión 25 con origen de replicación, proteína con membrana 2, fosfoproteína 32, proteína 57, factor de procesabilidad de polimerasa con ADN, proteína portal 54, ADN primasa, homólogo UL14 de proteína tegumentaria, homólogo UL21 de proteína tegumentaria, homólogo UL55 de proteína tegumentaria, homólogo UL33 de la subunidad de terminasa tripartita, homólogo de UL15 de la subunidad de terminasa tripartita, proteína 44 de unión con cápside, proteína 43 de empaque viriónico (virus de varicela zoster (VZV), Shingles (Herpes zoster)); homólogo de esteroide deshidrogenasa de 3-beta hidroxil-5-eno truncado, proteína A13 de membrana viriónica, proteína A19, proteína A31, homólogo A35 de proteína truncada, homólogo A37.5 de proteína, proteína A47, proteína A49, proteína A51, proteína A43 similar a semaforina, inhibidor 1 de serina proteinasa, inhibidor 2 de serina proteinasa, inhibidor 3 de serina proteinasa, proteína A6, proteína B15, proteína C1, proteína C5, proteína C6, 35 proteína F7, proteína F8, proteína F9, proteína F11, proteína F14, proteína F15, proteína F16 (Variola mayor o Variola menor, Smallpox (Variola)); adhesina/glucoproteína gp70, proteasas (Sporothrix schenckii, Sporotrichosis); proteína IsdB de unión hemo-hierro, colágeno adhesina Cna, factor A de aterronamiento ClfA, proteína MecA, proteína de unión A fibronectina FnbA, enterotoxina tipo A EntA, enterotoxina tipo B EntB, enterotoxina tipo C EntC1, enterotoxina tipo C EntC2, enterotoxina tipo D EntD, enterotoxina tipo E EntE, toxina-1 de síndrome de choque tóxico TSST-1, Stafiloquinasa, proteína 2a de unión penicilina PBP2a (MecA), 40 antígeno secretor SssA (Staphylococcus genus, envenenamiento Staphylococcal por alimentos); proteína IsdB de unión a hemo-hierro, colágeno adhesina Cna, factor A de aterronamiento ClfA, proteína MecA, proteína A de unión fibronectina FnbA, tipo A EntA de enterotoxina, tipo B EntB de enterotoxina, tipo C EntC1 de enterotoxina, tipo C EntC2 de enterotoxina, tipo D EntD de enterotoxina, tipo E EntE de enterotoxina, toxina 1 de síndrome de choque tóxico TSST-1, Stafiloquinasa, proteína 2a de unión a penicilina PBP2a (MecA), 45 antígeno secretorio SssA (Staphylococcus genus por ejemplo, infección por aureus, Staphylococcal); antígeno Ss-IR, antígeno NIE, estrogilastacina, Na<sup>+</sup>-K<sup>+</sup> ATPasa Sseat-6, tropomisina SsTmy-1, proteína LEC-5, antígeno de 41 kDa P5, proteína larval de 41-kDa, proteína larval de 31-kDa, proteína larval de 28-kDa (Strongyloides stercoralis, Strongyloidiasis); glicerofosfodiéster fosfodiesterasa GlpQ (Gpd), proteína TmpB de membrana externa, proteína Tp92, antígeno Tpf1, proteína de repetición Tpr, proteína de repetición F TprF, 50 proteína de repetición G TprG, proteína de repetición I TprI, proteína de repetición J TprJ, proteína de repetición K TprK, proteína A con membrana treponémica TmpA, lipoproteína, Tpp15 de 15 kDa, antígeno de membrana de 47 kDa, miniferitina Tpf1, adhesina Tp0751, lipoproteína TP0136, proteína TpN17, proteína TpN47, proteína de membrana externa TP0136, proteína de membrana externa TP0155, proteína de membrana externa TP0326, proteína de membrana externa TP0483, proteína de membrana externa TP0956 (Treponema pallidum, Sífilis); proteasas similares a catepsina L, antígeno de 53/25-kDa, miembros de la familia de 8kDa, proteína de cisticercos con una actividad TsAg5 similar a tripsina marginal, proteína TSOL18 oncosfera, proteína TSOL45-1A oncosfera, lactato deshidrogenasa A LDHA, lactato deshidrogenasa B LDHB (género Taenia, Taeniasis); toxina tetánica TetX, toxina tetánica C TTC, proteína con capa de 140 kDa S, subunidad CT3 de flavoproteína beta, fosfolipasa (lecitinas), proteína fosfoportadora HPr (Clostridium tetani, Tetanus (Lockjaw)); poliproteína genómica, proteína E, proteína M, proteína C con cápside (virus de encefalitis portada por garrapatas (TBEV), encefalitis portada por garrapatas); antígeno de 58-kDa, antígenos de 68-kDa, antígeno TES excretor-secretor de larva toxocara, glucoproteína de 32-kDa, glucoproteína TES-70, glucoproteína GP31, antígeno TcES-57 excretor-secretor, antígeno Pe de fluido perientérico, antígenos Ex de extracto soluble, antígenos ES de larva excretor/secretor, antígeno TES-120, alérgeno de poliproteína TBA-1, cisteína proteasa 65 c-cpl-1 similar a catepsina L, proteína de 26-kDa (Toxocara canis o Toxocara cati, Toxocaríasis (Larva Migrans

- Ocular (OLM) y Larva Migrans Visceral (VLM)); proteínas micronémicas (MIC1, MIC2, MIC3, MIC4, MIC5, MIC6, MIC7, MIC8), proteína rhoptry Rop2, proteínas rhoptry (Rop1, Rop2, Rop3, Rop4, Rop5, Rop6, Rop7, Rop16, Rjop17), proteína SR1, antígeno superficial P22, antígeno mayor p24, antígeno superficial mayor p30, proteínas de gránulo denso (GRA1, GRA2, GRA3, GRA4, GRA5, GRA6, GRA7, GRA8, GRA9, GRA10), antígeno de 28 kDa, antígeno superficial SAG1, antígeno relacionado con SAG2, nucleósidos-trifosfatasa 1, nucleósido-trifosfatasa 2, proteína Stt3, proteína que contiene un dominio similar a HesB, proteasa 5 similar a romboides, toxomepsin 1 (*Toxoplasma gondii*, *Toxoplasmosis*); glucoproteína secretada de 43 kDa, glucoproteína secretada de 53 kDa, paramiosina, antígeno Ts21, antígeno Ts87, antígeno p46000, antígenos TSL-1, caveolin-1 CAV-1, antígeno de larva recién nacida de 49 kDa, homólogo de prosaposina, serina proteasa, inhibidor de serina proteinasa, glucoproteína Gp45 de 45 kDa (*Trichinella spiralis*, *Trichinellosis*); factores transcripcionales similares a Myb (Myb1, Myb2, Myb3), proteína de adhesión AP23, proteína de adhesión AP33, proteína adhesina AP33-3, adhesins AP51, adhesina AP65, proteína de adhesión AP65-1, alfa-actinina, proteína asociada con cinesina, teneurina, proteinasa de 62 kDa, serine proteasa SUB1 similar a subtilisina, gen 3 CP3 de cisteína proteinasa, alfa-enolasa Eno1, cisteína proteinasa CP30, proteína de choque térmico (Hsp70, Hsp60), proteína inmunogénica P270, (*Trichomonas vaginalis*, *Trichomoniasis*); beta-tubulina, proteína de 47-kDa, proteinasa-1 similar a leucocito secretor SLP-1, proteína TT50 de 50-kDa, antígeno de 17 kDa, proteína de 43/47 kDa (*Trichuris trichiura*, *Trichuriasis* (infección por Whipworm)); proteína ESAT-6 (EsxA), antígeno filtrado EsxB de 10 kDa, antígeno secretado 85-B FBPB, proteína A de unión a fibronectina FbpA (Ag85A), serina proteasa PepA, proteína PPE18 de la familia PPE, proteína D de unión a fibronectina FbpD, proteína inmunogénica MPT64, proteína secretada MPT51, catalasa-peroxidasa-peroxinitritasa T KATG, lipoproteína PSTS3 de unión a fosfato periplásmico (PBP-3, Phos-1), hemaglutinina Hbha de unión a heparina hierro-regulada, proteína PPE14 de la familia PPE, proteína PPE68 de la familia PPE, proteína Mtb72F, proteína Apa, proteína inmunogénica MPT63, lipoproteína PSTS1 de unión a fosfato periplásmico (PBP-1), chaperona molecular DnaK, lipoproteína Mpt83 de superficie celular, lipoproteína P23, proteína permeasa pstA del sistema de transporte de fosfato, antígeno de 14 kDa, proteína C de unión a fibronectina FbpC1, alanina deshidrogenasa TB43, glutamina sintetasa 1, proteína ESX-1, proteína CFP10, proteína TB10.4, proteína MPT83, proteína MTB12, proteína MTB8, proteínas similares a Rpf, proteína MTB32, proteína MTB39, cristalina, proteína de choque térmico HSP65, proteína PST-S (usualmente *tuberculosis Mycobacterium*, *Tuberculosis*); proteína FobA de membrana externa, proteína FobB de membrana externa, locus IgIC1 de crecimiento intracelular, locus IgIC2 de crecimiento intracelular, aminotransferasa Wbtl, chaperonina GroEL, proteína TUL4 con membrana mayor de 17 kDa, lipoproteína LpnA, proteína de la familia de quitinasa 18, isocitrato deshidrogenasa, proteína de la familia Nif3, proteína de glucosilación pili tipo IV, proteína tolC de membrana externa, proteína de la familia de unión con FAD, proteína de membrana externa multimérica pilin tipo IV, proteína KdpD con sensor de dos componentes, proteína chaperona DnaK, proteína TolQ (*Francisella tularensis*, *Tularemia*); MB antígeno, ureasa, proteína GyrA, proteína GyrB, proteína ParC, proteína ParE, proteínas LAMP de membrana asociada con lípidos, timidina quinasa TK, fosfolipasa PL-A1, fosfolipasa PL-A2, fosfolipasa PL-C, antígeno de 96-kDa superficie expresado; (*Ureaplasma urealyticum*, infección por *Ureaplasma urealyticum*); poliproteína no estructural, poliproteína estructural, proteína CP con cápside, proteína E1, proteína E2, proteína E3, proteasa P1, proteasa P2, proteasa P3 (virus de encefalitis equina Venezolana, encefalitis equina Venezolana); glucoproteína GP, proteína matriz Z, polimerasa L, nucleoproteína N (virus Guanarito, fiebre hemorrágica Venezolana); poliproteína, proteína E, proteína M, proteína C con cápside, proteasa NS3, proteína NS1, proteína NS2A, proteína AS2B, proteína NS4A, proteína NS4B, proteína NS5 (virus del nilo Occidental, fiebre del Nilo Occidental); proteína CP con cápside, proteína E1, proteína E2, proteína E3, proteasa P2 (virus de encefalitis equina Occidental, encefalitis equina Occidental); poliproteína genómica, proteína E, proteína M, proteína C con cápside, proteasa NS3, proteína NS1, proteína NS2A, proteína AS2B, proteína NS4A, proteína NS4B, proteína NS5 (virus de fiebre amarilla, Fiebre amarilla); proteína YobB blanco Yop putativa, proteína efectora YopD, proteína efectora YopE, proteína YopH, proteína efectora YopJ, proteína de translocación proteínica YopK, proteína efectora YopT, proteína YpkA, proteína FlhA de biosíntesis flagelar, peptidasa M48, system de flujo de potasio KefA, regulatoer transcripcional RovA, adhesina lfp, proteína translocadora LcrV, proteína PcrV, invasina Inv, purina similar a OmpF de proteína con membrana externa, adhesina YadA, proteína quinasa C, fosfolipase C1, proteína PsaA, proteína WbyK similar manosiltransferasa, proteína YscU, antígeno YPMa (*Yersinia pseudotuberculosis*, infección por *Yersinia pseudotuberculosis*); proteína efectora YopB, chaperonina de 60 kDa, proteína WbcP, tirosina-proteína fosfatasa YopH, proteína YopQ, enterotoxina, Galactosido permeasa, reductasa NrdE, proteína YasN, Invasina Inv, adhesina YadA, porina F de membrana externa OmpF, proteína UspA1, proteína EibA, proteína Hia, proteína Ail de superficie celular, chaperona SycD, proteína LcrD, proteína LcrG, proteína LcrV, proteína SycE, proteína YopE, proteína TyeA reguladora, proteína YopM, proteína YopN, proteína YopO, proteína YopT, proteína YopD, proteasa ClpP, proteína MyfA, proteína FilA, y proteína PsaA (*Yersinia enterocolitica*, *Yersiniosis*).
- 60 (Entre paréntesis se encuentra el patógeno particular de la familia de patógenos de los cuales se deriva o derivan el antígeno o antígenos y la enfermedad infecciosa a la cual está asocia el patógeno)

Más específicamente, según la invención, el antígeno patogénico está asociado a una infección bacteriana, a una infección viral o a una infección protozoaria y donde el antígeno patogénico se selecciona de

- Antígeno p24 de VIH, proteínas de envoltura de VIH (Gp120, Gp41, Gp160), poliproteína GAG, proteína del factor negativo Nef, trans-activador de transcripción Tat si la enfermedad infecciosa en VIH, preferentemente una infección con el virus de inmunodeficiencia humana;
- 5 • proteína cápside C, proteína premembrana prM, proteína membrana M, proteína de envoltura E (dominio I, dominio II, dominio III), proteína NS1, proteína NS2A, proteína NS2B, proteína NS3, proteína NS4A, proteína 2K, proteína NS4B, proteína NS5, si la enfermedad infecciosa es fiebre del Dengue, preferentemente una infección con el virus del Dengue (DEN-1, DEN-2, DEN-3 y DEN-4)-Flavivirus;
- 10 • antígeno superficial de hepatitis B HBsAg, antígeno núcleo de Hepatitis B HbcAg polimerasa, proteína Hbx, proteína superficial media preS2, proteína superficial L, proteína larga S, proteína viral VP1, proteína viral VP2, proteína viral VP3, proteína viral VP4, si la enfermedad infecciosa es Hepatitis B, preferentemente una infección con el virus de la Hepatitis B (HBV);
- 15 • Hemaglutinina (HA), Neuraminidasa (NA), Nucleoproteína (NP), proteína M1, proteína M2, proteína NS1, proteína NS2 (proteína NEP: proteína de exportación nuclear), proteína PA, proteína PB1 (proteína de polimerasa básica 1), proteína PB1-F2 y proteína PB2 (familia Ortomixoviridae, virus de la gripe (flu));
- nucleoproteína N, proteína estructural grande L, fosfoproteína P, proteína de matriz M, glucoproteína G, si la enfermedad infecciosa es la rabia, preferentemente una infección con virus de la rabia;
- poliproteína genómica, proteína E, proteína M, proteína C de cápside, proteasa NS3, proteína NS1, proteína NS2A, proteína AS2B, proteína NS4A, proteína NS4B, proteína NS5, si la enfermedad infecciosa es fiebre amarilla, preferentemente una infección por el virus de la fiebre amarilla;
- 20 • proteína TRAP anónima relacionada con trombospondina, proteína superficial de SSP2 Sporozoite 2, antígeno de membrana apical 1 AMA1, antígeno de membrana rhoptry RMA1, antígeno de repetición básica ácida ABRA, proteína transversal celular PF, proteína Pvs25, proteína superficial de merozoito 1 MSP-1, proteína superficial de merozoito 2 MSP-2, antígeno superficial de eritrocitos anillo-infectados antígeno de etapa RESALiver 3 LSA-3, proteína Eba-175, antígeno 5 para repetición de serina SERA-5, proteína de circumsporozoitos CS, proteína 3 superficial de merozoito MSP3, proteína superficial 8 de merozoitos MSP8, enolasa PF10, proteína de hepatocitos-eritrocitos de 17 kDa HEP17, proteína con membrana 1 de eritrocitos EMP1, proteína superficial 4/5 de la proteína Kbetamerozoitos MSP 4/5, proteína de choque térmico Hsp90, proteína rica en glutamato GLURP, proteína superficial 4 de merozoitos MSP-4, proteína STARP, precursor antigénico relacionado con la proteína de circumsporozoitos CRA (Plasmodium genus, Malaria);
- 30 • proteína de fusión F, hemaglutinina-neuramidasa HN, glucoproteína G, proteína matriz M, fosfoproteína P, nucleoproteína N, polimerasa L, si la enfermedad infecciosa es la infección por el virus de la parainfluenza humana, de preferencia una infección con los virus de parainfluenza Humanos (HPIV);
- proteína de fusión F, nucleoproteína N, proteína de matriz M, proteína de matriz M2-1, proteína de matriz M2-2, fosfoproteína P, proteína hidrofóbica pequeña SH, glucoproteína G de superficie mayor, polimerasa L, proteína 1 no estructural NS1, proteína 2 no estructural NS2, si la enfermedad infecciosa es una infección por el virus sincitial respiratorio, preferentemente una infección por el virus sincitial respiratorio (RSV);
- 40 • glucoproteína L UL1, uracil-ADN glucosilasa UL2, proteína UL3, proteína UL4, proteína ADN de replicación UL5, proteína portal UL6, proteína para maduración de viriones UL7, ADN helicasa UL8, proteína de unión al origen de replicación UL9, glucoproteína M UL10, proteína UL11, exonucleasa alcalina UL12, proteína quinasa de serina-treonina UL13, proteína tegumentaria UL14, terminasa UL15, proteína tegumentaria UL16, proteína UL17, proteína con cápside VP23 UL18, proteína con cápside principal VP5 UL19, proteína con membrana UL20, proteína tegumentaria UL21, Glucoproteína H (UL22), Timidina quinasa UL23, proteína UL24, proteína UL25, proteína con cápside P40 (UL26, VP24, VP22A), glucoproteína B (UL27), proteína ICP18.5 (UL28), proteína principal de unión a ADN ICP8 (UL29), ADN polimerasa UL30, proteína de matriz nuclear UL31, glucoproteína con envoltura UL32, proteína UL33, proteína con membrana nuclear interna UL34, proteína con cápside VP26 (UL35), proteína tegumentaria larga UL36, proteína de ensamble cápside UL37, proteína VP19C (UL38), ribonucleótido reductasa (subunidad Largo) UL39, ribonucleótido reductasa (subunidad pequeña) UL40, proteína tegumentaria/proteína VHS para interrupción en el huésped de viriones (UL41), factor de procesamiento de ADN polimerasa UL42, proteína con membrana UL43, glucoproteína C (UL44), proteína con membrana UL45, proteína tegumentarias VP11/12 (UL46), proteína tegumentaria VP13/14 (UL47), proteína para maduración de viriones VP16 (UL48, Alfa-TIF), proteína con envoltura UL49, dUTP difosfatasa UL50, proteína tegumentaria UL51, proteína compleja de ADN helicasa/primasa UL52, glucoproteína K (UL53), proteína para regulación transcripcional IE63 (ICP27, UL54), proteína UL55, proteína UL56, proteína de replicación viral ICP22 (IE68, US1), proteína US2, proteína quinasa de serina/treonina US3, glucoproteína G (US4), glucoproteína J (US5), glucoproteína D (US6), glucoproteína I (US7), glucoproteína E (US8), proteína tegumentaria US9, proteína con cápside/tegumentaria US10, proteína Vmw21 (US11), proteína ICP47 (IE12, US12), activador transcriptional principal ICP4 (IE175, RS1), E3 ubiquitina ligasa ICP0 (IE110), proteína 1 latencia-relacionada LRP1, proteína 2 latencia-relacionada LRP2, factor de neurovirulencia RL1 (ICP34.5), transcripción LAT latencia-asociada (Virus de herpes simplex 1 y 2 (HSV-1 y HSV-2), Herpes simplex);
- 60



- proteína de replicación E1, proteína reguladora E2, proteína E3, proteína E4, proteína E5, proteína E6, proteína E7, proteína E8, proteína con cápside mayor L1, proteína con cápside menor L2 (papilomavirus humano (HPV), infección por papilomavirus humano (HPV));
- 5 • antígeno pp65, proteína de membrana pp15, proteína tegumental cápside-proximal pp150, proteína M45, ADN polimerasa UL54, helicasa UL105, glucoproteína gM, glucoproteína gN, glucoproteína H, glucoproteína B gB, proteína UL83, proteína UL94, proteína UL99 (Citomegalovirus (CMV), infección por Citomegalovirus);
- 10 • proteína IsdB de unión hemo-hierro, colágeno adhesina Cna, factor A de aterronamiento CifA, proteína MecA, proteína de unión A fibronectina FnbA, enterotoxina tipo A EntA, enterotoxina tipo B EntB, enterotoxina tipo C EntC1, enterotoxina tipo C EntC2, enterotoxina tipo D EntD, enterotoxina tipo E EntE, toxina-1 de síndrome de choque tóxico TSST-1, Stafiloquinasa, proteína 2a de unión penicilina PBP2a (MecA), antígeno secretor SssA (Staphylococcus genus, por ejemplo aureus, infección por estafilococos);
- 15 • poliproteína genómica, polimerasa 3D, proteína con cápside viral VP1, proteína con cápside viral VP2, proteína con cápside viral VP3, proteína con cápside viral VP4, proteasa 2A, proteasa 3C (Rhinovirus, infección por Rhinovirus);
- 20 • proteína ESAT-6 (EsxA), antígeno filtrado EsxB de 10 kDa, antígeno secretado 85-B FBPB, proteína A de unión a fibronectina FbpA (Ag85A), serina proteasa PepA, proteína PPE18 de la familia PPE, proteína D de unión a fibronectina FbpD, proteína inmunogénica MPT64, proteína secretada MPT51, catalasa-peroxidasa-peroxinitritasa T KATG, lipoproteína PSTS3 de unión a fosfato periplásmico (PBP-3, Phos-1), hemaglutinina Hbha de unión a heparina hierro-regulada, proteína PPE14 de la familia PPE, proteína PPE68 de la familia PPE, proteína Mtb72F, proteína Apa, proteína inmunogénica MPT63, lipoproteína PSTS1 de unión a fosfato periplásmico (PBP-1), chaperón molecular DnaK, lipoproteína Mpt83 de superficie celular, lipoproteína P23, proteína permeasa pstA del sistema de transporte de fosfato, antígeno de 14 kDa, proteína C de unión a fibronectina FbpC1, alanina deshidrogenasa TB43, glutamina sintetasa 1, proteína ESX-1, proteína CFP10, proteína TB10.4, proteína MPT83, proteína MTB12, proteína MTB8, proteínas similares a Rpf, proteína MTB32, proteína MTB39, cristalina, proteína de choque térmico HSP65, proteína PST-S (usualmente tuberculosis Mycobacterium, Tuberculosis);
- 25 • nucleoproteína N, polimerasa L, proteína de matriz Z, glucoproteína GP (Lassa virus, fiebre de Lassa);
- 30 • proteínas spike S, proteínas con envoltura E, proteínas de membrana M, proteínas nucleocápside N (por lo general rinovirus y coronavirus);
- glucoproteína GP, nucleoproteína NP, proteína matriz menor VP24, proteína matriz mayor VP40, activador de transcripción VP30, cofactor de polimerasa VP35, polimerasa L de ARN (Ebolavirus (EBOV), fiebre hemorrágica de Ebola);
- 35 • proteína con envoltura Gc, proteína con envoltura Gn, proteínas nucleocápside (virus de fiebre hemorrágica Crimean-Congo, fiebre hemorrágica Crimean-Congo (CCHF));
- glucoproteína G, proteína matriz M, nucleoproteína N, proteína de fusión F, polimerasa L, proteína W, proteína C, fosfoproteína P, proteína no estructural V (virus Henipavirus (Hendra virus Nipah), infecciones Henipavirus);
- 40 • glucoproteína GP con envoltura, polimerasa L, nucleoproteína N, proteína NSS no estructural (virus de fiebre Rift Valley, fiebre de Rift Valley (RVF)).
- proteína 6 no estructural NS6, proteína 2 no estructural NS2, proteína VP6 con cápside intermedia, proteína VP2 con cápside interna, proteína 3 no estructural NS3, polimerasa L de ARN-ARN-dirigida, proteína VP3, proteína 1 no estructural NS1, proteína 5 no estructural NS5, glucoproteína VP7 con cápside extra, glucoproteína 4 no estructural NS4, proteína VP4 con cápside externa (Rotavirus, infección por Rotavirus);
- 45 • proteína NS1, proteína hidrofóbica pequeña NS2, proteína SH, proteína de fusión F, glucoproteína G, proteína de matriz M, proteína de matriz M2-1, proteína de matriz M2-2, fosfoproteína P, nucleoproteína N, polimerasa L (metapneumovirus Humano (hMPV), infección por metapneumovirus Humano).

50 La región codificante del ácido nucleico inventivo de acuerdo con el primer aspecto de la presente invención se puede presentar como un ácido nucleico mono-, di-cistrónico o incluso multi-cistrónico; es decir, un ácido nucleico que porta las secuencias codificantes de una, dos o más proteínas o péptidos. Estas secuencias codificantes en ácidos di- o incluso multi-cistrónicos pueden estar separadas por al menos una secuencia con sitio de entrada de ribosoma interno (IRES), por ejemplo como se describe aquí, o por péptidos señal que inducen la segmentación del polipéptido resultante que comprende diversas proteínas o péptidos.

55 De acuerdo con el primer aspecto de la presente invención, la secuencia de ácido nucleico inventivo comprende una región codificante, que codifica para un péptido o proteína que comprende un antígeno patológico o un fragmento del mismo como se reivindica. Preferentemente, el antígeno patológico codificado no es una proteína histona. En el contexto de la presente invención, esta proteína de histona típicamente es una proteína fuertemente alcalina encontrada en núcleos celulares eucariotas, que empaqueta y ordena el ADN en unidades estructurales denominadas nucleosomas. Las proteínas histona son los componentes proteínicos líderes de la cromatina, actúan como bobinas alrededor de las cuales se enrolla el ADN y desempeñan un papel en la regulación génica. Sin las histonas, el ADN no enrollado en los cromosomas podría ser muy largo (una proporción longitud:anchura de más de 10 millones a uno en el ADN humano). Por ejemplo, cada célula humana

- tiene hasta aproximadamente 1,8 metros de ADN, aunque enrollado en las histonas tiene aproximadamente 90 milímetros de cromatina, la cual, cuando se duplica y condensa durante la mitosis, da como resultado aproximadamente 120 micrómetros de cromosomas. Con mayor preferencia, en el contexto de la presente invención, esta proteína de histona típicamente se define como una proteína altamente conservada
- 5 seleccionada de una de las siguientes cinco clases principales de histonas: H1/H5, H2A, H2B, H3 y H4, preferentemente seleccionadas de histonas de mamífero, con mayor preferencia de histonas humanas o proteínas histona. Estas histonas o proteínas histona típicamente están organizadas en dos superclases definidas como histonas núcleo, que comprenden las histonas H2A, H2B, H3 y H4, e histonas ligadoras, que comprenden las histonas H1 y H5.
- 10 En este contexto, las histonas ligadoras, preferentemente excluidas del alcance de la protección de la invención pendiente, de preferencia las histonas ligadoras de mamíferos, con mayor preferencia las histonas ligadoras humanas, típicamente se seleccionan de H1, incluyendo H1F, en particular incluyendo H1F0, H1FNT, H1FOO, H1FX, y H1H1, en particular incluyendo HIST1H1A, HIST1H1B, HIST1H1C, HIST1H1D, HIST1H1E, HIST1H1T; y
- 15 Además, las histonas núcleo, de preferencia excluidas del alcance de la protección de la invención pendiente, de preferencia las histonas núcleo de mamífero, con mayor preferencia las histonas núcleo de humanos, típicamente se seleccionan de H2A, incluyendo H2AF, en particular incluyendo H2AFB1, H2AFB2, H2AFB3, H2AFJ, H2AFV, H2AFX, H2AFY, H2AFZ, y H2A1, en particular incluyendo HIST1H2AA, HIST1H2AB, HIST1H2AC, HIST1H2AD, HIST1H2AE, HIST1H2AG, HIST1H2AI, HIST1H2AJ, HIST1H2AK, HIST1H2AL,
- 20 HIST1H2AM, y H2A2, en particular incluyendo HIST2H2AA3, HIST2H2AC; H2B, incluyendo H2BF, en particular incluyendo H2BFM, H2BFO, H2BFS, H2BFWT H2B1, en particular incluyendo HIST1H2BA, HIST1H2BB, HIST1H2BC, HIST1H2BD, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2BH, HIST1H2BI, HIST1H2BJ, HIST1H2BK, HIST1H2BL, HIST1H2BM, HIST1H2BN, HIST1H2BO, y H2B2, en particular incluyendo HIST2H2BE; H3, incluyendo H3A1, en particular incluyendo HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D,
- 25 HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I, HIST1H3J, y H3A2, en particular incluyendo HIST2H3C, y H3A3, en particular incluyendo HIST3H3; H4, incluyendo H41, en particular incluyendo HIST1H4A, HIST1H4B, HIST1H4C, HIST1H4D, HIST1H4E, HIST1H4F, HIST1H4G, HIST1H4H, HIST1H4I, HIST1H4J, HIST1H4K, HIST1H4L, y H44, en particular incluyendo HIST4H4, y H5.
- De acuerdo con el primer aspecto de la presente invención, la secuencia del ácido nucleico inventivo comprende una región codificante que codifica para un péptido o proteína que comprende un antígeno patológico o un fragmento del mismo como se reivindica. De preferencia, el antígeno patológico codificado no es una proteína reportera (por ejemplo luciferasa, proteína fluorescente verde (GFP), proteína fluorescente verde intensificada (EGFP), β-galactosidasa) y no es una proteína marcadora o de selección (por ejemplo alfa-globina, galactoquinasa y xantina:guanina fosforribosiltransferasa (GPT)). De preferencia, la secuencia de
- 30 ácido nucleico de la invención no contiene un gen de resistencia a antibióticos, por ejemplo una secuencia Neogénica (bacteriana) (gen de resistencia a neomicina) o una secuencia génica CAT (que codifica para cloranfenicol acetiltransferasa; un gen de resistencia a cloranfenicol).
- El ácido nucleico inventivo como se reivindica en general comprende o codifica para a) una región codificante que codifica para un péptido o proteína que comprende un antígeno patológico o un fragmento del mismo; b)
- 40 al menos un tallo-bucle de histona y c) una secuencia poli(A); preferentemente para aumentar la expresión de dicho péptido o proteína codificado, donde el péptido o proteína codificado preferentemente no es una proteína de histona, ni una proteína reportera ni tampoco una proteína marcadora o de selección, como se definió anteriormente. Los elementos b) a c) del ácido nucleico inventivo pueden estar presentes en el ácido nucleico inventivo en cualquier orden, es decir, los elementos a), b) y c) pueden tener el orden a), b) y c) o a), c) y b)
- 45 desde la dirección 5' hasta 3' en la secuencia de ácido nucleico inventivo, donde elementos adicionales como se describen aquí también pueden estar contenidos, como una estructura 5'-CAP, una secuencia poli(C), secuencias de estabilización, secuencias IRES, etc. Cada uno de los elementos a) a c) del ácido nucleico inventivo, en particular a) en las estructuras di-cistrónicas o multi-cistrónicas, y/o cada uno de los elementos b) y c), con mayor preferencia el elemento b), también se pueden repetir al menos una vez, preferentemente dos
- 50 o más veces, en el ácido nucleico inventivo.
- Como ejemplo, el ácido nucleico inventivo puede mostrar sus elementos de secuencia a), b) y c) en, por ejemplo, el siguiente orden:
- 5' - región codificante - tallo-bucle de histona - secuencia poli(A) - 3'; o
- 5' - región codificante - tallo-bucle de la histona - señal de poliadenilación - 3' (no reivindicado); o
- 55 5' - región codificante - secuencia poli(A) - tallo-bucle de histona - 3'; o
- 5' - región codificante - señal de poliadenilación - tallo-bucle de histona - 3' (no reivindicado); o
- 5' - región codificante - región codificante - tallo-bucle de la histona - señal de poliadenilación - 3' (no reivindicado); o
- 5' - región codificante - tallo-bucle de la histona - tallo-bucle de histona - secuencia poli(A) - 3'; o

5' - región codificante - tallo-bucle de la histona - tallo-bucle de la histona - señal de poliadenilación 3' (no reivindicado); etc.

En este contexto, la secuencia del ácido nucleico inventivo comprende o codifica para a) una región codificante, que codifica para un péptido o proteína que comprende un antígeno patogénico o un fragmento del mismo; b) al menos un tallo-bucle de histona y c) un secuencia poli(A) según la reivindicación 1; preferentemente para aumentar el nivel de expresión del péptido o proteína codificada, donde la proteína codificada preferentemente no es una proteína de histona, ni una proteína reportera (por ejemplo luciferasa, GFP, EGFP,  $\beta$ -galactosidasa, en particular EGFP) y/o no es una proteína marcadora de selección (por ejemplo alfa-globina, galactocinasa y xantina:guanina fosforibosiltransferasa (GPT)).

10 El ácido nucleico inventivo como se define aquí también puede estar presente en forma de un ácido nucleico modificado.

En este contexto, la secuencia de ácido nucleico inventivo como se define aquí se puede modificar para proporcionar un "ácido nucleico estabilizado", preferentemente un ARN estabilizado, con mayor preferencia un ARN esencialmente resistente a la degradación *in vivo* (por ejemplo mediante una exo- o endo-nucleasa). Un ácido nucleico estabilizado por ejemplo se puede obtener mediante la modificación del contenido de G/C de la región codificante de la secuencia de ácido nucleico inventivo, introduciendo análogos de nucleótidos (por ejemplo nucleótidos con modificaciones estructurales, modificaciones de azúcar o modificaciones de base) o introduciendo secuencias de estabilización de la región sin traducir 3'- y/o 5'- de la secuencia de ácido nucleico inventivo.

20 Como se mencionó anteriormente, la secuencia de ácido nucleico inventivo como se define aquí puede contener nucleótidos análogos/modificaciones, por ejemplo modificaciones estructurales, modificaciones de azúcar o modificaciones de base. Una modificación estructural en relación a la presente invención es una modificación donde se modifican químicamente los fosfatos del esqueleto de los nucleótidos contenidos en la secuencia de ácido nucleico inventivo como se define aquí. Una modificación del azúcar en relación a la presente invención es una modificación química del azúcar de los nucleótidos de la secuencia de ácido nucleico inventivo como se define aquí. Además, una modificación de bases en relación a la presente invención es una modificación química de la porción base de los nucleótidos de la molécula de ácido nucleico de la secuencia de ácido nucleico inventivo. En este contexto, preferentemente los análogos de nucleótidos o modificaciones se seleccionan de análogos de nucleótidos que sean aplicables para la transcripción y/o la traducción.

30 Los análogos/modificaciones de nucleótidos aquí definidos pueden seleccionarse preferentemente de modificaciones de bases que aumentan adicionalmente la expresión de la proteína codificada y que preferentemente se seleccionan de 2-amino-6-cloropurinribosido-5'-trifosfato, 2-aminoadenosin-5'-trifosfato, 2-tiocitidin-5'-trifosfato, 2-tiouridin-5'-trifosfato, 4-tiouridin-5'-trifosfato, 5-aminoallicitidin-5'-trifosfato, 5-aminoaliluridin-5'-trifosfato, 5-bromocitidin-5'-trifosfato, 5-bromouridin-5'-trifosfato, 5-iodocitidin-5'-trifosfato, 5-iodouridin-5'-trifosfato, 5-metilcitidin-5'-trifosfato, 5-metiluridin-5'-trifosfato, 6-azacitidin-5'-trifosfato, 6-azauridin-5'-trifosfato, 6-cloropurinribosido-5'-trifosfato, 7-deazaadenosin-5'-trifosfato, 7-deazaguanosin-5'-trifosfato, 8-azaadenosin-5'-trifosfato, 8-azidoadenosin-5'-trifosfato, benzimidazol-ribosido-5'-trifosfato, N1-metiladenosin-5'-trifosfato, N1-metilguanosin-5'-trifosfato, N6-metiladenosin-5'-trifosfato, O6-metilguanosin-5'-trifosfato, pseudouridin-5'-trifosfato, o puromicin-5'-trifosfato, xantosin-5'-trifosfato. Se da particular preferencia a los nucleótidos para las modificaciones de bases seleccionadas del grupo de nucleótidos base-modificados consiste en 5-metilcitidin-5'-trifosfato, 7-deazaguanosin-5'-trifosfato, 5-bromocitidin-5'-trifosfato y pseudouridin-5'-trifosfato.

La secuencia de ácido nucleico inventivo como se define aquí puede contener una modificación de lípidos. Este ácido nucleico lípido-modificado típicamente comprende un ácido nucleico como se define aquí. Esta molécula de ácido nucleico lípido-modificado de la secuencia de ácido nucleico inventivo como se define aquí típicamente comprende además al menos un ligador que se enlaza covalentemente a esa molécula de ácido nucleico y al menos un lípido enlazado covalentemente al ligador respectivo. Alternativamente, la molécula de ácido nucleico lípido-modificada comprende al menos una molécula de ácido nucleico como se define aquí y al menos un lípido (bifuncional) enlazado covalentemente (sin ligador) a esa molécula de ácido nucleico. De acuerdo con una tercera alternativa, la molécula de ácido nucleico lípido-modificada comprende una molécula de ácido nucleico como se define aquí, al menos un ligador enlazado covalentemente a esa molécula de ácido nucleico y al menos un lípido enlazado covalentemente al ligador respectivo y al menos un lípido (bifuncional) enlazado covalentemente (sin ligador) a esa molécula de ácido nucleico. En este contexto, se prefiere particularmente que la modificación de lípidos esté presente en los extremos terminales de una secuencia lineal del ácido nucleico inventivo.

La secuencia de ácido nucleico inventivo como se define aquí, en particular si se proporciona como un ARN(m), preferentemente puede por tanto estabilizarse frente a la degradación por RNAsas mediante la adición de la denominada estructura "5' CAP".

La secuencia de ácido nucleico inventivo como se define aquí preferentemente se puede modificar mediante una secuencia de al menos 10 citidinas, de preferencia al menos 20 citidinas, con mayor preferencia al menos 30 citidinas (también denominada "secuencia poli(C)"). En particular, la secuencia de ácido nucleico inventivo puede contener o codificar para una secuencia poli(C) de típicamente aproximadamente 10 a 200 nucleótidos de citidina, de preferencia de aproximadamente 10 a 100 nucleótidos de citidina, con mayor preferencia de aproximadamente 10 a 70 nucleótidos de citidina o incluso con mayor preferencia de aproximadamente 20 a 50 o incluso 20 a 30 nucleótidos de citidina. Esta secuencia poli(C) preferentemente se ubica en la dirección 3' de la región codificante comprendida en el ácido nucleico inventivo.

El contenido de G/C de la región codificante, que codifica para al menos un péptido o proteína que comprende un antígeno patogénico o un fragmento del mismo de la secuencia de ácido nucleico inventivo como se define aquí, preferentemente está modificado, en particular aumentado, en comparación con el contenido de G/C de su región codificante de tipo salvaje particular, es decir, la región codificante no modificada. La secuencia de aminoácidos codificados de la región codificante preferentemente no está modificada en comparación con la secuencia de aminoácidos codificados de la región codificante de tipo silvestre particular.

La modificación del contenido de G/C de la región codificante de la secuencia del ácido nucleico inventivo como se define aquí se basa en el hecho de que la secuencia de cualquier región de ARNm que se traducirá es importante para una traducción eficiente del ARNm. Así, son importantes la composición y la secuencia de diversos nucleótidos. En particular, las secuencias de ARNm que tienen un contenido aumentado de G (guanósina)/C (citosina) son más estables que las secuencias de ARNm que tienen un contenido aumentado de A (adenosina)/U (uracilo). De acuerdo con la invención, por tanto los codones de la región codificante varían en comparación con su región codificante de tipo silvestre, mientras que conservan la secuencia traducida de aminoácidos, de forma que incluyen una cantidad aumentada de nucleótidos G/C. Respecto al hecho de que diversos codones codifican para uno y el mismo aminoácido (denominado degeneración del código genético), se pueden determinar los codones más favorables para la estabilidad (denominado uso de codones alternativos).

Dependiendo del aminoácido que será codificado por la región codificante de la secuencia de ácido nucleico inventivo como se define aquí, existen diversas posibilidades para la modificación de la secuencia de ácido nucleico, por ejemplo de la región codificante, en comparación con su región codificante de tipo silvestre. En caso de que los aminoácidos que están codificados por codones que contienen exclusivamente nucleótidos G o C, no es necesaria ninguna modificación del codón. Así, los codones para Pro (CCC o CCG), Arg (CGC o CGG), Ala (GCC o GCG) y Gly (GGC o GGG) no requieren modificación, ya que no esta presente A o U.

Por el contrario, los codones que contienen los nucleótidos A y/o U se pueden modificar por la sustitución por otros codones que codifican para los mismos aminoácidos pero no contienen A y/o U. Ejemplos de éstos son:

los codones para Pro se pueden modificar de CCU o CCA a CCC o CCG;  
 los codones para Arg se pueden modificar de CGU o CGA o AGA o AGG a CGC o CGG;  
 los codones para Ala se pueden modificar de GCU o GCA a GCC o GCG;  
 los codones para Gly se pueden modificar de GGU o GGA a GGC o GGG.

En otros casos, aunque los nucleótidos A o U no se pueden eliminar de los codones, es posible sin embargo disminuir el contenido de A y U utilizando codones que tienen un contenido menor de nucleótidos A y/o U. Ejemplos de éstos son:

los codones para Phe se pueden modificar de UUU a UUC;  
 los codones para Leu se pueden modificar de UUA, UUG, CUU o CUA a CUC o CUG;  
 los codones para Ser se pueden modificar de UCU o UCA o AGU a UCC, UCG o AGC;  
 el codón para Tyr se pueden modificar de UAU a UAC;  
 el codón para Cys se pueden modificar de UGU a UGC;  
 el codón para His se pueden modificar de CAU a CAC;  
 el codón para Gln se pueden modificar de CAA a CAG;  
 los codones para Ile se pueden modificar de AUU o AUA a AUC;  
 los codones para Thr se pueden modificar de ACU o ACA a ACC o ACG;  
 el codón para Asn se pueden modificar de AAU a AAC;  
 el codón para Lys se pueden modificar de AAA a AAG;  
 los codones para Val se pueden modificar de GUU o GUA a GUC o GUG;  
 el codón para Asp se pueden modificar de GAU a GAC;  
 el codón para Glu se pueden modificar de GAA a GAG;  
 el codón de terminación UAA se puede modificar para UAG o UGA.

En el caso de los codones para Met (AUG) y Trp (UGG), por otro lado, no existe la posibilidad de modificación de secuencia.

5 Las sustituciones enumeradas anteriormente se pueden utilizar ya sea individualmente o en todas las combinaciones posibles para aumentar el contenido de G/C de la región codificante de la secuencia de ácido nucleico inventivo como se define aquí en comparación con su región codificante de tipo silvestre particular (es decir, la secuencia original). Así, por ejemplo, todos los codones para Thr que están presentes en la secuencia tipo silvestre se pueden modificar a ACC (o ACG).

En el contexto anterior, se muestran los codones presentes en el ARNm. Por ejemplo, la uridina presente en un ARNm también puede estar presente como timidina en el ADN respectivo que codifica para el ARNm particular.

10 De preferencia, el contenido de G/C de la región codificante de la secuencia de ácido nucleico inventivo como se define aquí aumenta en al menos un 7%, con mayor preferencia en al menos 15%, con particular preferencia en al menos 20%, en comparación con el contenido de G/C de la región codificante de tipo silvestre. Específicamente al menos un 5%, 10%, 20%, 30%, 40%, 50%, 60%, con mayor preferencia al menos un 70%, incluso con mayor preferencia al menos un 80% y con máxima preferencia al menos un 90%, 95% o incluso el 100% de los codones que se pueden sustituir en la región codificante que codifica para al menos un péptido o  
15 proteína que comprende un antígeno patogénico o un fragmento, variante o derivado del mismo se sustituyen, aumentando así el contenido de G/C de dicha región codificante.

En este contexto, es particularmente preferente aumentar el contenido de G/C de la región codificante de la secuencia de ácido nucleico inventivo como se define aquí al máximo (es decir, el 100% de los codones que se pueden sustituir) en comparación con la región codificante tipo silvestre.

20 De acuerdo con la invención, una modificación preferida adicional de la región codificante que codifica para al menos un péptido o proteína que comprende un antígeno patogénico o un fragmento del mismo de la secuencia de ácido nucleico inventivo como se define aquí se basa en el hallazgo de que la eficacia de traducción también está determinada por una frecuencia diferente en la presencia de los ARNt en las células. De esta forma, si los denominados “codones raros” están presentes en la región codificante de la secuencia de ácido nucleico  
25 inventivo como se define aquí en un grado aumentado, la secuencia de ácido nucleico modificado correspondiente se traduce en un grado significativamente más pobre que en el caso donde están presentes codones que codifican para los ARNt relativamente “frecuentes”.

30 En este contexto, la región codificante de la secuencia de ácido nucleico inventivo preferentemente se modifica en comparación con la región codificante tipo silvestre correspondiente de forma que al menos un codón de la secuencia de tipo silvestre que codifica para un ARNt que es relativamente raro en la célula se intercambia por un codón que codifica para un ARNt que relativamente es frecuente en la célula y porta el mismo aminoácido que el ARNt relativamente raro. Mediante esta modificación, la región codificante de la secuencia de ácido nucleico inventivo, como se define aquí se modifica de forma que los codones para los cuales los ARNt se presentan frecuentemente están disponibles e insertados. En otras palabras, de acuerdo con la invención,  
35 mediante esta modificación todos los codones de la región codificante tipo silvestre que codifican para un ARNt que es relativamente raro en la célula pueden en cada caso intercambiarse para un codón que codifica para un ARNt que es relativamente frecuente en la célula y el cual, en cada caso, porta el mismo aminoácido que el ARNt relativamente raro.

40 El experto en la técnica conoce qué ARNt se presentan con relativa frecuencia en la célula y cuáles, por el contrario, se presentan con relativa rareza; véase, por ejemplo Akashi, Curr. Opin. Genet. Prog. 2001, 11(6): 660-666. En particular se prefieren los codones que utilizan, para el aminoácido particular, el ARNt que se presenta con mayor frecuencia, por ejemplo el codón Gly que utiliza el ARNt que se presenta de forma más frecuente en la célula (humana).

45 De acuerdo con la invención, es particularmente preferente unir el contenido de G/C secuencial que aumenta, en particular al máximo, en la región codificante de la secuencia de ácido nucleico inventivo como se define aquí con los codones “frecuentes” sin modificar la secuencia de aminoácidos del péptido o proteína codificado por la región codificante de la secuencia de ácido nucleico. Esta realización preferente permite la provisión de una secuencia de ácido nucleico inventivo tal como se define aquí traducida y estabilizada (modificada) particularmente eficiente.

50 La secuencia de ácido nucleico inventivo como se define aquí preferentemente puede tener adicionalmente al menos una secuencia de estabilización 5' y/o 3'. Estas secuencias de estabilización en las regiones no traducidas 5' y/o 3' tienen el efecto de aumentar la vida media del ácido nucleico, en particular del ARNm en el citosol. Estas secuencias de estabilización pueden tener un 100% de identidad de secuencia con las secuencias de origen natural que presentes en virus, bacterias y eucariotas, pero también pueden ser parcial  
55 o totalmente sintéticas. Las secuencias no traducidas (UTR) del gen (alfa-)globina, por ejemplo de *Homo sapiens* o *Xenopus laevis*, se pueden mencionar como ejemplo de secuencias de estabilización que se pueden

- utilizar en la presente invención para un ácido nucleico estabilizado. Otro ejemplo de secuencia de estabilización tiene la fórmula general (C/U)CCAN<sub>x</sub>CCC(U/A)Py<sub>x</sub>UC(C/U)CC (SEC ID NO: 55), que está contenida en los 3'-UTR de los ARN muy estables que codifican para (alfa-)globina, colágeno tipo (I), 15-lipoxigenasa o para tirosina-hidroxilasa (véase Holcik *et al.*, Proc. Natl. Acad. Sci. USA 1997, 94: 2410 a 2414).
- 5 Estas secuencias de estabilización por supuesto se pueden utilizar individualmente o en combinación con otras y también en combinación con otras secuencias de estabilización conocidas del experto en la técnica. En este contexto, es particularmente preferente que la secuencia 3' UTR del gen alfa-globina esté ubicado 3' de la región codificante que codifica para al menos un péptido o proteína que comprende un antígeno patogénico o un fragmento del mismo comprendido en la secuencia de ácido nucleico inventivo de la invención.
- 10 Preferentemente, las sustituciones, adiciones o eliminaciones de bases se llevan a cabo con la secuencia de ácido nucleico inventivo como se define aquí, utilizando una matriz de ADN para preparar la secuencia de ácido nucleico por técnicas de mutagénesis sitio dirigida, bien conocidas, o con una estrategia de ligación de oligonucleótidos (véase, por ejemplo, Maniatis *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 3rd ed., Cold Spring Harbor, NY, 2001).
- 15 Cualquiera de las modificaciones anteriores se puede aplicar a la secuencia de ácido nucleico inventivo como se define aquí y además a cualquier ácido nucleico tal como se utiliza en el contexto de la presente invención y, si es adecuado o necesario, se pueden combinar entre sí en cualquier combinación, con la condición de que estas combinaciones o modificaciones no interfieran entre sí en el ácido nucleico respectivo. El experto en la técnica puede tomar esta decisión correspondientemente.
- 20 Las secuencias de ácido nucleico utilizadas de acuerdo con la presente invención como se definen aquí se pueden preparar utilizando cualquier método conocido en la técnica, incluyendo métodos sintéticos, por ejemplo síntesis en fase sólida, así como métodos *in vitro*, como reacciones de transcripción *in vitro* o *in vivo*, tales como propagación *in vivo* de plásmidos de ADN en bacterias.
- 25 En este proceso, para preparar la secuencia de ácido nucleico inventivo como se define aquí, especialmente si el ácido nucleico está en forma de un ARNm, una molécula de ADN correspondiente se puede transcribir *in vitro*. Esta matriz de ADN preferentemente comprende un promotor adecuado, por ejemplo un promotor T7 o SP6, para la transcripción *in vitro*, que es seguido por la secuencia de nucleótidos deseada para la molécula de ácido nucleico, por ejemplo ARNm, que será preparado y una señal de terminación para la transcripción *in vitro*. La molécula de ADN, que forma la matriz del al menos un ARN de interés, se puede preparar mediante proliferación fermentativa y aislamiento posterior como parte de un plásmido que puede replicarse en bacterias.
- 30 Los plásmidos que se pueden mencionar adecuados para la presente invención son, por ejemplo, los plásmidos pTTT (número de acceso GenBank U26404; Lai *et al.*, Development 1995, 121: 2349 a 2360), la serie pGEM®, por ejemplo pGEM®-1 (número de acceso GenBank X65300; de Promega) y pSP64 (número de acceso GenBank X65327); véase también Mezei y Storts, Purification of PCR Products, en: Griffin and Griffin (ed.), PCR Technology: Current Innovation, CRC Press, Boca Raton, FL, 2001.
- 35 La secuencia de ácido nucleico inventivo como se define aquí, así como las proteínas o péptidos codificados por esta secuencia de ácido nucleico puede comprender fragmentos o variantes de aquellas secuencias. Estos fragmentos o variantes típicamente pueden comprender una secuencia que tenga una identidad de secuencia con uno de los ácidos nucleicos mencionados anteriormente, o con una de las proteínas o péptidos o secuencias, si son codificados por la secuencia de ácido nucleico inventivo, de al menos un 5%, 10%, 20%, 30%, 40%, 50%, 60%, preferentemente al menos 70%, con mayor preferencia al menos 80%, igualmente con mayor preferencia al menos 85%, incluso con mayor preferencia al menos 90% y con la máxima preferencia al menos un 95% o incluso un 97%, 98% ó 99%, con la secuencia tipo silvestre total, ya sea a nivel de ácido nucleico o a nivel de aminoácidos.
- 40 Los "fragmentos" de proteínas o péptidos en el contexto de la presente invención (por ejemplo según se codifican por la secuencia de ácido nucleico inventivo como se define aquí) pueden comprender una secuencia de una proteína o péptido como se define aquí que, con respecto a su secuencia de aminoácidos (o a su molécula de ácido nucleico codificado), en comparación N-terminal, C-terminal y/o intra-secuencialmente está truncada/acortada con respecto a la secuencia de aminoácidos de la proteína original (natural) (o su molécula de ácido nucleico codificado). Este truncamiento se puede así presentar ya sea al nivel de aminoácidos o correspondientemente a nivel de ácido nucleico. Una identidad de secuencia con respecto a este fragmento como se define aquí, por tanto, preferentemente puede hacer referencia a la proteína o péptido total como se define aquí o a la molécula de ácido nucleico total (codificante) de esta proteína o péptido. Asimismo, los "fragmentos" de ácidos nucleicos en el contexto de la presente invención pueden comprender una secuencia
- 45 de un ácido nucleico como se define aquí que, con respecto a la molécula de ácido nucleico original (natural), su molécula de ácido nucleico está truncada/acortada 5'-, 3'- y/o intra-secuencialmente. Una identidad de secuencia con respecto a este fragmento como se define aquí, por tanto, de preferencia puede hacer referencia al ácido nucleico total como se define aquí y el nivel de identidad de secuencia preferente como se indica aquí.

5 Los fragmentos tienen la misma función biológica o actividad específica o al menos conservan la actividad de la proteína de longitud total natural en al menos un 50%, con mayor preferencia al menos 70%, incluso con mayor preferencia al menos un 90% (según se mide en un ensayo funcional adecuado, por ejemplo por cuantificación de la respuesta a células B del organismo) en comparación con el péptido o proteína natural de longitud total, por ejemplo su propiedad antigénica específica. Por consiguiente, el "fragmento" es una porción de la proteína antigénica de longitud completa que tiene propiedades antigénicas sobre el sistema inmunitario tal como se describe aquí.

10 Los fragmentos de proteínas o péptidos en el contexto de la presente invención (por ejemplo, según se codifican por la secuencia de ácido nucleico inventivo como se define aquí) pueden comprender además una secuencia de una proteína o péptido como se define aquí con una longitud de aproximadamente 6 a aproximadamente 20 o incluso más aminoácidos, por ejemplo los fragmentos según se procesan y presentan mediante las moléculas clase I de MHC, preferentemente con una longitud de aproximadamente 8 a aproximadamente 10 aminoácidos, por ejemplo 8, 9, ó 10 (o incluso 6, 7, 11, ó 12 aminoácidos), o los fragmentos según se procesan y presentan mediante las moléculas clase II de MHC preferentemente con una longitud de aproximadamente 13 o más aminoácidos, por ejemplo, 13, 14, 15, 16, 17, 18, 19, 20 o incluso más aminoácidos, pudiéndose seleccionar estos fragmentos a partir de cualquier parte de la secuencia de aminoácidos. Estos fragmentos típicamente son reconocidos por las células T en forma de un complejo consistente en el fragmento peptídico y una molécula de MHC, es decir, los fragmentos típicamente no son reconocidos en su forma nativa. Los fragmentos de proteínas o péptidos según se definen aquí pueden comprender al menos un epítipo de aquellas proteínas o péptidos. Además, también se puede entender que los dominios de una proteína, como el dominio extracelular, el dominio intracelular o el dominio transmembrana, y las versiones acortadas o truncadas de una proteína comprenden un fragmento de una proteína.

25 Los fragmentos de proteínas o péptidos como se definen aquí (por ejemplo, como son codificados por la secuencia de ácido nucleico inventivo como se define aquí) también pueden comprender epítopos de esas proteínas o péptidos. Los epítopos de células T o partes de proteínas en el contexto de la presente invención pueden comprender fragmentos con una longitud preferente de aproximadamente 6 a aproximadamente 20 o incluso más aminoácidos, por ejemplo los fragmentos según se procesan y presentan por las moléculas clase I de MHC con una longitud preferente de aproximadamente 8 a aproximadamente 10 aminoácidos, por ejemplo 8, 9, ó 10 (o incluso 6, 7, 11, ó 12 aminoácidos), o fragmentos según se procesan y presentan por las moléculas clase II de MHC con una longitud preferente de aproximadamente 13 o más aminoácidos, por ejemplo 13, 14, 15, 16, 17, 18, 19, 20 o incluso más aminoácidos, pudiéndose seleccionar estos fragmentos de cualquier parte de la secuencia de aminoácidos. Estos fragmentos típicamente son reconocidos por las células T en forma de un complejo consistente en el fragmento peptídico y una molécula de MHC, es decir, los fragmentos típicamente no se reconocen en su forma nativa.

35 Los epítopos de células B típicamente son fragmentos ubicados en la superficie externa de una proteína (nativa) o antígenos peptídicos como se definen aquí preferentemente con 5 a 15 aminoácidos, con mayor preferencia de 5 a 12 aminoácidos, incluso con mayor preferencia de 6 a 9 aminoácidos, que se puedan reconocer por anticuerpos, es decir, en su forma natural.

40 Estos epítopos de proteínas o péptidos además se pueden seleccionar de cualquiera de las variantes mencionadas aquí de estas proteínas o péptidos. En este contexto, los determinantes antigénicos pueden ser epítopos conformacionales o discontinuos que están compuestos de segmentos de las proteínas o péptidos como se definen aquí que son discontinuos en la secuencia de aminoácidos de las proteínas o péptidos como se definen aquí, pero están reunidos en la estructura tridimensional o son epítopos continuos o lineales que están compuestos de una única cadena de polipéptidos.

45 Los "fragmentos" según la invención tienen una longitud de al menos 6 residuos aminoácidos y tienen una propiedad antigénica específica del péptido o la proteína nativa de longitud total.

50 Las "variantes" de proteínas o péptidos como se definen en el contexto de la presente descripción pueden ser codificadas por la secuencia de ácido nucleico descrita como se define aquí. Por ello, se puede generar una proteína o péptido con una secuencia de aminoácidos que difiera de la secuencia original en una o más mutaciones (2, 3, 4, 5, 6, 7 o más), tal como uno o más aminoácidos sustituidos, insertados y/o suprimidos. El nivel preferido de identidad de secuencia de las "variantes" en vista de la secuencia proteica de longitud total típicamente es como se indica aquí. Preferentemente, estos fragmentos y/o variantes tienen la misma función biológica o actividad específica o al menos conservan una actividad de la proteína de longitud total y natural de al menos un 50%, con mayor preferencia al menos 70%, incluso con mayor preferencia al menos 90% (según se mide en un ensayo funcional adecuado, por ejemplo por cuantificación de la respuesta inmune a células B del organismo) en comparación con el péptido o proteína natural de longitud total, por ejemplo su propiedad antigénica específica. Por consiguiente, la "variante" es típicamente una variante de la proteína antigénica de longitud total que ejerce propiedades antigénicas sobre el sistema inmune al grado que se describe aquí.

Las “variantes” de proteínas o péptidos como se definen en el contexto de la presente descripción (por ejemplo, según son codificadas por un ácido nucleico como se define aquí) pueden comprender sustituciones de aminoácidos conservativas en comparación con su secuencia fisiológica natural, es decir no mutada. Aquellas secuencias de aminoácidos, así como sus secuencias de nucleótidos codificantes, en particular, caen bajo el término de variantes como se define aquí. Las sustituciones donde los aminoácidos que se originan a partir de la misma clase se intercambian entre sí se denominan sustituciones conservativas. En particular, estos son aminoácidos que tienen cadenas laterales alifáticas, cadenas laterales cargadas positiva o negativamente, grupos aromáticos en las cadenas laterales o aminoácidos, cuyas cadenas laterales pueden participar en puentes de hidrógeno, por ejemplo cadenas laterales con una función hidroxilo. Esto significa que, por ejemplo, un aminoácido que tenga una cadena lateral polar se reemplaza por otro aminoácido que también tenga una cadena lateral polar, o, por ejemplo, un aminoácido caracterizado por una cadena lateral hidrofóbica se sustituye por otro aminoácido también con una cadena lateral hidrofóbica (por ejemplo serina (treonina) por treonina (serina) o leucina (isoleucina) por isoleucina (leucina)). En particular, son posibles inserciones y sustituciones en esas posiciones de las secuencias que no provocan una modificación a la estructura tridimensional o no afectan a la región de unión. Las modificaciones de la estructura tridimensional mediante inserciones o supresiones se pueden determinar fácilmente, por ejemplo utilizando espectros CD (espectros de dicroísmo circular) (Urry, 1985, Absorption, Circular Dichroism and ORD of Polipeptides, en: Modern Physical Methods in Biochemistry, Neuberger *et al.* (ed.), Elsevier, Amsterdam).

Además, las variantes de proteínas o péptidos como se definen aquí, que pueden ser codificadas por la secuencia de ácido nucleico descrita como se define aquí también pueden comprender aquellas secuencias donde los nucleótidos del ácido nucleico se intercambian de acuerdo con la degeneración del código genético sin conducir a una alteración de la secuencia de aminoácidos respectiva de la proteína o péptido, es decir, la secuencia de aminoácidos o al menos parte de la misma no puede diferir de la secuencia original en una o más mutaciones dentro del significado anterior.

Para determinar el porcentaje al cual son idénticas las dos secuencias, por ejemplo las secuencias de ácido nucleico o de aminoácidos como se definen aquí, preferentemente las secuencias de aminoácidos codificadas por la secuencia de ácido nucleico como se define aquí o las secuencias de aminoácidos por sí mismas, las secuencias pueden alinearse con el fin de ser comparadas posteriormente entre sí. Por tanto, por ejemplo, una posición de la primera secuencia se puede comparar con la posición correspondiente de la segunda secuencia. Si una posición en la primera secuencia está ocupada por el mismo componente, como es el caso en una posición en la segunda secuencia, ambas secuencias son idénticas en esta posición. Si este no es el caso, las secuencias difieren en esta posición. Si están presentes inserciones en la segunda secuencia en comparación con la primera secuencia, se pueden insertar separaciones en la primera secuencia para permitir otro alineamiento. Si están presentes supresiones en la segunda secuencia en comparación con la primera secuencia, se pueden insertar separaciones en la segunda secuencia para permitir otro alineamiento. El porcentaje al cual las dos secuencias son idénticas es entonces función del número de posiciones idénticas dividido entre el número total de posiciones, incluyendo aquellas posiciones que estén ocupadas únicamente en una secuencia. El porcentaje al cual las dos secuencias son idénticas se puede determinar utilizando un algoritmo matemático. Un ejemplo preferido, pero no limitativo, de algoritmo matemático que se puede utilizar es el algoritmo de Karlin *et al.* (1993), PNAS USA, 90:5873-5877 o Altschul *et al.* (1997), Nucleic Acids Res., 25:3389-3402. Este algoritmo está integrado en el programa BLAST. Las secuencias que son idénticas a las secuencias de la presente invención en cierto grado pueden identificarse con este programa.

La secuencia de ácido nucleico descrita como se define aquí puede codificar para derivados de un péptido o proteína. Este derivado de un péptido o proteína es una molécula que se deriva de otra molécula, tal como el péptido o proteína. Un “derivado” típicamente contiene la secuencia de longitud completa del péptido o proteína natural y las características adicionales de la secuencia, por ejemplo en el terminal N o C, que pueden exhibir una función adicional para el péptido/proteína de longitud completa natural. Nuevamente estos derivados tienen la misma función biológica o actividad específica o al menos conservan una actividad de la proteína de longitud total natural de al menos el 50%, con mayor preferencia al menos 70%, incluso con mayor preferencia al menos el 90% (según se mide en un ensayo funcional adecuado), por ejemplo su propiedad antigénica específica. Así, un “derivado” también abarca proteínas/péptidos de fusión (quiméricos) que comprenden un péptido o proteína utilizado en la presente descripción o una proteína de longitud completa natural (o una variante o fragmento de la misma) fusionada a un péptido/proteína distinto, otorgando por ejemplo dos o más funciones biológicas al péptido/proteína de fusión. Por ejemplo, la fusión comprende una etiqueta, por ejemplo un epítipo, por ejemplo un epítipo FLAG o un epítipo V5 o un epítipo HA. Por ejemplo, el epítipo es un epítipo FLAG. Esta etiqueta es útil para, por ejemplo, purificar la proteína de fusión.

En este contexto, una “variante” de una proteína o péptido puede tener al menos un 70%, 75%, 80%, 85%, 90%, 95%, 98% o 99% de identidad de aminoácidos con respecto a un tramo de 10, 20, 30, 50, 75 o 100 aminoácidos de esta proteína o péptido. Análogamente, una “variante” o particularmente un “fragmento” de una secuencia de ácido nucleico pueden tener al menos un 70%, 75%, 80%, 85%, 90%, 95%, 98% o 99% de



identidad de nucleótidos con respecto a un tramo de 10, 20, 30, 50, 75 ó 100 nucleótidos de esta secuencia de ácido nucleico; típicamente, sin embargo, en referencia a las secuencias de longitud completa de origen natural. En el caso de los "fragmentos" típicamente la identidad de secuencia se determina para el fragmento sobre la longitud (del fragmento) de la porción de proteína de longitud completa (reflejando la misma longitud que el fragmento) que tiene el mayor nivel de identidad de secuencia.

Preferentemente, la secuencia de ácido nucleico inventivo puede estar asociada a un vehículo, agente de transfección o complejación para aumentar la eficiencia de transfección y/o las propiedades inmunoestimuladoras de la secuencia de ácido nucleico. Agentes adecuados particularmente preferentes en este contexto para aumentar la eficiencia de transfección son compuestos catiónicos o policatiónicos, incluyendo protamina, nucleolina, espermina o espermidina, u otros péptidos o proteínas catiónicos, tales como poli-L-lisina (PLL), poli-arginina, polipéptidos básicos, péptidos de penetración celular (CPP), incluyendo péptidos de unión a VIH, VIH-1 Tat (VIH), péptidos derivados de Tat, Penetratina, péptidos derivados o análogos de VP22 (Herpes simplex), MAP, KALA o dominios para transducción proteínica (PTD), PpT620, péptidos ricos en prolina, péptidos ricos en arginina, péptidos ricos en lisina, MPG-péptidos, Pep-1, L-oligómeros, péptidos de calcitonina, péptidos derivados de Antennapedia (en particular de *Drosophila antennapedia*), pAntp, plsl, FGF, Lactoferrina, Transportana, Buforin-2, Bac715-24, SynB, SynB(1), pVEC, péptidos derivados de hCT, SAP o histonas. Adicionalmente, proteínas o péptidos catiónicos o policatiónicos preferentes se pueden seleccionar de las siguientes proteínas o péptidos de la siguiente fórmula total:  $(Arg)_l; (Lys)_m; (His)_n; (Orn)_o; (Xaa)_x$ , donde  $l + m + n + o + x = 8-15$ , y  $l, m, n, o$  o independientemente entre si puede ser cualquier número seleccionado de 0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 ó 15, con la condición de que el contenido total de Arg, Lys, His y Orn represente al menos el 50% de todos los aminoácidos del oligopéptido; y Xaa puede ser cualquier aminoácido seleccionado de aminoácidos naturales (= de origen natural) o no naturales excepto Arg, Lys, His o Orn; y  $x$  puede ser cualquier número seleccionado de 0, 1, 2, 3 ó 4, con la condición de que el contenido total de Xaa no exceda el 50% de todos los aminoácidos del oligopéptido. Péptidos catiónicos particularmente preferentes en este contexto son, por ejemplo, Arg<sub>7</sub>, Arg<sub>8</sub>, Arg<sub>9</sub>, H<sub>3</sub>R<sub>9</sub>, R<sub>9</sub>H<sub>3</sub>, H<sub>3</sub>R<sub>9</sub>H<sub>3</sub>, YSSR<sub>9</sub>SSY, (RKH)<sub>4</sub>, Y(RKH)<sub>2</sub>R, etc. Compuestos catiónicos o policatiónicos preferentes adicionales que se pueden utilizar como agente de transfección pueden incluir polisacáridos catiónicos, por ejemplo quitosano, polibreno, polímeros catiónicos, por ejemplo polietilenimina (PEI), lípidos catiónicos, por ejemplo DOTMA: cloruro de [1-(2,3-sioleiloxi)propil]-N,N,N-trimetilamonio, DMRIE, di-C14-amidina, DOTIM, SAINT, DC-Chol, BGTC, CTAP, DOPC, DODAP, DOPE: Dioleil-fosfatidiletanol-amina, Dospa, DODAB, DOIC, DMEPC, DOGS: Dioctadecilamidoglicilespermina, DIMRI: bromuro de dimiristil-oxipropildimetilhidroxietilamonio, DOTAP: dioleoiloxi-3-(trimetilamonio)-propano DC-6-14: cloruro de O,O-ditetradecanoil-N-( $\alpha$ -trimetilamonioacetil)-dietanolamina, CLIP1: cloruro de rac-[(2,3-dioctadeciloxipropil)(2-hidroxietil)]-dimetilamonio, CLIP6: rac-[2(2,3-dihexadeciloxipropil-oximetiloxi)etil]-trimetilamonio, CLIP9: rac-[2(2,3-dihexadeciloxipropil-oxisucciniloxi)etil]-trimetilamonio, oligofectamina, o polímeros catiónicos o policatiónicos, por ejemplo poliaminoácidos modificados, tales como polímeros de  $\beta$ -aminoácidos o poliamidas inversas, etc., polietilenos modificados, como PVP (bromuro de poli(N-etil-4-vinilpiridinio)), etc., acrilatos modificados, como pDMAEMA (metilacrilato de poli(dimetilaminoetilo)), etc., amidoaminas modificadas como pAMAM (poli(amidoamina)), etc., polibeta-aminoéster modificado (PBAE), como polímeros de 1,4-butandioldiacrilato-co-5-amino-1-pentanol modificados en el extremo de diamina, etc., dendrímeros, como dendrímeros de polipropilamina o dendrímeros a base de pAMAM, etc., poliiiminas, como PEI: poli(etilenimina), poli(propilenimina), etc., polialilamina, polímeros basados en una estructura de azúcar, como polímeros basados en ciclodextrina, dextrano, quitosano, etc., polímeros basados en una estructura silano, como copolímeros PMOXA-PDMS, etc., polímeros en bloque consistentes en una combinación de uno o más bloques catiónicos (por ejemplo seleccionados de un polímero catiónico, como se mencionó anteriormente) y uno o más bloques hidrofílicos o hidrofóbicos (por ejemplo polietilenglicol); etc.

El ácido nucleico inventivo preferentemente puede estar complejado al menos parcialmente con un compuesto catiónico o policatiónico, preferentemente proteínas o péptidos catiónicos. Parcialmente significa que sólo una parte del ácido nucleico está complejado con un compuesto catiónico o policatiónico y que el resto del ácido nucleico está en forma no complejada ("libre"). Preferentemente, la proporción ácido nucleico complejado:ácido nucleico libre se selecciona en un rango de aproximadamente 5:1 (p/p) a aproximadamente de 1:10 (p/p), con mayor preferencia de aproximadamente 4:01 (p/p) a aproximadamente 1:8 (p/p), incluso con mayor preferencia de aproximadamente 3:01 (p/p) a aproximadamente 1:5 (p/p) o 1:3 (p/p) y con la máxima preferencia la proporción ácido nucleico complejado:ácido nucleico libre se selecciona de una proporción aproximadamente 1:1 (p/p).

La secuencia de ácido nucleico de la invención preferentemente puede proporcionarse en forma ya sea pura o complejada, por ejemplo con compuestos policatiónicos de cualquier estructura química, preferentemente (poli)péptidos policatiónicos o compuestos policatiónicos sintéticos. De preferencia, la secuencia de ácido nucleico no se proporciona junto con una célula de empaquetamiento.

En otro aspecto, la invención proporciona una composición o un kit o kit de partes que comprende una pluralidad o más de una, de preferencia 2 a 10, con mayor preferencia 2 a 5, con máxima preferencia 2 a 4 de las secuencias de ácido nucleico inventivo como se define aquí. Estas composiciones inventivas comprenden más de una secuencia de ácido nucleico inventivo, preferentemente codificando para diferentes péptidos o proteínas que preferiblemente comprenden diferentes antígenos patogénicos o fragmentos de los mismos.

De acuerdo con otro aspecto, la presente invención también proporciona un método *in vitro* o *ex vivo* para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de, por ejemplo, a) proporcionar la secuencia de ácido nucleico inventivo como se define aquí o la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí, b) aplicar o administrar la secuencia de ácido nucleico inventivo como se define aquí o la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí a un sistema de expresión, por ejemplo un sistema de expresión libre de células, una célula (por ejemplo una célula huésped para la expresión o una célula somática) o un tejido. El método *in vitro* o *ex vivo* se puede aplicar en laboratorio, para la investigación, para el diagnóstico, para la producción comercial de péptidos o proteínas y/o con fines terapéuticos. En este contexto, típicamente después de preparar la secuencia de ácido nucleico inventivo como se define aquí o la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí típicamente se aplica o administra a un sistema de expresión libre de células, a una célula (por ejemplo una célula huésped de expresión o una célula somática) o un tejido, por ejemplo en forma pura o complejada, o como una composición farmacéutica o vacuna como se describe aquí, de preferencia vía transfección o empleando cualquiera de los modos de administración aquí descritos. El método se lleva a cabo *in vitro* o *ex vivo*.

En este contexto, *in vitro* se define aquí como la transfección o transducción del ácido nucleico inventivo como se define aquí o de la composición de la invención que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí en células cultivadas fuera de un organismo; *in vivo* se define aquí como la transfección o transducción del ácido nucleico inventivo o de la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas en células mediante la aplicación del ácido nucleico inventivo o de la composición inventiva a todo el organismo o individuo y *ex vivo* se define en aquí como la transfección o transducción del ácido nucleico inventivo o de la composición de la invención que comprende una pluralidad de secuencias de ácido nucleico inventivas en células fuera de un organismo o individuo y posterior aplicación de las células transfectadas al organismo o individuo.

Asimismo, de acuerdo con otro aspecto, la presente invención también proporciona el uso de la secuencia de ácido nucleico inventivo como se define aquí o de la composición de la invención que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí, preferentemente con fines diagnósticos o terapéuticos o para aumentar la expresión de un péptido o proteína codificado *in vitro*, por ejemplo aplicando o administrando la secuencia de ácido nucleico inventivo como se define aquí o la composición de la invención que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí por ejemplo a un sistema de expresión libre de células, a una célula (por ejemplo una célula huésped de expresión o una célula somática) o a un tejido. El uso se puede aplicar para el laboratorio, en investigación, diagnóstico, para la producción comercial de péptidos o proteínas y/o con fines terapéuticos. En este contexto, típicamente después de preparar la secuencia de ácido nucleico inventivo como se define aquí o la composición de la invención que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí típicamente su uso requiere una aplicación o administración a un sistema de expresión libre de células, una célula (por ejemplo una célula huésped de expresión o una célula somática), un tejido o un organismo, de preferencia en forma pura o complejada, o como una composición farmacéutica o vacuna como se describe aquí, preferentemente vía transfección o utilizando cualquiera de los modos de administración aquí descritos. El uso se puede llevar a cabo *in vitro*, *in vivo* (para su uso como medicamento o vacuna) o *ex vivo*. El uso además se puede llevar a cabo en el contexto del tratamiento de una enfermedad específica, en particular en el tratamiento de enfermedades infecciosas preferentemente como se definen aquí.

Todavía en otro aspecto (no reivindicado), la presente descripción también se refiere a un sistema de expresión que comprende una secuencia de ácido nucleico o un vector o plásmido de expresión de acuerdo con el primer aspecto de la presente descripción. En este contexto, el sistema de expresión puede ser un sistema de expresión libre de células (por ejemplo un sistema de transcripción/traducción *in vitro*), un sistema de expresión celular (por ejemplo células CHO similares a células mamíferas, células de insecto, células de levadura, células bacterianas similares a *E. coli*) u organismos utilizados para la expresión de péptidos o proteínas (por ejemplo plantas o animales como vacas).

Adicionalmente, de acuerdo con otro aspecto, la presente invención también se refiere al uso del ácido nucleico inventivo como se define aquí o de la composición de la invención que comprende una pluralidad de secuencias de ácido nucleico inventivas (lo que significa típicamente más de 1, 2, 3, 4, 5, 6 o más de 10 ácidos nucleicos, por ejemplo 2 a 10, de preferencia 2 a 5 ácidos nucleicos) como se definen aquí para la preparación de una

composición farmacéutica para aumentar la expresión de un péptido o proteína codificado como un medicamento o vacuna, por ejemplo para tratar una enfermedad infecciosa, preferentemente tal como se define aquí, por ejemplo aplicando o administrando el ácido nucleico inventivo como se define aquí o la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí a una célula (por ejemplo una célula huésped de expresión o una célula somática), un tejido o un organismo, preferentemente en forma pura o complejada o como una composición farmacéutica o vacuna como se describe aquí, con mayor preferencia utilizando cualquiera de los modos de administración aquí descritos.

Por consiguiente, en otro aspecto, la presente invención también proporciona una composición farmacéutica que comprende un ácido nucleico inventivo como se define aquí o una composición de la invención que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí y opcionalmente un portador y/o vehículo farmacéuticamente aceptable.

Como un primer ingrediente, la composición farmacéutica de la invención comprende al menos un ácido nucleico inventivo como se define aquí.

Como un segundo ingrediente de la composición farmacéutica de la invención opcionalmente puede comprender al menos un componente farmacéuticamente activo adicional. Un componente farmacéuticamente activo a este respecto es un compuesto que tiene un efecto terapéutico de sanar, disminuir o evitar una indicación o enfermedad como se mencionan aquí, preferentemente enfermedades infecciosas. Estos compuestos incluyen, sin implicar ninguna limitación, péptidos o proteínas, de preferencia como se definen aquí, ácidos nucleicos, preferentemente tal como se definen aquí, compuestos orgánicos o inorgánicos de bajo peso molecular (terapéuticamente activos) (peso molecular inferior a 5.000, con preferencia inferior a 1.000), azúcares, antígenos o anticuerpos, de preferencia como se definen aquí, agentes terapéuticos ya conocidos en la técnica anterior, células antigénicas, fragmentos celulares antigénicos, fracciones celulares; componentes de la pared celular (por ejemplo polisacáridos), patógenos modificados, atenuados o desactivados (por ejemplo químicamente o por irradiación) (virus, bacterias, etc.), adyuvantes, de preferencia como se definen aquí, etc.

Además, la composición farmacéutica inventiva puede comprender un portador y/o vehículo farmacéuticamente aceptable. En el contexto de la presente invención, un portador farmacéuticamente aceptable típicamente incluye una base líquida o no líquida de la composición farmacéutica inventiva. Si la composición farmacéutica inventiva se proporciona en forma líquida, el portador típicamente será agua libre de pirógenos, solución salina isotónica o soluciones tampón (acuosas), por ejemplo fosfato, citrato, etc., soluciones tampón. El tampón de inyección puede ser hipertónico, isotónico o hipotónico en relación al medio de referencia específico, es decir, el tampón puede tener mayor, idéntico o menor contenido en sal en relación al medio de referencia específico, donde preferentemente estas concentraciones de sales mencionadas anteriormente se pueden utilizar pero sin conducir a daño celular debido a ósmosis u otros efectos de la concentración. Medios de referencia son, por ejemplo, líquidos que se emplean en métodos "in vivo", como sangre, linfa, líquidos citosólicos u otros líquidos corporales, o por ejemplo líquidos que se puedan utilizar como medios de referencia en métodos "in vitro", tales como tampones o líquidos comunes. Estos tampones o líquidos comunes son conocidos por el experto. En particular se prefiere como base líquida una solución Ringer-Lactato.

Sin embargo, se pueden utilizar uno o más materiales de carga o diluyentes sólidos o líquidos compatibles o compuestos encapsulantes para la composición farmacéutica inventiva, que sean adecuados para la administración a un paciente a tratar. El término "compatible" tal como se emplea aquí significa que estos constituyentes de la composición farmacéutica inventiva son capaces de mezclarse con el ácido nucleico inventivo como se define aquí de forma que no existe una interacción que pueda reducir sustancialmente la efectividad farmacéutica de la composición farmacéutica inventiva bajo las condiciones típicas de uso.

La composición farmacéutica inventiva preferentemente puede comprender un adyuvante. En este contexto, se puede entender como adyuvante cualquier compuesto adecuado para iniciar o aumentar una respuesta inmune del sistema inmune innato, es decir una respuesta inmune no específica. En otras palabras, cuando se administra, la composición farmacéutica inventiva preferentemente produce una respuesta inmune innata debido al adyuvante opcionalmente contenido en la misma. De preferencia, este adyuvante se puede seleccionar de un adyuvante conocido por el experto y adecuado para el caso presente, es decir que soporte la inducción de una respuesta inmune innata en un mamífero, por ejemplo una proteína adyuvante como se definió arriba o un adyuvante como se define en lo que sigue.

Como adyuvantes adecuados se prefieren en particular compuestos catiónicos o policatiónicos para el depósito y suministro como se describieron anteriormente para la secuencia de ácido nucleico inventivo como vehículo, agente de transfección o de complejación.

- La composición farmacéutica inventiva puede contener además una o más sustancias auxiliares para aumentar su inmunogenicidad o capacidad inmunoestimuladora, si se desea. Preferentemente así se alcanza una acción sinérgica de la secuencia de ácido nucleico inventivo como se define aquí y una sustancia auxiliar que puede estar contenida opcionalmente en la composición farmacéutica inventiva. Dependiendo de los diversos tipos de sustancias auxiliares, a este respecto entran en consideración diversos mecanismos. Por ejemplo, los compuestos que permitan la maduración de células dendríticas (DC), por ejemplo lipopolisacáridos, ligandos TNF-alfa o CD40, forman una primera clase de sustancias auxiliares adecuadas. En general, es posible utilizar como sustancia auxiliar cualquier agente que influye en el sistema inmune en forma de una "señal de peligro" (LPS, gp96, etc.) o citoquinas, como GM-CFS, que permiten que una respuesta inmune mejore y/o esté influida de forma dirigida. Sustancias auxiliares particularmente preferidas son citoquinas, como monoquinas, linfoquinas, interleuquinas o quimiocinas, que estimulan adicionalmente la respuesta inmune innata, tal como IL-1, IL-2, IL-3, IL-4, IL-5, IL-6, IL-7, IL-8, IL-9, IL-10, IL-12, IL-13, IL-14, IL-15, IL-16, IL-17, IL-18, IL-19, IL-20, IL-21, IL-22, IL-23, IL-24, IL-25, IL-26, IL-27, IL-28, IL-29, IL-30, IL-31, IL-32, IL-33, IFN-alfa, IFN-beta, IFN-gamma, GM-CSF, G-CSF, M-CSF, LT-beta o TNF-alfa, factores de crecimiento, tales como hGH.
- 5
- 10
- 15 Otros aditivos que pueden estar incluidos en la composición farmacéutica inventiva son emulsionantes, por ejemplo Tween®; agentes humectantes, por ejemplo laurilsulfato de sodio; agentes colorantes; agentes saborizantes, portadores farmacéuticos; agentes formadores de tabletas; estabilizantes; antioxidantes; conservantes.
- 20 La composición farmacéutica inventiva también puede contener adicionalmente cualquier compuesto adicional conocido por ser inmunoestimulador debido a su afinidad de unión (como ligando) a los receptores similares a Toll humanos TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, o debido a su afinidad de unión (como ligando) con los receptores similares a Toll-murino TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, TLR11, TLR12 o TLR13.
- 25 La composición farmacéutica inventiva se puede administrar vía oral, parenteral, inhalación de aerosol, tópica, rectal, nasal, bucal, vaginal o vía un depósito implantado. El término parenteral como se utiliza aquí incluye técnicas de inyección o infusión subcutáneas, intravenosas, intramusculares, intra-articulares, intra-sinoviales, intraesternales, intratecales, intrahepáticas, intralesionales, intracraneales, transdérmicas, intradérmicas, intrapulmonales, intraperitoneales, intracardiacas, intra-arteriales y sublinguales.
- 30 Preferentemente, la composición farmacéutica inventiva se puede administrar por inyección parenteral, con mayor preferencia por inyección subcutánea, intravenosa, intramuscular, intra-articular, intra-sinovial, intraesternal, intratecal, intrahepática, intralesional, intracraneal, transdérmica, intradérmica, intrapulmonal, intraperitoneal, intracardiaca, intraarterial, y sublingual o vía técnicas de infusión. Es particularmente preferente la inyección intradérmica e intramuscular. Las formas inyectables estériles de las composiciones farmacéuticas de la invención pueden ser una suspensión acuosa u oleosa. Estas suspensiones se pueden formular de acuerdo con técnicas conocidas en la técnica utilizando agentes de dispersión o humectantes adecuados y agentes de suspensión.
- 35
- La composición farmacéutica inventiva como se define aquí también se puede administrar vía oral en cualquier forma de dosificación oralmente aceptable, incluyendo, de manera no limitativa, cápsulas, tabletas, suspensiones o soluciones acuosas.
- 40 La composición farmacéutica inventiva también se puede administrar vía tópica, en especial cuando la diana de tratamiento incluye áreas u órganos a los que se tenga fácil acceso por una aplicación tópica, por ejemplo incluyendo enfermedades de la piel o de cualquier otro tejido epitelial accesible. Las formulaciones tópicas adecuadas se preparan fácilmente para cada una de estas áreas u órganos. Para aplicaciones tópicas, la composición farmacéutica inventiva se puede formular en un ungüento adecuado que contenga el ácido nucleico inventivo como se define aquí, suspendido o disuelto en uno o más portadores.
- 45
- La composición farmacéutica inventiva comprende típicamente una "cantidad segura y efectiva" de los componentes de la composición farmacéutica inventiva, en particular de las secuencias de ácido nucleico inventivas tal como se definen aquí. Tal como se emplea aquí, una "cantidad segura y efectiva" significa una cantidad de secuencias de ácido nucleico inventivas como se definen aquí que sea suficiente para inducir significativamente una modificación positiva de una enfermedad o trastorno como se define aquí. Sin embargo, al mismo tiempo, una "cantidad segura y efectiva" es lo suficientemente pequeña para evitar efectos secundarios graves, permitiendo una relación sensible entre ventaja y riesgo. La determinación de estos límites queda dentro del alcance del juicio médico sensible.
- 50
- 55 La composición farmacéutica inventiva se puede utilizar para un ser humano y también con fines médicos veterinarios, preferentemente con fines médicos humanos, como una composición farmacéutica en general o como una vacuna.

De acuerdo con otro aspecto particularmente preferente, la composición farmacéutica inventiva (o la secuencia de ácido nucleico inventivo como se define aquí o la composición de la invención que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí) se puede proporcionar o utilizar como una vacuna. Típicamente, esta vacuna es como se definió anteriormente para las composiciones farmacéuticas.  
 5 Adicionalmente, esta vacuna típicamente contiene el ácido nucleico inventivo como se define aquí o la composición de la invención que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí.

La vacuna inventiva también puede comprender un portador, adyuvante y/o vehículo farmacéuticamente aceptable como se definen aquí para la composición farmacéutica inventiva. En el contexto específico de la  
 10 vacuna inventiva, la selección de un portador farmacéuticamente aceptable se determina, en principio, por la forma en que se administra la vacuna. La vacuna inventiva se puede administrar, por ejemplo, sistémica o localmente. Las vías de administración sistémica en general incluyen, por ejemplo, las vías transdérmica, oral, parenteral, incluyendo subcutánea, intravenosa, intramuscular, intraarterial, intradérmica e inyecciones intraperitoneales y/o vías de administración intranasal. Las vías para la administración local en general incluyen,  
 15 por ejemplo, vías de administración tópica, aunque también intradérmica, transdérmica, subcutánea o inyecciones intramusculares o intralesionales, intracraneales, intrapulmonales, intracárdicas y sublinguales. Con mayor preferencia, las vacunas se pueden administrar por vía intradérmica, subcutánea o intramuscular. Por tanto, las vacunas inventivas preferentemente se formulan en forma líquida (o a veces en forma sólida).

La vacuna inventiva adicionalmente puede contener una o más sustancias auxiliares para aumentar su inmunogenicidad o capacidad inmunoestimuladora, si se desea. Son particularmente preferentes los adyuvantes como sustancias auxiliares o aditivos como se definen para la composición farmacéutica.  
 20

La presente invención proporciona además diversas aplicaciones y usos de la secuencia de ácido nucleico inventivo como se define aquí, de la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí, de la composición farmacéutica inventiva o de la vacuna inventiva, todas comprendiendo la secuencia de ácido nucleico inventivo como se define aquí.  
 25

De acuerdo con un aspecto específico, la presente invención se dirige al primer uso médico de la secuencia de ácido nucleico inventivo como se define aquí o de la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí como medicamento, preferentemente como una vacuna particularmente en el tratamiento de enfermedades infecciosas.

De acuerdo con otro aspecto, la presente invención se dirige al segundo uso médico de la secuencia de ácido nucleico inventivo como se define aquí o de la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí, para el tratamiento de enfermedades infecciosas como se definen aquí, de preferencia para el uso de la secuencia de ácido nucleico inventivo como se define aquí, de la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí, de una composición farmacéutica o vacuna que los comprende para la preparación de un medicamento para la profilaxis, tratamiento y/o mejora de enfermedades infecciosas como se definen aquí. De preferencia, la composición farmacéutica o una vacuna se utiliza o se administra a un paciente que lo necesite para este fin.  
 30  
 35

Preferentemente, las enfermedades infecciosas como se mencionan aquí se seleccionan de enfermedades infecciosas virales, bacterianas, protozoológicas y priónicas. Estas enfermedades infecciosas típicamente se seleccionan de la lista consistente en infecciones por Acinetobacter, enfermedad africana del sueño (tripanosomiasis africana), SIDA (síndrome de inmunodeficiencia adquirida), Amoebiasis, Anaplasmosis, Anthrax, Appendicitis, infecciones por Arcanobacterium haemolyticum, fiebre hemorragia Argentina, Ascariasis, Aspergillosis, infecciones por Astrovirus, pie de atleta, Babesiosis, infecciones por Bacillus cereus, meningitis  
 40 bacteriana, neumonía bacteriana, vaginosis bacteriana (BV), infecciones Bacteroides, Balantidiasis, infecciones por Baylisascaris, Bilharziosis, infecciones por virus BK, piedra negra, infecciones por Blastocystis hominis, Blastomycosis, fiebre hemorrágica boliviana, infecciones por Borrelia (Borreliosis), Botulismo (y botulismo infantil), cestodo bovino, fiebre hemorrágica brasileña, Brucellosis, infecciones por Burkholderia, úlcera de Buruli, infecciones por Calicivirus (Norovirus y Sapovirus), Campylobacteriosis, Candidiasis (Candidosis), infecciones por cestodo canino, enfermedad por arañazo de gato, enfermedad de Chagas (tripanosomiasis Americana), Chancroide, varicela, infecciones por clamida, infecciones por Chlamydia trachomatis, infecciones por Chlamyphila pneumoniae, Cólera, Chromoblastomycosis, bubón climático, Clonorchiasis, infecciones por Clostridium difficile, Coccidioidomycosis, catarro, fiebre por garrapatas de Colorado (CTF), resfriado común (rinofaringitis viral aguda; coryza agudo), Condyloma acuminata, conjuntivitis,  
 45  
 50 enfermedad de Creutzfeldt-Jakob (CJD), fiebre hemorrágica de Crimean-Congo (CCHF), Cryptococcosis, Cryptosporidiosis, larva migran cutánea (CLM), Leishmaniosis cutánea, Ciclosporiosis, Cisticercosis, infecciones por citomegalovirus, fiebre del dengue, dermatofitosis, Dientamoebiasis, Difteria, Difilobotriasis, Donavanosis, Dracunculiasis, meningoencefalitis de verano temprano (FSME), fiebre hemorrágica por Ebola,  
 55

Echinococcosis, Ehrlichiosis, Enterobiasis (infecciones por Pinworm), infecciones por Enterococcus, infecciones por Enterovirus, tífus epidémica, Epiglottitis, Mononucleosis infecciosa por el virus de Epstein-Barr, Eritema infeccioso (enfermedad Fifth), Exanthem subitum, Fasciolopsiasis, Fasciolosis, insomnio familiar fatal (FFI), enfermedad de Fifth, Filariasis, envenenamiento por pescado (Ciguatera), cestodo de pescado, gripe, 5 envenenamiento alimentario por Clostridium perfringens, cestodo de zorro, infecciones amebianas de crecimiento libre, infecciones por Fusobacterium, gangrena gaseosa, Geotricosis, síndrome de Gerstmann-Sträussler-Scheinker (GSS), Giardiasis, Glanders, Gnathostomiasis, Gonorrhea, Granuloma inguinal (Donovanosis), infecciones por estreptococos del Grupo A, infecciones por estreptococos Grupo B, infecciones por Haemophilus influenzae, enfermedad de manos, pies y boca (HFMD), síndrome pulmonar por Hantavirus 10 (HPS), infecciones por Helicobacter pylori, síndrome hemolítico-urémico (HUS), Fiebre hemorrágica con síndrome renal (HFRS), infecciones por Henipavirus, Hepatitis A, Hepatitis B, Hepatitis C, Hepatitis D, Hepatitis E, Herpes simplex, Herpes simplex tipo I, Herpes simplex tipo II, Herpes zoster, Histoplasmosis, verrugas huecas, infecciones por anquilostoma, infecciones por bocavirus humano, erliquiosis ewingii humana, anaplasmosis granulocítica humana (HGA), infecciones por metapneumovirus humano, erliquiosis monocítico humano, infecciones por papilomavirus Humano (HPV), infecciones por virus de parainfluenza humana, 15 Hymenolepiasis, Influenza, Iosporiasis, encefalitis Japonesa, enfermedad de Kawasaki, Keratitis, infecciones por Kingella kingae, Kuru, Lambliasis (Giardiasis), fiebre de Lassa, Legionellosis (enfermedad del Legionario, fiebre Pontiac), Leishmaniasis, Lepra, Leptospirosis, piojo, Listeriosis, Lyme borreliosis, enfermedad de Lyme, filariasis Linfática (Elefantiasis), coriomeningitis linfocítica, Malaria, fiebre hemorrágica por Marburg (MHF), virus de Marburg, Sarampión, Melioidosis (enfermedad de Whitmore), Meningitis, enfermedad Meningococcal, Metagonimiasis, Microsporidiosis, cestodo miniatura, miscarriage (inflamación de la próstata), Molluscum contagiosum (MC), Mononucleosis, paperas, tífus murino (tífus endémico), Mycetoma, Mycoplasma hominis, Mycoplasma pneumonia, Myiasis, dermatitis por pañal desechable/pañal de tela, conjuntivitis neonatal (Ophthalmia neonatorum), sepsis neonatal (Chorioamnionitis), Nocardiosis, Noma, infecciones por virus 20 Norwalk, Oncocercosis (ceguera de los ríos), Osteomielitis, Otitis media, Paracoccidioidomycosis (blastomicosis de Sudamérica), Paragonimiasis, Paratyphus, Pasteurellosis, Pediculosis capitis (Head lice), Pediculosis corporis (piojos del cuerpo), Pediculosis pubis (piojo púbico, liendre), enfermedad inflamatoria pélvica (PID), Pertussis (tosferina), fiebre glandular de Pfeiffer, plaga, infecciones por pneumococcus, Pneumocystis pneumonia (PCP), Pneumonia, Polio (cojera infantil), Poliomieltitis, cestodo porcino, infecciones por Prevoleta, meningoencefalitis amebiana primaria (PAM), leucoencefalopatía multifocal progresiva, Pseudo-croup, Psittacosis, fiebre Q, fiebre de conejo, Rabia, fiebre por mordida de rata, síndrome de Reiter, infección por virus sincitial respiratorio (RSV), Rhinosporidiosis, infecciones por Rinovirus, infecciones por Rickettsial, Rickettsialpox, fiebre amarilla (RVF), fiebre manchada de las montañas rocosas (RMSF), infecciones por Rotavirus, Rubeola, Salmonella paratyphus, Salmonella typhus, Salmonellosis, SARS (Síndrome Respiratorio 30 Agudo Grave) sarna, escarlatina, Esquistosomiasis (Bilharziosis), Scrub typhus, Sepsis, Shigellosis (disintiera basilar), Shingles, Smallpox (Viruela), chancro suave, Sporotricosis, envenenamiento alimentario por estafilococos, infecciones por estafilococos, Strongyloidiasis, sífilis, Taeniasis, Tétanos, fiebre de los tres días, encefalitis portada por garrpatas, Tinea barbae (picazón de Barber), Tiña del cuero cabelludo (Tiña del cuerpo), Tiña corporis (tiña del cuerpo), Tiña cruris (picazón Jock), Tiña de la mano (Tiña de la mano), Tiña nigra, Tiña pedis (pie de atleta), Tiña unguium (Onicomicosis), Tiña versicolor (Pityriasis versicolor), Toxocariasis (Larva Migrans Ocular (OLM) y Larva Migrans Visceral (VLM)), Toxoplasmosis, Triquinelosis, Tricomoniasis, Tricuriasis (infecciones del intestino grueso), Tripper, Trypanosomiasis (enfermedad del sueño), enfermedad Tsutsugamushi, Tuberculosis, Tularemia, Tifo, fiebre por tifo, infecciones por Ureaplasma urealyticum, Vaginitis (Colpitis), enfermedad de Creutzfeldt-Jakob variante (vCJD, nCJD), encefalitis equina 45 Venezolana, fiebre hemorrágica Venezolana, neumonía viral, Leishmaniosis visceral, verrugas, Fiebre del Nilo occidental, encefalitis equina occidental, piedra blanca (Tiña blanca), tosferina, puntos por hongos de levadura, fiebre amarilla, infecciones por Yersinia pseudotuberculosis, Yersiniosis y Zigomicosis.

En otro aspecto preferido, la secuencia de ácido nucleico inventivo como se define aquí o la composición inventiva que comprende una pluralidad de secuencias de ácido nucleico inventivas como se definen aquí se puede utilizar para la preparación de una composición farmacéutica o una vacuna, en particular para los fines aquí definidos. 50

La composición farmacéutica inventiva o la vacuna se pueden utilizar además para el tratamiento de una enfermedad o un trastorno, preferentemente de enfermedades infecciosas como se definen aquí.

De acuerdo con un aspecto final, la presente invención también proporciona kits, en particular kits de partes. 55 Estos kits, en particular kits de partes, típicamente comprenden como componentes solos o en combinación con componentes adicionales tal como se definen aquí al menos una secuencia de ácido nucleico inventivo como se define aquí, la composición farmacéutica inventiva o vacuna que comprende la secuencia de ácido nucleico inventivo. La al menos una secuencia de ácido nucleico inventivo como se define aquí está por ejemplo opcionalmente combinada con componentes adicionales tal como se definen aquí, donde el al menos un ácido 60 nucleico de la invención se proporciona por separado (primera parte del kit) de al menos una otra parte del kit que comprende uno o más de los otros componentes. La composición farmacéutica inventiva y/o la vacuna inventiva pueden por ejemplo presentarse en una o diferentes partes del kit. Como ejemplo, por ejemplo al

5 menos una parte del kit puede comprender al menos una secuencia de ácido nucleico inventivo como se define aquí y al menos otra parte del kit comprende al menos otro componente como se define aquí, por ejemplo al menos otra parte del kit puede comprender al menos una composición farmacéutica o vacuna o una parte de la misma, por ejemplo al menos una parte del kit puede comprender la secuencia de ácido nucleico inventivo como se define aquí, al menos una parte adicional del kit al menos otro componente como se define aquí, al menos otra parte del kit al menos un componente de la composición farmacéutica inventiva o vacuna o la composición farmacéutica inventiva o vacuna como un todo, y al menos otra parte del kit por ejemplo al menos un portador o vehículo farmacéutico, etc. En caso de que el kit o un kit de partes comprenda una pluralidad de secuencias de ácido nucleico inventivo (lo que significa típicamente más de 1, 2, 3, 4, 5, 6 o más de 10 ácidos nucleicos, por ejemplo 2 a 10, de preferencia 2 a 5 ácidos nucleicos), un componente del kit puede comprender sólo una, diversas o todas las secuencias de ácido nucleico inventivo comprendidas en el kit. En una realización alternativa, cada secuencia de ácido nucleico inventivo puede estar comprendida en un componente diferente/separado del kit, de forma que cada componente forme una parte del kit. Igualmente, más de un ácido nucleico puede estar comprendido en un primer componente como parte del kit, mientras que uno o más de otros (segundo, tercero, etc.) componentes (proporcionando uno o más partes del kit) pueden o bien contener uno o más de los ácidos nucleicos inventivos, que pueden ser idénticos o parcialmente idénticos o diferentes del primer componente. El kit o el kit de partes además puede contener instrucciones técnicas con información sobre la administración y dosificación de la secuencia de ácido nucleico inventivo, de la composición farmacéutica inventiva o de la vacuna inventiva o de cualquiera de sus componentes o partes, por ejemplo 20 cuando el kit se prepara como un kit de partes.

En general, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

25 donde dicho péptido o proteína comprende un antígeno patogénico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, de preferencia asociado a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, una infección por hongos o similares. La descripción proporciona además una composición o kit o kit de partes que comprende al menos una de estas secuencias de ácido nucleico. Además, 30 la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar tal secuencia de ácido nucleico o composición que contenga esta 35 secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

Además, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- 40 c) una secuencia poli(A) o una señal de poliadenilación;

45 donde dicho péptido o proteína comprende un antígeno patogénico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, de preferencia asociado a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, una infección por hongos o similar, con mayor preferencia donde el antígeno patogénico se deriva de patógenos seleccionados de virus de influenza, virus sincitial respiratorio (RSV), virus de herpes simplex (HSV), virus de papiloma humano (VPH), virus de inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, citomegalovirus (CMV), virus de hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y virus de la fiebre amarilla. La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar esta secuencia de 55 ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

Con más preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

donde el péptido o proteína comprende un antígeno patogénico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, de preferencia asociada a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, un infección por hongos o similar, con mayor preferencia donde el antígeno patogénico se puede derivar de patógenos seleccionados del virus de influenza, virus sincitial respiratorio (RSV), virus del herpes simplex (HSV), virus del papiloma humano (VPH), virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, citomegalovirus (CMV), virus de la Hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia, y virus de la fiebre amarilla, incluso de mayor preferencia el antígeno se selecciona del antígeno p24 del VIH, proteínas con envoltura del VIH (Gp120, gp41, Gp160), poliproteína GAG, proteína Nef de factor negativo, trans-activador de transcripción Tat si la enfermedad infecciosa es VIH, de preferencia una infección con el virus de la inmunodeficiencia humana. La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar una secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

donde dicho péptido o proteína comprende un antígeno patogénico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, con preferencia asociado a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, un infección por hongos o similar, con mayor preferencia donde el antígeno patogénico se puede derivar de patógenos seleccionados del virus de influenza, virus sincitial respiratorio (RSV), virus del herpes simplex (HSV), virus del papiloma humano (VPH), virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, citomegalovirus (CMV), virus de la hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y virus de la fiebre amarilla, incluso con mayor preferencia el antígeno se selecciona de proteína MOMP de membrana externa mayor, la proteína de membrana externa PMPC probable, proteína compleja de membrana externa B OmcB, proteínas de choque térmico Hsp60 HSP10, proteína IncA, proteínas provenientes del sistema de secreción tipo III, proteína de cadena pequeña de reductasa y ribonucleótido NrdB, proteína plasmídica Pgp3, proteína externa de clamidia N CopN, antígeno CT521, antígeno CT425, antígeno CT043, antígeno TC0052, antígeno TC0189, antígeno TC0582, antígeno TC0660, antígeno TC0726, antígeno TC0816, antígeno TC0828 si la enfermedad infecciosa es una infección por Chlamydia trachomatis. La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar esta secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;



donde dicho péptido o proteína comprende un antígeno patológico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, de preferencia asociado a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, un infección por hongos o similar, con mayor preferencia donde el antígeno patológico se puede derivar de patógenos seleccionados del virus de influenza, virus sincitial respiratorio (SV), virus del herpes simplex (HSV), virus del papiloma humano (VPH), virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, citomegalovirus (CMV), virus de la hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y virus de la fiebre amarilla, incluso con mayor preferencia el antígeno se selecciona del antígeno pp65, proteína de membrana pp15, proteína tegumental cápside-proximal pp150, proteína M45, ADN polimerasa UL54, helicasa UL105, glucoproteína gM, glucoproteína gN, glucoproteína H, glucoproteína B gB, proteína UL83, proteína UL94, proteína UL99 si la enfermedad infecciosa es una infección por Citomegalovirus, de preferencia una infección con citomegalovirus (CMV). La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como un medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar esta secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

donde dicho péptido o proteína comprende un antígeno patológico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, de preferencia asociado a una enfermedad que es una infección bacteriana, una infección viral, una infección protozoaria, una infección por hongos o similar, con mayor preferencia donde el antígeno patológico se puede derivar de patógenos seleccionados del virus de influenza, virus sincitial respiratorio (RSV), Virus de herpes simplex (HSV), virus de papiloma humano (HPV), Virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, Cytomegalovirus (CMV), virus de Hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y el Virus de fiebre amarilla, incluso con mayor preferencia el antígeno se selecciona de la proteína C con cápside, proteína premembrana prM, proteína membrana M, proteína con envoltura E (dominio I, dominio II, dominio II), proteína NS1, proteína NS2A, proteína NS2B, proteína NS3, proteína NS4A, proteína 2K, proteína NS4B, proteína NS5, de preferencia una infección con los virus de dengue (DEN-1, DEN-2, DEN-3 y DEN-4)-Flavivirus. La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar esta secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

donde dicho péptido o proteína comprende un antígeno patológico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, con preferencia asociado a una enfermedad que es una infección bacteriana, una infección viral, una infección protozoaria, una infección por hongos o similar, con mayor preferencia donde el antígeno patológico se puede derivar de patógenos seleccionados del virus de influenza, virus de influenza, virus sincitial respiratorio (RSV), Virus de herpes simplex (HSV), virus de papiloma humano (HPV), Virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, Cytomegalovirus (CMV), virus de Hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y el Virus de fiebre amarilla, incluso con

mayor preferencia el antígeno se selecciona del antígeno superficial hepatitis B HBsAg, antígeno núcleo de Hepatitis B HbcAg, polimerasa, proteína Hbx, proteína superficial media preS2, proteína superficial L, proteína larga S, proteína viral VP1, proteína viral VP2, proteína viral VP3, proteína viral VP4 si la enfermedad infecciosa es Hepatitis B, de preferencia una infección con virus de hepatitis B (HBV). La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar esta secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- 15 a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

donde dicho péptido o proteína comprende un antígeno patogénico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, de preferencia asociado a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, un infección por hongos o similar, con mayor preferencia donde el antígeno patogénico se puede derivar de patógenos seleccionados del virus de influenza, virus sincitial respiratorio (RSV), Virus de herpes simplex (HSV), virus de papiloma humano (HPV), Virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, Citomegalovirus (CMV), virus de Hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y el Virus de fiebre amarilla, incluso con mayor preferencia el antígeno se selecciona de la proteína de replicación de E1, proteína reguladora E2, proteína E3, proteína E4, proteína E5, proteína E6, proteína E7, proteína E8, proteína con cápside mayor L1, proteína con cápside menor L2 si la enfermedad infecciosa es infección por papilomavirus humano (HPV); de preferencia una infección con papilomavirus humano (HPV). La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar esta secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- 40 a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

donde dicho péptido o proteína comprende un antígeno patogénico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, de preferencia asociada a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, un infección por hongos o similar, con mayor preferencia donde el antígeno patogénico se puede derivar de patógenos seleccionados del virus de influenza, virus sincitial respiratorio (RSV), virus del herpes simplex (HSV), virus del papiloma humano (VPH), virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, citomegalovirus (CMV), virus de la hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y virus de la fiebre amarilla, incluso con mayor preferencia el antígeno se selecciona de proteína de fusión F, hemaglutinina-neuramidasa HN, glucoproteína G, proteína de matriz M, fosfoproteína P, nucleoproteína N, polimerasa L si la enfermedad infecciosa es una infección por el virus de la parainfluenza, preferentemente un virus de la parainfluenza humana (HPIV). La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los

pasos de proporcionar esta secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

5 Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

10 donde dicho péptido o proteína comprende un antígeno patogénico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, con preferencia asociado a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, un infección por hongos o similar, con mayor preferencia donde el antígeno patogénico se puede derivar de patógenos seleccionados del virus de influenza, virus sincitial respiratorio (RSV), virus del herpes simplex (HSV), virus del papiloma humano (VPH), virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, citomegalovirus (CMV), virus de la hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y virus de la fiebre amarilla, incluso con mayor preferencia el antígeno se selecciona de hemaglutinina (HA), neuraminidasa (NA), nucleoproteína (NP), proteína M1, proteína M2, proteína NS1, proteína NS2 (proteína NEP: proteína de exportación nuclear), proteína PA, proteína PB1 (proteína de polimerasa básica 1), proteína PB1-F2 y proteína PB2 (familia Ortomyxoviridae, virus de Influenza (gripe)); nucleoproteína N, proteína estructural grande L, fosfoproteína P, proteína de matriz M, glucoproteína G si la enfermedad infecciosa es la rabia, de preferencia una infección con el virus de la rabia; con mayor preferencia el antígeno se deriva de un virus de la Orthomyxoviridae, de mayor preferencia de un virus de influenza, con total preferencia de hemaglutinina (HA), neuraminidasa (NA), nucleoproteína (NP), uno o ambas de las proteínas de matriz (M1) y (M2), proteínas de la polimerasa (PB1), (PB2) y las proteínas NS1 y NS2. La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar esta secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

35 Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

40 donde el péptido o proteína comprende un antígeno patogénico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, con preferencia asociado a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, un infección por hongos o similar, de mayor preferencia donde el antígeno patogénico se puede derivar de patógenos seleccionados del virus de influenza, virus sincitial respiratorio (RSV), virus del herpes simplex (HSV), virus del papiloma humano (VPH), virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, citomegalovirus (CMV), virus de la hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y virus de la fiebre amarilla, incluso con mayor preferencia el antígeno se selecciona de proteína de fusión F, nucleoproteína N, proteína de matriz M, proteína de matriz M2-1, proteína de matriz M2-2, fosfoproteína P, proteína hidrofóbica pequeña SH, glucoproteína G de superficie mayor, polimerasa L, proteína 1 no estructural NS1, proteína 2 no estructural NS2 si la enfermedad infecciosa es la infección del virus sincitial respiratorio, de preferencia una infección con virus sincitial respiratorio (RSV).

Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

55 a) una región codificante, que codifica para al menos un péptido o proteína;  
 b) al menos un tallo-bucle de histona, y  
 c) una secuencia poli(A) o una señal de poliadenilación;

donde dicho péptido o proteína comprende un antígeno patológico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, de preferencia asociado a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, un infección por hongos o similar, con mayor preferencia donde el antígeno patológico se puede derivar de patógenos seleccionados del virus de influenza, virus sincitial respiratorio (RSV), virus del herpes simplex (HSV), virus del papiloma humano (VPH), virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, citomegalovirus (CMV), virus de la hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y virus de la fiebre amarilla, incluso con mayor preferencia el antígeno se selecciona de antígeno secretorio SssA (Staphylococcus genus; envenenamiento de alimentos por estafilococo) antígeno secretorio SssA (Staphylococcus genus; por ejemplo, infección por aureus, Staphylococcal); chaperona molecular DnaK, lipoproteína Mpt83 de superficie celular, lipoproteína P23, proteína permeasa pstA del sistema de transporte de fosfato, antígeno de 14 kDa, proteína C de unión a fibronectina FbpC1, alanina deshidrogenasa TB43, glutamina sintetasa 1, proteína ESX-1, proteína CFP10, proteína TB10.4, proteína MPT83, proteína MTB12, proteína MTB8, proteínas similares a Rpf, proteína MTB32, proteína MTB39, cristalina, proteína de choque térmico HSP65, proteína PST-S si la enfermedad infecciosa es tuberculosis, de preferencia una infección con Mycobacterium tuberculosis. La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar esta secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

Además con preferencia, la descripción proporciona una secuencia de ácido nucleico que comprende o codifica para:

- a) una región codificante, que codifica para al menos un péptido o proteína;
- b) al menos un tallo-bucle de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación;

donde dicho péptido o proteína comprende un antígeno patológico o un fragmento, variante o derivado del mismo, en particular un antígeno proveniente de un patógeno asociado a una enfermedad infecciosa, con preferencia asociado a una enfermedad infecciosa que es una infección bacteriana, una infección viral, una infección por protozoarios, un infección por hongos o similar, con mayor preferencia donde el antígeno patológico se puede derivar de patógenos seleccionados del virus de influenza, el virus sincitial respiratorio (RSV), virus del herpes simplex (HSV), virus del papiloma humano (VPH), virus de la inmunodeficiencia humana (VIH), Plasmodium, Staphylococcus aureus, virus del dengue, Chlamydia trachomatis, citomegalovirus (CMV), virus de la hepatitis B (HBV), tuberculosis Mycobacterium, virus de la rabia y virus de la fiebre amarilla, incluso con mayor preferencia el antígeno se selecciona de poliproteína genómica, proteína E, proteína M, proteína C con cápside, proteasa NS3, proteína NS1, proteína NS2A, proteína AS2B, proteína NS4A, proteína NS4B, proteína NS5, si la enfermedad infecciosa es la fiebre amarilla, de preferencia una infección por el virus de fiebre amarilla. La descripción proporciona además una composición o kit o un kit de partes que comprende al menos una de estas secuencias de ácidos nucleicos. Además, la descripción proporciona el uso de esta secuencia de ácido nucleico como medicamento, de preferencia para el tratamiento de enfermedades infecciosas, con mayor preferencia en una composición farmacéutica para el tratamiento de enfermedades infecciosas que comprende además un portador farmacéuticamente aceptable. Además, la descripción proporciona un método para aumentar la expresión de un péptido o proteína codificado que comprende los pasos de proporcionar esta secuencia de ácido nucleico o una composición que contenga esta secuencia de ácido nucleico y aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, a una célula, a un tejido o a un organismo.

En la presente invención, si no se indica de otra manera, las diferentes características de las alternativas y realizaciones se pueden combinar entre sí. Además, el término “que comprende” no se debe interpretar con el significado de “que consiste en”, si no se menciona específicamente. Sin embargo, en el contexto de la presente invención, el término “que comprende” se puede sustituir con el término “que consiste en”, cuando sea adecuado.

#### BREVE DESCRIPCIÓN DE LAS FIGURAS

Las siguientes figuras pretenden ilustrar la invención adicionalmente reflejando realizaciones reivindicadas y no reivindicadas y no se deben interpretar como limitantes de la presente invención.

Figura 1: muestra la secuencia consenso tallo-bucle de histona generada a partir de secuencias tallo-bucle de metazoarios y protozoarios (según se reporta por Dávila López, M., & Samuelsson, T. (2008), *RNA* (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308). Se alinearon 4001 secuencias tallo-bucle de histona provenientes de metazoarios y protozoarios y se indica el número de nucleótidos existentes para cada posición en la secuencia del tallo-bucle. La secuencia consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se proporciona utilizando el código de nucleótidos de una sola letra. Además de la secuencia consenso, se muestran secuencias que muestran al menos el 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

Figura 2: muestra la secuencia consenso tallo-bucle de histona generada a partir de secuencias del tallo-bucle de protozoarios (según se reporta por Dávila López, M., & Samuelsson, T. (2008), *RNA* (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308). Se alinearon 131 secuencias del tallo-bucle de histonas provenientes de protozoarios y se indica el número de nucleótidos presentes para cada posición en la secuencia del tallo-bucle. La secuencia consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se proporciona utilizando el código de nucleótidos de una sola letra. Además de la secuencia consenso, se muestran las secuencias que muestran al menos el 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

Figura 3: muestra la secuencia consenso del tallo-bucle de histona generada a partir de secuencias del tallo-bucle metazoarios (según se reporta por Dávila López, M., & Samuelsson, T. (2008), *RNA* (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308). Se alinearon 3870 secuencias del tallo-bucle de histona de metazoarios y se indica el número de nucleótidos presentes para cada posición en la secuencia del tallo-bucle. La secuencia consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se proporciona utilizando el código de nucleótidos de una sola letra. Además de la secuencia consenso, se muestran secuencias que representan al menos el 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

Figura 4: muestra la secuencia consenso del tallo-bucle de histona generada a partir de secuencias del tallo-bucle de vertebrados (según se reporta por Dávila López, M., & Samuelsson, T. (2008), *RNA* (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308). Se alinearon 1333 secuencias del tallo-bucle de histona de vertebrados y se indica el número de nucleótidos presentes para cada posición en la secuencia del tallo-bucle. La secuencia consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se proporciona utilizando el código de nucleótidos de una sola letra. Además de la secuencia consenso, se muestran secuencias que representan al menos el 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

Figura 5: muestra la secuencia consenso del tallo-bucle de histona generado a partir secuencias del tallo-bucle humana (*Homo sapiens*) (como se informa por Dávila López, M., & Samuelsson, T. (2008), *RNA* (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308). Se alinearon 84 secuencias del tallo-bucle de histona en seres humanos y se indica el número de nucleótidos presentes para cada posición en la secuencia del tallo-bucle. La secuencia consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se proporciona utilizando el código de nucleótidos de una sola letra. Además de la secuencia consenso, se muestran secuencias que representan al menos el 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

Figuras 6 a 19: muestran los ARNm de la transcripción *in vitro*. Se proporciona la designación y la secuencia de los ARNm obtenidos mediante transcripción *in vitro*. Se utilizan las siguientes abreviaturas:

ppLuc (GC):	secuencia de ARNm enriquecido con GC que codifica para <i>Photinus pyralis</i> luciferasa
ag:	región no traducida (UTR) 3' del gen alfa-globina
A64:	secuencia poli(A) con 64 adenilatos
A120:	secuencia poli(A) 120 adenilatos
histonaSL:	tallo-bucle de histona
aCPSL:	tallo-bucle que se ha seleccionado a partir de las biblioteca para su unión específica de la proteína $\alpha$ CP-2KL
PolioCL:	hoja de trébol 5' proveniente de ARN genómico del virus de polio
G30	secuencia poli(G) con 30 guanilatos
U30	secuencia poli(U) con 30 uridilatos
SL	tallo-bucle no específica/artificial
N32	secuencia no específica de 32 nucleótidos

Dentro de las secuencias, se resaltan los siguientes elementos: región codificante (ORF) (letras mayúsculas), ag (negrillas), histonaSL (subrayado), secuencias distintas adicionales probadas (cursiva).

- Figura 6: muestra la secuencia de ARNm de ppLuc(GC) - ag (SEC ID NO: 43). Por linearización del vector original en el sitio de restricción inmediatamente después del 3'-UTR de alfa-globina (ag), se obtiene el ARNm sin una secuencia poli(A).
- 5 Figura 7: muestra la secuencia de ARNm de ppLuc(GC) - ag - A64 (SEC ID NO: 44). Por linearización del vector original en el sitio de restricción inmediatamente después de la secuencia poli(A) A64, se obtiene el ARNm que termina con una secuencia poli(A) A64.
- Figura 8: muestra la secuencia de ARNm de ppLuc(GC) - ag - histonaSL (SEC ID NO: 45). La secuencia poli(A) A64 se reemplazó por una histonaSL. La secuencia del tallo-bucle de histona utilizada en los ejemplos se obtuvo a partir de Cakmakci *et al.* (2008). Molecular and Cellular Biology, 28(3), 1182-1194.
- 10 Figura 9: muestra la secuencia de ARNm de ppLuc(GC) - ag - A64 - histonaSL (SEC ID NO: 46). La histonaSL se agregó 3' de la poli(A) A64.
- Figura 10: muestra la secuencia de ARNm de ppLuc(GC) - ag - A120 (SEC ID NO: 47). La secuencia poli(A) A64 se reemplazó por una secuencia poli(A) A120.
- 15 Figura 11: muestra la secuencia ARNm de ppLuc(GC) - ag - A64 - ag (SEC ID NO: 48). Una segunda 3'-UTR de alfa-globina se agregó 3' de una poli(A) A64.
- Figura 12: muestra la secuencia de ARNm de ppLuc(GC) - ag - A64 - aCPSL (SEC ID NO: 49). Se agregó un tallo-bucle 3' de la poli(A) A64. El tallo-bucle se había seleccionado de una librería para su unión específica de la proteína  $\alpha$ CP-2KL (Thisted *et al.*, (2001), The Journal of Biological Chemistry, 276(20), 17484-17496)  $\alpha$ CP-2KL es una isoforma de  $\alpha$ CP-2, la proteína  $\alpha$ CP expresada con mayor fuerza (proteína de unión poli(C) de ARNm de alfa-globina) (Makeyev *et al.*, (2000), Genomics, 67(3), 301-316), un grupo de proteínas de unión a ARN que se unen a la 3'-UTR de alfa-globina (Chkheidze *et al.*, (1999), Molecular and Cellular Biology, 19(7), 4572-4581).
- 20 Figura 13: muestra la secuencia ARNm de ppLuc(GC) - ag - A64 - PolioCL (SEC ID NO: 50). La hoja de trébol 5' proveniente del ARN genómico del virus de Polio se agregó 3' de poli(A) A64.
- 25 Figura 14: muestra la secuencia ARNm de ppLuc(GC) - ag - A64 - G30 (SEC ID NO: 51). Se agregó un tramo de 30 guanilatos 3' de la poli(A) A64.
- Figura 15: muestra la secuencia ARNm de ppLuc(GC) - ag - A64 - U30 (SEC ID NO: 52). Se anexó un tramo de 30 uridilatos 3' de la poli(A) A64.
- 30 Figura 16: muestra la secuencia ARNm de ppLuc(GC) - ag - A64 - SL (SEC ID NO: 53). Se anexó un tallo-bucle 3' de poli(A) A64. La parte superior del tallo y el bucle se tomaron de Babendure *et al.*, (2006), RNA (New York, N.Y.), 12(5), 851-861. El tallo-bucle consiste en 17 pares de bases largas, un tallo rico en CG y un bucle largo de 6 bases.
- Figura 17: muestra ppLuc(GC) - ag - A64 - N32 (SEC ID NO: 54). Por linearización del vector original en un sitio de restricción alternativo, se obtuvo un ARNm con 32 nucleótidos adicionales después de la poli(A).
- 35 Figura 18: muestra la secuencia ARNm de HA (H1N1/PR8) (GC) - ag - A64 - C30 (SEC ID NO: 55).
- Figura 19: muestra la secuencia ARNm de HA (H1N1/PR8) (GC) - ag - A64 - C30 - histonaSL (SEC ID NO: 56).
- Figura 20: muestra la combinación de poli(A) y los aumentos de histonaSL de la expresión proteínica a partir de un ARNm de una forma sinérgica. Se examinó el efecto de la secuencia poli(A), histonaSL y la combinación de la poli(A) y la histonaSL sobre la expresión de luciferasa a partir del ARNm. Así, células HeLa se sometieron a electroporación con diferentes ARNm. Los niveles de luciferasa se midieron 6, 24 y 48 horas después de la transfección. Se expresa poca luciferasa proveniente del ARNm que no tiene ni una secuencia poli(A) ni una histonaSL. Tanto la secuencia poli(A) como la histonaSL aumentan el nivel de luciferasa. Sin embargo, extraordinariamente, la combinación de poli(A) e histonaSL aumenta además fuertemente el nivel de luciferasa, muchas veces por encima del nivel observado con cualquiera de los elementos individuales, actuando así sinérgicamente. Los datos se muestran como media RLU  $\pm$  SD (unidades de luz relativas  $\pm$  desviación estándar) para transfecciones por triplicado. En el ejemplo 11.2 se resumen los RLU específicos.
- 40 45
- Figura 21: muestra que la combinación de poli(A) y la histonaSL aumenta la expresión proteínica a partir del ARNm independientemente de su orden. Se examinó el efecto de la secuencia poli(A), la histonaSL, la combinación de poli(A) y la histonaSL, y su orden por la expresión de luciferasa a partir del ARNm. Así, se lipofectaron diferentes ARNm en células HeLa. Los niveles de luciferasa se midieron 6, 24 y 48 horas después
- 50

del inicio de la transfección. Tanto la secuencia poli(A) A64 o la histonaSL dieron lugar a niveles comparables de luciferasa. Aumentar la longitud de la secuencia poli(A) de A64 hasta A120 o hasta A300 aumenta el nivel de luciferasa moderadamente. Por el contrario, la combinación de poli(A) y la histonaSL aumenta el nivel de luciferasa mucho más que el alargamiento de la secuencia poli(A). La combinación de poli(A) e histonaSL actúa sinérgicamente, ya que aumenta el nivel de luciferasa muchas veces por encima del nivel observado con cualquiera de los elementos individuales. Se observa el nivel sinérgico de la combinación de poli(A) e histonaSL independientemente del orden de la poli(A) y la histonaSL e independientemente de la longitud de la poli(A) con el ARNm de A64-histonaSL o histonaSL-A250. Los datos se muestran como media RLU  $\pm$  SD para transfecciones por triplicado. En el ejemplo 11.3 se resumen los RLU específicos.

5  
10  
15  
20

Figura 22: muestra que el aumento en la expresión proteínica mediante la combinación de poli(A) y la histonaSL es específico. Se examinó el efecto de combinar la poli(A) y la histonaSL o la poli(A) y secuencias alternativas sobre la expresión de luciferasa a partir de un ARNm. Así, se sometieron a electroporación diferentes ARNm en células HeLa. Se midieron los niveles de luciferasa 6, 24, y 48 horas después de la transfección. Tanto la secuencia poli(A) o la histonaSL dieron lugar a niveles comparables de luciferasa. La combinación de poli(A) y la histonaSL aumenta en gran medida el nivel de luciferasa, muchas veces por encima del observado con cualquiera de los elementos individuales, actuando así sinérgicamente. Por el contrario, combinar la poli(A) con cualquiera de las otras secuencias no tiene efecto sobre el nivel de luciferasa en comparación con el ARNm que contiene sólo una secuencia poli(A). De esta forma, la combinación de la poli(A) y la histonaSL actúa específica sinérgicamente. Los datos se muestran como media RLU  $\pm$  SD para transfecciones por triplicado. En el ejemplo 11.4 se resumen los RLU específicos.

25  
30

Figura 23: muestra que la combinación de la poli(A) y la histonaSL aumenta la expresión proteínica a partir del ARNm de una manera sinérgica *in vivo*. Se examinó el efecto de la secuencia poli(A), la histonaSL y la combinación de la poli(A) y la histonaSL sobre la expresión de luciferasa a partir del ARNm *in vivo*. Así, se inyectaron intradérmicamente diferentes ARNm en ratones. Los ratones se sacrificaron 16 horas después de la inyección y se midieron los niveles de luciferasa en los sitios de inyección. La luciferasa se expresa a partir de un ARNm que tiene ya sea una histonaSL o una secuencia poli(A). Sin embargo, asombrosamente la combinación de poli(A) e histonaSL aumenta fuertemente el nivel de luciferasa, muchas veces por encima del nivel observado con cualquiera de los elementos individuales, actuando así sinérgicamente. Los datos se muestran como media RLU  $\pm$  SEM (unidades de luz relativa  $\pm$  error estándar de la media). En el ejemplo 11.5 se resumen los RLU específicos.

35

Figura 24: muestra que la combinación de la poli(A) y la histonaSL aumenta el nivel de anticuerpos producidos por la vacunación con un ARNm. Se examinó el efecto de la secuencia poli(A) y la combinación de la poli(A) y la histonaSL sobre la inducción de anticuerpos anti-HA producidos por la vacunación con un ARNm. Así, ratones Balb/c se vacunaron intradérmicamente con diferentes ARNm. El nivel de anticuerpos HA-específicos en ratones vacunados y de control se analizó mediante ELISA con diluciones en serie de sueros. El IgG1 anti-HA se indujo mediante ARNm que tiene sólo una secuencia poli(A). Sin embargo, sorprendentemente, la combinación de poli(A) e histonaSL aumenta fuertemente el nivel de IgG1 anti-HA, por encima del nivel observado con sólo una secuencia poli(A).

## Ejemplos

40

Los siguientes ejemplos se incluyen para ilustrar adicionalmente la invención en base a realizaciones reivindicadas y no reivindicadas y no se deben interpretar como limitantes de la presente invención.

### 1. Generación de secuencias consenso de tallo-bucle de histonas

45  
50

Antes de los experimentos, se determinaron las secuencias consenso del tallo-bucle de histonas en base a secuencias tallo-bucle de histonas de metazoarios y protozoarios. Las secuencias se tomaron del suplemento proporcionado por Lopez *et al.* (Dávila López, M., & Samuelsson, T. (2008), RNA (New York, N.Y.), 14(1), 1-10. doi:10.1261/rna.782308), quienes identificaron un gran número de secuencias del tallo-bucle de histonas naturales mediante la búsqueda de secuencias genómicas y etiquetas de secuencias expresadas. En primer lugar, todas las secuencias provenientes de metazoarios y protozoarios (4001 secuencias), o todas las secuencias provenientes de protozoarios (131 secuencias) o, alternativamente, de metazoarios (3870 secuencias), o de vertebrados (1333 secuencias) o de humanos (84 secuencias) se agruparon y alinearon. Después, se determinó la cantidad de los nucleótidos presentes para cada posición. Con base en las tablas así obtenidas, se generaron secuencias consenso para los 5 diferentes grupos de secuencias representando todos los nucleótidos presentes en las secuencias analizadas. Además, también se obtuvieron secuencias consenso restrictivas, enfatizando cada vez más los nucleótidos conservados.

### 55 2. Preparación de plantillas de ADN

Se construyó un vector para la transcripción *in vitro* que contenía un promotor T7 seguido por una secuencia enriquecida con GC que codifica para *Photinus pyralis* luciferasa (ppLuc(GQ), la parte central de la región no traducida 3' (UTR) de alfa-globina (AG) y una secuencia poli(A). La secuencia de poli(A) estaba seguida inmediatamente por un sitio de restricción utilizado para la linearización del vector antes de la transcripción *in vitro* para obtener el ARNm terminado en una secuencia poli(A) A64. El ARNm así obtenido de este vector mediante transcripción *in vitro* se denominó "ppLuc(GC) - ag - A64".

La linearización de este vector en los sitios de restricción alternativos antes de la transcripción *in vitro* permitió tener un ARNm ya sea extendido mediante nucleótidos adicionales 3' de A64 o que carecen de A64. Además, el vector original se modificó para incluir secuencias alternativas. En resumen, se obtuvieron los siguientes ARNm a partir de estos vectores mediante transcripción *in vitro* (en las figuras 6, 7, 8, 9, 11, 12, 13, 14, 15, 16 y 17 se proporcionan las secuencias de ARNm):

	ppLuc(GC) – ag	(SEC ID NO: 43)
	ppLuc(GC) – ag – A64	(SEC ID NO: 44)
	ppLuc(GC) – ag – histonaSL	(SEC ID NO: 45)
15	ppLuc(GC) – ag – A64 – histonaSL	(SEC ID NO: 46)
	ppLuc(GC) – ag – A120	(SEC ID NO: 47)
	ppLuc(GC) – ag – A64 – ag	(SEC ID NO: 48)
	ppLuc(GC) – ag – A64 – aCPSL	(SEC ID NO: 49)
	ppLuc(GC) – ag – A64 – PolioCL	(SEC ID NO: 50)
20	ppLuc(GC) – ag – A64 – G30	(SEC ID NO: 51)
	ppLuc(GC) – ag – A64 – U30	(SEC ID NO: 52)
	ppLuc(GC) – ag – A64 – SL	(SEC ID NO: 53)
	ppLuc(GC) – ag – A64 – N32	(SEC ID NO: 54)

Además, se prepararon como se describió anteriormente secuencias de plásmido de ADN que codifican para el antígeno patogénico HA (H1 N1/PR8).

En resumen, se obtuvieron los siguientes ARNm de estos vectores mediante transcripción *in vitro* (en las figuras 18 y 19 se proporciona las secuencias de ARN):

	HA (H1N1/PR8) (GC) – ag – A64 – C30	(SEC ID NO: 55)
	HA (H1N1/PR8) (GC) – ag – A64 – C30 - histonaSL	(SEC ID NO: 56)

### 30 3. Transcripción *in vitro*

La plantilla de ADN de acuerdo con el Ejemplo 2 se linearizó y se transcribió *in vitro* utilizando T7 polimerasa. La plantilla de ADN luego se dirigió mediante tratamiento con ADNasa. Todas las transcripciones de ARNm contenían una estructura 5'-CAP obtenida por adición de un exceso de N7-metil-guanosin-5'-trifosfato-5'-guanosina a la reacción de transcripción. El ARNm así obtenido se purificó y sesuspendió en agua.

### 35 4. Adenilación enzimática del ARNm

Dos ARNm se sometieron a adenilación enzimáticamente:

ppLuc(GC) - ag - A64 y ppLuc(GC) - ag - histonaSL.

Para ello, el ARN se incubó con *E. coli* poli(A)-polimerasa y ATP (Poli(A) Polymerase Tailing Kit, Epicentre, Madison, USA) siguiendo las instrucciones del fabricante. El ARNm con la secuencia poli(A) extendida se purificó y sesuspendió en agua. La longitud de la secuencia poli(A) se determinó vía electroforesis en gel de agarosa. Los ARNm de partida se extendieron en aproximadamente 250 adenilatos, los ARNm obtenidos se denominaron ppLuc(GC) - ag - A300 y ppLuc(GC) - ag - histonaSL - A250, respectivamente.

### 5. Expresión de luciferasa mediante electroporación de ARNm

Se trataron con tripsina células HeLa y se lavaron en opti-MEM. 1 x 10<sup>5</sup> células en 200 µl de opti-MEM cada una se sometieron a electroporación con 0,5 µg de ARNm que codifica para ppLuc. Como control, se sometió a electroporación por separado un ARNm no codificante para ppLuc. Las células sometidas a electroporación se sembraron en placas de 24 pocillos en 1 ml de medio RPMI 1640. 6, 24 ó 48 horas después de la transfección, el medio se aspiró y las células se lisaron en 200 µl de tampón de lisis (Tris 25 mM, pH 7.5 (HCl),



EDTA 2 mM, 10% de glicerol, 1% de Triton X-100, DTT 2 mM, PMSF 1 mM). Los lisados se almacenaron a -20°C hasta que se midió la actividad de ppLuc.

## 6. Expresión de luciferasa mediante lipofección de ARNm

5 Células HeLa se sembraron en placas de 96 pocillos a una densidad de  $2 \times 10^4$  células por pocillo. Al siguiente día, las células se lavaron en Opti-MEM y luego se transfectaron con 0,25 µg de ARNm que codifican para ppLuc-complejados con lipofectina en 150 µl de Opti-MEM. Como control, se sometió a lipofección por separado un ARNm no codificante para ppLuc. En algunos pocillos, el Opti-MEM se aspiró y las células se lisaron en 200 µl de tampón de lisis 6 horas después del inicio de la transfección. En los pocillos restantes, el Opti-MEM se intercambió por medio RPMI 1640 en ese momento. En estos pocillos, el medio se aspiró y las células se lisaron en 200 µl de tampón de lisis 24 ó 48 horas después del inicio de la transfección. Los lisados se almacenaron a -20°C hasta que se midió la actividad de ppLuc.

## 7. Medición de luciferasa

15 La actividad ppLuc se midió como unidades de luz relativa (RLU) en un lector de placas BioTek SynergyHT a un tiempo de medición de 5 segundos utilizando 50 µl del lisado y 200 µl de tampón de luciferina (Glicilglicina 25 mM, pH 7.8 (NaOH), MgSO<sub>4</sub> 15 mM, ATP 2 mM, luciferina 75 µM). Se calcularon los RLU específicos al restar el RLU del ARN control del RLU totales.

## 8. Expresión de luciferasa mediante inyección de ARNm intradérmico (expresión de luciferasa *in vivo*)

20 Se anestesiaron ratones con una mezcla Rompun y Ketavet. Cada ARNm que codifica para ppLuc se inyectó intradérmicamente (0,5 µg de ARNm en 50 µl por inyección). Como control, se inyectó por separado ARNm no codificante para ppLuc. 16 horas después de la inyección, los ratones se sacrificaron y se recogieron los tejidos. Las muestras de tejido se congelaron instantáneamente en nitrógeno líquido y se lisaron en un lisador de tejido (Qiagen) en 800 µl de tampón de lisis (Tris 25 mM, pH 7.5 (HCl), EDTA 2 mM, glicerol al 10%, Triton X-100 al 1%, DTT 2 mM, PMSF 1 mM). Posteriormente las muestras se sometieron a centrifugación a 13.500 rpm a 4°C durante 10 minutos. Los lisados se almacenaron a -80°C hasta que se midió la actividad de ppLuc (véase 7. Medición de luciferasa).

## 9. Detección de una respuesta inmune a linfocitos B antígeno-específica (anticuerpos)

30 Ratones BALB/c (8 ratones por grupo) se vacunaron dos veces en 7 días intradérmicamente con una vacuna que comprendía 10 µg de ARNm que codifica para HA (hemaglutinina de A/Puerto Rico/8/34, de acuerdo con la SEC ID NO: 55 y 56). Para el control negativo, los ratones se trataron un tampón.

35 La detección de una respuesta inmune antígeno-específica se llevó detectando los anticuerpos proteína-específicos HA. Así, se extrajeron muestras sanguíneas de ratones vacunados cuatro semanas después de la última vacunación y se prepararon los sueros. Placas MaxiSorp® (Nalgene Nunc International) se recubrieron con proteína HA (Charles River Laboratories). Después de bloquear con 1xPBS que contenía 0,05% de Tween-20 y 1% de BSA, las placas se incubaron con suero diluido de ratón (1:50). Posteriormente se agregó un anticuerpo secundario acoplado con biotina (anti-mouse-IgG Dianova, cat. #115035003). Después de lavar, la placa se incubó con peroxidasa estreptavidina de rábano picante y posteriormente se midió la conversión del sustrato ABTS (ácido 2,2'-azino-bis(3-etil-benzotiazolin-6-sulfónico). En la figura 24, se muestran los resultados de este experimento.

## 40 10. Detección de una respuesta inmune celular antígeno-específica (respuesta inmune a células T) mediante ELISPOT

45 Ratones C57BL/6 se vacunaron intradérmicamente con ARNm que codifica para HA (hemaglutinina de A/Puerto Rico/8/34, de acuerdo con la SEC ID NO: 55 y 56) complejada con protamina (2 veces en 7 días). Los ratones control se trataron con tampón. 1 semana después de la vacunación los ratones se sacrificaron, se retiraron los bazo y se aislaron los esplenocitos. Los esplenocitos se volvieron a estimular durante 7 días en presencia de los péptidos provenientes del antígeno anterior (biblioteca de péptidos) o se co-incubaron con células dendríticas generadas a partir de células de médula ósea de ratones singénicos naturales, que se sometieron a electroporación con un ARNm que codifica para el antígeno. Para determinar una respuesta inmune celular antígeno-específica, se midió la secreción INFgamma después de la re-estimulación. Para la detección de INFgamma, una placa de multi-selección recubierta (Millipore) se incubó durante la noche con tampón de recubrimiento de carbonato-bicarbonato 0,1 M, pH 9,6, 10,59 g/l de Na<sub>2</sub>CO<sub>3</sub>, 8,4 g/l de NaHCO<sub>3</sub>

que comprende el anticuerpo contra IFN $\gamma$  (BD Pharmingen, Heidelberg, Alemania). Se incubaron estimuladores y células efectoras conjuntamente en la placa en la proporción de 1:20 durante 24 h. La placa se lavó con 1 x PBS y se incubó con un anticuerpo secundario acoplado con biotina. Después del lavado con 1 x PBS/0.05% de Tween-20 al sustrato (5-bromo-4-cloro-3-indolyl fosfato/Nitro Blue Tetrazolium Liquid Substrate System de Sigma Aldrich, Taufkirchen, Alemania) se agregó a la placa y se podría detectar visualmente la conversión del sustrato.

## RESULTADOS

### 11.1 Secuencias tallo-bucle de histona

Para caracterizar las secuencias del tallo-bucle de histonas, se agruparon y alinearon las secuencias provenientes de metazoarios y protozoarios (4001 secuencias), o de protozoarios (131 secuencias) o alternativamente de metazoarios (3870 secuencias), o de vertebrados (1333 secuencias) o de humanos (84 secuencias). Luego, la cantidad de nucleótidos presentes se determinó para cada posición. Con base de las tablas así obtenidas, se generaron las secuencias consenso para los 5 diferentes grupos de secuencias representando todos los nucleótidos presentes en las secuencias analizadas. Dentro de la secuencia consenso de metazoarios y protozoarios combinados, se conservaron 3 nucleótidos, uno T/U en el asa y un G y un C en el tallo, formando un par de bases. Estructuralmente, típicamente se formó un tallo de 6 pares de bases y un asa de 4 nucleótidos. Sin embargo, son comunes estructuras de desviación: de 84 tallos\*-bucle de histona humanas, dos contenían un tallo de sólo 5 nucleótidos que comprende 4 pares de bases y un mal apareamiento. Otro tallo-bucle de histonas humanas contenía un tallo de sólo 5 pares de bases. Cuatro más tallos-bucle de histonas humanas contenían un tallo de 6 nucleótidos de largo, aunque incluían un mal apareamiento en tres posiciones diferentes, respectivamente. Además, cuatro tallos-bucle de histonas humanas contenían un par de bases inestables en dos posiciones diferentes, respectivamente. Con relación al bucle, una longitud de 4 nucleótidos parece no ser estrictamente necesaria, ya que un nucle de 5 nucleótidos se ha identificado en *D. discoideum*.

Además de las secuencias consenso que representan todos los nucleótidos presentes en las secuencias analizadas, también se obtuvieron más secuencias consenso restrictivas, enfatizando cada vez más los nucleótidos conservados. En resumen, se obtuvieron las siguientes secuencias:

(Cons):	representa todos los nucleótidos presentes
(99%):	representa al menos el 99% de todos los nucleótidos presentes
(95%):	representa al menos el 95% de todos los nucleótidos presentes
(90%):	representa al menos el 90% de todos los nucleótidos presentes

Los resultados de los análisis de las secuencias del tallo-bucle de histona se resumen en las siguientes Tablas 1 a 5 (véanse también las figuras 1, 2, 3, 4 y 5):











donde las abreviaturas empleadas se definen como sigue

Abreviatura	Base nucleótido	Anotación
R	G o A	Purina
Y	T/U o C	Pirimidina
M	A o C	Amino
K	G o T/U	Ceto
S	G o C	Fuerte (enlaces 3H)
W	A o T/U	Débil (enlaces 2H)
H	A o C o T/U	Sin G
B	G o T/U o C	Sin A
V	G o C o A	Sin T/U
D	G o A o T/U	Sin C
N	G o C o T/U o A	Cualquier base
*	Presente o no	La base puede estar presente o no

### 11.2 La combinación de la poli(A) y la histonaSL aumenta la expresión proteica de un ARNm de forma sinérgica

- 5 Para investigar el efecto de la combinación de la poli(A) y la histonaSL sobre la expresión proteica a partir de un ARNm, se sintetizaron ARNm con diferentes secuencias 3' de 3'-UTR de alfa-globina: ARNm terminado solo en 3' del 3'-UTR, careciendo así de la secuencia poli(A) como de la histonaSL, o conteniendo ya sea una secuencia poli(A) A64 o una histonaSL en su lugar, o conteniendo tanto la poli(A) A64 como la histonaSL 3' de la 3'-UTR. Los ARNm que codifican para luciferasa o el ARNm control se sometieron a electroporación en
- 10 células HeLa. Se midieron los niveles de luciferasa 6, 24, y 48 horas después de la transfección (véase la siguiente Tabla 6 y la figura 20).

Tabla 6

ARNm	RLU 6 horas	RLU 24 horas	RLU 48 horas
ppLuc(GC)-ag-A64-histonaSL	466553	375169	70735
ppLuc(GC)-ag-histonaSL	50947	3022	84
ppLuc(GC)-ag-A64	10471	19529	4364
ppLuc(GC)-ag	997	217	42

- 15 Se expresó poca luciferasa a partir de un ARNm sin una secuencia poli(A) ni una histonaSL. Tanto una secuencia poli(A) o la histonaSL aumentaron el nivel de luciferasa a un nivel similar. Cualquier ARNm dio lugar a un nivel de luciferasa mucho mayor de lo que lo hizo el ARNm que carece tanto de la poli(A) como de la histonaSL. Sin embargo, sorprendentemente, la combinación de la poli(A) y la histonaSL aumentaron fuerte y
- 20 adicionalmente el nivel de luciferasa, muchas veces por encima del nivel observado con cualquiera de los elementos individuales. La magnitud del aumento en el nivel de luciferasa debido a la combinación de la poli(A) y la histonaSL en el mismo ARNm demuestra que están actuando sinérgicamente.

La sinergia entre la poli(A) y la histonaSL se cuantificó dividiendo la señal proveniente del ARNm de poli(A)-histonaSL (+/+) entre la suma de las señales provenientes del ARNm de la histonaSL (-/+) más el ARNm de poli(A) (+/-) (véase la siguiente Tabla 7).

Tabla 7

	A64	histonaSL	RLU 6 horas	RLU 24 horas	RLU 48 horas
	+	+	466553	375169	70735
	-	+	50947	3022	84
	+	-	10471	19529	4364
Sinergia			7,6	16,6	15,9

25



El factor aquí calculado especifica cuánto mayor es el nivel de luciferasa proveniente de la combinación del ARNm de la poli(A) y la histonaSL de lo que se podría esperar si los efectos de la poli(A) y la histonaSL fueran puramente aditivos. El nivel de luciferasa a partir del ARNm que combina la poli(A) y la histonaSL era hasta 16,6 veces mayor que si sus efectos hubieran sido puramente aditivos. Este resultado confirma que la combinación de la poli(A) y la histonaSL provoca un aumento marcadamente sinérgico en la expresión de proteínas.

**11.3 La combinación de la poli(A) y la histonaSL aumenta la expresión proteica a partir de ARNm independientemente de su orden**

El efecto de la combinación de la poli(A) y la histonaSL podría depender de la longitud de la secuencia poli(A) y el orden de la poli(A) y la histonaSL. Así, se sintetizaron ARNm con una longitud de secuencia poli(A) cada vez mayor y un ARNm con la poli(A) y la histonaSL en orden inverso: los dos ARNm contenían 3' de la 3'-UTR y una secuencia poli(A) ya sea A120 o A300. Otro ARNm contenía 3' de la 3'-UTR en primer lugar una histonaSL, seguida por una secuencia poli(A) A250. Los ARNm que codifican para luciferasa o el ARNm control se sometieron a lipofección en células HeLa. Se midieron los niveles de luciferasa 6, 24, y 48 horas después del inicio de la transfección (véase la siguiente Tabla 8 y la figura 21).

Tabla 8

ARNm	RLU 6 horas	RLU 24 horas	RLU 48 horas
ppLuc(GC)-ag-histonaSL-A250	98472	734222	146479
ppLuc(GC)-ag-A64-histonaSL	123674	317343	89579
ppLuc(GC)-ag-histonaSL	7291	4565	916
ppLuc(GC)-ag-A300	4357	38560	11829
ppLuc(GC)-ag-A120	4371	45929	10142
ppLuc(GC)-ag-A64	1928	26781	537

Tanto la secuencia poli(A) A64 o la histonaSL dieron lugar a niveles comparables de luciferasa. De acuerdo con el experimento anterior, la combinación de A64 y la histonaSL aumentó fuertemente el nivel de luciferasa, muchas veces por encima del nivel observado con cualquiera de los elementos individuales. La magnitud del aumento en el nivel de luciferasa debido a la combinación de la poli(A) y la histonaSL en el mismo ARNm demuestra que están actuando sinérgicamente. La sinergia entre la A64 y la histonaSL se cuantificó como antes, en base a los niveles de luciferasa del ARNm del A64-histonaSL, A64 y la histonaSL (véase la siguiente Tabla 9). El nivel de luciferasa a partir de la combinación del ARNm A64 y la histonaSL era hasta 61,7 veces mayor que si los efectos de la poli(A) y la histonaSL fueran puramente aditivos.

Tabla 9

	A64	histonaSL	RLU 6 horas	RLU 24 horas	RLU 48 horas
	+	+	123674	317343	89579
	-	+	7291	4565	916
	+	-	1928	26781	537
Sinergia			13,4	10,1	61,7

Por el contrario, aumentar la longitud de la secuencia poli(A) de A64 hasta A120 o hasta A300 aumenta el nivel de luciferasa sólo moderadamente (véase la Tabla 8 y la figura 19). El ARNm con la secuencia poli(A) más larga, A300, también se comparó con el ARNm en el cual se combinó una secuencia poli(A) de longitud similar con la histonaSL, la histonaSL-A250. Además de tener una secuencia poli(A) larga, el orden de la histonaSL y la poli(A) se invierte en este ARNm con relación al ARNm de la A64-histonaSL. La combinación de A250 y la histonaSL aumenta fuertemente el nivel de luciferasa, muchas veces por encima del nivel observado con ya sea la histonaSL o A300. Nuevamente, se cuantificó la sinergia entre A250 y la histonaSL como antes en comparación con el RLU proveniente del ARNm de la histonaSL-A250 y el RLU proveniente del ARNm A300 más el ARNm de la histonaSL (véase la siguiente Tabla 10). El nivel de luciferasa a partir del ARNm que combina A250 y la histonaSL era hasta 17,0 veces mayor que si los efectos de la poli(A) y la histonaSL fueran puramente aditivos.

Tabla 10

	histonaSL	A250/A300	RLU 6 horas	RLU 24 horas	RLU 48 horas
	+	+	98472	734222	146479
	+	—	7291	4565	916
	—	+	4357	38560	11829
Sinergia			8,5	17,0	11,5

5 En resumen, se ha demostrado un efecto altamente sinérgico de la combinación de la histonaSL y la poli(A) sobre la expresión proteica a partir de ARNm para longitudes sustancialmente diferentes de la poli(A) e independientemente del orden de la poli(A) y la histonaSL.

**11.4 El aumento en la expresión proteica mediante la combinación de la poli(A) y la histonaSL es específico**

10 Para investigar si el efecto de la combinación de la poli(A) y la histonaSL sobre la expresión proteica a partir del ARNm es específico, se sintetizaron ARNm con secuencias alternativas en combinación con la poli(A): estos ARNm contenían 3' de A64 una de siete secuencias distintas, respectivamente. Los ARNm que codifican para luciferasa o el ARNm control se sometieron a electroporación en células HeLa. Se midieron los niveles de luciferasa 6, 24, y 48 horas después de la transfección (véase la siguiente Tabla 11 y la figura 22).

Tabla 11

ARNm	RLU 6 horas	RLU 24 horas	RLU 48 horas
ppLuc(GC)-ag-A64-N32	33501	38979	2641
ppLuc(GC)-ag-A64-SL	28176	20364	874
ppLuc(GC)-ag-A64-U30	41632	54676	3408
ppLuc(GC)-ag-A64-G30	46763	49210	3382
ppLuc(GC)-ag-A64-PolioCL	46428	26090	1655
ppLuc(GC)-ag-A64-aCPSL	34176	53090	3338
ppLuc(GC)-ag-A64-ag	18534	18194	989
ppLuc(GC)-ag-A64-histonaSL	282677	437543	69292
ppLuc(GC)-ag-histonaSL	27597	3171	0
ppLuc(GC)-ag-A64	14339	48414	9357

15 Tanto una secuencia poli(A) o la histonaSL dieron lugar a niveles comparables de luciferasa. Nuevamente, la combinación de la poli(A) y la histonaSL aumentó fuertemente el nivel de luciferasa, muchas veces por encima del nivel observado con cualquiera de los elementos individuales, actuando así sinérgicamente. Por el contrario, combinar la poli(A) con cualquiera de las secuencias alternativas no tuvo efecto sobre el nivel de luciferasa en comparación con el ARNm que contenía sólo una secuencia poli(A). De esta forma, la combinación de la poli(A) y la histonaSL aumenta la expresión proteica a partir del ARNm de una forma sinérgica y este efecto es específico.

**11.5 La combinación de la poli(A) y la histonaSL aumenta la expresión proteica a partir del ARNm de forma sinérgica *in vivo***

25 Para investigar el efecto de la combinación de la poli(A) y la histonaSL sobre la expresión proteica a partir de ARNm *in vivo*, ARNm que codifican para luciferasa con diferentes secuencias 3' de la 3'-UTR alfa-globina o del ARNm control se inyectaron intradérmicamente en ratones: los ARNm contenían ya sea una secuencia poli(A) A64 o una histonaSL en su lugar, o tanto una poli(A) A64 como una histonaSL 3' de la 3'-UTR. Se midieron los niveles de luciferasa 16 horas después de la inyección (véase la siguiente Tabla 12 y figura 23).

Tabla 12

ARNm	RLU 16 horas
ppLuc(GC)-ag-A64-histonaSL	38081
ppLuc(GC)-ag-histonaSL	137
ppLuc(GC)-ag-A64	4607

5 La luciferasa se expresó a partir del ARNm con ya sea una histonaSL o una secuencia poli(A). Sin embargo, sorprendentemente la combinación de la poli(A) y la histonaSL aumentó en gran medida adicionalmente el nivel de luciferasa, muchas veces por encima del nivel observado con cualquiera de los elementos individuales. La magnitud del aumento en el nivel de luciferasa debido a la combinación de la poli(A) y la histonaSL en el mismo ARNm demuestra que están actuando sinérgicamente.

10 La sinergia entre la poli(A) y la histonaSL se cuantificó dividiendo la señal proveniente del ARNm de poli(A) - histonaSL (+/+) entre la suma de las señales provenientes del ARNm de la histonaSL (-/+) más el ARNm de poli(A) (+/-) (véase la siguiente tabla 13).

Tabla 13

	A64	histonaSL	RLU 16 horas
	+	+	38081
	-	+	137
	+	-	4607
Sinergia			8.0

15 El factor así calculado especifica cuánto mayor es el nivel de luciferasa proveniente del ARNm que combina la poli(A) y la histonaSL es de lo que se podría esperar si los efectos de la poli(A) y la histonaSL fueran puramente aditivos. El nivel de luciferasa proveniente del ARNm al combinar la poli(A) y la histonaSL era 8 veces mayor que si sus efectos fueran puramente aditivos. Este resultado confirma que la combinación de la poli(A) y la histonaSL provoca un aumento marcadamente sinérgico en la expresión proteínica *in vivo*.

**11.6 La combinación de la poli(A) y la histonaSL aumenta el nivel de anticuerpos producidos por la vacunación con ARNm**

20 Para investigar el efecto de la combinación de la poli(A) y la histonaSL sobre la inducción de anticuerpos producidos por la vacunación con ARNm, se vacunaron intradérmicamente ratones Balb/c con los ARNm que codifican para HA con diferentes secuencias 3' de la 3'-UTR de alfa-globina mutada. Los ARNm contenían ya sea una secuencia poli(A) A64 o tanto una poli(A) A64 como la histonaSL 3' de la 3'-UTR. Mediante ELISA se analizó el nivel de anticuerpos HA-específicos en ratones vacunados y de control con diluciones en serie de suero (véase la figura 24).

25 Se indujo IgG1 anti-HA con un ARNm que tiene sólo una secuencia poli(A). Sin embargo, sorprendentemente la combinación de la poli(A) y la histonaSL aumentó fuertemente el nivel IgG1 anti-HA, por encima del nivel observado con sólo una secuencia poli(A).

**LISTADO DE SECUENCIAS**

- 30 <110> CUREVAC GMBH
- <120> Ácido nucleico que comprende o codifica para un tallo-bucle de histona y una secuencia poli(A) o una señal de poliadenilación para aumentar la expresión de un antígeno patogénico codificado
- 35 <130> CU01P127WO1
- <140> PCT/EP2012/000673
- <141> 2012-02-15
- <160> 56
- 40 <170> PatentIn version 3.5

# ES 2 701 652 T3

	<210>	1	
	<211>	16	
	<212>	ARN	
5	<213>	artificial	
	<220>		
	<223>	secuencia de tallo-bucle de histona de acuerdo con la fórmula (Ic): secuencia consenso de tallo-bucle de histona de metazoarios y protozoarios sin elementos frontera de tallo	
10	<220>		
	<221>	misc_feature	
	<222>	(1)..(1)	
15	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
	<220>		
	<221>	misc_feature	
	<222>	(3)..(8)	
20	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
	<220>		
	<221>	misc_feature	
25	<222>	(10)..(14)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
	<220>		
30	<221>	misc_feature	
	<222>	(16)..(16)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
35	<400>	1	16
		ngnnnnnnun nnnncn	
	<210>	2	
40	<211>	26	
	<212>	ARN	
	<213>	artificial	
	<220>		
45	<223>	secuencia de tallo-bucle de histona de acuerdo con la fórmula (IIc): secuencia consenso de tallo-bucle de histona de metazoarios y protozoarios con elementos frontera de tallo	
	<220>		
50	<221>	misc_feature	
	<222>	(1)..(6)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
55	<220>		
	<221>	misc_feature	
	<222>	(8)..(13)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
60	<220>		
	<221>	misc_feature	
	<222>	(15)..(19)	
65	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	

ES 2 701 652 T3

<220>  
 <221> misc\_feature  
 <222> (21)..(26)  
 5 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo

<400> 2  
 nnnnnngnnn nnnunnnnnc nnnnnn 26  
 10

<210> 3  
 <211> 16  
 <212> ARN  
 15 <213> artificial

<220>  
 <223> secuencia de tallo-bucle de histona de acuerdo con la fórmula (Id): sin elementos frontera de tallo  
 20

<220>  
 <221> misc\_feature  
 <222> (1)..(1)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 25

<220>  
 <221> misc\_feature  
 <222> (3)..(8)  
 30 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo

<220>  
 <221> misc\_feature  
 <222> (10)..(14)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 35

<220>  
 <221> misc\_feature  
 <222> (16)..(16)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 40

<400> 3  
 ncnnnnnnun nnnngn 16  
 45

<210> 4  
 <211> 26  
 <212> ARN  
 <213> artificial  
 50

<220>  
 <223> secuencia de tallo-bucle de histona de acuerdo con la fórmula (IId): con elementos frontera de tallo  
 55

<220>  
 <221> misc\_feature  
 <222> (1)..(6)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 60

<220>  
 <221> misc\_feature  
 <222> (8)..(13)  
 65

## ES 2 701 652 T3

- <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo
- 5 <220>  
 <221> misc\_feature  
 <222> (15)..(19)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo
- 10 <220>  
 <221> misc\_feature  
 <222> (21)..(26)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo
- 15 <400> 4  
 nnnnnncnnn nnnunnnng nnnnnn 26
- 20 <210> 5  
 <211> 16  
 <212> ARN  
 <213> artificial
- 25 <220>  
 <223> secuencia de tallo-bucle de histona de acuerdo con la fórmula (Ie): secuencia consenso de tallo-bucle de histona de protozoarios sin elementos frontera de tallo
- 30 <220>  
 <221> misc\_feature  
 <222> (3)..(8)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo
- 35 <220>  
 <221> misc\_feature  
 <222> (10)..(14)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo
- 40 <400> 5  
 dgnnnnnnun nnnnch 16
- 45 <210> 6  
 <211> 26  
 <212> ARN  
 <213> artificial
- 50 <220>  
 <223> secuencia de tallo-bucle de histona de acuerdo con la fórmula (Ile): secuencia consenso de tallo-bucle de histona de protozoarios con elementos frontera de tallo
- 55 <220>  
 <221> misc\_feature  
 <222> (1)..(5)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo
- 60 <220>  
 <221> misc\_feature  
 <222> (8)..(13)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo
- 65

## ES 2 701 652 T3

- <220>  
 <221> misc\_feature  
 <222> (15)..(19)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 5
- <220>  
 <221> misc\_feature  
 <222> (22)..(26)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 10
- <400> 6  
 nnnnndgnnn nnnunnnnc hnnnnn  
 15 26
- <210> 7  
 <211> 16  
 <212> ARN  
 <213> artificial  
 20
- <220>  
 <223> secuencia de tallo-bucle de histona de acuerdo con la fórmula (If): secuencia consenso de tallo-bucle de histona de metazoarios sin elementos frontera de tallo  
 25
- <220>  
 <221> misc\_feature  
 <222> (1)..(1)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 30
- <220>  
 <221> misc\_feature  
 <222> (3)..(3)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 35
- <220>  
 <221> misc\_feature  
 <222> (7)..(8)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 40
- <220>  
 <221> misc\_feature  
 <222> (10)..(10)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 45
- <220>  
 <221> misc\_feature  
 <222> (12)..(12)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 50
- <220>  
 <221> misc\_feature  
 <222> (14)..(14)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 55
- <220>  
 <221> misc\_feature  
 <222> (16)..(16)  
 <223> n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo  
 60

## ES 2 701 652 T3

	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
5	<400>	7	16
		ngnbyynnun vndncn	
10	<210>	8	
	<211>	26	
	<212>	ARN	
	<213>	artificial	
15	<220>		
	<223>	secuencia de tallo-bucle de histona de acuerdo con la fórmula (II <sub>f</sub> ): secuencia consenso de tallo-bucle de histona de metazoarios con elementos frontera de tallo	
20	<220>		
	<221>	misc_feature	
	<222>	(1)..(6)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
25	<220>		
	<221>	misc_feature	
	<222>	(8)..(8)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
30	<220>		
	<221>	misc_feature	
	<222>	(12)..(13)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
35	<220>		
	<221>	misc_feature	
	<222>	(15)..(15)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
40	<220>		
	<221>	misc_feature	
	<222>	(17)..(17)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
45	<220>		
	<221>	misc_feature	
	<222>	(19)..(19)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
50	<220>		
	<221>	misc_feature	
	<222>	(21)..(26)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
55	<220>		
	<221>	misc_feature	
	<222>	(21)..(26)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
60	<400>	8	26
		nnnnnngnby ynnunvndnc nnnnnn	
	<210>	9	
	<211>	16	
	<212>	ARN	
	<213>	artificial	
65			



## ES 2 701 652 T3

	<220>		
	<223>	secuencia de tallo-bucle de histona de acuerdo con la fórmula (Ig): secuencia consenso de tallo-bucle de histona de vertebrados sin elementos frontera de tallo	
5	<220>		
	<221>	misc_feature	
	<222>	(1)..(1)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
10	<220>		
	<221>	misc_feature	
	<222>	(8)..(8)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
15	<220>		
	<221>	misc_feature	
	<222>	(16)..(16)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
20	<220>		
	<221>	misc_feature	
	<222>	(16)..(16)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
	<400>	9	
25		nghyydnuh abrdcn	16
	<210>	10	
	<211>	26	
	<212>	ARN	
30	<213>	artificial	
	<220>		
	<223>	secuencia de tallo-bucle de histona de acuerdo con la fórmula (IIg): secuencia consenso de tallo-bucle de histona de vertebrados con elementos frontera de tallo	
35	<220>		
	<221>	misc_feature	
	<222>	(1)..(2)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
40	<220>		
	<221>	misc_feature	
	<222>	(4)..(6)	
45	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
	<220>		
	<221>	misc_feature	
	<222>	(13)..(13)	
50	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
	<220>		
	<221>	misc_feature	
55	<222>	(21)..(25)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
	<400>	10	
60		nnhnnghyy ydnuhabrdc nnnnnh	26
	<210>	11	
	<211>	16	
65	<212>	ARN	

ES 2 701 652 T3

	<213>	artificial	
	<220>		
5	<223>	secuencia de tallo-bucle de histona de acuerdo con la fórmula (Ih): secuencia consenso de tallo-bucle de histona de humano ( <i>Homo sapiens</i> ) sin elementos frontera de tallo	
	<400>	11 dghycudyuh asrrcc	16
10	<210>	12	
	<211>	26	
	<212>	ARN	
	<213>	artificial	
15	<220>		
	<223>	secuencia de tallo-bucle de histona de acuerdo con la fórmula (IIh): secuencia consenso de tallo-bucle de histona de humano ( <i>Homo sapiens</i> ) con elementos frontera de tallo	
20	<220>		
	<221>	misc_feature	
	<222>	(1)..(1)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
25	<220>		
	<221>	misc_feature	
	<222>	(25)..(25)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
30	<400>	12 nhaahdghyc udyuhasrrc cvhbnh	26
35	<210>	13	
	<211>	16	
	<212>	ADN	
	<213>	artificial	
40	<220>		
	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (Ic)	
45	<400>	13 vgyyyhhth rvrcb	16
	<210>	14	
	<211>	16	
50	<212>	ADN	
	<213>	artificial	
	<220>		
	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (Ic)	
55	<400>	14 sgyytytm arrrcs	16
60	<210>	15	
	<211>	16	
	<212>	ADN	
	<213>	artificial	
65	<220>		

## ES 2 701 652 T3

	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (1c)	
	<400>	15	
5		sgyyctttm agrrcs	16
	<210>	16	
	<211>	16	
	<212>	ADN	
10	<213>	artificial	
	<220>		
	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (1e)	
15	<220>		
	<221>	misc_feature	
	<222>	(3)..(5)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
20	<220>		
	<221>	misc_feature	
	<222>	(7)..(8)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
25	<220>		
	<221>	misc_feature	
	<222>	(12)..(14)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
30	<220>		
	<221>	misc_feature	
	<222>	(12)..(14)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
	<400>	16	
35		dgnnnbnnth vnnch	16
	<210>	17	
	<211>	16	
	<212>	ADN	
40	<213>	artificial	
	<220>		
	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (1e)	
45	<220>		
	<221>	misc_feature	
	<222>	(3)..(5)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
50	<220>		
	<221>	misc_feature	
	<222>	(13)..(14)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
55	<220>		
	<400>	17	
		rgnnnyhbth rdncy	16
60	<210>	18	
	<211>	16	
	<212>	ADN	
	<213>	artificial	
65	<210>	18	

## ES 2 701 652 T3

	<220>		
	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (Ie)	
5	<220>		
	<221>	misc_feature	
	<222>	(3)..(3)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
10	<220>		
	<221>	misc_feature	
	<222>	(14)..(14)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
15	<400>	18	
		rgndbyhyth rdhncy	16
20	<210>	19	
	<211>	16	
	<212>	ADN	
	<213>	artificial	
25	<220>		
	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (If)	
30	<400>	19	
		vgyyytyhth rvrrcb	16
35	<210>	20	
	<211>	16	
	<212>	ADN	
	<213>	artificial	
40	<220>		
	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (If)	
	<400>	20	
		sgyyctytm agrrcs	16
45	<210>	21	
	<211>	16	
	<212>	ADN	
	<213>	artificial	
50	<220>		
	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (If)	
55	<400>	21	
		sgyyctttm agrrcs	16
60	<210>	22	
	<211>	16	
	<212>	ADN	
	<213>	artificial	
65	<220>		
	<223>	secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) ccording to formula (Ig)	
	<400>	22	

# ES 2 701 652 T3

	ggyycttyth agrcc	16
5	<210> 23 <211> 16 <212> ADN <213> artificial	
10	<220> <223> secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (lg)	
	<400> 23 ggcycttytm agrgcc	16
15	<210> 24 <211> 16 <212> ADN <213> artificial	
20	<220> <223> secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (lg)	
25	<400> 24 ggctctttm agrgcc	16
30	<210> 25 <211> 16 <212> ADN <213> artificial	
35	<220> <223> secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (lh)	
	<400> 25 dghyctdyth asrrcc	16
40	<210> 26 <211> 16 <212> ADN <213> artificial	
45	<220> <223> secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (lh)	
50	<400> 26 ggcyctttth agrgcc	16
55	<210> 27 <211> 16 <212> ADN <213> artificial	
60	<220> <223> secuencia de tallo-bucle de histonas (sin elementos frontera de tallo) de acuerdo con la fórmula (lh)	
	<400> 27 ggcyctttm agrgcc	16
65	<210> 28 <211> 26	

ES 2 701 652 T3

	<212>	ADN	
	<213>	artificial	
5	<220>		
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (Ilc)	
10	<220>		
	<221>	misc_feature	
	<222>	(25)..(26)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
15	<400>	28	
		hhhhvgyyy yhhthrvrc bvhhnn	26
	<210>	29	
	<211>	26	
20	<212>	ADN	
	<213>	artificial	
25	<220>		
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (Ilc)	
	<400>	29	
		mhmhmsgyyy tlytmarrc smchhh	26
30	<210>	30	
	<211>	26	
	<212>	ADN	
	<213>	artificial	
35	<220>		
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (Ilc)	
40	<400>	30	
		mmmmmsgyyc tttmagrrc sachmh	26
45	<210>	31	
	<211>	26	
	<212>	ADN	
	<213>	artificial	
50	<220>		
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (Ile)	
55	<220>		
	<221>	misc_feature	
	<222>	(1)..(5)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
60	<220>		
	<221>	misc_feature	
	<222>	(8)..(10)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
65	<220>		

## ES 2 701 652 T3

	<221>	misc_feature	
	<222>	(12)..(13)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
5	<220>		
	<221>	misc_feature	
	<222>	(17)..(19)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
10	<220>		
	<221>	misc_feature	
	<222>	(22)..(22)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
15	<220>		
	<221>	misc_feature	
	<222>	(24)..(26)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
20	<220>		
	<221>	misc_feature	
	<222>	(24)..(26)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
25	<400>	31	26
		nnnnndgnnn bnrthvnnnc hnhnnn	
30	<210>	32	
	<211>	26	
	<212>	ADN	
	<213>	artificial	
35	<220>		
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (Ile)	
40	<220>		
	<221>	misc_feature	
	<222>	(1)..(2)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
45	<220>		
	<221>	misc_feature	
	<222>	(5)..(5)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
50	<220>		
	<221>	misc_feature	
	<222>	(8)..(10)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
55	<220>		
	<221>	misc_feature	
	<222>	(18)..(19)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
60	<220>		
	<221>	misc_feature	
	<222>	(25)..(26)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
65	<220>		

ES 2 701 652 T3

	<400>	32		
		nnhhnrgnnc yhbthrdnnc ydhnn	26	
5	<210>	33		
	<211>	26		
	<212>	ADN		
	<213>	artificial		
10	<220>			
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (Ile)		
15	<220>			
	<221>	misc_feature		
	<222>	(1)..(1)		
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo		
20	<220>			
	<221>	misc_feature		
	<222>	(8)..(8)		
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo		
25	<220>			
	<221>	misc_feature		
	<222>	(19)..(19)		
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo		
30	<220>			
	<221>	misc_feature		
	<222>	(19)..(19)		
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo		
35	<400>	33		
		nhhhvrngdb yhythrdhnc yrhhhh	26	
	<210>	34		
	<211>	26		
	<212>	ADN		
40	<213>	artificial		
	<220>			
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (IIf)		
45	<220>			
	<221>	misc_feature		
	<222>	(26)..(26)		
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo		
50	<400>	34		
		hhmhmvgyyy tyhthrvrc bvmhnn	26	
55	<210>	35		
	<211>	26		
	<212>	ADN		
	<213>	artificial		
60	<220>			
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (IIf)		
	<400>	35		
		mmmmmsgyyc ttytmgrrc smchhh	26	
65				



ES 2 701 652 T3

	<210>	36	
	<211>	26	
	<212>	ADN	
5	<213>	artificial	
	<220>		
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (If)	
10	<400>	36	26
		mmmmmsgyyc tttmagrrc sachmh	
	<210>	37	
15	<211>	26	
	<212>	ADN	
	<213>	artificial	
	<220>		
20	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (Ifg)	
	<220>		
	<221>	misc_feature	
25	<222>	(24)..(25)	
	<223>	n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
	<400>	37	
30		hhmamggyyc ttythagrrc cvhnm	26
	<210>	38	
	<211>	26	
35	<212>	ADN	
	<213>	artificial	
	<220>		
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (Ifg)	
40	<400>	38	
		hhaamggcyc ttytmagrgc cvchhm	26
45	<210>	39	
	<211>	26	
	<212>	ADN	
	<213>	artificial	
50	<220>		
	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (Ifg)	
55	<400>	39	26
		mmaamggctc tttmagrgc cmcymm	
	<210>	40	
60	<211>	26	
	<212>	ADN	
	<213>	artificial	
	<220>		

## ES 2 701 652 T3

	<223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (IIh)	
5	<220> <221> <222> <223>	misc_feature (1)..(1) n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
10	<220> <221> <222> <223>	misc_feature (25)..(25) n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
15	<400>	40 nhaahdghyc tdythasrrc cvhbnh	26
20	<210> <211> <212> <213>	41 26 ADN artificial	
25	<220> <223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (IIh)	
30	<220> <221> <222> <223>	misc_feature (25)..(25) n se selecciona a partir de un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo	
35	<400>	41 hhaamggcyc tttthagrgc cvmynm	26
40	<210> <211> <212> <213>	42 26 ADN artificial	
45	<220> <223>	secuencia de tallo-bucle de histona (con elementos frontera de tallo) de acuerdo con la fórmula (IIh)	
50	<400>	42 hmaaaggcyc tttmagrgc crmyhm	26
55	<210> <211> <212> <213>	43 1747 ARN artificial	
	<220> <223>	secuencia de ARNm de ppLuc(GC)-ag	
60	<400>	43 gggagaaagc uugaggaugg aggacccaa gaacaucaag aagggcccgg cgcccuuca	60
		cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccu	120
65		ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga	180

## ES 2 701 652 T3

guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 ccaccggauc guggugugcu cggagaacag ccugcaguuc ucaugccgg ugcuggggcg 300  
 5 ccucuucauc ggcguggccg ucgcccggc gaacgacauc uacaacgagc gggagcugcu 360  
 gaacagcaug gggauacagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420  
 10 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggcg 540  
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 15 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660  
 cugcgugcgc uucucgacg cccgggacc caucuucggc aaccagauca ucccggacac 720  
 20 cgccauccug agcggggugc cguuccacca cggcuucggc auguucacga ccugggcua 780  
 ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugc 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugc ucagcuucu 900  
 25 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcg 960  
 gggcgccccg cugagcaagg aggugggcca ggccguggcc aagcgguucc accuccggg 1020  
 30 cauccgccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca cccccgagg 1080  
 ggacgacaag ccggcgccg ugggcaaggu gguccgguuc uucgaggcca agguggugga 1140  
 ccuggacacc ggcaagacc ugggugugaa ccagcggggc gagcugugcg ugcggggg 1200  
 35 gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgcccua ucgacaagga 1260  
 cggcugcgug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucg 1320  
 40 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccg ccgagcugga 1380  
 gagcauccug cuccagcacc ccaacaucu cgacgccggc guggccggg ugccggacga 1440  
 cgacgccggc gagcugccg ccgcguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 45 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaagaagc ugcggggcg 1560  
 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620  
 50 ccgcgagauc cugaucaagg ccaagaagg cggaagauc gccguguaag acuaguuaa 1680  
 agacugacua gcccgauagg ccucccaacg ggcccuccuc ccucuuugc accgagaua 1740  
 55 auagauc 1747  
 <210> 44  
 <211> 1806  
 <212> ARN  
 60 <213> artificial  
 <220>  
 <223> secuencia de ARNm de ppLuc(GC)-ag-A64  
 65 <400> 44

ES 2 701 652 T3

gggagaaagc uugaggauagg aggacgcaa gaacaucaag aagggcccg cgcccuucia 60  
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccu 120  
 5 ggugccgggc acgauccgu ucaccgacgc ccacaucgag gucgacauc ccaucgagg 180  
 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 10 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggggc 300  
 ccucucauc ggcguggccg ucgcccggc gaacgacauc uacaacgagc gggagcugcu 360  
 gaacagcaug gggauccagc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420  
 15 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggggg 540  
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 20 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660  
 cugcgugcgc uucucgcacg cccgggacc caucuucggc aaccagauca ucccggacac 720  
 25 cgccauccug agcguugguc cguuccacca cggcuucggc auguucacga ccuggggcu 780  
 ccuacucgc ggcuccggg ugguccugau guaccgguuc gaggaggagc uguuccugc 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugc ucagcuucu 900  
 30 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgcccagg 960  
 gggcgcccc cugagcaagg aggugggcca ggccguggcc aagcgguucc accuccggg 1020  
 35 cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgagg 1080  
 ggacgacaag ccggcgccg ugggcaaggu ggucccuuc uucgaggcca agguggugga 1140  
 ccuggacacc ggcaagacc ugggugugaa ccagcggggc gacgugugc ugcggggggc 1200  
 40 gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgccuca ucgacaagga 1260  
 cggcuggcug cacagcggc acaucgccua cugggacgag gacgagcacu ucuucaucg 1320  
 45 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccg ccgagcugga 1380  
 gagcauccug cuccagcacc ccaacaucu cgacgcccgc guggccggc ugcgggacga 1440  
 cgacccggc gacgugccg ccgcguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 50 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcg 1560  
 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620  
 55 ccgaggauc cugaucaagg ccaagaagg cggcaagauc gccuguaag acuaguuaa 1680  
 agacugacua gccgauggg ccuccaacg ggcccuccc cccuccuugc accgagaua 1740  
 60 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800  
 aaaaaa 1806  
 <210> 45  
 65 <211> 1772

## ES 2 701 652 T3

<212> ARN  
 <213> artificial  
  
 <220>  
 5 <223> secuencia de ARNm de ppLuc(GC)-ag-histonaSL  
  
 <400> 45  
 gggagaaagc uugaggaugg aggacgcaa gaacaucaag aagggcccgg cgcccuucia 60  
 10 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccu 120  
 ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga 180  
 15 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggcg 300  
 ccucucauc ggcguggccg ucgcccggc gaacgacauc uacaacgagc gggagcugcu 360  
 20 gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420  
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggcg 540  
 25 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660  
 30 cugcugcgcg uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac 720  
 cgccauccug agcugggugc cguuccacca cggcuucggc auguucacga ccugggcua 780  
 ccuacucgc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugc 840  
 35 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugc ucagcuucu 900  
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960  
 40 gggcgccccg cugagcaagg aggugggcca ggccguggcc aagcgguucc accuccggg 1020  
 cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca ccccggagg 1080  
 ggacgacaag ccggcgccg ugggcaaggu ggucccuuc uucgaggcca agguggugga 1140  
 45 ccuggacacc ggcaagacc uggcgugaa ccagcggggc gagcugugcg ugcggggg 1200  
 gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga 1260  
 50 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucg 1320  
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga 1380  
 gagcauccug cuccagcacc ccaacaucuu cgacgccggc guggccgggc ugccggacga 1440  
 55 cgacgccggc gagcugccgg ccgcguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcg 1560  
 60 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620  
 ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccuguaag acuaguuaa 1680  
 agacugacua gccgauggg ccuccaacg ggccuccuc ccucuuugc accgagaua 1740  
 65

# ES 2 701 652 T3

auagaucuca aaggcucuuu ucagagccac ca
1772

5 <210> 46  
 <211> 1835  
 <212> ARN  
 <213> artificial  
  
 10 <220>  
 <223> secuencia de ARNm de ppLuc(GC)-ag-A64-histonaSL  
  
 <400> 46  
 gggagaaagc uugaggauagg aggacgcaa gaacaucaag aagggcccgg cgcccuucua 60  
 15 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccu 120  
 ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauc ccuacgcgga 180  
 20 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggcg 300  
 ccucucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu 360  
 25 gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420  
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 30 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccgccggg 540  
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660  
 35 cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac 720  
 cgccauccug agcguggugc cguuccacca cggcuucggc auguucacga ccugggcua 780  
 40 ccucaucugc ggcuuccggg ugguccugau guaccguuc gaggaggagc uguuccugc 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugcu ucagcuucuu 900  
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960  
 45 gggcgccccg cugagcaagg aggugggcga ggccguggcc aagcgguuucc accuuccggg 1020  
 cauccgccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca cccccgaggg 1080  
 50 ggacgacaag ccgggcgccc ugggcaaggu gguccguuc uucgaggcca agguggugga 1140  
 ccuggacacc ggcaagaccc ugggcgugaa ccagcggggc gagcugugcg ugcgggggccc 1200  
 gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga 260  
 55 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320  
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcggccg ccgagcugga 1380  
 60 gagcauccug cuccagcacc ccaacaucuu cgacgcccgc guggccgggc ugccggacga 1440  
 cgacgccggc gagcugccgg ccgcguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcgg 1560  
 65 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccgaagau 1620

# ES 2 701 652 T3

cgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguuua 1680  
 agacugacua gcccgauggg ccucccaacg ggcccuccuc cccuccuugc accgagauua 1740  
 5 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800  
 aaaaaaugca ucaaaggcuc uuuucagagc cacca 1835  
 10 <210> 47  
 <211> 1869  
 <212> ARN  
 <213> artificial  
 15 <220>  
 <223> secuencia de ARNm de ppLuc(GC)-ag-A120  
 <400> 47  
 20 gggagaaagc uagaggauagg aggacgcaa gaacaucaag aagggcccgg cgcccuucua 60  
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccu 120  
 ggugccggg acgaucccu ucaccgagc ccacaucgag gucgacauc cuuacgagg 180  
 25 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggcg 300  
 30 ccucucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu 360  
 gaacagcaug gggauagcc agccgaccgu gguguucgug agcaagaag gccugcagaa 420  
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 35 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggggg 540  
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 40 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660  
 cugcgugcg uucucgcacg cccgggacc caucuucggc aaccagauca ucccggacac 720  
 cgccauccug agcggugugc cguuccacca cggcuucggc auguucacga ccuggggcu 780  
 45 ccucaucugc ggcuccggg ugguccugau guaccguuc gaggaggagc uguuccugc 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccuugu ucagcuucu 900  
 50 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcagcaga ucgccagcgg 960  
 gggcgcccc cugagcaagg aggugggcga ggccguggcc aagcgguuucc accuccggg 1020  
 cauccgccag ggcuaaggcc ugaccgagac cagagcgcg auccugauca ccccggagg 1080  
 55 ggacgacaag ccgggcgccc ugggcaaggu gguccguuc uucgaggcca agguggugga 1140  
 ccuggacacc ggcaagacc ugggcgugaa ccagcggggc gagcugugcg ugcgggggccc 1200  
 60 gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgccuca ucgacaagga 1260  
 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320  
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcggccg ccgagcugga 1380  
 65

## ES 2 701 652 T3

gagcauccug cuccagcacc ccaacaucuu cgacgccggc guggccgggc ugccggacga 1440  
 cgacgccggc gagcugccgg ccgcgguugu ggugcuggag cacggcaaga ccaugacgga 1500  
 5 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaagaagc ugcggggcgg 1560  
 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620  
 10 ccgcgagauc cugaucagg ccaagaagg cggaagauc gccguguaag acuaguuaa 1680  
 agacugacua gcccgauggg ccuccaacg ggcccuccuc cccuccuugc accgagaua 1740  
 auagaucuaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1800  
 15 aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1860  
 aaaaaaaaa 1869  
  
 20 <210> 48  
 <211> 1858  
 <212> ARN  
 <213> artificial  
  
 25 <220>  
 <223> secuencia de ARNm de ppLuc(GC)-ag-A64-ag  
  
 <400> 48  
 30 gggagaaagc uugaggauug aggacgcaa gaacaucaag aagggcccgg cgccuucua 60  
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccu 120  
 ggugcccggc acgaucgcu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga 180  
 35 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggc 300  
 ccucucauc ggcguggccg ucgcccggc gaacgacauc uacaacgagc gggagcugcu 360  
 40 gaacagcaug gggauacgac agccgaccgu gguguucgug agcaagaagc gccugcagaa 420  
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 45 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggcgg 540  
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660  
 50 cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac 720  
 cgcauccug agcguugguc gguuccacca cgguucggc auguucacga cccugggcu 780  
 55 ccuacucgc ggcuuccggg ugguccugau guaccguuc gaggaggagc uguuccugc 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugc ucagcuucuu 900  
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcagcaga ucgaccgagc 960  
 60 gggcgcccc cugagcaagg aggugggcga ggccguggcc aagcguucc accuccggg 1020  
 cauccgccag ggcucggcc ugaccgagac cacgagcgcg auccugauca ccccggagg 1080  
 65 ggacgacaag ccggcgccg ugggcaaggu ggucccuuc uucgaggcca agguggugga 1140



## ES 2 701 652 T3

ccuggacacc ggcaagaccc ugggcgugaa ccagcggggc gagcugugcg ugcgggggccc 1200  
 5 gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccuca ucgacaagga 1260  
 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320  
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga 1380  
 10 gagcauccug cuccagcacc ccaacaucuu cgacgccggc guggccgggc ugccggacga 1440  
 cgacgccggc gagcugccgg ccgcguggu ggugcuggag cacggcaaga ccaugacgga 500  
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcgg 1560  
 15 cgugguguuc guggacgagg ucccgaaggg ccugaccggg aagcucgacg cccggaagau 1620  
 ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguuaa 1680  
 20 agacugacua gcccgauggg ccucccaacg ggcccuccuc cccuccuugc accgagauua 1740  
 aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1800  
 aaaaaaugca uccgcccga ugggccuccc aacgggccc cccuccucc ugcaccg 1858  
 25 <210> 49  
 <211> 1894  
 <212> ARN  
 <213> artificial  
 30 <220>  
 <223> secuencia de ARNm de ppLuc(GC)-ag-A64-aCPSL  
 <400> 49  
 35 gggagaaagc uugaggauagg aggcgcca gaacaucaag aagggcccgg cgcccuucua 60  
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccc 120  
 ggugcccggc acgaucgcu ucaccgacgc ccacaucgag gucgacauc ccaucgagg 180  
 40 guacuucgag augagcguvc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggccc 300  
 45 ccucuaucau ggcguggccg ucgcccggc gaacgacauc uacaacgagc gggagcugcu 360  
 gaacagcaug gggauagcc agccgaccgu gguguucgug agcaagaag gccugcagaa 420  
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 50 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggcggg 540  
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 55 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660  
 cugcgugcgc uucucgcacg cccgggacc caucuucggc aaccagauca ucccggacac 720  
 cgccauccug agcgguguc cguuccacca cggcuucggc auguucacga cccugggcu 780  
 60 ccucaucugc ggcuccggg ugguccugau guaccgguuc gaggaggagc uguuccugc 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccuugu ucagcuucuu 900  
 65 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgcccagg 960

## ES 2 701 652 T3

5  
 10  
 15  
 20  
 25  
 30  
 35  
 40  
 45  
 50  
 55  
 60  
 65

gggcgccccg cugagcaagg agggugggcga ggccguggcc aagcgguucc accucccggg 1020  
 cauccgccag ggcuacggcc ugaccgagac caccgagcgc auccugauca cccccgaggg 1080  
 ggacgacaag ccgggcccgc ugggcaaggu ggucccgauu uucgaggcca aggguggugga 1140  
 ccuggacacc ggcaagaccc ugggcgugaa ccagcggggc gagcugugcg ugcggggggc 1200  
 gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga 1260  
 cggcuggcug cacagcggcg acaucgcca cugggacgag gacgagcacu ucuucaucgu 1320  
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga 1380  
 gagcauccug cuccagcacc ccaacaucuu cgacgcccgc guggccgggc ugccggacga 1440  
 cgacgcccgc gagcugcccg ccgcgguugu ggugcuggag cacggcaaga ccaugacgga 1500  
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaagaagc ugcggggcgg 1560  
 cgugguguuc guggacgagg uccgaaggg ccugaccggg aagcucgacg cccggaagau 1620  
 ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguuaa 1680  
 agacugacua gcccgauggg ccucccaacg ggcccuccuc cccuccuugc accgagauua 1740  
 auaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1800  
 aaaaaaugca ucaauuccua cacgugaggc gcugugauuc ccuaucuccc uucauuccu 1860  
 auacauuagc acagcgccau ugcauguagg aaau 1894

<210> 50  
 <211> 1909  
 <212> ARN  
 <213> artificial

<220>  
 <223> secuencia de ARNm de ppLuc(GC)-ag-A64-PolioCL

<400> 50  
 gggagaaagc uugaggauug aggacgcaa gaacaucaag aagggcccgg cgcccuucua 60  
 cccgucggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccc 120  
 ggugcccggc acgaurcgc ucaccgacgc ccacaucgag gucgacauc cuacgcccga 180  
 guacuucgag augagcugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugcccg ugcuggggcg 300  
 ccucuucauc ggcguggccg ucgcccggc gaacgacauc uacaacgagc gggagcugcu 360  
 gaacagcaug gggauccagc agccgaccgu gguguucgug agcaagaagc gccugcagaa 420  
 gauccugaac gugcagaaga agcugcccua cauccagaag aucaucauca uggacagcaa 480  
 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggcggg 540  
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 caugaacagc agcggcagca ccggccugcc gaagggggug gccugcccgc accggaccgc 660  
 cugcugcgc uucucgacg cccgggacc caucuucgc aaccagauc ucccggacac 720

## ES 2 701 652 T3

cgccauccug agcguggugc cguuccacca cggcuucggc auguucacga cccugggcu 780  
 5 ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgacccugu ucagcuucuu 900  
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960  
 10 gggcgccccg cugagcaagg aggugggcga ggccguggcc aagcgguucc accucccggg 1020  
 cauccgccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca cccccgaggg 1080  
 ggacgacaag cggggcgccg ugggcaaggu ggucccguuc uucgaggcca agguggugga 1140  
 15 ccuggacacc ggcaagacc ugggcgugaa ccagcggggc gagcugugcg ugcggggggc 1200  
 gaugaucaug agcgguacg ugaacaacc ggaggccacc aacgccuca ucgacaagga 1260  
 20 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320  
 cgaccggcug aagucgcuga ucaaguaca gggcuaccag guggcgccgg cggagcugga 1380  
 25 gagcauccug cuccagcacc ccaacauuu cgacgcggc guggccggg ugcgggacga 1440  
 cgacgcggc gagcugccg cgcgguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcg 1560  
 30 cgugguguuc guggacgagg uccgaaggg ccugaccggg aagcucgacg cccggaagau 1620  
 ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguuaa 1680  
 35 agacugacua gcccgaugg ccucccaacg ggcccuccc cccuccuugc accgagaua 1740  
 auaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1800  
 aaaaaaugca ucauuucua aacagcucug gguuguuacc cccccagag gccacgugg 1860  
 40 cggcuaguac uccgguauug cgguaccuu guacgccugu uuuagaau 1909

<210> 51  
 <211> 1841  
 45 <212> ARN  
 <213> artificial

<220>  
 <223> secuencia de ARNm de ppLuc(GC)-ag-A64-G30

50 <400> 51  
 gggagaaagc uugaggauug aggacgcaa gaacaucaag aagggcccgg cgccuucua 60  
 cccgucggag gacgggaccg cggcgagca gcuccacaag gccaugaagc gguacgccu 120  
 55 ggugccggc acgauccu ucaccgacgc ccacaucgag gucgacauc cuacgcgga 180  
 guacuucgag augagcugc gccuggccga ggccaugaag cgguacggc ugaacaccaa 240  
 60 ccaccggauc guggugugc cggagaacag ccugcaguuc uucaugccg ugcugggcg 300  
 ccucucauc ggcguggccg ucgcccggc gaacgacauc uacaacgagc gggagcugc 360  
 gaacagcaug gggauccagc agccgaccg gguguucgug agcaagaag gccucagaa 420  
 65

## ES 2 701 652 T3

gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccgccggg 540  
 5 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 caugaacagc agcggcagca cgggccugcc gaagggggug gccugccgc accggaccgc 660  
 10 cugcugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac 720  
 cgccaucug agcguugugc cguuccacca cggcuucggc auguucacga ccuggggcu 780  
 ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugc 840  
 15 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugug ucagcuucu 900  
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcg 960  
 20 gggcgccccg cugagcaagg aggugggcca ggccgugcc aagcguucc accuccggg 1020  
 cauccgccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca cccccgagg 1080  
 ggacgacaag cggggcggc ugggcaaggu ggucccuuc uucgaggcca agguggugga 1140  
 25 ccuggacacc ggcaagacc ugggugugaa ccagcggggc gacugugcg ugcggggcc 1200  
 gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgccuca ucgacaagga 1260  
 cggcuggcug cacagcggc acaucgccua cugggacgag gacgagcacu ucuucaucg 1320  
 30 cgaccggcug aagucgcuga ucaaguacaa gggcuaccg guggcggcg ccgagcugga 1380  
 gagcaucug succagcacc ccaacaucu cgacgccgc guggccggc ugccggacga 1440  
 35 cgacccggc gacugcccg ccgcguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 gaaggagauc gucgaucag uggccagcca ggugaccacc gccagaagc ugcggggcg 1560  
 cgugguuc guggacgagg ucccgaagg ccugaccggg aagcucgac cccggaagau 1620  
 40 ccgaggauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguuaa 1680  
 agacugacua gccgaugg ccuccaacg ggcccuccc ccucuuugc accgagaua 1740  
 45 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800  
 aaaaaaugca uggggggggg gggggggggg gggggggggg g 1841

50 <210> 52  
 <211> 1841  
 <212> ARN  
 <213> artificial

55 <220>  
 <223> secuencia de ARNm de ppLuc(GC)-ag-A64-U30

<400> 52  
 60 gggagaaagc uugaggauagg aggacgcca gaacaucaag aagggcccgg cgccuucua 60  
 cccgucggag gacgggacc cggcgagca gcuccacaag gccaugaagc gguacgccu 120  
 ggugccggc acgaucgcc ucaccgagc ccacaucgag gucgaucua ccuacgagg 180  
 65 guacuucgag augagcugc gccuggccga ggccaugaag cgguacggc ugaacaccaa 240

## ES 2 701 652 T3

ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggcg 300  
 5 ccucucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu 360  
 gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420  
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 10 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggcg 540  
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 caugaacagc agcggcagca cggccugcc gaagggggug gccugccgc accggaccgc 660  
 15 cugcugcgcg uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac 720  
 cgccaucug agcggugguc cguuccacca cggcuucggc auguucacga cccugggcu 780  
 20 ccucaucugc ggcuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugcu ucagcuucu 900  
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960  
 25 gggcgccccg cugagcaagg aggugggcca ggccguggcc aagcgguucc accuccggg 1020  
 cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgagg 1080  
 30 ggacgacaag ccggcgccg ugggcaaggu ggucccuuc uucgaggcca agguggugga 1140  
 ccuggacacc ggcaagacc ugggugugaa ccagcggggc gagcugugcg ugcggggg 1200  
 gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgcccua ucgacaagga 1260  
 35 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucg 1320  
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga 1380  
 40 gagcauccug cuccagcacc ccaacaucu cgacgccggc guggccgggc ugccggacga 1440  
 cgacgccggc gagcugccgg ccgcguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcg 1560  
 45 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620  
 ccgagagauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguuaa 1680  
 50 agacugacua gccgauggg ccucccaacg ggcccuccc cccuccuugc accgagaua 740  
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800  
 55 aaaaaaugca uuuuuuuuuu uuuuuuuuuu uuuuuuuuuu u 1841  
 <210> 53  
 <211> 1857  
 <212> ARN  
 60 <213> artificial  
 <220>  
 <223> secuencia de ARNm de ppLuc(GC)-ag-A64-SL  
 65 <400> 53

ES 2 701 652 T3

gggagaaagc uugaggaugg aggacgcaa gaacaucaag aagggcccg cgcccuucia 60  
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccu 120  
 5 ggugccgggc acgaucgcu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga 180  
 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 10 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggggc 300  
 ccucucauc ggcguggccg ucgcccggc gaacgacauc uacaacgagc gggagcugcu 360  
 gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420  
 15 gauccugaac gugcagaaga agcugccau cauccagaag aucaucauca uggacagcaa 480  
 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggcggg 540  
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 20 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660  
 cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac 720  
 25 cgccauccug agcguugugc cguuccacca cggcuucggc auguucacga ccuggggcu 780  
 ccuacucgc ggcuccggg ugguccugau guaccgguuc gaggaggagc uguuccugc 840  
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugc ucagcuucu 900  
 30 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960  
 gggcgccccg cugagcaagg aggugggcga ggccguggcc aagcgguucc accuccggg 1020  
 35 cauccgccag ggcucggcc ugaccgagac cacgagcgcg auccugauca ccccgaggg 1080  
 ggacgacaag ccggcgccg ugggcaaggu ggucccuuc uucgaggcca agguggugga 1140  
 ccuggacacc ggcaagacc uggcgugaa ccagcggggc gacgugugc ugcgggggccc 1200  
 40 gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgccuca ucgacaagga 1260  
 cggcuggcug cacagcggc acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320  
 45 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga 1380  
 gagcauccug cuccagcacc ccaacaucuu cgacgccggc guggccgggc ugccggacga 1440  
 cgacgccggc gacgugccgg ccgcguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 50 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcgg 1560  
 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620  
 55 ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguuaa 1680  
 agacugacua gcccgauagg ccuccaacg ggcccuccc cccuccuugc accgagaua 1740  
 60 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800  
 aaaaaaugca uuaugcggc cguguccacc acggauauca ccguggugga cgcgcc 1857  
 <210> 54  
 65 <211> 1838

# ES 2 701 652 T3

<212> ARN  
 <213> artificial  
  
 <220>  
 5 <223> ppLuc(GC)-ag-A64-N32  
  
 <400> 54  
 gggagaaagc uugaggaugg aggacgcaa gaacaucaag aagggcccgg cgcccuucia 60  
 10 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccu 120  
 ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga 180  
 15 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240  
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggcg 300  
 ccucucauc ggcguggccg ucgcccggc gaacgacauc uacaacgagc gggagcugcu 360  
 20 gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420  
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480  
 gaccgacuac caggguucc agucgaugua cacguucgug accagccacc ucccggcggg 540  
 25 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600  
 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660  
 30 cugcugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac 720  
 cgccauccug agcguugugc cguuccacca cggcuucggc auguucacga ccugggcua 780  
 ccuacucgc ggcuuaccggg ugguccugau guaccgguuc gaggaggagc uguuccugc 840  
 35 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugc ucagcuucu 900  
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960  
 40 gggcgccccg cugagcaagg aggugggcca ggccguggcc aagcgguucc accuccggg 1020  
 cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgagg 1080  
 ggacgacaag ccggcgccg ugggcaaggu ggucccuuc uucgaggcca agguggugga 1140  
 45 ccuggacacc ggcaagacc ugggugugaa ccagcggggc gagcugugcg ugcgggggccc 1200  
 gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga 1260  
 50 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320  
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga 1380  
 gagcauccug cuccagcacc ccaacaucuu cgacgccggc guggccgggc ugccggacga 1440  
 55 cgacgccggc gagcugccgg ccgcguggu ggugcuggag cacggcaaga ccaugacgga 1500  
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcgg 1560  
 60 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620  
 ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccuguaag acuaguuaa 1680  
 agacugacua gcccgauagg ccuccaacg ggccuccuc ccucuuugc accgagaua 1740  
 65

## ES 2 701 652 T3

aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800  
 aaaaaaugca uccccucua gacaauugga auuccaua 1838  
 5  
   <210> 55  
   <211> 1902  
   <212> ARN  
   <213> artificial  
 10  
   <220>  
   <223> secuencia de ARNm de HA (H1N1/PR8) (GC)-ag-A64-C30  
   <400> 55  
 15 gggagaaagc uuaccaugaa ggccaaccug cucgugcugc ugugcgcccu cgcggccgcc 60  
   gagccgaca ccaucugcau cggcuaccac gccacaaca gcaccgacac ggucgacacc 120  
   gugcuggaga agaacgugac cgucaccac uccgugaacc ugcucgagga cagccacaac 180  
 20 ggggaagcugu gccggcugaa gggcaucgag cccuccagc uggggaagug caaucugcc 240  
   ggcuggcugc ucgggaacc ggagugcgac cccugcugc cggugcguc cuggagcuac 300  
 25 aucgucgaga cgccaacuc cgagaacggc aucugcuacc cggcgacuuc caucgacuac 360  
   gaggagcucc gggagcagcu gagcuccgug agcuccuucg agcgcucgga gaucuuccc 420  
   aaggagagcu ccuggcccaa ccacaacacc aacgggguga ccgccgcug cagccacgag 480  
 30 ggcaauguca gcuucuaccg gaaccugcuc uggcugaccg agaaggagg guccuacc 540  
   aagcugaaga acagcuacgu caacaagaag ggcaaggagg ugcucgugcu gugggggaur 600  
 35 caccaccgc ccaacuccaa ggagcagcag aaccuguacc agaacgagaa cgcuacguc 660  
   agcuggguga cguccaacua caaccgccg uucaccccg agaucgccga gcgcccag 720  
   guccgggacc aggccggccg caugaacuac uacuggacc uccugaagcc gggcgacacc 780  
 40 aucaucuucg aggccaacgg gaaccugauc gcccgaugu acgcuucgc ccucagccg 840  
   ggcuucggga gcggcaucau cacguccaac gccagcaugc acgagugcaa caccaaguc 900  
 45 cagaccccc uggcgccau caacuccagc cugccuacc agaacaucca cccgugacc 960  
   aucggggagu gcccagaa cgugcguc gccaagcucc ggauggucac gggccugcgc 1020  
   aacaaccga gcauccaguc ccgggggugc uucggcgca ucgcccgguu caucgagggc 1080  
 50 ggcuggaccg ggaugaucga cggcugguac gguaccacc accagaacga gcagggcagc 1140  
   ggguacgcc cggaccagaa guccaccag aacgccauca acggcaucac caacaaggug 1200  
 55 aacacgguga ucgagaagau gaacaucagc uucaccgccg ucgcaagga guucaacaag 1260  
   cucgagaagc gcauggagaa ccugaacaag aagguggacg acggguuccu ggacaucugc 1320  
   accuacaacg ccgagcuccu ggugcugcuc gagaacgagc ggaccucgga cuuccacgac 1380  
 60 agcaacguca agaaccugua cgagaaggug aagucccagc ucaagaaca cgccaaggag 1440  
   aucggcaacg ggugcuucga guucuaccac aagugcgaca acgagugcau ggagagcugc 1500  
 65 cgcaacggca cguacgacua cccaaguac uccgaggaga gcaagcugaa ccgggagaag 1560



## ES 2 701 652 T3

guggacgggg ugaagcugga guccaugggc aucuaccaga uccucgccau cuacagcacc 1620  
 5 gucgccucca gccuggugcu gcuggugucc cucggcgcgga ucagcuucug gaugugcagc 1680  
 aacggguccc ugcagugccg caucugcauc ugaccacuag uuuaaagacu gacuagcccc 1740  
 augggccucc caacgggccc uccuccccuc cuugcaccga gauuaaauaa aaaaaaaaaa 1800  
 10 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa auauuccccc 1860  
 ccccccccc ccccccccc cccccucuag acaauuggaa uu 1902  
  
 15 <210> 56  
 <211> 1915  
 <212> ARN  
 <213> artificial  
  
 20 <220>  
 <223> secuencia de ARNm de HA (H1N1/PR8) (GC)-ag-A64-C30-histonaSL  
 <400> 56  
 25 gggagaaagc uuaccaugaa ggccaaccug cucgugcugc ugugcgcccu cgcgggccgc 60  
 gacgccgaca ccaucugcau cggcuaccac gccaacaaca gcaccgacac ggucgacacc 120  
 gugcuggaga agaacgugac cgucaccac uccgugaacc ugcucgagga cagccacaac 180  
 30 gggaagcugu gccggcugaa gggcaucgcg cccucccagc uggggaagug caacaucgcc 240  
 gccuggcugc ucggaacc ggagugcgc cccugcugc ccgugcgcuc cuggagcuac 300  
 aucgucgaga cgccaacuc cgagaacggc aucugcuacc cgggcgacuu caucgacuac 360  
 35 gaggagcucc gggagcagcu gagcuccgug agcuccuucg agcgcuucga gaucuuccc 420  
 aaggagagcu ccuggcccaa ccacaacacc aacgggguga ccgccgccug cagccacgag 480  
 40 ggcaagucca gcuucuaccg gaaccugcuc uggcugaccg agaaggagg guccuacccc 540  
 aagcugaaga acagcuacgu caacaagaag ggcaaggagg ugcucgugcu gugggggauc 600  
 caccaccgc ccaacuccaa ggagcagcag aaccuguacc agaacgagaa cgcguacguc 660  
 45 agcuggguga cguccaacia caaccgccg uucaccccc agaucgccga gcgccccaaag 720  
 guccgggacc aggccggccg caugaacuac uacuggacc uccugaagcc gggcgacacc 780  
 50 aucaucuucg aggccaacgg gaaccugauc gcccgaugu acgcuucgc ccucagccgg 840  
 ggcuucggga gcggcaucau cacguccaac gccagcaugc acgagugcaa caccaagugc 900  
 cagaccccc ugggcgccau caacuccagc cugcccuacc agaacaacca cccgugacc 960  
 55 aucggggagu gcccgaagua cgugcgcucc gccaagcucc ggauggucac gggccugcgc 1020  
 aacaaccca gcauccaguc ccggggcgug uucggcgcgga ucgcccgguu caucgagggc 1080  
 60 gccuggaccg ggaugaucga cggcugguac gguuaccacc accagaacga gcagggcagc 1140  
 gguuacgcc cgcaccagaa guccaccag aacgccauca acggcaucac caacaaggug 1200  
 aacacgguga ucgagaagau gaacaucag uucaccgccg ucggcaagga guucaacaag 1260  
 65

# ES 2 701 652 T3

cucgagaagc gcauggagaa ccugaacaag aagguggacg acggguuccu ggacaucugg 1320  
accuacaacg cggagcuccu ggugcugcuc gagaacgagc ggacccugga cuuccacgac 1380  
5 agcaacguca agaaccugua cgagaaggug aagucccagc ucaagaacaa cgccaaggag 1440  
aucggcaacg ggugcuucga guucuaccac aagugcgaca acgagugcau ggagagcguc 1500  
10 cgcaacggca cguacgacua cccaaguac uccgaggaga gcaagcugaa ccgggagaag 1560  
guggacgggg ugaagcugga guccaugggc aucuaccaga uccucgcca cuacagcacc 1620  
gucgccucca gccuggugcu gcuggugucc cucggcgca ucagcuucug gaugugcagc 1680  
15 aacggguccc ugcagugccg caucugcauc ugaccacuag uuuaagacu gacuagcccg 1740  
augggccucc caacggggcc uccuccccuc cuugcaccga gauuaaaaa aaaaaaaaaa 1800  
20 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa augcaucucc 1860  
ccccccccc cccccccc ccccccaaa ggcucuuuuc agagccacca gaauu 1915

25

**Reivindicaciones**

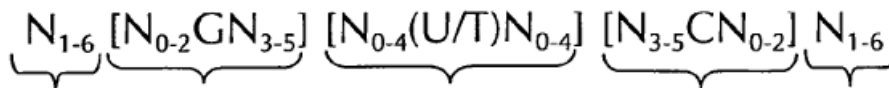
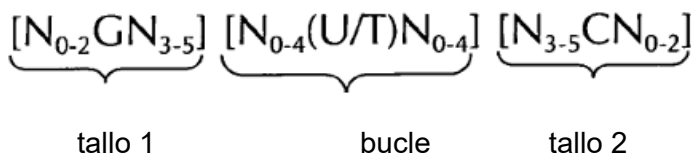
1. Secuencia de ácido nucleico que comprende o codifica en la dirección 5' → 3' para:

- 5 i) – una región codificante, que codifica para al menos un péptido o proteína;  
 – al menos un tallo-bucle de histona, y  
 – una secuencia poli(A),  
 o  
 10 ii) – una región codificante, que codifica para al menos un péptido o proteína;  
 – una secuencia poli(A), y  
 – al menos un tallo-bucle de histona,

donde el al menos tallo-bucle de histona en i) o ii) se selecciona de entre las siguientes fórmulas (I) o (II):

fórmula (I) (secuencia tallo-bucle sin elementos frontera de tallo):

15



20

donde

25

los elementos frontera de tallo 1 o tallo 2  $N_{1-6}$  son una secuencia consecutiva de 1 a 6, de preferencia de 2 a 6, con mayor preferencia de 2 a 5, incluso con mayor preferencia de 3 a 5, con máxima preferencia de 4 a 5 o 5N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos;

30

Tallo 1  $[N_{0-2}GN_{3-5}]$  es complementario inverso o complementario parcialmente inverso con el elemento de tallo 2 y es una secuencia consecutiva de entre 5 a 7 nucleótidos; donde  $N_{0-2}$  es una secuencia consecutiva de 0 a 2, de preferencia 0 a 1, con mayor preferencia 1N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; donde  $N_{3-5}$  es una secuencia consecutiva de 3 a 5, de preferencia 4 a 5, con mayor preferencia 4 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; y donde G es guanosina o un análogo de la misma y se puede reemplazar opcionalmente por una citidina o un análogo de la misma, con la condición de que su nucleótido citidina complementario en el tallo 2 se reemplace por guanosina;

35

40

Secuencia bucle  $[N_{0-4}(U/T)N_{0-4}]$  está ubicada entre los elementos tallo 1 y tallo 2 y es una secuencia consecutiva de 3 a 5 nucleótidos, con mayor preferencia de 4 nucleótidos, donde cada  $N_{0-4}$  es independientemente una secuencia consecutiva de 0 a 4, de preferencia de 1 a 3, con mayor preferencia 1 a 2 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; y donde U/T representa uridina u opcionalmente timidina;

45

Tallo 2  $[N_{3-5}CN_{0-2}]$  es complementaria inversa o complementaria parcialmente inversa con el elemento tallo 1, y es una secuencia consecutiva de 5 a 7 nucleótidos; donde  $N_{3-5}$  es una secuencia consecutiva de 3 a 5, de preferencia 4 a 5, con mayor preferencia 4 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; donde  $N_{0-2}$  es una secuencia consecutiva de 0 a 2, de preferencia 0 a 1, con mayor preferencia 1 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C, o un nucleótido análogo de los mismos; y donde C es citidina o un análogo de la misma y se puede reemplazar opcionalmente por

una guanosina o un análogo de la misma, con la condición de que su nucleótido guanosina complementario en el tallo 1 se reemplace por citidina;

donde

5 el tallo 1 y tallo 2 son capaces de realizar un apareamiento de bases entre sí formando una secuencia complementaria inversa, donde el apareamiento de bases se puede presentar entre el tallo 1 y el tallo 2, o formando una secuencia complementaria inversa parcial, donde el apareamiento de bases incompleto se puede presentar entre el tallo 1 y el tallo 2

donde al menos un tallo-bucle se enlaza a una proteína de unión a tallo-bucle (SLBP); y

10 donde dicho péptido o proteína comprende un antígeno patogénico o un fragmento del mismo con una longitud de al menos seis residuos aminoácidos y tiene la propiedad antigénica específica del péptido o proteína nativa de longitud completa,

donde el antígeno patogénico está asociado a una infección bacteriana, una infección viral o una infección protozoaria y donde el antígeno patogénico se selecciona de un antígeno patogénico del grupo consistente en

- 15
- Virus sincitial respiratorio (RSV),
  - Virus de la inmunodeficiencia humana (VIH),
  - Virus del herpes simplex (HSV),
  - Virus de papiloma humano (HPV),
  - Virus de parainfluenza humana (HPIV),

20

  - Virus del Dengue,
  - Virus de la Hepatitis B (HBV),
  - Virus de Influenza,
  - Virus de fiebre amarilla,
  - Virus de la rabia,

25

  - Plasmodium,
  - Cytomegalovirus (CMV),
  - Staphylococcus,
  - Mycobacterium tuberculosis,
  - Chlamydia trachomatis,

30

  - Rotavirus,
  - Metapneumovirus Humano (hMPV),
  - Virus de fiebre hemorrágica de Crimea-Congo (CCHF),
  - Virus de ébola,
  - Henipavirus,

35

  - Norovirus,
  - Virus Lassa,
  - Coronavirus,
  - Rinovirus,
  - Flavivirus,

40

  - Virus de fiebre Rift Valley y
  - Virus de la enfermedad de manos, pies y boca.

2. Secuencia de ácido nucleico según la reivindicación 1, donde el antígeno patogénico se selecciona del grupo consistente en:

- 45
- Antígeno p24 de VIH, proteínas de envoltura de VIH (Gp120, Gp41, Gp160), poliproteína GAG, proteína del factor negativo Nef, trans-activador de transcripción Tat si la enfermedad infecciosa en VIH, preferentemente una infección con el virus de inmunodeficiencia humana;
  - proteína cápside C, proteína premembrana prM, proteína membrana M, proteína de envoltura E (dominio I, dominio II, dominio III), proteína NS1, proteína NS2A, proteína NS2B, proteína NS3, proteína NS4A, proteína 2K, proteína NS4B, proteína NS5, si la enfermedad infecciosa es fiebre del Dengue, preferentemente una infección con el virus del Dengue (DEN-1, DEN-2, DEN-3 y DEN-4)—Flavivirus;
  - antígeno superficial de hepatitis B HBsAg, antígeno núcleo de Hepatitis B HbcAg polimerasa, proteína Hbx, proteína superficial media preS2, proteína superficial L, proteína larga S, proteína viral VP1, proteína viral VP2, proteína viral VP3, proteína viral VP4, si la enfermedad infecciosa es Hepatitis B, preferentemente una infección con el virus de la Hepatitis B (HBV);

55

- Hemaglutinina (HA), Neuraminidasa (NA), Nucleoproteína (NP), proteína M1, proteína M2, proteína NS1, proteína NS2 (proteína NEP: proteína de exportación nuclear), proteína PA, proteína PB1 (proteína de polimerasa básica 1), proteína PB1-F2 y proteína PB2 (familia Ortomixoviridae, virus de la gripe (flu));
  - 5 • nucleoproteína N, proteína estructural grande L, fosfoproteína P, proteína de matriz M, glucoproteína G, si la enfermedad infecciosa es la rabia, preferentemente una infección con virus de la rabia;
  - poliproteína genómica, proteína E, proteína M, proteína C de cápside, proteasa NS3, proteína NS1, proteína NS2A, proteína AS2B, proteína NS4A, proteína NS4B, proteína NS5, si la enfermedad
  - 10 • infecciosa es fiebre amarilla, preferentemente una infección por el virus de la fiebre amarilla;
  - proteína TRAP anónima relacionada con trombospondina, proteína superficial de SSP2 Sporozoite 2, antígeno de membrana apical 1 AMA1, antígeno de membrana rhoptry RMA1, antígeno de repetición
  - 15 • básica ácida ABRA, proteína transversal celular PF, proteína Pvs25, proteína superficial de merozoito 1 MSP-1, proteína superficial de merozoito 2 MSP-2, antígeno superficial de eritrocitos anillo-
  - infectados antígeno de etapa RESALiver 3 LSA-3, proteína Eba-175, antígeno 5 para repetición de serina SERA-5, proteína de circumsporozoitos CS, proteína 3 superficial de merozoito MSP3, proteína
  - 20 • superficial 8 de merozoitos MSP8, enolasa PF10, proteína de hepatocitos-eritrocitos de 17 kDa HEP17, proteína con membrana 1 de eritrocitos EMP1, proteína superficial 4/5 de la proteína Kbetamerozoitos MSP 4/5, proteína de choque térmico Hsp90, proteína rica en glutamato GLURP, proteína superficial 4 de merozoitos MSP-4, proteína STARP, precursor antigénico relacionado con la proteína de circumsporozoitos CRA (Plasmodium genus, Malaria);
3. Secuencia de ácido nucleico según la reivindicación 1, donde el antígeno patógeno se selecciona del grupo consistente en:
- 25 • proteína de fusión F, hemaglutinina-neuramidasa HN, glucoproteína G, proteína matriz M, fosfoproteína P, nucleoproteína N, polimerasa L, si la enfermedad infecciosa es la infección por el virus de la parainfluenza humana, de preferencia una infección con los virus de parainfluenza Humanos (HPIV);
  - proteína de fusión F, nucleoproteína N, proteína de matriz M, proteína de matriz M2-1, proteína de
  - 30 • matriz M2-2, fosfoproteína P, proteína hidrofóbica pequeña SH, glucoproteína G de superficie mayor, polimerasa L, proteína 1 no estructural NS1, proteína 2 no estructural NS2, si la enfermedad infecciosa es una infección por el virus sincitial respiratorio, preferentemente una infección por el virus sincitial respiratorio (RSV);
  - glucoproteína L UL1, uracil-ADN glucosilasa UL2, proteína UL3, proteína UL4, proteína ADN de
  - 35 • replicación UL5, proteína portal UL6, proteína para maduración de viriones UL7, ADN helicasa UL8, proteína de unión al origen de replicación UL9, glucoproteína M UL10, proteína UL11, exonucleasa alcalina UL12, proteína quinasa de serina-treonina UL13, proteína tegumentaria UL14, terminasa UL15, proteína tegumentaria UL16, proteína UL17, proteína con cápside VP23 UL18, proteína con cápside principal VP5 UL19, proteína con membrana UL20, proteína tegumentaria UL21, Glucoproteína H (UL22), Timidina quinasa UL23, proteína UL24, proteína UL25, proteína con cápside
  - 40 • P40 (UL26, VP24, VP22A), glucoproteína B (UL27), proteína ICP18.5 (UL28), proteína principal de unión a ADN ICP8 (UL29), ADN polimerasa UL30, proteína de matriz nuclear UL31, glucoproteína con envoltura UL32, proteína UL33, proteína con membrana nuclear interna UL34, proteína con cápside VP26 (UL35), proteína tegumentaria larga UL36, proteína de ensamble cápside UL37, proteína VP19C (UL38), ribonucleótido reductasa (subunidad Largo) UL39, ribonucleótido reductasa (subunidad
  - 45 • pequeña) UL40, proteína tegumentaria/proteína VHS para interrupción en el huésped de viriones (UL41), factor de procesamiento de ADN polimerasa UL42, proteína con membrana UL43, glucoproteína C (UL44), proteína con membrana UL45, proteína tegumentarias VP11/12 (UL46), proteína tegumentaria VP13/14 (UL47), proteína para maduración de viriones VP16 (UL48, Alfa-TIF), proteína con envoltura UL49, dUTP difosfatasa UL50, proteína tegumentaria UL51, proteína compleja de ADN helicasa/primasa UL52, glucoproteína K (UL53), proteína para regulación transcripcional IE63 (ICP27, UL54), proteína UL55, proteína UL56, proteína de replicación viral ICP22 (IE68, US1), proteína
  - 50 • US2, proteína quinasa de serina/treonina US3, glucoproteína G (US4), glucoproteína J (US5), glucoproteína D (US6), glucoproteína I (US7), glucoproteína E (US8), proteína tegumentaria US9, proteína con cápside/tegumentaria US10, proteína Vmw21 (US11), proteína ICP47 (IE12, US12), activador transcripcional principal ICP4 (IE175, RS1), E3 ubiquitina ligasa ICP0 (IE110), proteína 1 latencia-relacionada LRP1, proteína 2 latencia-relacionada LRP2, factor de neurovirulencia RL1 (ICP34.5), transcripción LAT latencia-asociada (Virus de herpes simplex 1 y 2 (HSV-1 y HSV-2), Herpes simplex);
  - 55 • proteína de replicación E1, proteína reguladora E2, proteína E3, proteína E4, proteína E5, proteína E6, proteína E7, proteína E8, proteína con cápside mayor L1, proteína con cápside menor L2 (papilomavirus humano (HPV), infección por papilomavirus humano (HPV));
  - antígeno pp65, proteína de membrana pp15, proteína tegumentaria cápside-proximal pp150, proteína
  - 60 • M45, ADN polimerasa UL54, helicasa UL105, glucoproteína gM, glucoproteína gN, glucoproteína H, glucoproteína B gB, proteína UL83, proteína UL94, proteína UL99 (Citomegalovirus (CMV), infección por Citomegalovirus);
  - 65

- 5 • proteína IsdB de unión hemo-hierro, colágeno adhesina Cna, factor A de aterronamiento ClfA, proteína MecA, proteína de unión A fibronectina FnbA, enterotoxina tipo A EntA, enterotoxina tipo B EntB, enterotoxina tipo C EntC1, enterotoxina tipo C EntC2, enterotoxina tipo D EntD, enterotoxina tipo E EntE, toxina-1 de síndrome de choque tóxico TSST-1, Stafiloquinasa, proteína 2a de unión penicilina PBP2a (MecA), antígeno secretor SssA (Staphylococcus genus, por ejemplo aureus, infección por estafilococos);
  - 10 • poliproteína genómica, polimerasa 3D, proteína con cápside viral VP1, proteína con cápside viral VP2, proteína con cápside viral VP3, proteína con cápside viral VP4, proteasa 2A, proteasa 3C (Rhinovirus, infección por Rhinovirus);
  - 15 • proteína ESAT-6 (EsxA), antígeno filtrado EsxB de 10 kDa, antígeno secretado 85-B FBPB, proteína A de unión a fibronectina FbpA (Ag85A), serina proteasa PepA, proteína PPE18 de la familia PPE, proteína D de unión a fibronectina FbpD, proteína inmunogénica MPT64, proteína secretada MPT51, catalasa-peroxidasa-peroxinitritasa T KATG, lipoproteína PSTS3 de unión a fosfato periplásmico (PBP-3, Phos-1), hemaglutinina Hbha de unión a heparina hierro-regulada, proteína PPE14 de la familia PPE, proteína PPE68 de la familia PPE, proteína Mtb72F, proteína Apa, proteína inmunogénica MPT63, lipoproteína PSTS1 de unión a fosfato periplásmico (PBP-1), chaperón molecular DnaK, lipoproteína Mpt83 de superficie celular, lipoproteína P23, proteína permeasa pstA del sistema de transporte de fosfato, antígeno de 14 kDa, proteína C de unión a fibronectina FbpC1, alanina deshidrogenasa TB43, glutamina sintetasa 1, proteína ESX-1, proteína CFP10, proteína TB10.4, proteína MPT83, proteína MTB12, proteína MTB8, proteínas similares a Rpf, proteína MTB32, proteína MTB39, cristalina, proteína de choque térmico HSP65, proteína PST-S (usualmente tuberculosis Mycobacterium, Tuberculosis);
  - 20 • nucleoproteína N, polimerasa L, proteína de matriz Z, glucoproteína GP (Lassa virus, fiebre de Lassa);
  - 25 • proteínas spike S, proteínas con envoltura E, proteínas de membrana M, proteínas nucleocápside N (por lo general rinovirus y coronavirus);
  - glucoproteína GP, nucleoproteína NP, proteína matriz menor VP24, proteína matriz mayor VP40, activador de transcripción VP30, cofactor de polimerasa VP35, polimerasa L de ARN (Ebolavirus (EBOV), fiebre hemorrágica de Ebola);
  - 30 • proteína con envoltura Gc, proteína con envoltura Gn, proteínas nucleocápside (virus de fiebre hemorrágica Crimean-Congo, fiebre hemorrágica Crimean-Congo (CCHF));
  - glucoproteína G, proteína matriz M, nucleoproteína N, proteína de fusión F, polimerasa L, proteína W, proteína C, fosfoproteína P, proteína no estructural V (virus Henipavirus (Hendra virus Nipah), infecciones Henipavirus);
  - 35 • glucoproteína GP con envoltura, polimerasa L, nucleoproteína N, proteína NSS no estructural (virus de fiebre Rift Valley, fiebre de Rift Valley (RVF)).
4. Secuencia de ácido nucleico según la reivindicación 1, donde el antígeno patogénico se selecciona del grupo consistente en:
    - 40 • proteína 6 no estructural NS6, proteína 2 no estructural NS2, proteína VP6 con cápside intermedia, proteína VP2 con cápside interna, proteína 3 no estructural NS3, polimerasa L de ARN-ARN-dirigida, proteína VP3, proteína 1 no estructural NS1, proteína 5 no estructural NS5, glucoproteína VP7 con cápside extrna, glucoproteína 4 no estructural NS4, proteína VP4 con cápside externa (Rotavirus, infección por Rotavirus);
  5. Secuencia de ácido nucleico según la reivindicación 1, donde el antígeno patogénico se selecciona del grupo consistente en:
    - 45 • proteína NS1, proteína hidrofóbica pequeña NS2, proteína SH, proteína de fusión F, glucoproteína G, proteína de matriz M, proteína de matriz M2-1, proteína de matriz M2-2, fosfoproteína P, nucleoproteína N, polimerasa L (metapneumovirus Humano (hMPV)).
  6. Secuencia de ácido nucleico según las reivindicaciones 1 a 5, donde el al menos un tallo-bucle de histona es heterólogo a la región codificante que codifica para al menos un péptido o proteína, preferentemente donde la regióncodificante no codifica para una proteína de histona o un fragmento, derivado o variante de la misma que tenga una función histona o similar a histona.
  7. Secuencia de ácido nucleico según cualquiera de las reivindicaciones 1 a 6, donde el ácido nucleico es un ARN, preferentemente un ARNm.
  - 55 8. Secuencia de ácido nucleico según cualquiera de las reivindicaciones 1 a 7, donde el al menos un tallo-bucle de histona se selecciona de al menos una de las siguientes fórmulas (Ia) o (IIa):

$$\underbrace{[N_{0-1}GN_{3-5}]}_{\text{tallo 1}} \underbrace{[N_{1-3}(U/T)N_{0-2}]}_{\text{bucle}} \underbrace{[N_{3-5}CN_{0-1}]}_{\text{tallo 2}}$$

fórmula (Ia) (secuencia tallo-bucle sin elementos frontera de tallo):

5

$$\underbrace{N_{2-5}}_{\text{Elemento frontera tallo 1}} \underbrace{[N_{0-1}GN_{3-5}]}_{\text{tallo 1}} \underbrace{[N_{1-3}(U/T)N_{0-2}]}_{\text{bucle}} \underbrace{[N_{3-5}CN_{0-1}]}_{\text{tallo 2}} \underbrace{N_{2-5}}_{\text{Elemento frontera tallo 2}}$$

10 fórmula (IIa) (secuencia tallo-bucle con elementos frontera de tallo):o una de las siguientes fórmulas (Ib) o (IIb):

fórmula (Ib) (secuencia de tallo-bucle sin elementos frontera de tallo):

15

$$\underbrace{[N_1GN_4]}_{\text{tallo 1}} \underbrace{[N_2(U/T)N_1]}_{\text{bucle}} \underbrace{[N_4CN_1]}_{\text{tallo 2}}$$

fórmula (IIb) (secuencia de tallo-bucle con elementos frontera de tallo):

20

$$\underbrace{N_{4-5}}_{\text{Elemento frontera tallo 1}} \underbrace{[N_1GN_4]}_{\text{tallo 1}} \underbrace{[N_2(U/T)N_1]}_{\text{bucle}} \underbrace{[N_4CN_1]}_{\text{tallo 2}} \underbrace{N_{4-5}}_{\text{Elemento frontera tallo 2}}$$

25

Elemento frontera tallo 1      tallo 1      bucle      tallo 2      Elemento frontera tallo 2

30

- 5      **9.** Secuencia de ácido nucleico según cualquiera de las reivindicaciones 1 a 8, donde la secuencia poli(A) comprende una secuencia de aproximadamente 25 a aproximadamente 400 nucleótidos de adenosina, de preferencia una secuencia de aproximadamente 50 a 400 nucleótidos de adenosina, con mayor preferencia una secuencia de aproximadamente 50 a aproximadamente 300 nucleótidos de adenosina, incluso con mayor preferencia una secuencia de aproximadamente 50 a aproximadamente 250 nucleótidos de adenosina, con la máxima preferencia una secuencia de aproximadamente 60 a aproximadamente 250 nucleótidos de adenosina.
- 10     **10.** Composición o kit o kit de partes que comprende una pluralidad o de más de una secuencia de ácido nucleico de conformidad con cualquiera de las reivindicaciones 1 a 9.
- 15     **11.** Secuencia de ácido nucleico como se define según cualquiera de las reivindicaciones 1 a 8 o composición como se define según la reivindicación 10 para su uso como un medicamento o vacuna.
- 15     **12.** Secuencia de ácido nucleico como se define según cualquiera de las reivindicaciones 1 a 9 o composición como se define según la reivindicación 10 para su uso en el tratamiento o la profilaxis de enfermedades infecciosas, preferentemente mediante inyección intradérmica o intramuscular.
- 20     **13.** Secuencia de ácido nucleico como se define según cualquiera de las reivindicaciones 1 a 9 o composición como se define según la reivindicación 10 para su uso según la reivindicación 12, donde el tratamiento comprende incrementar la expresión de dicho péptido o proteína codificado.
- 20     **14.** Composición farmacéutica que comprende una secuencia de ácido nucleico como se define según cualquiera de las reivindicaciones 1 a 9 o una composición como se define según la reivindicación 10 y opcionalmente un portador farmacéuticamente aceptable.
- 25     **15.** Uso de una secuencia de ácido nucleico como se define según cualquiera de las reivindicaciones 1 a 9 o de una composición o kit o kit de partes como se define según la reivindicación 10, para aumentar la expresión de dicho péptido o proteína codificado *in vitro*.
- 25     **16.** Método *in vitro* o *ex vivo* para aumentar la expresión de un péptido o proteína codificado comprendiendo los pasos de:
- 30        a) proporcionar la secuencia de ácido nucleico como se define según cualquiera de las reivindicaciones 1 a 9 o la composición como se define según la reivindicación 10;
- 30        b) aplicar o administrar la secuencia de ácido nucleico o la composición a un sistema de expresión libre de células, una célula o un tejido.



Figura 1

# A	2224	1586	3075	2872	1284	<	<	<	<	<	<	<	<	<	<	<	<	<	>	>	>	>	>	>	>	>	771	2012	2499
# T	172	188	47	205	19	6	0	569	1620	199	3947	3030	3704	4001	182	1	21	15	11	0	0	179	8	64	557	201	690		
# C	1537	2211	875	918	2675	270	0	3394	2342	3783	51	119	227	0	3140	7	50	31	16	4001	3543	154	3870	2636	1744	674			
# G	25	16	4	6	23	3541	4001	25	27	10	2	5	11	0	4	175	3735	2359	3451	0	265	112	4	37	43	138			
Comis	N*	N*	N	N	N	N	G	N	N	N	N	N	N	T	N	N	N	N	N	C	N	N	N	N	N	N*	N*		
99%	H*	H*	H	H	V	V	G	Y	Y	Y	Y	H	H	T	H	R	V	V	R	C	B	V	H	H*	N*	N*			
95%	M*	H*	M	H	M	S	G	Y	Y	Y	T	T	Y	T	M	A	R	R	R	C	S	M	C	H*	H*	H*			
90%	M*	M*	M	M	M	S	G	Y	Y	C	T	T	T	T	M	A	G	R	R	C	S	A	C	H*	M*	H*			

Tallo 1
Bucle
Tallo 2



Figura 3

#A	2172	1554	3004	2790	1208	171	0	1	0	0	0	1	56	0	600	3736	142	1517	503	0	10	3633	44	736	1938	2443
	#T	152	156	10	184	11	3	0	548	1535	141	3861	3639	3870	154	0	4	2	1	1	0	164	1	33	525	181
#C	1512	2152	855	893	2637	270	0	3308	2334	3729	9	106	169	0	3113	5	44	0	6	3870	3431	149	3788	2578	1714	634
#G	11	8	1	3	14	3426	3870	13	1	0	0	3	6	0	3	129	3680	2351	3360	0	265	87	3	31	36	131
Coni	N*	N*	N	N	N	N	G	N	B	Y	Y	N	N	T	N	V	N	D	N	C	N	N	N	N*	N*	N*
99%	H*	H*	M	H	M	V	G	Y	Y	Y	T	Y	H	T	H	R	V	R	R	C	B	V	M	H*	N*	N*
95%	M*	M*	M	M	M	S	G	Y	Y	C	T	Y	T	T	M	A	G	R	R	C	S	M	C	H*	H*	H*
90%	M*	M*	M	M	M	S	G	Y	Y	C	T	T	T	T	M	A	G	R	R	C	S	A	C	H*	M*	H*

Tallo 1 { < < < < < < < < < < < < }
 Bucle { \* \* \* \* \* }
 Tallo 2 { > > > > > > > }



Figura 5

#A	10	17	84	84	76	1	0	1	0	0	0	0	0	0	0	0	12	84	0	65	3	0	0	0	0	69	5	0	10	64
#T	8	6	0	0	2	2	0	1	67	0	84	80	81	84	5	0	0	5	0	0	0	0	0	0	0	0	4	25	24	3
#C	62	61	0	0	6	0	0	82	17	84	0	0	3	0	0	67	0	0	1	0	0	0	84	84	5	75	57	44	17	
#C	4	0	0	0	0	81	84	0	0	0	0	3	0	0	0	0	83	19	81	0	0	0	0	0	10	0	2	6	0	
Cons	N*	H*	A	A	H	D	G	H	Y	C	T	D	Y	T	H	A	S	R	R	C	R	C	C	C	V	H	B*	N*	H*	
99%	N*	H*	A	A	H	D	G	H	Y	C	T	D	Y	T	H	A	S	R	R	C	R	C	C	C	V	H	B*	N*	H*	
95%	H*	H*	A	A	M	G	G	C	Y	C	T	T	T	T	H	A	G	R	G	C	C	C	C	C	V	M	Y*	N*	M*	
90%	H*	M*	A	A	A	G	G	C	Y	C	T	T	T	T	M	A	G	R	G	C	C	C	C	C	R	M	Y*	H*	M*	

Tallo 2

Bucle

Tallo 1

**ppLuc(GC)-ag**  
 gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCCGGCGCCCUUCUA  
 CCCGUCGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCAUGCCGGUGCUGGGCGC  
 CCUCUUCAUCGGCGUGGCCGUCGCCCGGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCCGCCGGG  
 CUUCAACGAGUACGACUUCGUCGCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCCUGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC  
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUCGUCGUGCCGACCCUGUUCAGCUUCUU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUUGUCGAACCUGCACGAGAUCCAGCGG  
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCCCGGG  
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGCAAGACCCUGGGCGUGAACCAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
 GAUGAUCAUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA  
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGU  
 CGACCGGCUGAAGUCGUGAUCAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCGUGCCGGACGA  
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUUCGUGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGGCGG  
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
 CCGCGAGAUCCUGAUCAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA<sup>A</sup>gacua<sup>g</sup>uuua  
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCCUCCUUGCACCG**agauua  
 auagauc-3'

**Figura 6**

**ppLuc(GC)-ag-A64**

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA  
CCCGCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU  
GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUAUGCCGGUGCUGGGCGC  
CCUCUUAUCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCCGCCGGG  
CUUCAACGAGUACGACUUCGUUCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUCGAU  
CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC  
CGCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
GAGCCUGCAGGACUACAAGAUCCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU  
CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCAGAGAUCCAGCGG  
GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCCGGG  
CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGGAUCCUGAUCACCCCGAGGG  
GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUUCCGUUCUUCGAGGCCAAGGUGGUGGA  
CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
GAUGAUC AUGAGCGGCUACGUGAACACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA  
CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUAUCGU  
CGACCGGCUAAGUCGUGAUC AAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA  
GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA  
CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA  
GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
CCGCGAGAUCCUGAUC AAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAgacuaguaua  
agacugacua**GCCC GAUGGGCCUCCCAACGGGCCCCUCCUCCCUUGCACCG**agauua  
auAA  
AAAAA-3'

**Figura 7**

**ppLuc(GC)-ag-histonaSL**  
 gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCCGGCGCCCUUCUA  
 CCCGUCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUC AUGCCGGUGCUGGGCGC  
 CCUCUUC AU CGGCUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG  
 CUUCAACGAGUACGACUUCGUCCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGCAACCAGAUCAUCCGGACAC  
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCUU  
 CGCCAAGAGCACCCUGAU CGACAAGUACGACCUGUCGAACCUGCACGAGAU CGCCAGCGG  
 GGGCGCCCCGUCGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGG  
 CAUCCGCCAGGGCUACGGCCUGACCAGACCACGAGCGGAUCCUGAUACCCCCGAGGG  
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGCAAGACCCUGGGCGUGAACCAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
 GAUGAUCAUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA  
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCUUCGU  
 CGACCGGCUGAAGUCGCUGAUCAAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCCAACUUCUUCGACGCCGGCGUGGCCGGGCGGCCGACGA  
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAU CGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
 CCGCGAGAUCCUGAUCAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA *Agacuaguuua*  
*agacugacua* **GCCCGAUGGGCCUCCCAACGGGCCUCCUCCCUUGCACCG** *agauua*  
*auagauca* CAAAGGCUCUUUUCAGAGCCACCA - 3'

**Figura 8**



**ppLuc(GC)-ag-A64-histonaSL**

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA  
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUAUGCCGGUGCUGGGCGC  
 CCUCUUAUCGGCGUGGCCGUCGCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCCGCCGGG  
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC  
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCCAGCGG  
 GGGCGCCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCCGGG  
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
 GAUGAUC AUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA  
 CGGCUGGCUGCACAGCGGCACAUCGCCUACUGGGACGAGGACGAGCACUUCUUAUCGU  
 CGACCGGCUAAGUCGUGAUC AAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA  
 GAGCAUCCUGCUC CAGCACCCCAACUUCGACGCCGGCGUGGCCGGGCGUGCCGGACGA  
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
 CCGCGAGAUCUGAUC AAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA Agacuaguua  
agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCUUGCACC**gagauua  
auAA  
 AAAAAAugcauCAAAGGCUCUUUCAGAGCCACCA-3'

**Figura 9**

**ppLuc(GC)-ag-A120**

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCCUUCUA  
 CCCGCUGGAGGACGGGACCGCCGGCGAGCAGCUCACAAAGGCCAUGAAGCGGUACGCCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUCUCGGAGAACAGCCUGCAGUUCUUCUUCGCGGUGCGUGGGCGC  
 CCUCUUCUUCGCGUGGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCCGCCGGG  
 CUUCAACGAGUACGACUUCGUCCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUUGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC  
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCCUGUCGAACCUGCAGAGAUCCAGCGG  
 GGGCGCCCCGCUAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCCCGGG  
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCCGAGGG  
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
 GAUGAUCAUGAGCGGCUACGUGAACACCCGGAGGCCACCAACGCCUUCUUCGACAAGGA  
 CGGCUUGGUCGACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCUUCG  
 CGACCGGCUAAGUCGUGAUCUAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA  
 GAGCAUCCUGCUCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCGGCCGACGA  
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
 CCGCGAGAUCCUGAUCUAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA  
 Agacuaguua  
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCUUGCACCG**agauua  
 auagaucuAA  
 AA  
 AAAAAAAAA-3'

**Figura 10**

**ppLuc(GC)-ag-A64-ag**

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCCGGCGCCCUUCUA  
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUACCCUACGCGGA  
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUC AUGCCGGUGCUGGGCGC  
 CCUCUUC AUCCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAAGCUGCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCGGGG  
 CUUCAACGAGUACGACUUCGUUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGAACCAGAUCAUCCCGGACAC  
 CGCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCCGACGGG  
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCGG  
 CAUCCGCCAGGGCUACGGCCUGACCAGACACGAGCGCGAUCCUGAUCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGCAAGACCCUGGGCGUGAACCAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
 GAUGAUCAUGAGCGGCUACGUGAACAAACCCGGAGGCCACCAACGCCUUC AUCCGACAAGGA  
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUC AUUCGU  
 CGACCGGCUGAAGUCGUGAUCAGUACAAGGGCUACCAGGUGGCGCCGCGGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCGUGCCGGACGA  
 CGACCGCGGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUUCGUGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
 CGUGGUGUUCGUGGACGAGGUCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
 CCGCGAGAUCCUGAUCAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAgacuaguaua  
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCCUCCUUGCACCG**agauua  
 auAA  
 AAAAAAugcau**CCUGCCCGAUGGGCCUCCCAACGGGCCCUCUCCCCUCCUUGCACCG**3'

**Figura 11**

**ppLuc(GC)-ag-A64-aCPSL**

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCCGGCGCCCUUCUA  
 CCCGCUUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
 GUACUUCGAGAUAGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCUUCGCGGUGCUGGGCGC  
 CCUCUUCUUCGCGGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCGGGG  
 CUUCAACGAGUACGACUUCGUCGCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCGGACAC  
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCAGGCUUCGGCAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCUU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCAGAGAUCCGCGAGCGG  
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCGG  
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCAGAGCGCGAUCCUGAUCACCCCCGAGGG  
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGCAAGACCCUGGGCGUGAACCAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
 GAUGAUCAUGAGCGGCUACGUGAACAAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA  
 CGGCUUGGUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCUUCGU  
 CGACCGGCGUGAAGUCGUGAUCAGUACAAGGGCUACCAAGGUGGCGCCGGCCGAGCUGGA  
 GAGCAUCCUGCUCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCGGCCGACGA  
 CGACCGCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
 CCGCGAGAUCUGAUCUAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAgacuaguaua  
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCUCCUCCCCUCCUUGCACCG**agauua  
 auAA  
 AAAAAAugcauCAAUCCUACACGUGAGGCGCUGUGAUUCCUAUCCCCCUUCAUUCUU  
 AUACAUUAGCACAGCGCCAUUGCAUGUAGGAAUU-3'

**Figura 12**

**ppLuc(GC)-ag-A64-polioCL**

gggagaaagcuugaggaUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA  
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCUAUGCCGGUGCUGGGCGC  
 CCUCUUCUUCGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCCGACCUGGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG  
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCGGACAC  
 CGCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCGCCAGCGG  
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGGG  
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCCGAGGG  
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGCAAGACCCUGGGCGUGAACCAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
 GAUGAUCAUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCUUCGACAAGGA  
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCUUCG  
 CGACCGGCUGAAGUCGUGAUCAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACCGCGGUGGCCGGGCGUGCCGGACGA  
 CGACCGCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUUCGUGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
 CGUGGUGUUCGUGGACGAGGUCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
 CCGCGAGAUCCUGAUCAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAgacuaguua  
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCCUCCUCCUCCUUGCACCG**agauua  
 auAA  
 AAAAAAugcauCAAUUCUAAAACAGCUCUGGGGUUGUACCCACCCAGAGGCCACGUGG  
 CGGCUAGUACUCCGGUAUUGCGGUACCCUUGUACGCCUGUUUAGAAUU-3'

**Figura 13**



**ppLuc(GC)-ag-A64-U30**

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA  
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUAUGCCGGUGCUGGGCGC  
 CCUCUUAUCGGCGUGGCCGUCGCCCGGCCAAGCACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUUAUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCCGGG  
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUCGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC  
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCCGACGG  
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCGG  
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG  
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
 GAUGAUC AUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCUCAUCGACAAGGA  
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUAUCGU  
 CGACCGGCUGAAGUCGUGAUC AAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA  
 GAGCAUCCUGCUC CAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA  
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
 CCGCGAGAUCCUGAUC AAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA *Agacuaguua*  
*agacugacua* **GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCCUCCUUGCACCG** *agauua*  
*au*AA  
 AAAAAA *ugca* UU-3'

**Figura 15**

**ppLuc(GC)-ag-A64-SL**

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA  
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU  
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUAUGCCGGUGCUGGGCGC  
 CCUCUUAUCGGCGUGGCCGUCGCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUUGGACAGCAA  
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG  
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCGGACAC  
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCUU  
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCCAGCGG  
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGG  
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCCGAGGG  
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
 GAUGAUCAUGAGCGGCUACGUGAACAAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA  
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUAUCGU  
 CGACCGGCUGAAGUCGUGAUCAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA  
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCGGCCGACGA  
 CGACCGCGGCGAGCUGCCGGCCGCGGUGGUGGUCUGGAGCACGGCAAGACCAUGACGGA  
 GAAGGAGAUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
 CCGCGAGAUCCUGAUCAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAgacuaquua  
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCUCCUCCCUCCUUGCACCG**agauua  
 auAA  
 AAAAAAugcauUAUGGCGGCCGUGUCCACCACGGAUUAUACCGUGGUGGACGCGGCC-3'

**Figura 16**



**ppLuc(GC)-ag-A64-N32**

gggagaaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA  
CCCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU  
GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA  
GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA  
CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUC AUGCCGGUGCUGGGCGC  
CCUCUUC AUGCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU  
GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA  
GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA  
GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG  
CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU  
CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC  
CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCGGACAC  
CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA  
CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG  
GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCUU  
CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCCAGCGG  
GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGG  
CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG  
GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA  
CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC  
GAUGAUC AUGAGCGGCUACGUGAACAAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA  
CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCUUCGU  
CGACCGGCUGAAGUCGUGAUC AAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA  
GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCGUGCCGGACGA  
CGACCGCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA  
GAAGGAGAU CGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG  
CGUGGUGUUCGUGGACGAGGUCCC GAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU  
CCGCGAGAUCCUGAUC AAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA *Agacuaguua*  
*agacugacua* **GCCCGAUGGGCCUCCCAACGGGCCUCCUCCCCUCCUUGCACCG** *agauua*  
*au*AA  
AAAAA *Augcau* **CCCCUCUAGACAAUUGGAAUCCAUA** -3'

**Figura 17**

**H1N1(PR8)-ag-A64-C30**

GGGAGAAAGCUUACCAUGAAGGCCAACCGUCUCGUGCUGCUGUGCGCCUCGCGGCCGCC  
 GACGCCGACACCAUCUGCAUCGGCUACCACGCCAACACAGCACCGACACGGUCGACACC  
 GUGCUGGAGAAGAACGUGACCGUCACCCACUCCGUGAACCUGCUCGAGGACAGCCACAAC  
 GGGAAAGCUGUGCCGGCUGAAGGGCAUCGCGCCCCUCCAGCUGGGGAAGUGCAACAUCCG  
 GGCUGGCUGCUCGGGAACCCGGAGUGCGACCCCGUCUGCCCGUGCGCUCCUGGAGCUAC  
 AUCGUCGAGACGCCAACUCCGAGAACGGCAUCUGCUACCCGGGCGACUUAUCGACUAC  
 GAGGAGCUCGGGAGCAGCUGAGCUCGUGAGCUCUUCGAGCGCUUCGAGAUCUCCCC  
 AAGGAGAGCUCUCCGGCCAAACCACAACACCAACGGGGUGACCGCCGCCUGCAGCCACGAG  
 GGCAAGUCCAGCUUCUACCGGAACCUGCUCUGGCUGACCGAGAAGGAGGGGUCCUACCC  
 AAGCUGAAGAACAGCUACGUACAACAAGAAGGGCAAGGAGGUGCUCGUGCUGUGGGGAUC  
 CACCACCCGCCAACUCCAAGGAGCAGCAGAACCUGUACCAGAACGAGAACGCGUACGUC  
 AGCGUGGUGACGUCCAACUACAACCGCCGUUCACCCCGAGAUCGCCGAGCGCCCAAG  
 GUCCGGGACCAGGCCGGCCGAUGAACUACUACUGGACCCUCCUGAAGCCGGGCGACACC  
 AUAUCUUCGAGGCCAACGGGAACCUGAUCGCCCCGAUGUACGCGUUCGCCUCAGCCGG  
 GGCUUCGGGAGCGGCAUCAUCACGUCCAACGCCAGCAUGCACGAGUGCAACACCAAGUGC  
 CAGACCCCGUGGGCGCAUCAACUCCAGCCUGCCUACCAGAACAUCACCCGGUGACC  
 AUCGGGGAGUGCCCCAAGUACGUGCGCUCGCCAAGCUCGGGAUGGUCACGGGCCUGCGC  
 AACACCCCGAGCAUCCAGUCCCGGGGGCUGUUCGGCGCGAUCGCCGGGUUCAUCGAGGGC  
 GGCUGGACCGGGAUGAUCGACGGCUGGUACGGGUACCACCACCAGAACGAGCAGGGCAGC  
 GGGUACGCCCGGACCAGAAGUCCACCCAGAACGCCAUAACGGCAUCACCAACAAGGUG  
 AACACGGUGAUCGAGAAGAUGAACAUCCAGUUCACCGCGGUCGGCAAGGAGUUAACAAG  
 CUCGAGAAGCGCAUGGAGAACCUGAACAAAGAAGGUGGACGACGGGUUCCUGGACAUCCGG  
 ACCUACAACGCCGAGCUCUCCUGGUGCUGCUCGAGAACGAGCGGACCCUGGACUUCACGAC  
 AGCAACGUACAAGAACCUGUACGAGAAGGUGAAGUCCAGCUCAGAACAACGCCAAGGAG  
 AUCGGCAACGGGUGCUUCGAGUUCUACCACAAGUGCGACAACGAGUGCAUGGAGAGCGUC  
 CGCAACGGCACGUACGACUACCCCAAGUACUCCGAGGAGAGCAAGCUGAACCAGGAGAAAG  
 GUGGACGGGGUGAAGCUGGAGUCCAUGGGCAUCUACCAGAUCUCCGCAUCUACAGCACC  
 GUCGCCUCCAGCCUGGUGCUGCUGGUGUCCUCGGCGCGAUCAGCUUCUGGAUGUGCAGC  
 AACGGGUCCUGCAGUGCCGCAUCUGCAUCUGAaccacuaguuaaagacugacua**GCCCG**  
**AUGGGCCUCCCAACGGGGCCUCCUCCCUCCUUGCACC**GagauuaauAAAAAAAAAAAAA  
 AAUUCCCCC  
 CCCCCCCCCCCCCCCCCCCCCCCCCCUAGACAAUUGGAAUU-3'

**Figura 18**

**H1N1(PR8)-ag-A64-C30-histonaSL**

GGGAGAAAGCUUACCAUGAAGGCCAACCGUCUCGUGCUGCUGUGCGCCUCGCGGCCGCC  
GACGCCGACACCAUCUGCAUCGGCUACCACGCCAACAACAGCACCGACACGGUCGACACC  
GUGCUGGAGAAGAACGUGACCGUCACCCACUCCGUGAACCUGCUCGAGGACAGCCACAAC  
GGGAAGCUGUGCCGGCUGAAGGGCAUCGCGCCCUCCAGCUGGGGAAGUGCAACAUCGCC  
GGCUGGCUGCUCGGGAACCCGGAGUGCGACCCCGUGCUGCCCGUGCGCUCUCCUGGAGCUAC  
AUCGUCGAGACGCCAACUCCGAGAACGGCAUCUGCUACCCGGGCGACUUCAUCGACUAC  
GAGGAGCUCGGGAGCAGCUGAGCUCGUGAGCUCCUUCGAGCGCUUCGAGAUUUCCCC  
AAGGAGAGCUCUCCUGGCCAACCAACAACACCAACGGGGUGACCGCCGCCUCGACCCACGAG  
GGCAAGUCCAGCUUCUACCGGAACCGUCUCUGGCUGACCGAGAAGGAGGGGUCCUACCCC  
AAGCUGAAGAACAGCUACGUAACAAGAAGGGCAAGGAGGUGCUCGUGCUGUGGGGGAUC  
CACCACCCGCCAACUCCAAGGAGCAGCAGAACCUGUACCAGAACGAGAACGCGUACGUC  
AGCGUGGUGACGUCCAACUACAACCGCCGUUCACCCCGAGAUCCGCGAGCGCCCAAG  
GUCCGGGACCAGGCCGGCCGAUGAACUACUACUGGACCCUCCUGAAGCCGGGCGACACC  
AUCAUCUUCGAGGCCAACGGGAACCGUAUCGCCCCGAUGUACGCGUUCGCCUCAGCCGG  
GGCUUCGGGAGCGGCAUCAUCACGUCCAACGCCAGCAUGCACGAGUGCAACACCAAGUGC  
CAGACCCCGCCCAUCUCCAAGGAGCAGCAGAACCUGUACCAGAACAUCACCCGGUGACC  
AUCGGGAGUGCCCCAAGUACGUGCGUCCGCCAAGCUCGGAUGGUCACGGGCCUGCGC  
AACAAACCCAGCAUCCAGUCCCGGGGGCUGUUCGGCGCGAUCGCCGGGUUCAUCGAGGGC  
GGCUGGACCGGGAUGAUCGACGGCUGGUACGGGUACCACCACCAGAACGAGCAGGGCAGC  
GGGUACGCCCGACCAGAAGUCCACCAGAACGCCAUCAACGGCAUCACCAACAAGGUG  
AACACGGUGAUCGAGAAGAUGAACAUCCAGUUCACCGCGGUCGGCAAGGAGUUAACAAG  
CUCGAGAAGCGCAUGGAGAACCUGAACAGAAGGUGGACGACGGGUUCCUGGACAUCUGG  
ACCUACAACGCCGAGCUCUCCUGGUGCUGCUCGAGAACGAGCGGACCCUGGACUCCACGAC  
AGCAACGUACAAGAACCUGUACGAGAAGGUGAAGUCCAGCUACAAGAACAACGCCAAGGAG  
AUCGGCAACGGGUGCUUCGAGUUCUACCACAAGUGCGACAACGAGUGCAUGGAGAGCGUC  
CGCAACGGCAGCUACGACUACCCCAAGUACUCCGAGGAGAGCAAGCUGAACCGGGAGAAG  
GUGGACGGGGUGAAGCUGGAGUCCAUGGGCAUCUACCAGAUCUCCGCCAUUCAGCACC  
GUCGCCUCCAGCCUGGUGCUGCUGGUGUCCUCGGCGCGAUCAGCUUCUGGAUGUGCAGC  
AACGGGUCCUGCAGUGCCGAUCUGCAUCUGAccacuaguuaaagacugacua**GCCCG**  
**AUGGGCCUCCCAACGGGCCUCCUCCCUCCUUGCACC**gagauuaauAAAAAAAAAAAAA  
AAAUGCAUCCCC  
CCCCCCCCCCCCCCCCCCCCCCCCCAAGGCUCUUUCAGAGCCACCAGaauu

**Figura 19**

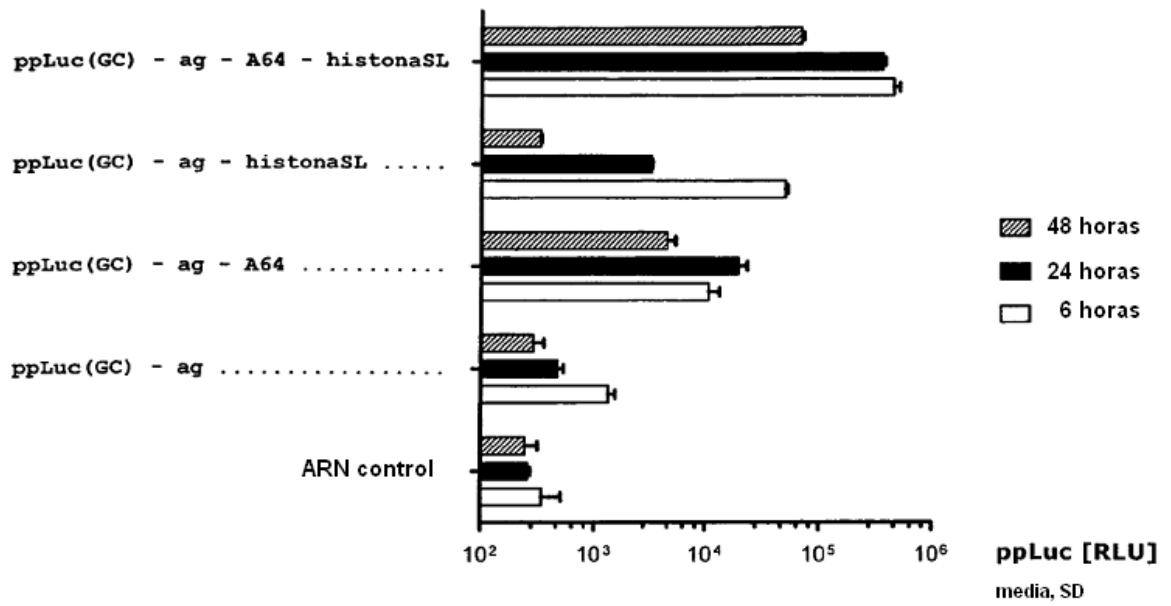


Figura 20

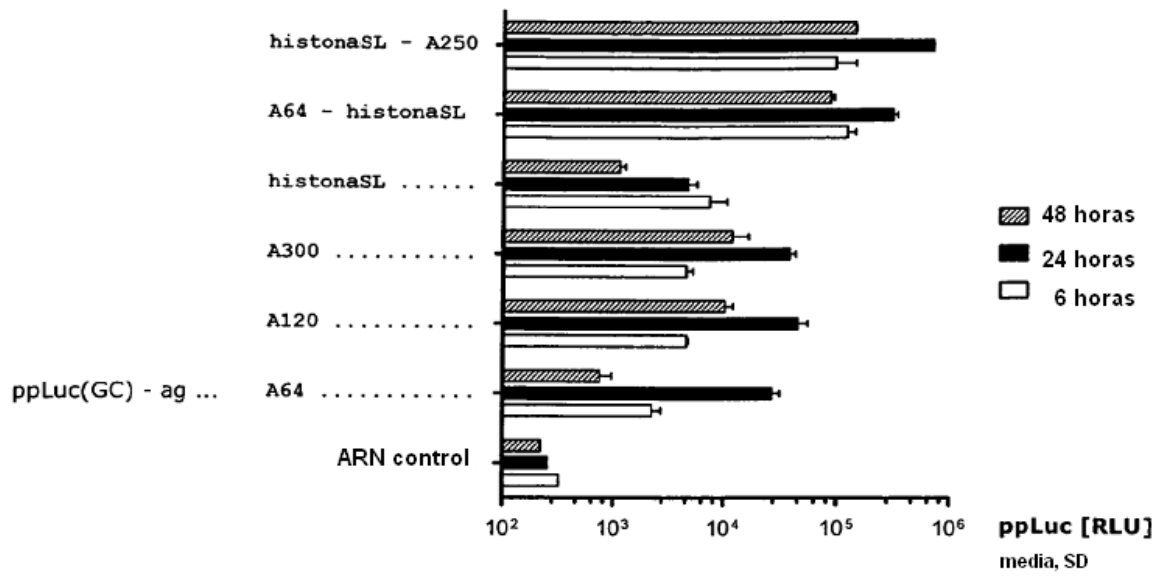


Figura 21

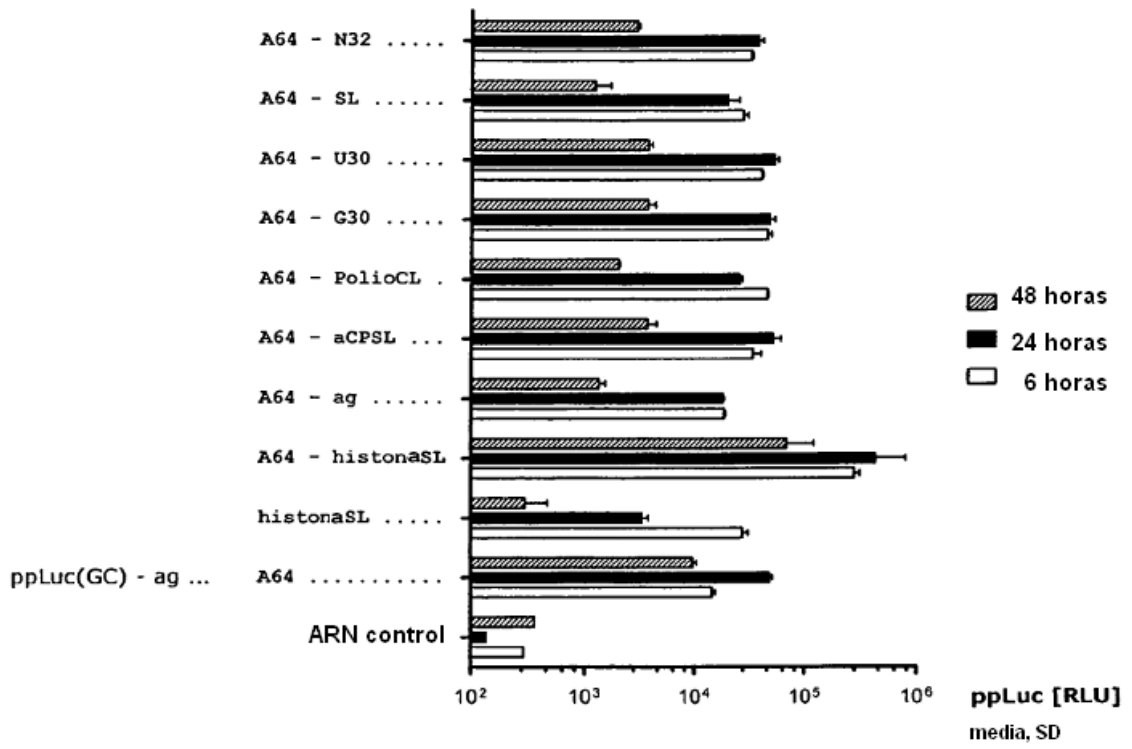


Figura 22

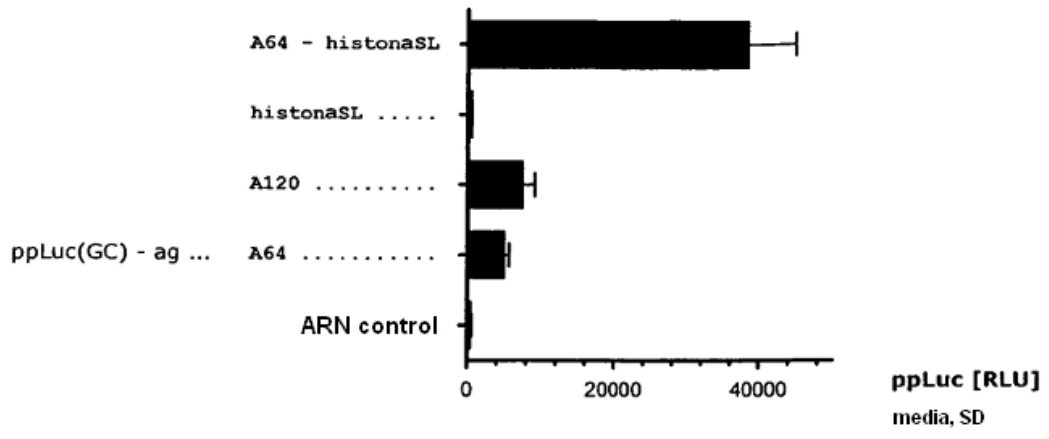


Figura 23

Inducción de anticuerpos IgG1 contra HA del virus PR8 de la Influenza  
4 semanas después de la inmunización

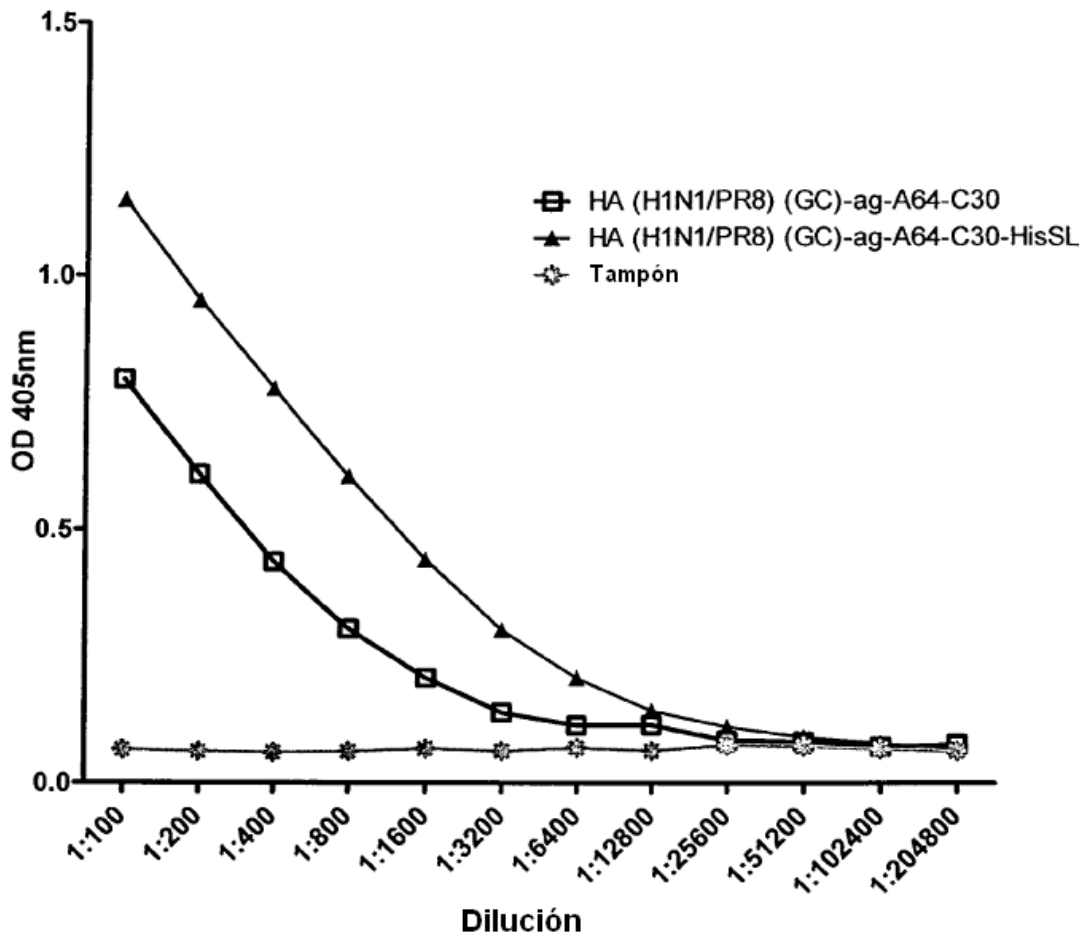


Figura 24