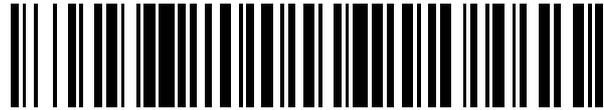


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 525 556**

51 Int. Cl.:

C12N 15/67 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **12.08.2011 E 11745701 (0)**

97 Fecha y número de publicación de la concesión europea: **10.09.2014 EP 2603590**

54 Título: **Ácido nucleico que comprende o codifica para una secuencia de tallo-lazo de histona y poli(A) o una señal de poliadenilación para incrementar la expresión de una proteína**

30 Prioridad:

13.08.2010 WO PCT/EP2010/004998

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

26.12.2014

73 Titular/es:

**CUREVAC GMBH (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 2 525 556 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 una secuencia de tallo-lazo de histona y poli(A) o una señal de poliadenilación para incrementar la expresión de una proteína codificada

5

La presente solicitud describe una secuencia de ácido nucleico codificante, en particular un ARN mensajero (ARNm), que comprende o codifica para una secuencia tallo-lazo de histona y poli(A) o una señal de poliadenilación, así como a su uso para incrementar la expresión de una proteína codificada. Se describe su uso para la preparación de una composición farmacéutica, en especial una vacuna, por ejemplo para su utilización en el tratamiento de enfermedades cancerosas, enfermedades cardiovasculares, enfermedades infecciosas, enfermedades autoinmunes o enfermedades genéticas, o en la terapia génica.

10

Además de las enfermedades cardiovasculares e infecciosas, la ocurrencia de tumores y enfermedades cancerosas es una de las causas de muerte más frecuentes en la sociedad moderna y, en la mayoría de los casos, está asociada a un coste considerable en términos de terapia y medidas rehabilitadoras posteriores. El tratamiento de tumores y enfermedades cancerosas depende en gran medida, por ejemplo, del tipo de tumor que se presente, de la edad, la distribución de las células cancerosas en el paciente a tratar, etc. Actualmente, la terapia del cáncer se lleva a cabo convencionalmente mediante el uso de terapias de radiación o quimioterapias, además de operaciones invasivas. Sin embargo, estas terapias convencionales típicamente provocan un estrés extraordinario en el sistema inmunológico y en algunos casos sólo pueden aplicarse a un nivel limitado. Además, la mayoría de estas terapias convencionales requieren largos intervalos entre los tratamientos individuales para permitir la regeneración del sistema inmunológico.

15

20

Así, en los últimos años se han investigado estrategias complementarias a estos "tratamientos convencionales" para evitar o al menos reducir el impacto de estos tratamientos sobre el sistema inmunológico. En particular, uno de estos tratamientos complementarios incluye enfoques de terapia génica o vacunación genética, que se ha encontrado son muy prometedores para el tratamiento o para soportar estas terapias convencionales.

25

La terapia génica y la vacunación genética son métodos de medicina molecular que ya han sido probados en la terapia y prevención de enfermedades y generalmente provocan un efecto considerable en la práctica médica diaria, en particular en el tratamiento de enfermedades como las anteriormente mencionadas. La terapia génica también se usa en otros campos de la medicina, por ejemplo en caso de enfermedades genéticas, es decir enfermedades (heredadas), que son causadas por un defecto génico definido y se heredan de acuerdo con las leyes de Mendel. Ejemplos bien conocidos de estas enfermedades genéticas incluyen, entre otras, mucoviscidosis (fibrosis quística) y anemia de células falciformes. Ambos métodos, terapia génica y vacunación genética, se basan en la introducción de ácidos nucleicos en las células o tejidos del paciente y procesamiento subsecuente de la información codificada por el ácido nucleico que ha sido introducido en las células o tejido, es decir la expresión (proteínica) de los polipéptidos deseados.

30

35

En los enfoques de terapia génica, típicamente se usa ADN incluso a pesar de que el ARN también se ha desarrollado recientemente. De forma importante, en todos estos enfoques de terapia génica el ARNm funciona como mensajero para la información de secuencia de la proteína codificada, independientemente de si se usa ADN, ARN viral o ARNm.

40

En general el ARN se considera una molécula inestable: las ARNasas son ubicuas y notoriamente difíciles de inactivar. Además, el ARN también es químicamente más lábil que el ADN. Así, tal vez sorprenda que el "estado por defecto" de un ARNm en una célula eucariótica se caracterice por una estabilidad relativa y sean necesarias señales específicas para acelerar el decaimiento de los ARNm individuales. La principal razón de este descubrimiento parece ser que el decaimiento de ARNm dentro de las células es catalizado casi exclusivamente por exonucleasas. Sin embargo, los extremos de las moléculas de ARNm eucarióticas están protegidos contra estas enzimas por estructuras terminales específicas y sus proteínas asociadas: una m⁷GpppN CAP en el extremo 5' y típicamente una secuencia poli(A) en el extremo 3'. La eliminación de estas dos modificaciones terminales se considera así limitadora de la velocidad para el decaimiento del ARNm. Aunque se haya caracterizado un elemento estabilizador en el 3' UTR del ARNm de alfa-globina, las secuencias de ARN que afectan el cambio de las moléculas de ARNm eucarióticas actúan típicamente como un promotor de decaimiento, normalmente acelerando la desadenilación (revisado en Meyer, S., C. Temme, y col. (2004), Crit Rev Biochem Mol Biol 39(4): 197-216).

45

50

Como se mencionó anteriormente, los extremos 5' de los ARNm eucarióticos son típicamente modificados post-transcripcionalmente para llevar una estructura CAP metilada, por ejemplo m⁷GpppN. Además de los papeles en el empalme, estabilización y transporte del ARN, la estructura CAP incrementa significativamente el reclutamiento de la subunidad ribosómica 40S al extremo 5' del ARNm durante el inicio de la traducción. Esta última función requiere el reconocimiento de la estructura CAP por el complejo de factor de inicio eucariótico eIF4F. La secuencia poli(A) además estimula la traducción mediante el reclutamiento de la

55

subunidad 40S incrementado a moléculas de ARNm, un efecto que requiere la intervención de una proteína de unión poli(A) (PABP). PABP, a su vez, ha demostrado recientemente interactuar físicamente con eIF4G, que es parte del complejo eIF4F unido a CAP. Así, se postuló un modelo de lazo cerrado de inicio de traducción en moléculas de ARNm poliadeniladas de extremos bloqueados (Michel, Y. M., D. Poncet, y col. 5 (2000), *J Biol Chem* 275(41): 32268-76).

Prácticamente todos los ARNm eucarióticos concluyen con esta secuencia poli(A) que es añadida a su extremo 3' por la maquinaria ubicua de corte/poliadenilación. La presencia de una secuencia poli(A) en el extremo 3' es una de las características más reconocibles de los ARN eucarióticos. Después del corte, la mayoría de los pre-ARNm, excepto 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 promueve el transporte del ARNm del núcleo al citoplasma y afecta a la estabilidad y la traducción de las moléculas de ARNm. La formación de este extremo 3' ocurre en una reacción de dos etapas dirigida por la maquinaria de corte/poliadenilación y depende de la presencia de dos elementos de secuencia en los precursores de ARNm (pre-ARNm); un hexanucleótido altamente conservado AAUAAA (señal de poliadenilación) y una secuencia rica en G/U aguas abajo. En una primera etapa, los pre-ARNm son cortados entre estos dos elementos. En una segunda etapa estrechamente acoplada a la primera etapa, el extremo 3' recién formado es ampliado por la adición de una secuencia poli(A) consistente en 200-250 adenilatos, que afecta posteriormente a todos los aspectos del metabolismo del ARNm, incluyendo su exportación, estabilidad y traducción (Dominski, Z. y W. F. Marzluff (2007), *Gene* 396(2): 373-90).

Levy y col. proporciona una secuencia y una caracterización funcional del exón terminal del gen receptor de la insulina humana y la identificación de cuatro dominios de poliadenilación responsables del procesamiento del extremo 3' del ARNm del receptor de la insulina humano (Levy y col., (1995), *Biochimica et Biophysica Acta* 1263(3):253-257).

La WO01/12824 A1 proporciona métodos y medios para reducir la expresión de un ácido nucleico de interés, alterando con ello el fenotipo de un organismo, en particular de una planta, proporcionando un ARN diana específico no poliadenilado al núcleo de la célula huésped.

La única excepción conocida a esta regla son las moléculas de ARNm de histona dependientes de la replicación que concluyen con un tallo-lazo de histona en lugar de una secuencia poli(A). Secuencias tallo-lazo de histona ilustrativas se describen en López y col. (Dávila López, M., & Samuelsson, T. (2008), *RNA* (Nueva York, N. Y.), 14(1), 1-10. Doi:10.1261/ma.782308).

Las estructuras tallo-lazo en pre-ARNm de histona van seguidas típicamente por una secuencia rica en purina conocida como el elemento aguas abajo de histona (HDE). Estos pre-ARNm son procesados en el núcleo por un solo corte endonucleolítico de aproximadamente 5 nucleótidos aguas abajo del tallo-lazo, catalizado por el snRNP U7 mediante apareo de bases del U7 snARN con el HDE.

Debido a la necesidad de empaquetar el ADN recién sintetizado en la cromatina, la síntesis de histona es regulada en conjunto con el ciclo celular. Se consigue aumentar la síntesis de proteínas histona durante la fase S mediante la activación transcripcional de genes de histona, así como por la regulación post-transcripcional de los niveles de ARNm de histona. Podría demostrarse que el tallo-lazo de la histona es esencial para todas las etapas post-transcripcionales de la regulación de la expresión de histonas. Para un procesamiento eficiente, es necesario exportar la molécula de ARNm en el citoplasma, cargarla en polirribosomas y regular la estabilidad del ARNm.

En este contexto, se identificó una proteína de 32 kDa, la cual está asociada al tallo-lazo 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 a tallo-lazo (SLBP) es regulado en el ciclo celular y es el más alto durante la fase S, cuando los niveles de ARNm de histona son mayores. La SLBP es necesaria para un procesamiento 3'-terminal eficiente del pre-ARNm de histona por el U7 snRNP. Después del procesamiento, la SLBP permanece asociada al tallo-lazo en el extremo de los ARNm de histona maduros y estimula su traducción en proteínas histona en el citoplasma (Dominski, Z. y W. F. Marzluff (2007), *Gene* 396(2): 373-90). De forma interesante, el dominio de unión a ARN de la SLBP se conserva a través de metazoarios y protozoarios (Dávila López, M., y Samuelsson, T. (2008), *RNA* (Nueva York, N.Y.), 14(1), 1-10. Doi:10.1261/ma.782308) y podría demostrarse que su unión a la secuencia tallo-lazo de la histona es dependiente de la estructura tallo-lazo y que el sitio de unión mínimo contiene al menos 3 nucleótidos 5' y 2 nucleótidos 3' del tallo-lazo (Pandey, N.B., y col., (1994), *Molecular and Cellular Biology*, 14(3), 1709-1720 y Williams, A. S., y Marzluff, W. F., (1995), *Nucleic Acids Research*, 23(4), 654-662).

Incluso aunque los genes de histona generalmente se clasifican bien como "dependientes de la replicación", dando origen a que el ARNm concluya en un tallo-lazo de histona, o bien "tipo de reemplazo", dando origen a un ARNm que porta una cola poli(A) en su lugar, en algunos casos muy raros se han identificado ARNm de

origen natural que contienen tanto tallo-lazo de histona como poli(A) u oligo(A) 3' de las mismas. Sánchez y col. examinaron el efecto de las colas oligo(A) de origen natural anexas a 3' del tallo-lazo de histona del ARNm de histona durante la oogénesis de *Xenopus* usando luciferasa como proteína reportera y encontraron que la cola oligo(A) es una parte activa del mecanismo de represión de la traducción que silencia el ARNm de histona durante la oogénesis y su eliminación es parte del mecanismo que activa la traducción de los ARNm de histona (Sánchez, R. y W. F. Marzluff (2004), *Mol Cell Biol* 24(6): 2513-25).

Zhong y col. describen una proteína enlazante de ARN de doble hebra necesaria para la activación de los mensajes reprimidos en células germinales de mamíferos (Zhong, J. y col.,(1999), 22(2) 1 Junio:171-174).

Además, los requisitos para la regulación de histonas dependientes de replicación al nivel de procesamiento de pre-ARNm y estabilidad de ARNm han sido investigados usando constructos artificiales que codifican para la proteína marcadora alfa-globina, tomando ventaja del hecho de que el gen de globina contiene intrones a diferencia de los genes de histona, sin intrones. Para ello, se generaron constructos donde la secuencia de codificación de alfa-globina estaba seguida por una señal tallo-lazo de histona (tallo-lazo de histona seguida por el elemento aguias abajo de histona) y una señal de poliadenilación (Whitelaw, E., y col., (1986). *Nucleic Acids Research*, 14(17), 7059-7070; Pandey, N. B., y Marzluff, W. F. (1987). *Molecular and Cellular Biology*, 7(12), 4557-4559; Pandey, N.B., y col., (1990). *Nucleic Acids Research*, 18(11), 3161-3170).

En otro enfoque, Lüscher y col. investigaron la regulación dependiente del ciclo celular de un gen H4 de histona recombinante. Se generaron constructos donde la secuencia de codificación de H4 estaba seguida por una señal tallo-lazo de histona y una señal de poliadenilación, ambas señales de procesamiento incidentalmente separadas por una secuencia de codificación de galactoquinasa (Lüscher, B. y col., (1985). *Proc. Natl. Acad. Sci. USA*, 82(13), 4389-4393).

Además, Stauber y col. identificaron la secuencia mínima requerida para conferir regulación de ciclo celular en los niveles de ARNm de la histona H4. Para estas investigaciones se usaron constructos que comprendían una secuencia de codificación para el marcador de selección Xantina:guanina-fosforribosil-transferasa (GPT), que precede una señal de tallo-lazo de histona seguida por una señal de poliadenilación (Stauber, C. y col., (1986). *EMBO J*, 5(12), 3297-3303).

Al examinar el procesamiento de pre-ARNm de histona, Wagner y col. identificaron los factores requeridos para el corte de las moléculas de pre-ARNm de histona usando un constructo reporter que situando EGFP entre una señal de tallo-lazo de histona y una señal de poliadenilación, de manera que el EGFP se expresara sólo en caso de que el procesamiento de pre-ARNm de histona fuera interrumpido (Wagner, E. J. y col. (2007), *Mol Cell* 28(4), 692-9).

Debe mencionarse que la traducción de ARNm poliadenilado normalmente requiere que la secuencia poli(A) 3' esté cercana a la CAP 5'. Esto es mediado por una interacción proteína-proteína entre la proteína de unión poli(A) y el factor de inicio eucariótico eIF4G. Con respecto a las moléculas de ARNm de histona dependientes de replicación, se ha descubierto un mecanismo análogo. En este contexto, Gallie y col. demuestran que el tallo-lazo de histona es funcionalmente similar a una secuencia poli(A), toda vez que incrementa la eficiencia de traducción y es co-dependiente de un 5'-CAP para establecer un nivel de traducción eficiente. Demostraron que el tallo-lazo de histona es suficiente y necesario para incrementar la traducción de un ARNm reporter en células de ovario de hámster Chino transfectadas, debiendo situarse en el extremo 3' para funcionar óptimamente. Así, de forma similar a la cola poli(A) en otras moléculas de ARNm, el extremo 3' de estas moléculas de 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., y Marzluff, W. F. (1996), *Nucleic Acids Research*, 24(10), 1954-1962).

Ling y col. demostraron que la proteína de histona de tallo-lazo 3' terminal mejora la traducción mediante una interacción funcional y física con el factor de iniciación eucariótico 4G (eIF4G) y eIF3 (Ling, J. y col. (2002), *Molecular and Cellular Biology*, Vol. 22, No. 22, 7853-7867).

Además, podría demostrarse que la SLBP está unido al ARNm de histona citoplásmico y se requiere para su traducción. Incluso a pesar de que SLBP no interactúa directamente con eIF4G, el dominio requerido para la traducción de ARNm de histona interactúa con la proteína SLIP1, recientemente identificada. En un paso más, SLIP1 interactúa con eIF4G y permite circularizar el ARNm de histona y soportar una traducción eficiente de ARNm de histona por un mecanismo similar a la traducción de moléculas de ARNm poliadeniladas.

Como se mencionó arriba, los enfoques de la terapia génica normalmente usan ADN para transferir la información de codificación en la célula, que luego es transcrita en ARNm, llevando los elementos de origen

natural de un ARNm, particularmente la estructura 5'-CAP y la secuencia 3' poli(A) para asegurar la expresión de la proteína terapéutica codificada.

5 Sin embargo, muchos casos sistemas de expresión basados en la introducción de estos ácidos nucleicos en las células o tejido del paciente y la expresión subsecuente de los polipéptidos deseados codificados por estos ácidos nucleicos no exhiben el nivel deseado o incluso requerido de expresión que puede permitir una terapia eficiente, independientemente de si se usa ADN o ARN.

10 En la técnica anterior, se han hecho hasta la fecha diferentes intentos para incrementar la producción de la expresión de una proteína codificada, en particular empleando sistemas de expresión mejorados, *in vitro* y/o *in vivo*. Los métodos para incrementar la expresión descritos generalmente en la técnica anterior se basan en el uso convencional de vectores o casetes de expresión que contienen promotores específicos y elementos de regulación correspondientes. Debido a que estos vectores de expresión o casetes están típicamente limitados a sistemas celulares particulares, estos sistemas de expresión tienen que ser adaptados para su uso en diferentes sistemas celulares. Estos vectores o casetes de expresión adaptados son luego transfectedos normalmente en las células y tratados típicamente dependiendo de la línea celular específica.

15 Por tanto, se da preferencia principalmente a las moléculas de ácido nucleico que son capaces de expresar las proteínas codificadas en una célula diana por sistemas inherentes en la célula, independientemente de promotores y elementos de regulación que son específicos para tipos celulares particulares. En este contexto, se puede distinguir entre elementos de estabilización de ARN y elementos que incrementan la eficiencia de traducción del ARNm.

20 La US 2007/172949 A9 describe vectores y vectores virales capaces de expresar genes exógenos o secuencias de ácido nucleico de exógena en una célula diana de interés.

25 En la solicitud WO 02/098443 (CureVac GmbH) se describen ARNm que son optimizados en su secuencia de codificación y que en general son adecuados para este propósito. Por ejemplo, la WO 02/098443 describe moléculas de ARNm que son estabilizadas en forma general y optimizadas para traducción en sus regiones de codificación. La WO 02/098443 describe además un método para determinar modificaciones de secuencia. La WO 02/098443 describe además posibilidades para sustituir nucleótidos adenina y uracilo en secuencias de ARNm para incrementar el contenido en guanina/citosina (G/C) de las secuencias. De acuerdo con la WO 02/098443, estas sustituciones y adaptaciones para incrementar el contenido en G/C pueden usarse para aplicaciones de terapia génica, pero también en vacunas genéticas para el tratamiento del

30 cáncer o de enfermedades infecciosas. En este contexto, la WO 02/098443 menciona secuencias en general como secuencias base para estas modificaciones, donde el ARNm modificado codifica para al menos un péptido o polipéptido biológicamente activo que es traducido en el paciente a tratar, por ejemplo tanto no en absoluto o inadecuadamente o con fallas. Alternativamente, la WO 02/098443 propone moléculas de ARNm que codifican para antígenos, por ejemplo antígenos tumorales o virales, como una secuencia base para estas modificaciones.

35

En otro enfoque para incrementar la expresión de una proteína codificada, la solicitud WO 2007/036366 describe el efecto positivo de secuencias poli(A) largas (particularmente de longitud superior a 120 pb) y la combinación de al menos dos regiones no traducidas 3' del gen de beta-globina en la estabilidad y actividad de traducción de ARNm.

40 Sin embargo, incluso aunque todos estos documentos de la técnica anterior ya intentaron proporcionar herramientas bastante eficientes para enfoques de terapia génica y además mejoraron la estabilidad y actividad de traducción del ARNm, sigue existiendo el problema de una estabilidad generalmente más baja de las aplicaciones basadas en ARN frente a las vacunas de ADN y enfoques terapéuticos genéticos basados en ADN. En consecuencia, sigue existiendo la necesidad en la técnica de proporcionar herramientas mejoradas para enfoques de terapia génica y vacunación genética o como terapia complementaria para tratamientos convencionales como los descritos arriba, que permitan una mejor provisión de proteínas codificadas *in vivo*, por ejemplo mediante una actividad de traducción y/o estabilidad de ARNm mejorada, preferiblemente para la terapia génica.

45

50 El objetivo subyacente a la presente invención es así proporcionar métodos adicionales y/o alternativos para incrementar la expresión de una proteína codificada, preferentemente mediante estabilización adicional del ARNm y/o por un incremento de la eficiencia de traducción de este ARNm con respecto a tales ácidos nucleicos conocidos de la técnica anterior, para el uso en aplicaciones terapéuticas (por ejemplo, terapia génica y vacunación genética).

55 Este objetivo se resuelve por el objeto de las reivindicaciones anexas. En particular, el objetivo de la presente invención se resuelve de acuerdo con una primera realización mediante una secuencia de ácido nucleico que comprende o codifica para

- 5 a) una región de codificación, que codifica preferentemente para un péptido o proteína terapéuticamente activos, una proteína adyuvante, un antígeno, un antígeno tumoral, un antígeno patogénico, un antígeno animal, un antígeno viral, un antígeno protozoal, un antígeno bacteriano, un antígeno alergénico, un antígeno autoinmune, un alérgeno, un anticuerpo, un péptido o proteína inmunoestimulador o un antígeno específico del receptor de células T;
- b) al menos un tallo-lazo de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación,

10 para su uso con el fin de incrementar el nivel de expresión de dicha proteína codificada en el tratamiento de enfermedades cancerosas, enfermedades infecciosas, preferentemente enfermedades infecciosas (virales, bacterianas o protozoológicas), enfermedades autoinmunes, alergias o enfermedades alérgicas, enfermedades monogénicas o enfermedades genéticas, enfermedades cardiovasculares, enfermedades neuronales, enfermedades del sistema respiratorio, del sistema digestivo, de la piel, trastornos musculoesqueléticos, trastornos del tejido conectivo, inmunodeficiencias, enfermedades endocrinas, nutricionales y metabólicas, enfermedades oculares o del oído, donde la proteína codificada no es una proteína histona, no es una proteína reportera seleccionada entre EGFP y luciferasa, y no es una proteína marcadora o de selección seleccionada entre alfa-globina, galactoquinasa y xantina:guanina-fosforribosil-transferasa (GPT).

15 En este contexto es particularmente preferente que el ácido nucleico de la invención de acuerdo con la primera realización de la presente invención se produzca al menos en partemediante síntesis de ADN o ARN o sea un ácido nucleico aislado.

20 La presente invención se basa en el sorprendente descubrimiento de los presentes inventores de que la combinación de una secuencia poli(A) o una señal de poliadenilación y al menos un tallo-lazo de histona, incluso aunque ambos representen mecanismos alternativos en la naturaleza, actúa sinérgicamente, ya que esta combinación incrementa la expresión de proteínas 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-lazo de histona se observa independientemente del orden de poli(A) y tallo-asa de histona y de la longitud de la secuencia poli(A).

30 Por tanto, la molécula de ácido nucleico empleada comprende o codifica para a) una región de codificación, que codifica para un péptido o proteína terapéuticamente activos, una proteína adyuvante, un antígeno, un antígeno tumoral, un antígeno patogénico, un antígeno animal, un antígeno viral, un antígeno protozoal, un antígeno bacteriano, un antígeno alergénico, un antígeno autoinmune, un alérgeno, un anticuerpo, un péptido o proteína inmunoestimulador o un antígeno específico del receptor de células T; b) al menos un tallo-lazo de histona y c) una secuencia poli(A) o una secuencia de poliadenilación; para su uso con el fin de incrementar el nivel de expresión de dicha proteína codificada en el tratamiento de enfermedades cancerosas, enfermedades infecciosas, preferentemente enfermedades infecciosas (virales, bacterianas o protozoológicas), enfermedades autoinmunes, alergias o enfermedades alérgicas, enfermedades monogénicas o enfermedades genéticas, enfermedades cardiovasculares, enfermedades neuronales, enfermedades del sistema respiratorio, del sistema digestivo, de la piel, trastornos musculoesqueléticos, trastornos del tejido conectivo, inmunodeficiencias, enfermedades endocrinas, nutricionales y metabólicas, enfermedades oculares o del oído, donde la proteína codificada no es una proteína histona, no es una proteína reportera seleccionada entre EGFP y luciferasa, y no es una proteína marcadora o de selección seleccionada entre alfa-globina, galactoquinasa y xantina:guanina-fosforribosil-transferasa (GPT). La proteína codificada puede también no ser GFP o β -galactosidasa.

40 En otro aspecto alternativo de la primera realización de la presente invención, el ácido nucleico empleado no comprende el elemento aguas abajo de histona (HDE).

45 En este contexto, es preferente que el ácido nucleico empleado comprenda o codifique en la dirección 5' a 3':

a) una región de codificación, que codifica preferentemente para un péptido o proteína terapéuticamente activos, una proteína adyuvante, un antígeno, un antígeno tumoral, un antígeno patogénico, un antígeno animal, un antígeno viral, un antígeno protozoal, un antígeno bacteriano, un antígeno alergénico, un antígeno autoinmune, un alérgeno, un anticuerpo, un péptido o proteína inmunoestimulador o un antígeno específico del receptor de células T;

50 b) al menos un tallo-lazo de histona, opcionalmente sin un elemento 3' de histona aguas abajo al tallo-lazo de histona,

c) una secuencia poli(A) o una señal de poliadenilación,

55 para su uso con el fin de incrementar el nivel de expresión de dicha proteína codificada en el tratamiento de enfermedades cancerosas, enfermedades infecciosas, preferentemente enfermedades infecciosas (virales, bacterianas o protozoológicas), enfermedades autoinmunes, alergias o enfermedades alérgicas, enfermedades monogénicas o enfermedades genéticas, enfermedades cardiovasculares, enfermedades neuronales, enfermedades del sistema respiratorio, del sistema digestivo, de la piel, trastornos musculoesqueléticos, trastornos del tejido conectivo, inmunodeficiencias, enfermedades endocrinas, nutricionales y metabólicas, enfermedades oculares o del oído, donde la proteína codificada no es una proteína histona, no es una proteína reportera seleccionada entre EGFP y luciferasa, y no es una proteína marcadora o de selección seleccionada entre alfa-globina, galactoquinasa y xantina:guanina-fosforribosil-transferasa (GPT).

esqueléticos, trastornos del tejido conectivo, inmunodeficiencias, enfermedades endocrinas, nutricionales y metabólicas, enfermedades oculares o del oído, donde la región codificante no codifica para proteínas histona, proteínas reporteras y proteínas marcadoras o de selección tial como se han mencionado anteriormente.

5 El término “elemento aguas abajo de histona (HDE)” se refiere a un tramo de polinucleótido rico en purina de aproximadamente 15 a 20 nucleótidos 3’ de tallos-lazos de origen natural, que representa el sitio de unión para el U7 snRNA implicado en el procesamiento de pre-ARNm de histona en el ARNm de histona maduro. Por ejemplo en erizos de mar el HDE es CAAGAAAGA (Dominski, Z. y W. F. Marzluff (2007), Gene 396(2): 373-90).

10 Además, es preferible que el ácido nucleico empleado de acuerdo con la primera realización de la presente invención no comprenda un intrón.

En otra realización preferente, la secuencia de ácido nucleico para su uso como se ha indicado anteriormente según la primera realización de la presente invención comprende o codifica para desde 5’ a 3’:

- 15 a) una región de codificación, que codifica preferentemente para un péptido o proteína terapéuticamente activos, una proteína adyuvante, un antígeno, un antígeno tumoral, un antígeno patogénico, un antígeno animal, un antígeno viral, un antígeno protozoal, un antígeno bacteriano, un antígeno alergénico, un antígeno autoinmune, un alérgeno, un anticuerpo, un péptido o proteína inmunoestimulador o un antígeno específico del receptor de células T;
- 20 b) al menos un tallo-lazo de histona,
c) una secuencia poli(A); y

donde la región codificante no codifica para proteínas histona, proteínas reporteras o proteínas de selección como se ha referido anteriormente.

25 El uso de la invención se basa en una secuencia de ácido nucleico de acuerdo con la primera realización que comprende cualquier ácido nucleico adecuado, seleccionado por ejemplo entre cualquier ADN (monohebra o de doble hebra), de preferencia, sin limitarse a, por ejemplo ADN genómico, moléculas de ADN de una sola hebra, moléculas de ADN de doble hebra, o puede seleccionarse por ejemplo de cualquier PNA (ácido nucleico péptido), o puede seleccionarse, por ejemplo, de cualquier ARN (de una sola hebra o de doble hebra), preferentemente de un ARN mensajero (ARNm); etc. El ácido nucleico empleado también puede 30 comprender un ARN viral (ARNv). Sin embargo, la secuencia de ácido nucleico empleado puede no ser un ARN viral o puede no contener un ARN viral. Más específicamente, la secuencia de ácido nucleico empleado puede no contener elementos de secuencia virales, por ejemplo potenciadores virales o promotores virales (por ejemplo ningún promotor viral inactivado o elemento de secuencia, más específicamente no activado por estrategias de reemplazo), u otros elementos de secuencia viral, o secuencias de ácido nucleico virales o 35 retrovirales. Más específicamente, la secuencia de ácido nucleico empleado puede no ser un vector retroviral o viral o un vector retroviral o viral modificado.

40 En cualquier caso, la secuencia de ácido nucleico empleado puede o no contener una secuencia potenciadora y/o promotora, que puede estar modificada o no o activada o no. El potenciador y/o promotor puede ser expresable o no expresable en eucariotes y/o expresable o no expresable en procariontes. La molécula de ácido nucleico empleado puede contener una secuencia que codifique para una ribozima (de auto-empalme) o no.

De preferencia, la molécula de ácido nucleico empleado es un ARN.

45 En aspectos particulares de la primera realización de la presente invención, el ácido nucleico empleado es una secuencia de ácido nucleico comprendida en un ácido nucleico adecuado para la transcripción *in vitro*, particularmente 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*, tales como promotores T3, T7 o Sp6).

50 En aspectos preferentes particulares adicionales de la primera realización de la presente invención, el ácido nucleico empleado está comprendido en un ácido nucleico adecuado para la transcripción y/o traducción en un sistema de expresión (por ejemplo en un vector o plásmido de expresión), particularmente un sistema de expresión procariótico (por ejemplo bacterias como *E. coli*) o eucariótico (por ejemplo células de mamífero como células CHO, células de levadura o células de insecto u organismos completostales como plantas o animales).

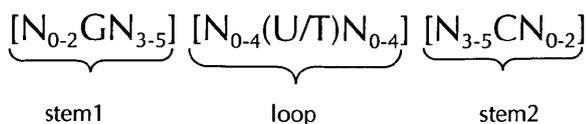
55 El término “sistema de expresión” significa un sistema (cultivo celular u organismos completos) que es adecuado para la producción de péptidos, proteínas o ARN, particularmente ARNm.

El uso de la invención está basado en una secuencia de ácido nucleico de acuerdo con la primera realización de la presente invención que comprende o codifica para al menos un tallo-lazo de histona. En el contexto de la presente invención, este tallo-lazo de histona se deriva típicamente de genes de histona y comprende un apareamiento de bases intramolecular de dos secuencias complementarias completa o parcialmente inversas adyacentes, formando así un tallo-lazo. El tallo-lazo puede estar presente en ADN de una sola hebra o más comúnmente en ARN. La estructura también se conoce como horquilla u horquilla y asa y consiste normalmente en un tallo y un lazo (terminal) dentro de una secuencia consecutiva, donde el tallo está formado por dos secuencias complementarias completa o parcialmente inversas adyacentes separadas por una secuencia corta a modo de separador que construye el lazo de la estructura tallo-lazo. Las dos secuencias complementarias completa o parcialmente inversas adyacentes pueden definirse por ejemplo como elementos tallo y lazo stem1 y stem2. El tallo-lazo se forma cuando estas dos secuencias complementarias inversas completa o parcialmente adyacentes, por ejemplo, elementos de tallo-lazo stem1 y stem2, forman pares de bases entre sí, llevando a una secuencia de ácido nucleico de doble hebra que comprende un lazo no apareado en su extremo terminal formado por la secuencia corta ubicada entre los elementos tallo-lazo stem1 y stem2 de la secuencia consecutiva. Así, el lazo no apareado representa típicamente una región del ácido nucleico que no es capaz del apareamiento de bases con cualquiera de estos elementos de tallo y lazo. La estructura en forma de piruleta resultante es un bloque de construcción clave de muchas estructuras secundarias de ARN. La formación de una estructura tallo-lazo depende entonces de la estabilidad de las regiones de tallo y lazo resultantes, donde el primer prerrequisito es típicamente la presencia de una secuencia que pueda plegarse sobre sí misma para formar una doble hebra apareada. La estabilidad de los elementos tallo y lazo apareados se determina por la longitud, el número de no coincidencias o bultos que contiene (un pequeño número de no coincidencias es típicamente tolerable, especialmente en una doble hebra larga) y la composición de bases de la región apareada. En el contexto de la presente invención, la longitud de lazo óptima es 3-10 bases, muy preferiblemente de 3 a 8, 3 a 7, 3 a 6 o incluso más preferiblemente de 4 a 5 bases, y con total preferencia de 4 bases.

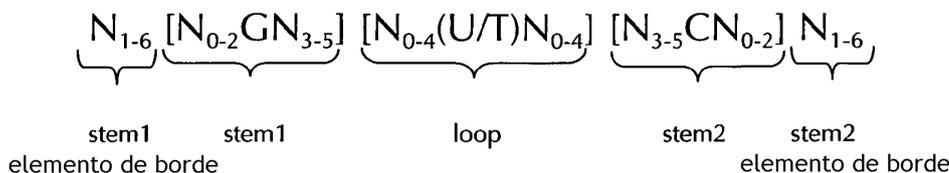
De acuerdo con la presente invención, la secuencia de tallo y lazo de histona de acuerdo con el componente (b) de la reivindicación 1 puede no derivarse de una proteína de histona de ratón. Más específicamente, la secuencia de tallo y lazo de histona puede no derivarse del gen de histona de ratón H2A614. Asimismo, el ácido nucleico empleado puede no contener una secuencia de tallo y lazo de histona de ratón ni un gen de histona de ratón H2A614. Además, la secuencia de ácido nucleico empleado puede no contener una señal de procesamiento de tallo-lazo, más específicamente, una señal de procesamiento de histona de ratón y, en particular puede no contener una señal de procesamiento de tallo-lazo de ratón H2kA614. Asimismo, la molécula de ácido nucleico empleado puede contener al menos un gen de histona de mamífero. Sin embargo, el al menos un gen de histona de mamífero puede no ser de Seq. ID No. 7 de WO 01/12824.

De acuerdo con un aspecto preferente de la primera realización, la secuencia de ácido nucleico empleado comprende o codifica para al menos una secuencia de tallo-lazo de histona, preferentemente de acuerdo con al menos una de las siguientes fórmulas (I) o (II):

fórmula (I) (secuencia de tallo-lazo sin elementos de borde de tallo):



fórmula (II) (secuencia de tallo-lazo con elementos de borde de tallo):



donde:

los elementos de borde stem1 o stem2 N_{1-6} son una secuencia consecutiva de 1 a 6, de preferencia de 2 a 6, en especial de 2 a 5, en particular de 3 a 5, con especial preferencia de 4 a 5 ó 5 N, donde cada N se selecciona independientemente de un nucleótido seleccionado entre A, U, T, G y C, o un análogo nucleótido del mismo.

stem1 [N₀₋₂GN₃₋₅]

Es complementario inverso o complementario parcialmente inverso con el elemento stem2, y es una secuencia consecutiva de entre 5 a 7 nucleótidos; donde N₀₋₂ es una secuencia consecutiva de 0 a 2, de preferencia de 0 a 1, en especial de 1 N, seleccionándose cada N independientemente de un nucleótido seleccionado entre A, U, T, G y C o un análogo nucleótido del mismo; donde N₃₋₅ es una secuencia consecutiva de 3 a 5, de preferencia de 4 a 5, en especial de 4 N, seleccionándose cada N independientemente de un nucleótido seleccionado entre A, U, T, G y C o un análogo nucleótido del mismo, y donde G es guanosina o un análogo de la misma, y puede ser reemplazado opcionalmente por una citidina o un análogo de la misma, siempre que su citidina nucleótido complementaria en stem2 sea reemplazada por guanosina;

la secuencia asa (loop) [N₀₋₄(U/T)N₀₋₄]

Se ubica entre los elementos stem1 y stem2 y es una secuencia consecutiva de 3 a 5 nucleótidos, preferiblemente de 4 nucleótidos; donde cada N₀₋₄ es independientemente del otro una secuencia consecutiva de 0 a 4, de preferencia de 1 a 3, en especial de 1 a 2 N, seleccionándose cada N independientemente de un nucleótido seleccionado entre A, U, T, G y C o un análogo de nucleótido del mismo; y donde U/T representa uridina, u opcionalmente timidina;

stem2 [N₃₋₅CN₀₋₂]

Es complementario inverso o parcialmente complementario inverso con el elemento stem1 y es una secuencia consecutiva de entre 5 a 7 nucleótidos; donde N₃₋₅ es una secuencia consecutiva de 3 a 5, de preferencia de 4 a 5, en especial de 4 N, seleccionándose cada N independientemente de un nucleótido seleccionado entre A, U, T, G y C o un análogo nucleótido del mismo; donde N₀₋₂ es una secuencia consecutiva de 0 a 2, de preferencia de 0 a 1, en especial de 1 N, seleccionándose cada N independientemente de un nucleótido seleccionado entre A, U, T, G o C o un análogo de nucleótido del mismo; y donde C es citidina o un análogo del mismo, y puede ser reemplazado opcionalmente por una guanosina o un análogo de la misma siempre que su guanosina nucleótido complementaria en stem1 sea reemplazada por citidina;

Donde stem1 y stem2 son capaces del apareo de bases entre sí formando una secuencia complementaria inversa, pudiendo ocurrir el apareo de bases entre stem1 y stem2, por ejemplo por apareo de bases de Watson-Crick de nucleótidos A y U/T o G y C o por apareo de bases no de Watson-Crick, por ejemplo, apareo de bases tambaleo, apareo de bases de Watson-Crick inverso, apareo de bases de Hoogsteen, apareo de bases de Hoogsteen inverso o son capaces del apareo de bases entre sí formando una secuencia complementaria parcialmente inversa, pudiendo ocurrir un apareo de bases incompleto entre stem1 y stem2 en base en que una o más bases en un tallo no tengan una base complementaria en la secuencia complementaria inversa del otro tallo.

En el contexto anterior, un apareo de bases tambaleo(wobble) es típicamente un apareo de bases no de Watson-Crick entre dos nucleótidos. Los cuatro pares de bases wobble principales en el presente contexto que pueden usarse 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).

En consecuencia, en el contexto de la presente invención, una base wobble es una base que forma un par de bases de tambaleo con una base adicional como la descrita arriba. Así, el apareo de bases no de Watson-Crick, por ejemplo apareo de bases wobble puede ocurrir en el tallo de la estructura tallo-lazo de histona de acuerdo con la presente invención.

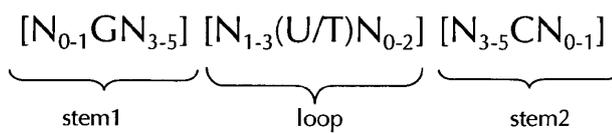
En el contexto anterior una secuencia complementaria parcialmente inversa comprende como máximo 2, de preferencia sólo una no coincidencia en la estructura de tallo de la secuencia tallo-lazo formada por el apareo de bases de stem1 y stem2. En otras palabras, preferentemente stem1 y stem2 son capaces del apareo de bases (completo) entre sí a través de la secuencia completa de stem1 y stem2 (100% de apareos de bases de Watson-Crick o no Watson-Crick correctos posibles), formando así una secuencia complementaria inversa donde cada base tiene su base de Watson-Crick o no Watson-Crick correcta pendiente como un socio de unión complementario. Alternativamente, stem1 y stem2 son preferentemente capaces del apareo de bases parcial entre sí a lo largo de la secuencia completa de stem1 y stem2, donde al menos aproximadamente el 70%, 75%, 80%, 85%, 90% o 95% del 100% de los apareos de bases de Watson-Crick o no de Watson-Crick correctos posibles son ocupados con los apareos de bases de Watson-Crick o no de Watson-Crick correctos y como máximo aproximadamente un 30%, 25%, 20%, 15%, 10% o 5% de las bases restantes no están apareadas.

De acuerdo con un aspecto preferente de la primera realización de la invención, la al menos una secuencia de tallo-lazo de histona (con elementos de borde de tallo) de la secuencia de ácido nucleico empleado como la definida aquí comprende una longitud de aproximadamente 15 a alrededor de 25 nucleótidos, preferentemente una longitud de aproximadamente 15 a alrededor de 45 nucleótidos, en especial una longitud de aproximadamente 15 a alrededor de 40 nucleótidos, en particular una longitud de aproximadamente 15 a alrededor de 35 nucleótidos, con especial preferencia una longitud de aproximadamente 15 a alrededor de 30 nucleótidos y con particular preferencia una longitud de aproximadamente 20 a alrededor de 30 nucleótidos, con total preferencia una longitud de aproximadamente 24 a alrededor de 28 nucleótidos.

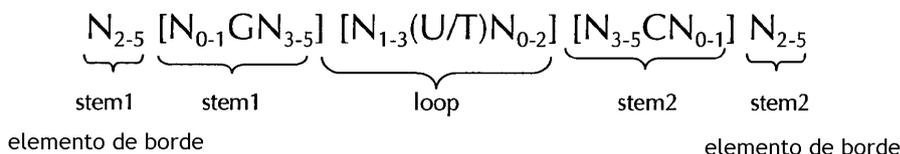
- 5
- 10 De acuerdo con otro aspecto preferente de la primera realización de la invención, la al menos una secuencia de tallo-lazo de histona (sin elementos de borde de tallo) de la secuencia de ácido nucleico empleado como la definida aquí comprende una longitud de alrededor de 10 a aproximadamente 30 nucleótidos, preferiblemente una longitud de alrededor de 10 a aproximadamente 20 nucleótidos, en especial una longitud de alrededor de 12 a aproximadamente 20 nucleótidos, en particular una longitud de alrededor de 14 a aproximadamente 20 nucleótidos, y con especial preferencia una longitud de alrededor de 16 a aproximadamente 17 y con total preferencia una longitud de alrededor de 16 nucleótidos.

De acuerdo con otro aspecto preferente de la primera realización de la invención, la secuencia de ácido nucleico empleado de acuerdo con la primera realización de la presente invención puede comprender o codificar para al menos una secuencia de tallo-lazo de histona de acuerdo con al menos una de las siguientes fórmulas específicas (Ia) o (IIa):

20 fórmula (Ia) (secuencia de tallo-lazo sin elementos de borde de tallo):



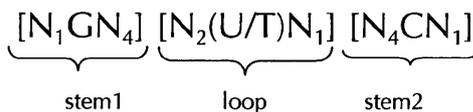
fórmula (IIa) (secuencia de tallo-lazo con elementos de borde de tallo):



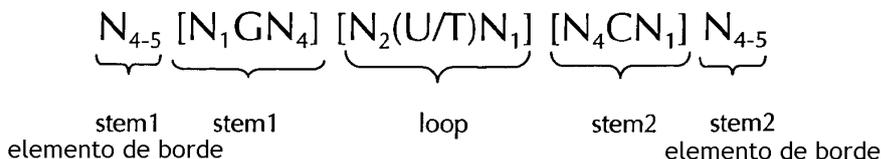
25 donde N, C, G, T y U son como se definió arriba.

De acuerdo con otro aspecto más particularmente preferente de la presente realización, la secuencia de ácido nucleico emplea puede comprender o codificar para al menos una secuencia de tallo-lazo de histona de acuerdo con al menos una de las siguientes fórmulas específicas (Ib) o (IIb):

30 fórmula (Ib) (secuencia de tallo-lazo sin elementos de borde de tallo):



fórmula (IIb) (secuencia de tallo-lazo con elementos de borde de tallo):



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

5 De acuerdo con un aspecto especialmente preferente de la primera realización de la invención, la secuencia de ácido nucleico empleado de acuerdo con la primera realización de la presente invención puede comprender o codificar para al menos una secuencia de tallo-lazo 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 de tallo-lazo y como una secuencia lineal que representa secuencias tallo-lazo de histona como las generadas de acuerdo con el ejemplo 1:

fórmula (Ic): (secuencia de consenso de tallo-lazo de histona de metazoarios y protozoarios sin elementos de borde del tallo):



fórmula (IIc): (secuencia de consenso de tallo-lazo de histona de metazoarios y protozoarios con elementos de borde de tallo):



fórmula (Id): (sin elementos de borde de tallo):



NCNNNNNNUNNNNNGN
 (secuencia lineal) (SEQ ID NO: 3)

Fórmula (IIId): (con elementos de borde de tallo)



N*N*NNNN-NNNN*N*N* (estructura de tallo-asa)

N*N*NNNNCNNNNNNUNNNNNGNNNN*N*N*
 (secuencia lineal) (SEQ ID NO: 4)

5 fórmula (IIe): (secuencia de consenso de tallo-lazo de histona de protozoarios sin elementos de borde de tallo)



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

10 fórmula (IIe): (secuencia de consenso de tallo-lazo de histona de protozoarios con elementos de borde de tallo)

N U
 N N
 N-N
 N-N
 N-N
 N-N
 G-C

N*N*NNND-HNNN*N*N* (estructura de tallo-asa)

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

fórmula (If): (secuencia de consenso de tallo-lazo de histona de metazoarios sin elementos de borde de tallo)

N U
 N N
 Y-V
 Y-N
 B-D
 N-N
 G-C
 N-N (estructura de tallo-asa)

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

5

fórmula (If): (secuencia consenso de tallo-lazo de histona de metazoarios con elementos de borde de tallo)

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

N*N*NNNN-NNNN*N*N* (estructura de tallo-asa)

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

fórmula (Ig): (secuencia de consenso de tallo-lazo de histona de vertebrados sin elementos de borde de tallo)

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

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

fórmula (IIg): (secuencia de consenso de tallo-lazo de histona de vertebrados con elementos de borde de tallo)

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

N*N*HNNN-NNNN*N*H* (estructura de tallo-asa)

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

5

fórmula (IIh): (secuencia de consenso de tallo-lazo de histona humana (*Homo sapiens*) sin elementos de borde de tallo)

Y U
D H
U-A
C-S
Y-R
H-R
G-C
D-C (estructura de tallo-asa)

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

10 fórmula (IIh): (secuencia de consenso de tallo-lazo de histona humana (*Homo sapiens*) con elementos de borde de tallo)

Y U
D H
U-A
C-S
Y-R
H-R
G-C

N*H*AAHD-CVHB*N*H* (estructura de tallo-asa)

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

donde en cada una de las anteriores fórmulas (Ic) a (Ih) o (IIc) a (IIh):

N, C, G, A, T y U son como se definió arriba;

5 cada U puede reemplazarse por T;

cada G o C (altamente) conservado en los elementos de tallo 1 y 2 puede reemplazarse por su base nucleótido C o G complementaria, siempre y cuando este nucleótido complementario en el tallo correspondiente sea reemplazado por su nucleótido complementario en paralelo; y/o

10 G, A, T, U, C, R, Y, M, K, S, W, H, B, V, D y N son bases de nucleótidos como las definidas en la siguiente tabla:

Abreviatura	Bases de nucleótido	Aclaración
G	G	Guanina
A	A	Adenina
T	T	Timina
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 (enlaces 2H)
H	A o C o T/U	No G
B	G o T/U o C	No A
V	G o C o A	No T/U
D	G o A o T/U	No C
N	G o C o T/U o A	Cualquier base
*	Presente o no	La base puede estar presente o no

15 En este contexto es particularmente preferente que la secuencia de tallo-lazo de histona de acuerdo con al menos una de las fórmulas (I) o (Ia) a (Ih) o (II) o (IIa) a (IIh) empleada según la presente invención se seleccione de una secuencia de tallo y lazo de histona de origen natural, en particular de una secuencia tallo-lazo de histona de protozoarios o metazoarios y en especial de una secuencia tallo-lazo de histona de vertebrados, con total preferencia de mamíferos, especialmente secuencias tallo-lazo de histona humana.

20 De acuerdo con un aspecto particularmente preferente de la primera realización, la secuencia tallo-lazo de histona de acuerdo con al menos una de las siguientes fórmulas específicas (I) o (Ia) a (Ih) o (II) o (IIa) a (IIh) empleadas según la presente invención es una secuencia tallo-lazo de histona que comprende en cada posición de nucleótido el nucleótido frecuencia de ocurrencia mayor, o tanto el nucleótido de mayor ocurrencia como el segundo de mayor frecuencia de ocurrencia de secuencias tallo-lazo de histona de origen natural en metazoarios y protozoarios (Fig. 5) como se muestra en las figura 1-5. En este contexto se prefiere particularmente que al menos el 80%, de preferencia al menos el 85% o muy preferiblemente al menos el 90% de todos los nucleótidos correspondan al nucleótido que ocurra más frecuentemente en las secuencias tallo-lazo de histona de origen natural.

25

En un aspecto particular adicional de la primera realización, la secuencia tallo-lazo de histona de acuerdo con al menos una de las fórmulas específicas (I) o (Ia) a (Ih) empleada según la presente invención se selecciona

de las siguientes secuencias tallo-lazo de histona (sin elementos de borde de tallo) que representa secuencias de tallo-lazo de histona como las generadas de acuerdo con el ejemplo 1:

- 5 VGYYYHHHTRWRCB (SEQ ID NO: 13 de acuerdo a la fórmula (Ic))
 SGYYTTYTMARRRCS (SEQ ID NO: 14 de acuerdo a la fórmula (Ic))
 SGYYCTTTMAGRRCS (SEQ ID NO: 15 de acuerdo a la fórmula (Ic))
- DGNNBNNTHVNNCH (SEQ ID NO: 16 de acuerdo a la fórmula (Ie))
 RGNNNYHBTHRDNNCY (SEQ ID NO: 17 de acuerdo a la fórmula (Ie))
 RGNDBYHYTHRDHNCY (SEQ ID NO: 18 de acuerdo a la fórmula (Ie))
- 10 VGYYYHHRVRCB (SEQ ID NO: 19 de acuerdo a la fórmula (If))
 SGYYCTTYMAGRRCS (SEQ ID NO: 20 de acuerdo a la fórmula (If))
 SGYYCTTTMAGRRCS (SEQ ID NO: 21 de acuerdo a la fórmula (If))
- GGYYCTTYTHAGRCC (SEQ ID NO: 22 de acuerdo a la fórmula (Ig))
 GGCYCTTYMAGRGCC (SEQ ID NO: 23 de acuerdo a la fórmula (Ig))
 15 GGCTCTTTMAGRGCC (SEQ ID NO: 24 de acuerdo a la fórmula (Ig))
- DGHYCTDYTHASRRCC (SEQ ID NO: 25 de acuerdo a la fórmula (Ih))
 GGCYCTTTTHAGRCC (SEQ ID NO: 26 de acuerdo a la fórmula (Ih))
 GGCYCTTTMAGRGCC (SEQ ID NO: 27 de acuerdo a la fórmula (Ih))

Además, en este contexto las siguientes secuencias tallo-lazo de histona (con elementos de borde de tallo) generadas de acuerdo con el ejemplo 1 de acuerdo con una de las fórmulas específicas (II) o (IIa) a (IIh) son particularmente preferentes:

- H*H*HHVVGYYHHHTRVRCBVHH*N*N* (SEQ ID NO: 28 según la fórmula. (IIc))
 M*H*MHMSGYYTTYTMARRRCSMCH*H*H* (SEQ ID NO: 29 según la fórm. (IIc))
 M*M*MMMSGYYCTTTMAGRRCSACH*M*H* (SEQ ID NO: 30 según la fórm. (IIc))
- 25 N*N*NNNDGNNBNNTHVNNCHNHN*N*N* (SEQ ID NO: 31 según la fórm. (IIe))
 N*N*HHNRGNNNYHBTHRDNNCYDHH*N*N* (SEQ ID NO: 32 según la fórmula (IIe))
 N*H*HHVRGNDBYHYTHRDHNCYRHH*H*H* (SEQ ID NO: 33 según la fórmula (IIe))
- H*H*MHMVGYYTYHTRVRCBVMH*H*N* (SEQ ID NO: 34 según la fórmula (IIf))
 M*M*MMMSGYYCTTYTMAGRRCSMCH*H*H* (SEQ ID NO: 35 según la fórmula (IIf))
 30 M*M*MMMSGYYCTTTMAGRRCSACH*M*H* (SEQ ID NO: 36 según la fórmula (IIf))
- H*H*MAMGGYCTTYTHAGRCCVHN*N*M* (SEQ ID NO: 37 según la fórmula (IIg))
 H*H*AAMGGCYCTTYMAGRGCCVCH*H*M* (SEQ ID NO: 38 según la fórmula (IIg))
 M*M*AAMGGCTCTTTMAGRGCCMCY*M*M* (SEQ ID NO: 39 según la fórmula (IIg))
- N*H*AAHDGHYCTDYTHASRRCCVHB*N*H* (SEQ ID NO: 40 según la fórmula (IIh))
 35 H*H*AAMGGCYCTTTTHAGRCCVMY*N*M* (SEQ ID NO: 41 según la fórmula (IIh))
 H*M*AAAGGYCTTTMAGRGCCRMV*H*M* (SEQ ID NO: 42 según la fórmula (IIh))

De acuerdo con otro aspecto preferente de la primera realización inventiva, la secuencia de ácido nucleico empleado comprende o codifica para al menos una secuencia tallo-lazo de histona que muestra al menos aproximadamente 80%, de preferencia al menos alrededor de 85%, muy preferiblemente al menos aproximadamente 90% o todavía más preferiblemente al menos alrededor de 95% de identidad de secuencia con los no al 100% nucleótidos conservados en las secuencias de tallo-lazo de histona de acuerdo con al menos una de las fórmulas específicas (I) o (Ia) a (Ih) o (II) o (IIa) o (IIh) o con una secuencia tallo-lazo de histona de origen natural y además comprende o codifica para una secuencia poli(A). Cuando está presente, esta secuencia poli(A) comprende una secuencia de alrededor de 25 a aproximadamente 400 nucleótidos adenosina, de preferencia una secuencia de alrededor de 50 a aproximadamente 400 nucleótidos adenosina, muy preferiblemente una secuencia de alrededor de 50 a aproximadamente 300 nucleótidos adenosina, aún más preferiblemente una secuencia de alrededor de 50 a aproximadamente 250 nucleótidos adenosina, todavía más preferiblemente una secuencia de alrededor de 60 a alrededor de 250 nucleótidos adenosina. En este contexto el término "aproximadamente" se refiere a una desviación de $\pm 10\%$ de los valores a los que se refiere.

40

45

50

Preferentemente, el uso de la invención puede estar basado en un ácido nucleico que no contiene uno o dos o al menos uno o todos pero uno o todos de los componentes del grupo que consiste 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-asa de histona, en particular una secuencia de procesamiento de tallo-asa de histona derivada de gen H2A614 de histona de ratón, un gen Neo, una secuencia promotora inactivada y una secuencia mejoradora inactivada. De forma especialmente preferente, el ácido nucleico usado de acuerdo con esta invención no contiene una ribozima, de preferencia una ribozima de auto-empalme, y uno del grupo consistente en: un gen Neo, una secuencia promotora inactivada, una secuencia mejoradora inactivada, una secuencia de procesamiento de tallo-asa de histona, en particular una secuencia de procesamiento de tallo-asa de histona derivada del gen H2A614 de histona de ratón. En consecuencia, el ácido nucleico usado, en un modo preferente, no puede contener ni una ribozima, de preferencia una ribozima de auto-empalme, ni un gen Neo o, como alternativa, ni una ribozima, de preferencia una ribozima de auto-empalme, ni ningún gen de resistencia (por ejemplo, aplicado normalmente para selección). En otra modalidad preferida, el ácido nucleico usado de acuerdo con la invención no puede ni contener una ribozima, de preferencia una ribozima de auto-empalme, ni una señal de procesamiento de tallo-lazo de histona, en particular una secuencia de procesamiento de tallo-lazo de histona derivada de gen H2A614 de histona de ratón.

Alternativamente, de acuerdo con la primera realización de la presente invención, la secuencia de ácido nucleico usada de acuerdo con la invención comprende opcionalmente una señal de poliadenilación, que se define aquí como una señal que transmite poliadenilación a un ARNm (transcrito) por factores de proteína específicos (por ejemplo, factor de especificidad de corte y poliadenilación (CPSF), factor de estimulación de corte (CstF), factores de corte I y II (CF I y CF II), poli(A) polimerasa (PAP)). En este contexto se prefiere una señal de poliadenilación de consenso que comprende la secuencia de consenso NNUANA. En un aspecto preferido particular, la señal de poliadenilación comprende una de las siguientes secuencias: AAUAAA o AUUAAA.

La secuencia de ácido nucleico para su uso de acuerdo con la primera realización de la presente invención codifica además para una proteína o un péptido, que se selecciona de proteínas o péptidos terapéuticamente activos, incluyendo proteínas adyuvantes, antígenos, antígenos tumorales, antígenos patógenos (por ejemplo, seleccionados de antígenos animales, virales, protozoarios, bacterianos), antígenos alergénicos, antígenos autoinmunes, de alérgenos, de anticuerpos, de proteínas o péptidos inmunoestimuladores, y de receptores de células T específicos de antígenos, donde el ácido nucleico usado puede ser transportado en una célula (por ejemplo una célula huésped de expresión o una célula somática), un tejido o un organismo y la proteína puede ser expresada posteriormente en esta célula, tejido u organismo.

La región de codificación del ácido nucleico de acuerdo con la primera realización de la presente invención se puede presentar como un ácido nucleico mono-, di- o incluso multicistrónico, es decir un ácido nucleico que porte las secuencias de codificación de uno, dos o más proteínas o péptidos. Estas secuencias de codificación en ácidos nucleicos di- o incluso multicistrónicos pueden estar separadas por al menos una secuencia de sitio de entrada de ribosoma interna (IRES), por ejemplo como la aquí definida o por péptidos de señal que inducen el corte del polipéptido resultante que comprende varias proteínas o péptidos.

En aspectos particularmente preferidos de la primera realización de la presente invención, los péptidos o proteínas codificados se seleccionan de entre proteínas o péptidos humanos, virales, bacterianos y protozoarios.

En el contexto de la presente invención, las proteínas terapéuticamente activas, codificadas por la molécula de ácido nucleico usada se pueden seleccionar, sin estar restringidas a las mismas, de proteínas que tengan un efecto curativo, prevengan profilácticamente o traten terapéuticamente una enfermedad, preferiblemente como la aquí definida, o sean proteínas de las cuales un individuo tenga necesidad. Estas proteínas se pueden seleccionar de cualquier proteína recombinante o aislada que se diseñe sintéticamente o de origen natural conocida por el experto en la técnica. Sin limitarse a las mismas, las proteínas terapéuticamente activas pueden comprender proteínas capaces de estimular o inhibir la transducción de señales en la célula, por ejemplo citoquinas, linfoquinas, monoquinas, factores de crecimiento, receptores, moléculas de transducción de señales, factores de transcripción, etc.; anticoagulantes; antitrombinas; proteínas antialérgicas; factores apoptóticos o proteínas relacionadas con la apoptosis, enzimas terapéuticas activas y cualquier proteína relacionada con cualquier enfermedad adquirida o hereditaria.

De preferencia, una proteína terapéuticamente activa, que puede ser codificada por la molécula de ácido nucleico empleada, también puede ser una proteína adyuvante. En este contexto, una proteína adyuvante se debe entender preferentemente como cualquier proteína capaz de desarrollar una respuesta inmune innata como la aquí definida. Preferiblemente, esta respuesta inmune innata comprende la activación de un receptor de reconocimiento de patrones, tal como un receptor seleccionado de la familia de los receptores tipo Toll (TLR), incluyendo por ejemplo un receptor tipo Toll seleccionado de TLR1 a TLR10 humano o de receptores tipo Toll murinos TLR1 a TLR13. Más preferiblemente, la proteína adyuvante se selecciona de proteínas adyuvantes humanas o de proteínas adyuvantes patógenas seleccionadas del grupo consistente en, sin

limitarse a, proteínas bacterianas, proteínas protozoarias, proteínas virales o proteínas fúngicas, proteínas animales, en particular proteínas adyuvantes bacterianas. Además, también pueden usarse ácidos nucleicos que codifican para proteínas humanas implicadas en efectos adyuvantes (por ejemplo ligandos de receptores de reconocimiento de patrones, receptores de reconocimiento de patrones, proteínas de las vías de transducción de señales, factores de transcripción o citoquinas).

La molécula de ácido nucleico para su uso de acuerdo con la invención puede alternativamente codificar para un antígeno. De acuerdo con la presente invención, el término "antígeno" se refiere a una sustancia que es reconocida por el sistema inmunológico y es capaz de desencadenar una respuesta inmunológica específica de antígenos, por ejemplo mediante la formación de anticuerpos o células T específicas de antígenos, como parte de la respuesta inmunológica adaptativa. En este contexto, un epítipo antigénico, fragmento o péptido de una proteína significa particularmente epítipos de células B y células T que pueden ser reconocidos por células B, anticuerpos o células T respectivamente.

En el contexto de la presente invención, los antígenos que pueden ser codificados mediante el uso de la molécula de ácido nucleico comprenden típicamente cualquier antígeno, epítipo antigénico, fragmento antigénico o péptido antigénico que esté bajo la anterior definición, preferiblemente antígenos de proteínas y péptidos, por ejemplo antígenos tumorales, antígenos alergénicos o alérgenos, auto-antígenos autoinmunes, antígenos patógenos, etc.

En particular los antígenos como los codificados por la molécula de ácido nucleico usada según la invención pueden ser antígenos generados fuera de la célula, típicamente antígenos no derivados del propio organismo anfitrión (por ejemplo un humano) (es decir no auto-antígenos), así como derivados de células huésped fuera del organismo anfitrión, por ejemplo antígenos virales, antígenos bacterianos, antígenos fúngicos, antígenos protozoológicos, antígenos animales, antígenos alergénicos, etc. Los antígenos alergénicos (antígenos de alergias o alérgenos) son típicamente antígenos que causan una alergia en un humano y pueden derivarse ya sea de un humano o de otras fuentes. Además, los antígenos como los codificados por la molécula de ácido nucleico usada pueden ser antígenos generados dentro de la célula, el tejido o el cuerpo. Estos antígenos incluyen antígenos derivados del propio organismo anfitrión (por ejemplo humano), por ejemplo antígenos tumorales, antígenos propios o auto-antígenos, tales como antígenos propios autoinmunes, pero también (no auto) antígenos como los definidos aquí, los cuales se han derivado originalmente de células huésped fuera del organismo anfitrión, pero han sido fragmentados o degradados dentro del cuerpo, tejido o célula, por ejemplo mediante degradación (proteasas), metabolismo, etc.

Una clase de antígenos que pueden ser codificados por la molécula de ácido nucleico usado de acuerdo con la invención comprende antígenos tumorales. Los "antígenos tumorales" se ubican preferentemente sobre la superficie de la célula (tumor). Los antígenos tumorales también se pueden seleccionar de proteínas que están sobre-expresadas en células tumorales en comparación con una célula normal. Además, los antígenos tumorales también incluyen antígenos expresados en células que a su vez no son (fueron) (u originalmente no ellas mismas) degeneradas sino que están asociadas con el tumor supuesto. Los antígenos que están relacionados con los vasos de suministro al tumor o la reformación de los mismos, en particular aquellos antígenos que están asociados con neovascularización, por ejemplo factores de crecimiento, tales como VEGF, bFGF, etc., también se incluyen aquí. Los antígenos relacionados con un tumor incluyen además antígenos de células o tejidos que típicamente están embebidos en un tumor. Además, algunas sustancias (normalmente proteínas o péptidos) son expresados en pacientes que sufren (con o sin conocimiento) de una enfermedad cancerosa y se presentan en concentraciones incrementadas en los fluidos corporales de dichos pacientes. Estas sustancias también se conocen como "antígenos tumorales", sin embargo no son antígenos en el estricto significado de una sustancia inductora de una respuesta inmunológica. La clase de antígenos tumorales puede dividirse además en antígenos específicos de tumor (TSAs) y antígenos asociados a tumor (TAAs). Los TSAs sólo pueden estar presentes en células tumorales y nunca en células "sanas" normales. Típicamente resultan de una mutación específica de tumor. Los TAAs, más comunes, están presentes normalmente tanto en células tumorales como sanas. Estos antígenos son reconocidos y la célula presentadora de antígenos puede ser destruida por células T citotóxicas. Además, los antígenos tumorales también se pueden presentar sobre la superficie del tumor en forma de, por ejemplo, un receptor mutado. En este caso, pueden ser reconocidos por anticuerpos.

De acuerdo con otra alternativa, una clase más de antígenos que pueden ser codificados por la molécula de ácido nucleico usada comprende antígenos alergénicos. Estos antígenos alergénicos se pueden seleccionar de antígenos derivados de diferentes fuentes, por ejemplo de animales, plantas, hongos, bacterias, etc. Los alérgenos en este contexto incluyen por ejemplo hierbas, pólenes, mohos, fármacos o numerosos activadores ambientales, etc. Los antígenos alergénicos pertenecen típicamente a diferentes clases de compuestos, tales como ácidos nucleicos y sus fragmentos, proteínas o péptidos y sus fragmentos, carbohidratos, polisacáridos, azúcares, lípidos, fosfolípidos, etc. De particular interés en el contexto de la presente invención son los antígenos que pueden ser codificados por la molécula de ácido nucleico usada tal como se define aquí, es decir antígenos de proteínas o péptidos y sus fragmentos o epítipos, o ácidos nucleicos y sus fragmentos,

particularmente ácidos nucleicos y sus fragmentos, que codifican para estos antígenos de proteínas o péptidos y sus fragmentos o epítomos.

De acuerdo con otra alternativa, la molécula de ácido nucleico usada de acuerdo con la invención puede codificar para un anticuerpo o un fragmento de anticuerpo. Según la presente invención, este anticuerpo se puede seleccionar de cualquier anticuerpo, por ejemplo cualquier anticuerpo de origen natural o producido de forma recombinante, conocido en la técnica, en particular anticuerpos adecuados para propósitos terapéuticos, de diagnóstico o científicos, o anticuerpos que han sido identificados con relación a enfermedades cancerosas específicas. Aquí, el término "anticuerpo" se usa en su sentido más amplio y abarca específicamente anticuerpos monoclonales y policlonales (incluyendo anticuerpos agonistas, antagonistas y de bloqueo o neutralización) y especies de anticuerpos con especificidad poliepitópica. De acuerdo con la invención, el término "anticuerpo" comprende típicamente cualquier anticuerpo conocido en la técnica (por ejemplo, anticuerpos IgM, IgD, IgG, IgA e IgE), tales como anticuerpos de origen natural, anticuerpos generados por inmunización en un organismo anfitrión, anticuerpos que fueron aislados e identificados a partir de anticuerpos de origen natural o anticuerpos generados por inmunización en un organismo anfitrión y producidos recombinantemente por métodos biomoleculares conocidos en la técnica, así como anticuerpos quiméricos, anticuerpos humanos, anticuerpos humanizados, anticuerpos biespecíficos, intracuerpos, es decir anticuerpos expresados en células y ubicados opcionalmente en compartimientos celulares específicos, y fragmentos y variantes de los anticuerpos mencionados. En general, un anticuerpo consiste en una cadena ligera y una cadena pesada ambas con dominios variables y constantes. La cadena ligera consiste en un dominio variable N-terminal, V_L , y un dominio constante C-terminal, C_L . Por el contrario, la cadena pesada del anticuerpo IgG, por ejemplo, comprende un dominio variable N-terminal, V_H , y tres dominios constantes, C_{H1} , C_{H2} y C_{H3} .

En el contexto de la presente invención, los anticuerpos codificados por la molécula de ácido nucleico de la invención pueden comprender, de preferencia, anticuerpos de longitud completa, es decir anticuerpos compuestos de las cadenas pesada y ligera completas, como se describió arriba. Sin embargo, también pueden ser codificados por la molécula de ácido nucleico usada según la invención derivados de anticuerpos, tales como fragmentos de anticuerpo, variantes o aductos. Preferentemente, los fragmentos de anticuerpo se seleccionan de fragmentos Fab, Fab', F(ab')₂, Fc, Facb, pFc', Fd y Fv de los anticuerpos (de longitud completa) mencionados arriba. En general, los fragmentos de anticuerpo son conocidos en la técnica. Por ejemplo, un fragmento Fab ("fragmento, unión a antígeno") está compuesto por un dominio constante y uno variable de cada una de las cadenas pesada y ligera. Los dos dominios variables se unen al epítomo en antígenos diferentes. Las dos cadenas están conectadas por un enlace disulfuro. Un fragmento scFv ("fragmento variable de cadena individual") consiste típicamente en los dominios variables de las cadenas ligera y pesada. Los dominios están enlazados por un enlace artificial, en general un enlace polipéptido, tal como un péptido compuesto de 15-25 residuos de glicina, prolina y/o serina.

En el presente contexto es preferible que las diferentes cadenas del anticuerpo o fragmento de anticuerpo sean codificadas por una molécula de ácido nucleico multicistrónica. Alternativamente, las diferentes cepas del anticuerpo o fragmento de anticuerpo son codificadas por varias (secuencias) de ácido nucleico monocistrónicas.

De acuerdo con la primera realización de la presente invención, la secuencia de ácido nucleico usada según la invención comprende una región de codificación, preferentemente que codifica para un péptido o proteína terapéuticamente activo, una proteína adyuvante, un antígeno, un antígeno tumoral, un antígeno patogénico, un antígeno animal, un antígeno viral, un antígeno protozoal, un antígeno bacteriano, un antígeno alérgico, un antígeno autoinmune, un alérgeno, un anticuerpo, un péptido o proteína inmunoestimulador o un antígeno específico del receptor de células T. La proteína codificada no es una proteína histona. En el contexto de la presente invención, esta proteína histona típicamente es una proteína fuertemente alcalina que se encuentra en los núcleos de las células eucarióticas, que empaquetan y ordenan al ADN en unidades estructurales llamadas nucleosomas. Las proteínas de histona son los componentes proteínicos claves de la cromatina, actúan como carretes alrededor de los cuales se devana el ADN y juegan un papel en la regulación génica. Sin histonas, el ADN desdevanado en cromosomas sería muy largo (una relación longitud:ancho de más de 10 millones a uno en ADN humano). Por ejemplo, cada célula humana tiene aproximadamente 1,8 metros de ADN, pero devanados en las histonas tiene aproximadamente 90 milímetros de cromatina, la cual, cuando se duplica y condensa durante la mitosis, se traduce en aproximadamente 120 micras de cromosomas. Muy preferiblemente, en el contexto de la presente invención esta proteína de histona se define típicamente como una proteína altamente conservada seleccionada de una de las siguientes cinco clases principales de histonas: H1/H5, H2A, H2B, H3 y H4", de preferencia seleccionadas de histona de mamífero, muy preferiblemente de histonas de humanos o proteínas de histona. Estas histonas o proteínas de histona se organizan típicamente en dos súper clases definidas como histonas de núcleo, que comprenden las histonas H2A, H2B, H3 y H4, e histonas enlazadoras, que comprenden las histonas H1 y H5.

En este contexto, las histonas enlazadoras, preferentemente excluidas del alcance de protección de la invención, de preferencia histonas enlazadoras de mamífero, muy preferiblemente histonas enlazadoras de humanos, se seleccionan típicamente de H1, incluyendo H1F, incluyendo particularmente H1F0, H1FNT, H1FOO, H1FX y H1H1, incluyendo particularmente HIST1H1A, HIST1H1B, HIST1H1C, HIST1H1D, HIST1H1E, HIST1H1T.

5

Además, las histonas nucleares, preferentemente excluidas del alcance de protección de la invención, de preferencia histonas de núcleo de mamífero, muy preferiblemente histonas nucleares de humano, se seleccionan típicamente de H2A, incluyendo H2AF, incluyendo particularmente H2AFB1, H2AFB2, H2AFB3, H2AFJ, H2AFV, H2AFX, H2AFY, H2AFY2, H2AFZ y H2A1, incluyendo particularmente HIST1H2AA, HIST1H2AB, HIST1H2AC, HIST1H2AD, HIST1H2AE, HIST1H2AG, HIST1H2AI, HIST1H2AJ, HIST1H2AK, HIST1H2AL, HIST1H2AM y H2A2, incluyendo particularmente HIST2HAA3, HIST2H2AC; H2B, incluyendo H2BF, incluyendo particularmente H2BFM, H2BFO, H2BFS, H2BFWT H2B1, incluyendo particularmente HIST1H2BA, HIST1H2BB, HIST1H2BC, HIST1H2BD, HIST1H2BE, HIST1H2BF, HIST1H2BG, HIST1H2BH, HIST1H2BI, HIST1H2BJ, HIST1H2BK, HIST1H2BL, HIST1H2BM, HIST1H2BN, HIST1H2BO y H2B2, incluyendo particularmente HIST2H2BE; H3, incluyendo H3A1, incluyendo particularmente HIST1H3A, HIST1H3B, HIST1H3C, HIST1H3D, HIST1H3E, HIST1H3F, HIST1H3G, HIST1H3H, HIST1H3I, HIST1H3J y H3A2, incluyendo particularmente HIST2H3C, y H3A3, incluyendo particularmente HIST3H3; H4, incluyendo H41, incluyendo particularmente HIST1H4A, HIST1H4B, HIST1H4C, HIST1H4D, HIST1H4E, HIST1H4F, HIST1H4G, HISTH4H, HIST1H4I, HIST1H4J, HIST1H4K, HIST1H4L y H44, incluyendo particularmente HIST4H4 y H5.

20

Preferentemente, la secuencia de ácido nucleico de la invención no contiene una secuencia génica (bacteriana) Neo (gen de resistencia a la neomicina).

El ácido nucleico usado como se define aquí comprende o codificada para a) una región de codificación, que codifica para un péptido o proteína terapéuticamente activo, una proteína adyuvante, un antígeno, un antígeno tumoral, un antígeno patogénico, un antígeno animal, un antígeno viral, un antígeno protozoal, un antígeno bacteriano, un antígeno alérgico, un antígeno autoinmune, un alérgeno, un anticuerpo, un péptido o proteína inmunoestimulador o un antígeno específico del receptor de células T; b) al menos un tallo-lazo de histona y c) una secuencia poli(A) o señal de poliadenilación; para su uso con el fin de incrementar el nivel de expresión de dicha proteína codificada, donde dicha proteína codificada no es una proteína de histona, no es una proteína reportera y/o no es una proteína marcadora o de selección como las definidas arriba. Los elementos b) a c) del ácido nucleico usado pueden estar presentes en el ácido nucleico en cualquier orden, es decir, los elementos a), b) e c) pueden presentarse en el orden a), b) y c) o a), c) y b) de la dirección 5' a 3' en la secuencia de ácido nucleico usado, donde los elementos adicionales aquí descritos también pueden estar contenidos, tales 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 usado, particularmente a) en constructos di- o multicistronicos y/o cada uno de los elementos b) y c), muy preferiblemente el elemento b) también puede repetirse al menos una vez, de preferencia dos veces o más en el ácido nucleico usado. Como un ejemplo, el ácido nucleico usado puede mostrar sus elementos de secuencia a), b) y c) por ejemplo en el siguiente orden:

25

30

35

40

45

50

- 5'-región de codificación-tallo-lazo de histona-3'; o
- 5'-región de codificación-región de codificación-tallo-lazo de histona-3'; o
- 5'-región de codificación-IRES-región de codificación-tallo-lazo de histona-3'; o
- 5'-región de codificación-tallo-lazo de histona-secuencia poli(A)-3'; o
- 5'-región de codificación-tallo-lazo de histona-señal de poliadenilación-3'; o
- 5'-región de codificación-región de codificación-tallo-lazo de histona-señal de poliadenilación-3'; o
- 5'-región de codificación-tallo-lazo de histona-tallo-lazo de histona-3'; o
- 5'-región de codificación-tallo-lazo de histona-tallo-lazo de histona-secuencia poli(A)-3'; o
- 5'-región de codificación-tallo-lazo de histona-tallo-lazo de histona-señal de poliadenilación-3'; o
- 5'-región de codificación-tallo-lazo de histona-secuencia poli(A)-tallo-lazo de histona-3'; o
- 5'-región de codificación-secuencia poli(A)-tallo-lazo de histona-3'; o
- 5'-región de codificación-secuencia poli(A)-tallo-lazo de histona-tallo-lazo de histona-3'; etc.

55

En este contexto se prefiere particularmente que la molécula de ácido nucleico usada comprenda o codifique para a) una región de codificación, que codifique para un péptido o proteína; b) al menos un tallo-lazo de histona y c) una secuencia poli(A) o secuencia de poliadenilación; para su uso con el fin de incrementar el nivel de expresión de una proteína codificada, donde la proteína codificada no es una proteína de histona, no es una proteína reportera seleccionada entre luciferasa y EGFP (y por ejemplo β -galactosidasa), y/o no es proteína marcadora o de selección seleccionada entre alfa-globina, galactoquinasa y xantina:guanina-fosforribosil-transferasa (GPT).

60

En otro aspecto preferente de la primera realización, la molécula de ácido nucleico usada como se define aquí también puede estar en forma de un ácido nucleico modificado.

De acuerdo con un aspecto de la primera realización, la molécula de ácido nucleico de la invención como se define aquí puede proporcionarse como un “ácido nucleico estabilizado”, de preferencia como un ARN estabilizado, en especial como un ARN que es esencialmente resistente a degradación *in vivo* (por ejemplo, por una exo o endo-nucleasa).

5 En este contexto, la molécula de ácido nucleico usada como se define aquí puede contener análogos/modificaciones de nucleótidos, por ejemplo modificaciones de esqueleto, modificaciones de azúcar o modificaciones de bases. Una modificación de esqueleto en relación con la presente invención es una modificación en la que los fosfatos del esqueleto de los nucleótidos contenidos en la molécula de ácido nucleico de la invención como se define aquí se modifican químicamente. Una modificación de azúcar en
10 relación con la presente invención es una modificación química del azúcar de los nucleótidos de la molécula de ácido nucleico usada como se define aquí. Además, una modificación de base en relación con la presente invención es una modificación química de la parte de base de los nucleótidos de la molécula de ácido nucleico usada. En este contexto, los análogos o modificaciones de nucleótidos se seleccionan preferentemente de análogos de nucleótidos que son aplicables para la transcripción y/o la traducción.

15 En un aspecto preferente particular de la primera realización de la presente invención, los análogos/modificaciones de nucleótidos definidos aquí se seleccionan de modificaciones de bases que incrementan además la expresión de la proteína codificada y que se seleccionan preferentemente de entre 2-amino-6-cloropurínribósido-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-yodocitidin-5'-trifosfato, 5-yodouridin-5'-trifosfato, 5-metilcitidin-5'-trifosfato, 5'-metiluridin-5'-trifosfato, 6-azacitidin-5'-trifosfato, 6-azauridin-5'-trifosfato, 5'-metiluridin-5'-trifosfato, 6-azacitidin-5'-trifosfato, 6-azauridin-5'-trifosfato, 6-cloropurínribósido-5'-trifosfato, 7-desazaadenosin-5'-trifosfato, 7-desazaguanosin-5'-trifosfato, 8-azaadenosin-5'-trifosfato, 8-azoadenosin-5'-trifosfato, bencimidazol-ribosido-5'-trifosfato, N1-metiladenosin-5'-trifosfato, N1-metilguanosin-5'-trifosfato, N6-metiladenosin-5'-trifosfato, O6-metilguanosin-5'-trifosfato, pseudouridin-5'-trifosfato, o puomicin-5'-trifosfato, xantosin-5'-trifosfato. Se da particular preferencia a nucleótidos para modificaciones de bases seleccionados del grupo de nucleótidos modificados en base que consisten en 5-metilcitidin-5'-trifosfato, 7-desazaguanosin-5'-trifosfato, 5-bromocitidin-5'-trifosfato y pseudouridin-5'-trifosfato.

20 De acuerdo con otro aspecto, la molécula de ácido nucleico usada como se define aquí puede contener una modificación de lípidos. Este ácido nucleico modificado en lípidos comprende típicamente un ácido nucleico como el aquí definido. Esta molécula de ácido nucleico modificada en lípidos de la molécula de ácido nucleico usada como se define aquí comprende típicamente además al menos un enlazador unido covalentemente a esa molécula de ácido nucleico y al menos un lípido enlazado covalentemente con el enlazador respectivo. Alternativamente, la molécula de ácido nucleico modificada por lípidos comprende al menos una molécula de
35 ácido nucleico como la definida aquí y al menos un lípido (bifuncional) enlazado covalentemente (sin un enlazador) a esa molécula de ácido nucleico. De acuerdo con una tercera alternativa, la molécula de ácido nucleico modificada por lípidos comprende una molécula de ácido nucleico como la definida aquí, al menos un enlazador unido covalentemente con la molécula de ácido nucleico y al menos un lípido enlazado covalentemente con el enlazador respectivo, así como también al menos un lípido (bifuncional) enlazado
40 covalentemente (sin un enlazador) con esa molécula de ácido nucleico. En este contexto, se prefiere particularmente que la modificación de lípido esté presente en los extremos terminales de una secuencia de ácido nucleico lineal como se define aquí.

45 De acuerdo con otro aspecto preferente de la primera realización de la invención, la molécula de ácido nucleico usado como se define aquí, particularmente si se proporciona como un ARN(m), puede por lo tanto estar estabilizada contra la degradación por ARNasas mediante la adición de una llamada estructura “5' CAP”.

De acuerdo con otro aspecto preferente de la primera realización de la invención, la molécula de ácido nucleico usada como se define aquí puede estar modificada con una secuencia de al menos 10 citidinas, de preferencia al menos 20 citidinas, muy preferiblemente al menos 30 citidinas (la llamada “secuencia poli(C)”). Particularmente, la molécula de ácido nucleico usada puede contener o codificar para una secuencia poli(C) típicamente de alrededor de 10 a 200 nucleótidos de citidina, de preferencia aproximadamente 10 a 100 nucleótidos de citidina, muy preferiblemente alrededor de 10 a 70 nucleótidos de citidina o todavía más preferiblemente alrededor de 20 a 50 o incluso 20 a 30 nucleótidos de citidina. Esta secuencia poli(C) se ubica preferentemente en 3' de la región de codificación comprendida en el ácido nucleico usado de acuerdo
55 con la primera realización de la presente invención.

De acuerdo con otro aspecto preferente de la primera realización de la invención, la molécula de ácido nucleico usada como se define aquí preferentemente tiene 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 incrementar la vida media del ácido nucleico en el citosol. Estas secuencias estabilizadoras pueden tener un 100% de

identidad de secuencia con secuencias de origen natural que se presentan en virus, bacterias y eucariontes, pero también pueden ser parcial o completamente sintéticas. Las secuencias no traducidas (UTR) del gen de (alfa)-globina, por ejemplo de *Homo sapiens* o *Xenopus Laevis*, pueden mencionarse como ejemplos de secuencias estabilizadoras que pueden usarse en la presente invención para un ácido nucleico estabilizado.

5 Otro ejemplo de secuencia estabilizadora tiene la fórmula general (C/U)CCAN,CCC(U/A)Py_xJC(C/U)CC (SEQ ID NO: 55), la cual está contenida en las 3'-UTRs de las moléculas de ARN muy estables que codifican para (alfa)-globina, colágeno tipo(I), 15-lipooxigenasa o para tirosina hidroxilasa (véase Holcik y col., Proc. Natl. Acad. Sci. E.U.A. 1997, 94: 2410 a 2414). Estas secuencias estabilizadoras pueden por supuesto usarse individualmente o en combinación unas con otras y también en combinación con otras secuencias

10 estabilizadoras conocidas por el experto en la técnica. En este contexto se prefiere particularmente que la secuencia 3' UTR del gen de alfa-globina se ubique 3' de la secuencia de codificación comprendida en el ácido nucleico usado de acuerdo con la primera realización de la presente invención.

Preferentemente, las sustituciones, adiciones o eliminaciones de bases se realizan en la molécula de ácido nucleico usada como se define aquí, empleando una matriz de ADN para preparar la molécula de ácido nucleico mediante técnicas de mutagénesis dirigida a sitio, bien conocida, o mediante una estrategia de ligación de oligonucleótidos (véase, por ejemplo, Maniatis y col., Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, 3ª edición, Cold Spring Harbor, NY, 2001). En ese proceso, para preparar la molécula de ácido nucleico usada como se define aquí, especialmente si el ácido nucleico está en forma de ARNm, una molécula de ADN correspondiente puede ser transcrita *in vitro*. Esta matriz de ADN

20 comprende preferentemente 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, a preparar 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 puede prepararse por proliferación fermentativa y posterior aislamiento como parte de un plásmido que puede ser replicado en bacterias. Plásmidos que pueden citarse como adecuados para la presente invención son, por ejemplo, los plásmidos pT7Ts (número de registro GenBank U26404; Lai y col., Development 1995, 121: 2349 a 2360), la serie pGEM[®], por ejemplo pGEM[®]-1 (número de registro GenBank X65300; de Promega) y pSP64 (número de registro GenBank X65327); véase también Mezei y Storts, Purification of PCR Products, en: Griffin and Griffin (ed.), PCR Technology: Current Innovation, CRC Press, Boca Ratón, FL, 2001.

25

30 Las moléculas de ácido nucleico usadas de acuerdo con la presente invención como se define aquí pueden prepararse por cualquier método conocido en la técnica, incluyendo métodos sintéticos tales como, por ejemplo, síntesis en fase sólida, así como métodos *in vitro*, tales como reacciones de transcripción *in vitro* o reacciones *in vivo*, tal como propagación *in vivo* de plásmidos de ADN en bacterias.

Cualquiera de las modificaciones anteriores puede aplicarse a la molécula de ácido nucleico usada como se define aquí y además a cualquier ácido nucleico tal como se emplean en el contexto de la presente invención y pueden ser, si es adecuado o necesario, combinadas entre sí en cualquier combinación, siempre y cuando estas combinaciones de modificaciones no interfieran unas con otras en el ácido nucleico relevante. El experto en la técnica será capaz de adoptar esta elección en consecuencia.

35

La molécula de ácido nucleico usada como se define aquí, así como las proteínas o péptidos como son codificados por esta molécula de ácido nucleico, pueden comprender fragmentos o variantes de esas secuencias. Esos fragmentos o variantes pueden comprender típicamente una secuencia que tenga una identidad de secuencia con uno de los ácidos nucleicos mencionados arriba, o con una de las proteínas o péptidos o secuencias, si son codificados por la al menos una molécula de ácido nucleico, de al menos un 5%, 10%, 20%, 30%, 40%, 50%, 60%, preferiblemente al menos 70%, muy preferiblemente al menos 80%,

40 especialmente al menos 85%, aún más preferiblemente al menos 90% y con total preferencia al menos un 95% o incluso 97%, 98% o 99%, con la secuencia tipo silvestre completa, ya sea a nivel de ácido nucleico o a nivel de aminoácidos.

45

En otro aspecto preferente de la primera realización de la presente invención la secuencia de ácido nucleico usado está asociada a un vehículo, agente de transfección o formación de complejos para incrementar la eficiencia de transfección de la secuencia de ácido nucleico de la invención. Agentes particularmente adecuados preferentes en este contexto, para incrementar 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 penetradores de células (CPPs), incluyendo péptidos de unión a VIH, VIH-1 Tat (VIH), péptidos derivados de Tat, penetratina, péptidos derivados de VP22 o análogos, VP22 de HSV (herpes simple), MAP, KALA o dominios de transducción de proteínas (PTDs), 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 antenapedia (particularmente de *Drosophila antennapedia*), pAntp, plsl, FGF, lactoferrina, Transportan, Buforin-2, Bac715-24, SynB, Synb(1), pVEC, péptidos derivados de hCT, SAP, o histonas. Además,

50

55

60 proteínas o péptidos catiónicos o policatiónicos también preferentes pueden seleccionarse de las siguientes

proteínas o péptidos que tienen la siguiente fórmula total: (Arg)_l; (Lys)_m; (His)_n; (Orn)_o; (Xaa)_x, en donde l + m + n + o + x = 8-15 y "l", "m", "n" u "o", independientemente unos de otros, pueden ser cualquier número seleccionado de 0, 1, 2, 3, 4,, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14 ó 15, siempre y 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 nativos (= de origen natural) o no nativos excepto Arg, Lys, His u Orn; y x puede ser cualquier número seleccionado de 0, 1, 2, 3 ó 4, siempre que el contenido total de Xaa no exceda el 50% de todos los aminoácidos del oligopéptido. Son péptidos catiónicos particularmente preferentes en este contexto, por ejemplo, Arg₇, Arg₈, Arg₉, H₃R₉, R₉H₃, H₃R₉H₃, YSSR₉SSY, (RKH)₄, Y(RKH)₂R, etc. Otros compuestos catiónicos o policatiónicos preferentes adicionales que se pueden usar 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 (dioctadecilamidoglicilcilespermina), DIMRI (bromuro de dimiristo-oxipropil-dimetilhidroxietyl-amonio), DOTAP (dioleiloxi-3-(trimetilamonio)propano), DC-6-14 (cloruro de O,O-ditetradecanoil-N-(α-trimetilamonioacetil)dietanolamina), CLIP1 (cloruro de rac-[(2,3-dioctadeciloxipropil)(2-hidroxietyl)-dimetilamonio), CLIP9 (rac-[2(2,3-dihexadeciloxipropil-oxisucciniloxi)etyl]trimetilamonio), oligofectamina o polímeros catiónicos o policatiónicos, por ejemplo poliaminoácidos modificados, tales como polímeros de β-aminoácidos o poliamidas inversas, etc., polietilenos modificados, tales como bromuro de PVP (bromuro de poli(N-etyl-4-vinilpiridinio)), etc., acrilatos modificados, tales como pDMAEMA (metacrilato de poli(dimetilaminoetyl), etc., amidoaminas modificadas tales como pAMAM (poli(amidoamina)), etc., poli-beta-aminoésteres modificados (PBAE), tal como polímeros de acrilato de 1,4-butanodiol-co-5-amino-1-pentanol modificados en el extremo con diamina, etc., dendrímeros, tales como dendrímeros de polipropilamina o dendrímeros a base de pAMAM, etc., poliiminas tales como PEI (poli(etilenimina), poli(propilenimina), etc.), polialilamina, polímeros a base de esqueleto de azúcar, tales como polímeros a base de ciclodextrina, polímeros a base de dextrano, quitosano, etc., polímeros a base de esqueleto de silano, tales como copolímeros PMOXA-PDMS, etc., polímeros en bloque que consisten en una combinación de uno o más bloques catiónicos (por ejemplo seleccionados de un polímero catiónico como el mencionado arriba) y de uno o más bloques hidrófilos o hidrófobos (por ejemplo, polietilenglicol); etc.

Preferentemente la secuencia de ácido nucleico usada de acuerdo con la invención se proporciona en forma desnuda o complejada, por ejemplo con compuestos policatiónicos de cualquier estructura química, de preferencia (poli)péptidos policatiónicos o compuestos policatiónicos sintéticos. Preferiblemente, la secuencia de ácido nucleico no se proporciona junto con una célula de empaque.

Además, de acuerdo con otra realización, la presente invención también se refiere al uso del ácido nucleico empleado como se define aquí para la preparación de una composición farmacéutica o de una composición farmacéutica para su uso con el fin de incrementar el nivel de expresión de una proteína/péptido codificado, para tratar una enfermedad tal como se define aquí, por ejemplo aplicar o administrar el ácido nucleico como se define aquí a una célula (por ejemplo una célula huésped de expresión o una célula somática), a un tejido o a un organismo, preferentemente en forma desnuda o complejada, o como una composición farmacéutica o vacuna tal como se describe aquí, en particular empleando cualquiera de las formas de administración aquí descritas.

Así, en una realización preferente particular, la presente invención también proporciona una composición farmacéutica que comprende el ácido nucleico como se define aquí y opcionalmente un portador y/o vehículo farmacéuticamente aceptable para el uso médico aquí definido.

Como un primer ingrediente, la composición farmacéutica de la invención comprende el ácido nucleico tal como se define aquí.

Como un segundo ingrediente, la composición farmacéutica de la invención puede comprender al menos un componente farmacéuticamente activo adicional. Un componente farmacéuticamente activo en este contexto es un compuesto que tiene un efecto terapéutico para sanar, mejorar o prevenir una indicación o enfermedad particular como las mencionadas aquí, preferentemente enfermedades cancerosas, enfermedades autoinmunes, alergias o enfermedades infecciosas, enfermedades cardiovasculares, enfermedades del sistema respiratorio, enfermedades del sistema digestivo, enfermedades de la piel, trastornos musculoesqueléticos, trastornos del tejido conectivo, neoplasmas, deficiencias inmunológicas, enfermedades endocrinas, nutricionales y metabólicas, enfermedades neurales, enfermedades oculares, enfermedades ópticas y enfermedades hereditarias. Estos compuestos incluyen, sin limitación, péptidos o proteínas, de preferencia como los definidos aquí, ácidos nucleicos, de preferencia como los definidos aquí, compuestos orgánicos o inorgánicos de bajo peso molecular (terapéuticamente activos) (peso molecular inferior a 5.000, de preferencia inferior a 1.000), azúcares, antígenos o anticuerpos, de preferencia como los definidos aquí, agentes terapéuticos ya conocidos en la técnica anterior, células antigénicas, fragmentos celulares antigénicos, fracciones celulares; componentes de pared celular (por ejemplo polisacáridos), patógenos

modificados, atenuados o desactivados (por ejemplo químicamente o por irradiación) (virus, bacterias, etc.), adyuvantes, de preferencia como los aquí definidos, etc.

Además, la composición farmacéutica usada puede comprender un portador y/o vehículo farmacéuticamente aceptable. En el contexto de la presente invención, un portador farmacéuticamente aceptable incluye típicamente una base líquida o no líquida de la composición farmacéutica de la invención. Si la composición farmacéutica empleada se proporciona en forma líquida, el portador típicamente será agua libre de pirógenos; solución salina isotónica o soluciones tamponadas (acuosas), por ejemplo soluciones tampón fosfato, citrato, etc. El tampón de inyección puede ser hipertónico, isotónico o hipotónico con relación al medio de referencia específico, es decir, el tampón puede tener un contenido en sal más alto, idéntico o más bajo con relación al medio de referencia específico, donde preferentemente las concentraciones de las sales mencionadas antes que pueden usarse no lleven al daño celular debido a ósmosis o a otros efectos de concentración. Medios de referencia son, por ejemplo, líquidos que se presentan en métodos "in vivo", tales como sangre, linfa, líquidos citosólicos u otros líquidos corporales, por ejemplo líquidos que pueden usarse como medios de referencia en métodos "in vitro", tales como tampones o líquidos comunes. Estos tampones o líquidos comunes son conocidos del experto. La solución de lactato de Ringer es particularmente preferente como base líquida.

Sin embargo, pueden usarse también para la composición farmacéutica empleada una o más cargas sólidas o líquidas compatibles o diluyentes o compuestos encapsulantes que sean adecuados para su administración a un paciente a tratar. El término "compatible" tal como se usa aquí significa que estos constituyentes de la composición farmacéutica empleada son capaces de ser mezclados con el ácido nucleico usado como se define aquí de manera que no se produzca una interacción que pudiera reducir sustancialmente la efectividad farmacéutica de la composición farmacéutica usada bajo condiciones de uso típicas.

De acuerdo con un aspecto específico, la composición farmacéutica usada puede comprender un adyuvante. En este contexto, un adyuvante puede entenderse como cualquier compuesto que es capaz de iniciar o incrementar una respuesta inmune del sistema inmunológico innato, es decir una respuesta inmunológica no específica. En otras palabras, cuando se administra, la composición farmacéutica usada desarrolla preferiblemente una respuesta inmune innata debida al adyuvante contenido opcionalmente en la misma. Preferentemente, tal adyuvante puede seleccionarse de los adyuvantes conocidos por el experto y adecuado para el presente caso, es decir, apoyar la inducción de una respuesta inmune innata en un mamífero, por ejemplo una proteína adyuvante como la definida arriba o un adyuvante como el definido a continuación.

Son particularmente preferentes como adyuvantes adecuados para depósito y suministro compuestos catiónicos o policatiónicos como los definidos arriba para la secuencia de ácido nucleico usada como vehículo, agente de transfección o formador de complejos.

La composición farmacéutica usada puede contener además una o más sustancias auxiliares para incrementar así su inmunogenicidad o capacidad inmunoestimuladora, si se desea. Preferentemente, de esta forma se consigue una acción sinérgica del ácido nucleico usado tal como se define aquí y de una sustancia auxiliar, que puede estar contenida opcionalmente en la composición farmacéutica usada. Dependiendo de los diferentes tipos de sustancias auxiliares, a este respecto pueden entrar en consideración varios mecanismos. Por ejemplo, los compuestos que permiten la maduración de las células dendríticas (DCs), por ejemplo lipopolisacáridos, TNF-alfa o ligando CD40, forman una primera clase de sustancias auxiliares adecuadas. En general, es posible emplear como sustancia auxiliar cualquier agente que influya el sistema inmunológico a la manera de una "señal de peligro" (LPS, GP96, etc.) o citoquinas, tales como GM-CSF, que permitan incrementar y/o influir en una respuesta inmunológica de forma predeterminada. Sustancias auxiliares particularmente preferentes son citoquinas, tales como monoquinas, linfoquinas, interleuquinas o quimioquinas, que promueven además la respuesta inmune innata, tales 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.

Otros aditivos que pueden incluirse en la composición farmacéutica usada son emulsionantes, por ejemplo Tween®; agentes humectantes, por ejemplo lauril-sulfato de sodio; agentes colorantes; agentes saborizantes, vehículos farmacéuticos; agentes formadores de tabletas; estabilizadores; antioxidantes y conservantes.

La composición farmacéutica usada también puede contener además cualquier compuesto adicional conocido como inmunoestimulador debido a su afinidad de unión (como ligandos) a los receptores tipo Toll humanos TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, o gracias a su afinidad de unión (como ligandos) a los receptores tipo Toll de murino TLR1, TLR2, TLR3, TLR4, TLR5, TLR6, TLR7, TLR8, TLR9, TLR10, TLR11, TLR12 o TLR13.

La composición farmacéutica empleada se puede administrar vía oral, parenteral, conespráis de inhalación, vía tópica, rectal, nasal, bucal, vaginal, o mediante un depósito implantado. El término parenteral según se

usa aquí incluye técnicas de inyección o infusión subcutánea, intravenosa, intramuscular, intraarticular, intrasinovial, intraesternal, intratecal, intrahepática, intralesional, intracraneal, transdérmica, intradérmica, intrapulmonar, intraperitoneal, intracardiaca, intraarterial y sublingual.

5 Preferentemente, la composición farmacéutica usada puede administrarse mediante inyección parenteral, en especial con técnicas de inyección subcutánea, intravenosa, intramuscular, intraarticular, intrasinovial, intraesternal, intratecal, intrahepática, intralesional, intracraneal, transdérmica, intradérmica, intrapulmonar, intraperitoneal, intracardiaca, intraarterial y sublingual o por medio de técnicas de infusión. Son particularmente preferentes la inyección transdérmica e intramuscular. Formas inyectables estériles de las composiciones farmacéuticas usadas pueden ser en suspensión acuosa u oleosa. Estas suspensiones
10 pueden formularse de acuerdo con técnicas conocidas usando agentes de dispersión o humectación adecuados y agentes de suspensión.

La composición farmacéutica usada como se define aquí también se puede administrar vía oral en cualquier forma de dosificación oralmente aceptable, incluyendo, pero no limitada a, cápsulas, tabletas, suspensiones o soluciones acuosas.

15 La composición farmacéutica usada también se puede administrar vía tópica, especialmente cuando el objetivo de tratamiento incluya áreas u órganos fácilmente accesibles mediante 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 de la invención puede formularse en una pomada adecuada que contenga el ácido
20 nucleico como se define aquí suspendido o disuelto en uno o más vehículos.

La composición farmacéutica usada comprende típicamente una "cantidad segura y efectiva" de los componentes, particularmente del ácido nucleico usado como se define aquí. Tal como se usa aquí, una "cantidad segura y efectiva" significa una cantidad de ácido nucleico como se define aquí que sea suficientemente significativa para inducir una modificación positiva de una enfermedad o trastorno como el
25 definido aquí. Al mismo tiempo, sin embargo, una "cantidad segura y efectiva" es lo suficientemente pequeña como para evitar serios efectos secundarios y permitir una relación sensible entre ventaja y riesgo. La determinación de estos límites se encuentra típicamente dentro del alcance del juicio médico.

La composición farmacéutica de la invención puede usarse para propósitos médicos humanos y también veterinarios, preferentemente para propósitos médicos en humanos, como una composición farmacéutica en
30 general o como una vacuna.

De acuerdo con otra realización particularmente preferente, la composición farmacéutica (o el ácido nucleico como se define aquí) puede proporcionarse o usada como una vacuna. Típicamente, tal vacuna es como se definió arriba para las composiciones farmacéuticas. Además, tal vacuna contiene típicamente el ácido nucleico como se define aquí, que preferentemente codifica para un antígeno como el definido arriba.
35 Alternativamente, esta vacuna puede contener el ácido nucleico usado como se define aquí y un antígeno adicional, preferentemente como una proteína o péptido o como un ácido nucleico que codifica para un antígeno, por ejemplo como se define aquí, o como cualquier formato antigénico como se define aquí o todas las combinaciones posibles de los mismos.

La vacuna como se proporciona para su uso aquí también puede comprender un portador, adyuvante y/o vehículo farmacéuticamente aceptable como se define aquí para la composición farmacéutica usada. En el contexto específico de la vacuna usada, la selección de un portador farmacéuticamente aceptable se determina en principio por la manera en la cual la vacuna usada es administrada. La vacuna usada puede administrarse, por ejemplo, vía sistémica o local. Las vías de administración sistémica en general incluyen, por
40 ejemplo, las vías transdérmica, oral, parenteral, incluyendo inyecciones subcutáneas, intravenosas, intramusculares, intraarteriales, intradérmicas e intraperitoneales y/o vías de administración intranasal. Las vías para la administración local en general incluyen, por ejemplo, vías de administración tópica, pero también inyecciones intradérmicas, transdérmicas, subcutáneas o intramusculares o inyecciones intralesionales, intracraneales, intrapulmonares, intracardiacas y sublinguales. Muy preferiblemente, las vacunas pueden administrarse por vía intradérmica, subcutánea o intramuscular. Las vacunas de la invención se formulan por
45 tanto, de preferencia, en forma líquida (o algunas veces en forma sólida).

La vacuna usada de acuerdo con la invención puede contener además una o más sustancias auxiliares para incrementar su inmunogenicidad o su capacidad inmunoestimuladora, si se desea. Son particularmente preferentes adyuvantes como sustancias auxiliares o aditivos como los definidos para la composición farmacéutica. La presente invención se dirige además a los usos médicos del ácido nucleico como se define
55 aquí, a la composición farmacéutica, a la vacuna, ambas comprendiendo el ácido nucleico como se define aquí, por ejemplo en el campo de la terapia génica. De acuerdo con la primera realización, la presente invención está dirigida al segundo uso médico del ácido nucleico usado como se define aquí para el

tratamiento de las enfermedades aquí definidas, en particular al uso del ácido nucleico como se define aquí para la profilaxis, el tratamiento y/o la mejora de las diversas enfermedades aquí definidas. La composición farmacéutica o la vacuna se usa o se administrará a un paciente que la requiera para este propósito.

5 Las enfermedades mencionadas aquí se seleccionan de entre cáncer o enfermedades tumorales, enfermedades infecciosas, de preferencia enfermedades infecciosas (virales, bacterianas o protozoológicas), enfermedades autoinmunes, alergias o enfermedades alérgicas, enfermedades monogénicas, es decir, enfermedades (hereditarias), o enfermedades genéticas en general, enfermedades que tengan un antecedente genético heredado y que sean causadas típicamente por un defecto génico definido y se hereden de acuerdo con las leyes de Mendel, enfermedades cardiovasculares, enfermedades neuronales, 10 enfermedades del sistema respiratorio, enfermedades del sistema digestivo, enfermedades de la piel, trastornos musculoesqueléticos, trastornos del tejido conectivo, neoplasmas, deficiencias inmunológicas, enfermedades endocrinas, nutricionales y metabólicas, enfermedades oculares, enfermedades ópticas y cualquier enfermedad que pueda ser influenciada por la presente invención.

15 Las enfermedades cancerosas o tumorales como las mencionadas arriba incluyen preferentemente, por ejemplo, carcinomas de colon, melanomas, carcinomas renales, linfomas, leucemia mieloide aguda (AML), leucemia linfocítica aguda (ALL), leucemia mieloide crónica (CML), leucemia linfocítica crónica (CLL), tumores gastrointestinales, carcinomas pulmonares, gliomas, tumores de tiroides, carcinomas mamarios, tumores de próstata, hepatomas, varios tumores inducidos por virus tales como carcinomas inducidos por virus de papiloma (por ejemplo, carcinoma cervical), adenocarcinomas, tumores inducidos por virus del herpes (por 20 ejemplo, linfoma de Burkitt, linfoma de células B inducido por EBV), tumores inducidos por hepatitis B (carcinoma hepatocelular), linfomas inducidos por HTLV-1 y HTLV-2, neuromas/neurinomas acústicos, cáncer cervical, cáncer pulmonar, cáncer faríngeo, carcinomas anales, glioblastomas, linfomas, carcinomas rectales, astrocitomas, tumores cerebrales, cáncer de estómago, retinoblastomas, basaliomas, metástasis cerebrales, meduloblastomas, cáncer vaginal, cáncer pancreático, cáncer testicular, melanomas, carcinomas de tiroides, 25 cáncer de vejiga, síndrome de Hodgkin, meningiomas, enfermedad de Schneeberger, carcinomas bronquiales, tumor de hipófisis, micosis fungoides, cáncer esofágico, cáncer de mama, carcinoides, neurinomas, espinaliomas, linfomas de Burkitt, cáncer de laringe, cáncer renal, timomas, carcinomas de cuerpo, cáncer de hueso, linfomas no Hodgkin, cáncer uretral, síndrome CUP, tumores de cuello/cabeza, oligodendrogliomas, cáncer vulvar, cáncer intestinal, carcinomas de colon, carcinomas esofágicos, 30 implicaciones de verrugas, tumores del intestino delgado, craneofaringeomas, carcinomas ováricos, tumores/sarcomas de tejidos blandos, cáncer ovárico, cáncer hepático, carcinomas pancreáticos, carcinomas cervicales, carcinomas endometriales, metástasis de hígado, cáncer penil, cáncer de lengua, cáncer de vesícula biliar, leucemia, plasmocitomas, cáncer uterino, tumor de párpado, cáncer de próstata, etc.

35 Además, en el contexto anterior, las enfermedades infecciosas se seleccionan preferentemente de entregripe, malaria, SARS, fiebre amarilla, SIDA, borreliosis de Lyme, Leishmaniasis, meningitis por ántrax, enfermedades infecciosas virales tales como SIDA, condiloma acuminata, verrugas huecas, fiebre de Dengue, fiebre de tres días, virus del Ébola, resfriado, meningoencefalitis de verano temprana (FSME), gripe, herpes zóster, hepatitis, herpes simple tipo I, herpes simple tipo II, herpes zoster, influenza, encefalitis Japonesa, fiebre de Lassa, virus de Marburg, sarampión, enfermedad de fiebre aftosa, mononucleosis, 40 paperas, infección por virus de Norwalk, fiebre glandular de Pfeiffer, viruela, polio (cojera en la niñez), pseudo-crup, enfermedad quinta, rabia, verrugas, fiebre del Nilo Occidental, viruela, virus citomegálico (CMV), enfermedades infecciosas bacterianas tales como aborto (inflamación de la próstata), ántrax, apendicitis, borreliosis, botulismo, *Camphylobacter*, *Chlamydia trachomatis* (inflamación de la uretra, conjuntivitis), cólera, difteria, donavanosis, epiglotitis, fiebre de tifo, gangrena gaseosa, gonorrea, fiebre de conejo, *Helicobacter pylori*, tos silvante, bubo climático, osteomielitis, enfermedad del Legionario, lepra, listeriosis, neumonía, 45 meningitis, meningitis bacteriana, ántrax, otitis media, micoplasma hominis, sepsis neonatal (corioamnionitis), noma, paratifo, plaga, síndrome de Reiter, fiebre manchada de las Montañas Rocallosas, paratifo por salmonella, tifo por salmonella, fiebre escarlata, sífilis, tétanos, enfermedad de tsutsugamushi, tuberculosis, tífus, vaginitis (colpitis), chancro suave y de enfermedades infecciosas causadas por parásitos, protozoarios u hongos, tales como amibiásis, bilharziosis, enfermedad de Chagas, pie de atleta, machas por hongos de levadura, escabiosis, malaria, oncocercosis (ceguera de río) o enfermedades fúngicas, toxoplasmosis, tricomoniasis, tripanosomiasis (enfermedad del sueño), Leishmaniosis visceral, dermatitis por braga/pañal, esquistosomiasis, envenenamiento por pescado (Ciguatera), candidosis, Leishmaniosis cutánea, lambliasis (giardiasis) o enfermedad del sueño, o de enfermedades infecciosas causadas por *Equinococcus*, tenia de 55 pescado, tenia de zorro, tenia de caninos, piojos, tenia de bovino, tenia de porcina, tenia en miniatura.

Además, en el contexto anterior, las alergias resultan normalmente en una respuesta inflamatoria local o sistémica a estos antígenos o alérgenos y lleva a inmunidad en el cuerpo contra estos alérgenos. Los alérgenos en este contexto incluyen, por ejemplo, hierbas, pólenes, mohos, fármacos o numerosos desencadenadores ambientales, etc. Sin estar limitados a teoría, se supone que están implicados varios 60 mecanismos de enfermedad diferentes en el desarrollo de las alergias. De acuerdo con un esquema de clasificación de P. Gell y R. Coombs, la palabra "alergia" fue restringida a hipersensibilidades tipo I, las cuales

son causadas por el mecanismo IgE clásico. La hipersensibilidad tipo I se caracteriza por excesiva activación de mastocitos y basófilos por IgE, dando como resultado una respuesta inflamatoria sistémica que puede traducirse en síntomas tan benignos como nariz irritada, como hasta un choque anafiláctico que amenace la vida y la propia muerte. Tipos conocidos de alergia incluyen, sin estar limitados a éstos, asma alérgica (que lleva a hinchazón de la mucosa nasal), conjuntivitis alérgica (que lleva a enrojecimiento y comezón de la conjuntiva), rinitis alérgica (“fiebre del heno”), anafilaxis, angioderma, dermatitis atópica (eczema), urticaria, eosinofilia, alergias a picaduras de insectos, alergias de la piel (que llevan a e incluyen erupciones, tales como eczema, urticaria y dermatitis por contacto), alergias a alimentos, alergias a medicinas, etc. Con respecto a la presente invención, por ejemplo, se proporciona un ácido nucleico, una composición farmacéutica o una vacuna que codifica o contiene un alérgeno (por ejemplo un alérgeno de gato, de polvo, antígeno de ácaro, un antígeno vegetal (por ejemplo un antígeno de abedul), etc.), ya sea como una proteína, como un ácido nucleico que codifica para ese alérgeno de proteína en combinación con un ácido nucleico de la invención como se define arriba o como un ácido nucleico como se usa aquí. Una composición farmacéutica usada de acuerdo con la presente invención puede desplazar la respuesta inmunológica (excedente) a una respuesta TH1 más fuerte, suprimiendo o atenuando así la respuesta IgE indeseada.

Además, las enfermedades autoinmunes pueden dividirse ampliamente en trastornos autoinmunes sistémicos y específicos de órganos o localizados, dependiendo de las características clínico-patológicas principales de cada enfermedad. Las enfermedades autoinmunes pueden dividirse en las categorías de síndromes sistémicos, incluyendo LES, síndrome de Sjögren, escleroderma, artritis reumatoide y polimiositis o síndromes locales que pueden ser endocrinológicos (DM tipo 1, tiroiditis de Hashimoto, enfermedad de Addison), dermatológicos (pénfigo vulgar), hematológicos (anemia hemolítica autoinmune), neurales (esclerosis múltiple) o pueden implicar casi cualquier masa circunscrita de tejido corporal. Las enfermedades autoinmunes a tratar pueden seleccionarse del grupo consistente en enfermedades autoinmunes tipo I o enfermedades autoinmunes tipo II o enfermedades autoinmunes tipo III o enfermedades autoinmunes tipo IV, por ejemplo esclerosis múltiple (MS), artritis reumatoide, diabetes, diabetes tipo I (diabetes mellitus), lupus eritematoso sistémico (LES), poliartritis crónica, enfermedad de Basedow, formas autoinmunes de hepatitis crónica, colitis ulcerosa, enfermedades de alergia tipo I, enfermedades de alergia tipo II, enfermedades de alergias tipo III, enfermedades de alergias tipo IV, fibromialgia, caída de pelo, enfermedad de Bechterew, enfermedad de Crohn, miastenia grave, neurodermitis, polimialgia reumática, esclerosis sistémica progresiva (PSS), psoriasis, síndrome de Reiter, artritis reumática, psoriasis, vasculitis, etc., o diabetes tipo II. Aunque el modo exacto de por qué el sistema inmunológico induce una reacción inmune contra autoantígenos no ha sido elucidado hasta el momento, existen varios descubrimientos con respecto a la etiología. Así, la autorreacción puede deberse a un Bypass de células T. Un sistema inmunológico normal requiere la activación de las células B por las células T antes de que las primeras puedan producir anticuerpos en grandes cantidades. Este requerimiento de una célula T puede estar derivado en raros casos, tales como infección por organismos que produzcan superantígenos, los cuales son capaces de iniciar la activación policlonal de células B, o incluso de células T, al unirse directamente a una subunidad de los receptores de células T de una forma no específica. Otra explicación deduce enfermedades autoinmunes a partir de un mimetismo molecular. Un antígeno exógeno puede compartir similitudes estructurales con ciertos antígenos huéspedes; así, cualquier anticuerpo producido contra este antígeno (que imite a los auto-antígenos) también puede, en teoría, unirse a los antígenos huésped y amplificar la respuesta inmunológica. La forma más desconcertante de mimetismo molecular se observa en estreptococos beta-hemolíticos del grupo A, los cuales comparten antígenos con el miocardio humano, y son responsables de las manifestaciones cardíacas de fiebre reumática. La presente invención permite por tanto la provisión de un ácido nucleico o de una composición farmacéutica o de una vacuna como se definen aquí que codifican o contienen por ejemplo un autoantígeno (como proteína, ARNm o ADN que codifica para una proteína de autoantígeno) que permite típicamente que el sistema inmunológico sea desensibilizado.

De acuerdo con una realización más, la presente invención está dirigida al segundo uso médico del ácido nucleico como se define aquí, para el tratamiento de enfermedades como las aquí definidas mediante terapia génica.

En una realización preferente, el ácido nucleico como se define aquí puede usarse para la preparación de una composición farmacéutica o de una vacuna para los propósitos aquí definidos.

La composición farmacéutica o vacuna de la invención se puede usar además para el tratamiento de una enfermedad o trastorno como se definen aquí.

En la presente invención, si no se indica lo contrario, las características diferentes de alternativas y realizaciones pueden combinarse unas con otras. Además, el término “que comprende” no se deberá considerar como significando “que consiste en”, si no se menciona específicamente. Sin embargo, en el contexto de la presente invención, el término “que comprende” puede sustituirse por el término “que consiste en”, cuando sea aplicable.

Figuras

Las siguientes figuras intentar ilustrar la invención y no deberán considerarse como limitativas de la presente invención.

5 Figura 1: muestra la secuencia de consenso de tallo-lazo de histona generada a partir de secuencias de tallo y lazo de metazoarios y protozoarios (como se reporta por Dávila López, M., y Samuelsson, T. (2008), RNA (Nueva York, N.Y.), 14(1), 1-10. doi: 10.1261/rna.782308). 4.001 secuencias de tallo-lazo de histona de metazoarios y protozoarios fueron alineadas y se indica la cantidad de los nucleótidos que se presentaban para cada posición en la secuencia de tallo-lazo. La secuencia de consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se da usando el código de nucleótidos de una sola letra. Además de la secuencia de consenso, las secuencias se muestran representando al menos 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

15 Figura 2: muestra la secuencia de consenso de tallo-lazo de histona generada a partir de secuencias de tallo y lazo de protozoarios (como se reporta por Dávila López, M., y Samuelsson, T. (2008), RNA (Nueva York, N.Y.), 14(1), 1-10. doi: 10.1261/rna.782308). 131 secuencias de tallo-lazo de histona de protozoarios fueron alineadas y se indica la cantidad de los nucleótidos que se presentan para cada posición en la secuencia de tallo-lazo. La secuencia de consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se da usando el código de nucleótido de una sola letra. Además de la secuencia de consenso, las secuencias que se muestran representando al menos 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

20 Figura 3: muestra la secuencia de consenso de tallo-lazo de histona generada a partir de secuencias de tallo y lazo de metazoarios (como se reporta por Dávila López, M., y Samuelsson, T. (2008), RNA (Nueva York, N.Y.), 14(1), 1-10. doi: 10.1261/ma.782308). 3.870 secuencias de tallo-lazo de histona de metazoarios fueron alineadas y se indica la cantidad de los nucleótidos que se presentaban para cada posición en la secuencia de tallo-lazo. La secuencia de consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se da usando el código de nucleótidos de letra individual. Además de la secuencia de consenso, se muestran secuencias que representan al menos 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

30 Figura 4: muestra la secuencia de consenso de tallo-lazo de histona generada a partir de secuencias de tallo-lazo de vertebrados (como se reporta por Dávila López, M., y Samuelsson, T. (2008), RNA (Nueva York, N.Y.), 14(1), 1-10. doi: 10.1261/ma.782308). 3.870 secuencias de tallo-lazo de histona de vertebrados fueron alineadas y se indica la cantidad de los nucleótidos que se presentaban para cada posición en la secuencia de tallo-lazo. La secuencia de consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se da usando el código de nucleótidos de letra individual. Además de la secuencia de consenso, se muestran secuencias que representan al menos 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

40 Figura 5: muestra la secuencia de consenso de tallo-lazo de histona generada a partir de secuencias de tallo y lazo de humano (*Homo sapiens*) (como se reporta por Dávila López, M., y Samuelsson, T. (2008), RNA (Nueva York, N.Y.), 14(1), 1-10. doi: 10.1261/ma.782308). 84 Secuencias de tallo-lazo de histona de humanos fueron alineadas y se indica la cantidad de los nucleótidos que se presentaban para cada posición en la secuencia de tallo-lazo. La secuencia de consenso generada que representa todos los nucleótidos presentes en las secuencias analizadas se da usando el código de nucleótidos de letra individual. Además de la secuencia de consenso, se muestran secuencias que representan al menos 99%, 95% y 90% de los nucleótidos presentes en las secuencias analizadas.

45 Figuras 6 a 17: muestran moléculas de ARNm de transcripción *in vitro*. Se muestra la designación y la secuencia de las moléculas de ARNm obtenidas mediante transcripción *in vitro*. Se usan las siguientes abreviaturas:

ppLuc (GC): secuencia de ARNm enriquecida con GC que codifica para luciferasa de *Photinus pyralis*
 ag: región no traducida 3' (UTR) del gen de alfa globina
 50 A64: secuencia poli(A) con 64 adenilados
 A120: secuencia poli(A) con 120 adenilados
 histonaSL: tallo-lazo de histona
 aCPSL: tallo y lazo que ha sido seleccionado de una biblioteca para su unión específica a la proteína α CP-2KL
 55 polioCL: hoja de trébol 5' del ARN genómico del virus de la polio
 G30: secuencia poli(G) con 30 guanilados
 U30: secuencia poli(U) con 3e0 uridilados
 SL: tallo-asa no específico/artificial

N32: secuencia no específica de 32 nucleótidos

dentro de las secuencias, se resaltan los siguientes elementos: ppLuc(GC) ORF (letras mayúsculas, ag (negritas), histonaSL (subrayado), secuencias probadas distintas adicionales (cursivas).

5 Figura 6: muestra la secuencia de ARNm de ppLuc(GC)-ag (SEQ ID NO: 43). Mediante linearización del vector original en el sitio de restricción inmediatamente después del 3'-UTR (ag) de alfa-globina se obtiene un ARNm que carece de una secuencia poli(A).

Figura 7: muestra la secuencia de ARNm de ppLuc(GC)-ag-A64 (SEQ ID NO: 44). Mediante linearización del vector original en el sitio de restricción inmediatamente después de la secuencia poli(A) A64, se obtiene un ARNm que concluye en una secuencia poli(A) A64.

Figura 8: muestra la secuencia de ARNm de ppLuc(GC)-ag-histonaSL (SEQ ID NO: 45). La secuencia poli(A) A64 se reemplazó por una histonaSL. La secuencia de tallo-lazo de histona usada en los ejemplos se obtuvo de Cakmakci y col., (2008). *Molecular and Cellular Biology*, 28(3), 1182-1194.

Figura 9: muestra la secuencia de ARNm de ppLuc(GC)-ag-A64-histonaSL (SEQ ID NO: 46). La histonaSL se anexó 3' de A64 poli(A).

Figura 10: muestra la secuencia de ARNm de ppLuc(GC)-ag-A120 (SEQ ID NO: 47). La secuencia A64 poli(A) se reemplazó por una secuencia A120 poli(A).

Figura 11: muestra la secuencia de ARNm de ppLuc(GC)-ag-A64-ag (SEQ ID NO: 48). Una segunda 3'-UTR de alfa-globina se anexó 3' de A64 poli(A).

Figura 12: muestra la secuencia de ARNm de ppLuc(GC)-ag-A64-aCPSL (SEQ ID NO: 49). Un tallo y lazose anexó 3' en A64 poli(A). El tallo y lazose seleccionó de una biblioteca para su unión específica de la proteína α CP-2KL (Thisted y col., (2001), *The Journal of Biological Chemistry*, 276(20), 17484-17496). α CP-2KL es una isoforma de α CP-2, la proteína α CO más fuertemente expresada (proteína de unión a poli(C) de ARNm de alfa-globina) (Makeyev y col., (2000), *Genomics*, 67(3), 301-316), un grupo de proteínas de unión a ARN que se unen a 3'-UTR de alfa-globina (Chkheidze y col., (1999), *Molecular and Cellular Biology*, 19(7), 4572-4581).

Figura 13: muestra la secuencia de ARNm de ppLuc(GC)-ag-A64-PolioCL (SEQ ID NO: 50). La hoja de trébol de 5' del ARN genómico del virus de la polio se anexó 3' de A64 poli(A).

Figura 14: muestra la secuencia de ARNm de ppLuc(GC)-ag-A64-G30 (SEQ ID NO: 51). Un tramo de 30 guanilados se anexó 3' de A64 poli(A).

Figura 15: muestra la secuencia de ARNm de ppLuc(GC)-ag-A64-U30 (SEQ ID NO: 52). Un tramo de 30 uridilados se anexó 3' de A64 poli(A).

Figura 16: muestra la secuencia de ARNm de ppLuc(GC)-ag-A64-SL (SEQ ID NO: 53). Un tallo y lazo se anexaron 3' de A64 poli(A). La parte superior del tallo y lazo se tomaron de Babendure y col., (2006), *RNA* (Nueva York, N. Y.), 12(5), 851-861. El tallo y lazo consiste en un tallo rico en CG de 17 pares de bases de longitud y un lazo de 6 pares de bases de longitud.

Figura 17: muestra ppLuc(GC)-ag-A64-N32 (SEQ ID NO: 54). Mediante linearización del vector original en el sitio de restricción alternativo se obtuvo un ARNm con 32 nucleótidos adicionales después del poli(A).

Figura 18: muestra que la combinación de poli(A) e histonaSL incrementa la expresión de proteínas a partir de ARNm de una manera sinérgica. Se examinó el efecto de la secuencia poli(A), histonaSL y la combinación de poli(A) e histonaSL en la expresión de luciferasa de ARNm. Así, diferentes ARNm fueron electroporados en células HeLa. Los niveles de luciferasa se midieron 6, 24 y 48 horas después de la transfección. Se expresó poca luciferasa a partir de ARNm que no tiene ni la secuencia poli(A) ni histonaSL. Tanto la secuencia poli(A) como la histonaSL incrementan el nivel de luciferasa. Sorprendentemente, sin embargo, la combinación de poli(A) e histonaSL incrementa más fuertemente el nivel de luciferasa, varias veces por encima del nivel observado con cualquiera de los elementos individuales, actuando así sinérgicamente. Los datos se indican como media RLU \pm SD (unidades de luz relativas \pm desviación estándar) para transfecciones triplicadas. RLU específicas se resumen en el Ejemplo 11.2.

Figura 19: muestra que la combinación de poli(A) e histonaSL incrementa la expresión de proteínas a partir de ARNm sin importar su orden. Se examinó el efecto de la secuencia poli(A), histonaSL, la combinación

de kpoli(A) e histonaSL, y su orden en la expresión de luciferasa de ARNm. Por tanto, diferentes moléculas de ARNm fueron lipofectadas usando células HeLa. Los niveles de luciferasa se midieron 6, 24 y 48 horas después del inicio de la transfección. Tanto una secuencia A64 poli(A) como la histonaSL dieron origen a niveles de luciferasa comparables. Incrementar la longitud de la secuencia poli(A) de A64 a A120 o a A300 incrementa el nivel de luciferasa moderadamente. En contraste, la combinación de poli(A) e histonaSL incrementa el nivel de luciferasa mucho más que el alargamiento de la secuencia poli(A). La combinación de poli(A) e histonaSL actúa sinérgicamente toda vez que incrementa el nivel de luciferasa varias veces por encima del nivel observado con cualquiera de los elementos individuales. El efecto sinérgico de la combinación de poli(A) e histonaSL observado es no obstante del orden de poli(A) e histonaSL y no obstante la longitud de poli(A) con ARNm de histonaSL A64 o histonaSL A250. Se muestran los datos como RLU media \pm SD para transfecciones triplicadas. Las RLU específicas se resumen en el Ejemplo 11.3.

Figura 20: muestra que la elevación en la expresión de proteínas por la combinación de poli(A) e histonaSL es específica. Se estudió el efecto de combinar poli(A) e histonaSL o poli(A) y secuencias alternativas en la expresión de luciferasa de ARNm. Así, diferentes moléculas de ARN fueron electroporadas en células HeLa. Los niveles de luciferasa se midieron 6, 24 y 48 horas después de la transfección. Tanto una secuencia poli(A) como la histonaSL dieron origen a niveles de luciferasa comparables. La combinación de poli(A) e histonaSL incrementa fuertemente el nivel de luciferasa varias veces por encima del nivel observado con cualquiera de los elementos individuales, actuando así sinérgicamente. En contraste, combinar poli(A) con cualquiera de las demás secuencias no tiene efectos en el nivel de luciferasa en comparación con ARNm que contiene sólo una secuencia poli(A). Así, la combinación de poli(A) e histonaSL actúa específicamente y sinérgicamente. Se muestran los datos como media RLU \pm SD para transfecciones triplicadas. Las RLU específicas se resumen en el Ejemplo 11.4.

Figura 21: muestra que la combinación de poli(A) e histonaSL incrementa la expresión de proteínas a partir de ARNm de una manera sinérgica *in vivo*. Se estudió el efecto de la secuencia poli(A), histonaSL y la combinación de poli(A) e histonaSL en la expresión de luciferasa de ARNm *in vivo*. Por tanto, diferentes moléculas de ARNm se inyectaron intradérmicamente en ratones. Los ratones se sacrificaron 16 horas después de la inyección y los niveles de luciferasa en los sitios de inyección fueron medidos. La luciferasa es expresada a partir de ARNm que tiene ya sea una histonaSL o una secuencia poli(A). Sorprendentemente, sin embargo, la combinación de poli(A) e histonaSL incrementa fuertemente el nivel de luciferasa varias veces por encima del nivel observado con cualquiera de los elementos individuales, actuando así sinérgicamente. Se muestran los datos como RLU media \pm SEM (unidades de luz relativas \pm error estándar de la media). Las RLU específicas se resumen en el Ejemplo 11.5.

Figura 22: muestra que la combinación de poli(A) e histonaSL incrementa la expresión de proteínas NY-ESO-1 a partir de ARNm. Se estudió el efecto de la secuencia poli(A) y la combinación de poli(A) e histonaSL en la expresión de NY-ESO-1 de ARNm. Por tanto, diferentes moléculas de ARNm fueron electroporadas en células HeLa. Los niveles de NY-ESO-1 fueron medidos 24 horas después de la transfección mediante citometría de flujo. NY-ESO-1 se expresa a partir de ARNm que sólo tiene una secuencia poli(A). Sorprendentemente, sin embargo, la combinación de poli(A) e histonaSL incrementa fuertemente el nivel de NY-ESO-1, varias veces por encima del nivel observado sólo con una secuencia poli(A). Los datos se muestran como recuentos frente a intensidad de fluorescencia. Las intensidades de fluorescencia medias (MFI) se resumen en el Ejemplo 11.6.

Figura 23: muestra que la combinación de poli(A) e histonaSL incrementa el nivel de anticuerpos desarrollados por vacunación con ARNm. Se estudió el efecto de la secuencia poli(A) y la combinación de poli(A) e histonaSL en la inducción de anticuerpos anti NY-ESO-1 desarrollados por vacunación con ARNm. Así, ratones C57BL/6 fueron vacunados intradérmicamente con diferentes moléculas de ARNm complejadas con protamina. El nivel de anticuerpos específicos para NY-ESO-1 en los ratones vacunados y de control se analizan por ELISA con varias diluciones de sueros. Anti-NY-ESO-1 IgG2a[b] se induce por ARNm que tiene sólo una secuencia poli(A). Sorprendentemente, sin embargo, la combinación de poli(A) e histonaSL incrementa fuertemente el nivel de anti NY-ESO-1 IgG2a[b], varias veces por encima del nivel observado sólo con una secuencia poli(A). Los datos se grafican como los títulos de criterio promedio. Los títulos de criterio promedio se resumen en el Ejemplo 11.7.

Ejemplos

Los siguientes ejemplos intentan ilustrar la invención y no deberá considerarse que limitan la presente invención a los mismos.

1. Generación de secuencias de consenso de histona-tallo-lazo

Antes de los experimentos, se determinaron secuencias de consenso de histona de tallo-lazo en base a las secuencias de tallo-lazo de histona de metazoarios y protozoarios. Las secuencias se tomaron del

suplemento proporcionado por López y col. (Dávila López, M., y Samuelsson, T. (2008), RNA (Nueva York, N. Y.), 14(1), 1-10. doi: 10.1261/rna.782308), quien identificó un gran número de secuencias de tallo-lazo de histona naturales al investigar secuencias genómicas y marcadores de secuencia expresados. Primero, todas las secuencias de metazoarios y protozoarios (4.001 secuencias), o todas las secuencias de protozoarios (131 secuencias) o como alternativa de metazoarios (3.870 secuencias), o de vertebrados (1.333 secuencias) o de humanos (84 secuencias) se agruparon y alinearon. Luego, se determinó la cantidad de los nucleótidos que se originaban para cada posición. Con base en las tablas así obtenidas, se generaron secuencias de consenso para los 5 grupos de secuencias diferentes que representaban todos los nucleótidos presentes en las secuencias analizadas. Además, también se obtuvieron secuencias de consenso más restrictivas, enfatizando cada vez más los nucleótidos conservados.

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 GC que codificaba para luciferasa de *Photinus pyralis* (ppLuc(GC)), la parte central de la región 3' no traducida (UTR) de alfa-globina (ag), y una secuencia poli(A). La secuencia poli(A) fue inmediatamente seguida por un sitio de restricción usado para linealización del vector antes de la transcripción *in vitro* para obtener así ARNm que concluía en una secuencia A64 poli(A). ARNm obtenido de este vector en consecuencia por transcripción *in vitro* se designa como "ppLuc(GC)-ag-A64".

La linealización de este vector en sitios de restricción alternativos antes de la transcripción *in vitro* permitió obtener ARNm ya sea extendido por nucleótidos adicionales 3' de A64 o que carecían de A64. Además, el vector original se modificó para incluir secuencias alternativas. En resumen, se obtuvieron las siguientes secuencias de ARNm de estos vectores mediante transcripción *in vitro* (las secuencias de ARNm se dan en las Figuras 6 a 17):

ppLuc(GC)-ag (SEQ ID NO: 43)
 ppLuc(GC)-ag-A64 (SEQ ID NO: 44)
 ppLuc(GC)-ag-histonaSL (SEQ ID NO: 45)
 ppLuc(GC)-ag-A64-histonaSL (SEQ ID NO: 46)
 ppLuc(GC)-ag-A120 (SEQ ID NO: 47)
 ppLuc(GC)-ag-A64-ag (SEQ ID NO: 48)
 ppLuc(GC)-ag-A64-Acpsi (SEQ ID NO: 49)
 ppLuc(GC)-ag-A64-PolioCL (SEQ ID NO: 50)
 ppLuc(GC)-ag-A64-G30 (SEQ ID NO: 51)
 ppLuc(GC)-ag-A64-U30 (SEQ ID NO: 52)
 ppLuc(GC)-ag-A64-SL (SEQ ID NO: 53)
 ppLuc(GC)-ag-A64-N32 (SEQ ID NO: 54)

3. Transcripción *in vitro*

La plantilla de ADN de acuerdo con el Ejemplo 2 se linealizó y transcribió *in vitro* usando T7-polimerasa. La plantilla de ADN se digirió después mediante tratamiento con ADNasa. Todos los transcritos de ARNm contenían una estructura 5'-CAP obtenida al añadir un exceso de N7-metil-guanosin-5'-trifosfato-5'-guanosina a la reacción de transcripción. El ARNm así obtenido se purificó y resuspendió en agua.

4. Adenilación enzimática de ARNm

Se adenilaron enzimáticamente dos moléculas de ARNm: ppLuc(GC)-ag-A64 y ppLuc(GC)-ag-histonaSL.

Para este fin, ARN se incubó con el *E. coli* Poly(A)-polimerase and ATP (Poly(A) Polimerase Tailing Kit, Epicentre, Madison, E.U.A.) siguiendo las instrucciones del fabricante. Se purificó y resuspendió en agua ARNm con secuencia poli(A) extendida. La longitud de la secuencia poli(A) se determinó mediante electroforesis en gel de agarosa. Las moléculas de ARN partidas se extendieron en aproximadamente 250 adenilados, las moléculas de ARN obtenidas se designan como ppLuc(GC)-ag-A300 y ppLuc(GC)-ag-histonaSL-A250, respectivamente.

5. Expresión de luciferasa por electroporación de ARNm

Células HeLa fueron tripsinizadas y lavadas en opti-MEM. 1×10^5 células en 200 μ l de opti-MEM cada una fueron electroporadas con 0,5 μ g de ARNm que codificaba para ppLuc. Como un control, ARNm que no codificaba para ppLuc fue electroporado por separado. Las células electroporadas fueron sembradas en placas de 24 pocillos en 1 ml de medio RPMI 1640. 6, 24 ó 48 Horas después de la transfección, el medio

fue aspirado y las células fueron lisadas en 200 µl de regulador de pH de lisis (25 mM de Tris, pH 7.5 (HCl), 2 mM de EDTA, 10% de glicerol, 1% de Triton X-100, 2 mM de DTT, 1 mM de PMSF). Los lisados se almacenaron a -20°C hasta que se midiera la actividad ppLuc.

5 6. Expresión de luciferasa por lipofección de ARNm

Células HeLa fueron sembradas en placas de 96 pocillos a una densidad de 2×10^4 células por pocillo. Al día siguiente, las células fueron lavadas en opti-MEM y luego transfectadas con 0,25 µg de ARNm que codificaba para ppLuc acompañado a lipofectina en 150 µl de opti-MEM. Como control, ARNm que no codificaba para ppLuc fue lipofectado por separado. En algunos pocillos, opti-MEM fue aspirado y las células fueron lisadas en 200 µl de tampón de lisis 6 horas después del inicio de la transfección. En los pocillos restantes, el opti-MEM se cambió por medio RPMI 1640 en ese momento. En estos pocillos, el medio fue aspirado y las células fueron lisadas 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 midiera la actividad ppLuc.

15 7. Medición de luciferasa

La actividad ppLuc se midió como unidades de luz relativas (RLU) en un lector de placa BioTek SynergyHT a 5 segundos de tiempo de medición usando 50 µl de lisado y 200 µl de tampón de luciferina (25 mM de glicilglicina, pH 7,8 (NaOH), 15 mM de $MgSO_4$, 2 mM de ATP, 75 µM de luciferina). Las RLU específicas fueron calculadas al restar RLU del ARN de control a partir de RLU total.

20 8. Expresión de luciferasa por inyección de ARNm intradérmica (expresión de luciferasa *in vivo*)

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

30 9. Expresión de NY-ESO-1 por electroporación de ARNm

Células HeLa fueron tripsinizadas y lavadas con opti-MEM. 2×10^5 células en 200 µl de opti-MEM fueron electroporadas con 10 µg de ARNm que codificaba para NY-ESO-1. Las células de las tres electroporaciones fueron combinadas y sembradas en una placa de 6 pocillos en 2 ml de medio RPMI 1640. 24 Horas después de la transfección, las células fueron cosechadas y transfectadas en una placa de fondo en V de 96 pocillos (2 pocillos por ARNm). Las células fueron lavadas con solución salina tamponada con fosfato (PBS) y permeabilizadas con 200 µl por pocillo de Cytofix/Cytoperm (Becton Dickinson (BD)). Después de 15 minutos, las células fueron lavadas con PermWash (BD). Luego, las células fueron incubadas durante 1 hora a temperatura ambiente ya sea con IgG1 anti-NY-ESO-1 de ratón o un control de isotipo (20 µg/ml). Las células fueron lavadas dos veces con PermWash de nuevo. Después, las células fueron incubadas durante 1 hora a 4°C con una dilución 1:500 de Alexa-647 acoplada a IgG de anti-ratón de cabra. Finalmente, las células fueron lavadas dos veces con PermWash. Las células fueron resuspendidas en 200 µl de tampón (PBS, 2% de FCS, 2 mM de EDTA, 0.01% de azida de sodio). La expresión de NY-ESO-1 se cuantificó mediante citometría de flujo como intensidad de fluorescencia media (MFI).

45 10. Inducción de anticuerpos anti NY-ESO-1 por vacunación con ARNm

Ratones C57BL/6 fueron vacunados intradérmicamente con ARNm que codificaba para NY-ESO-1 acompañado con protamina (5 veces en 14 días). Los ratones de control fueron tratados con tampón. El nivel de anticuerpos específicos para NY-ESO-1 en los ratones vacunados y de control se analizó 8 días después de la última vacunación mediante ELISA: Placas por ELISA de 96 pocillos (Nunc) fueron recubiertas con 100 µl por pocillo de 10 µg/ml de proteína NY-ESO-1 recombinante durante 16 horas a 4°C. Las placas se lavaron dos veces con tampón de lavado (PBS, 0,05% de Tween-20). Para bloquear la unión no específica, las placas se incubaron después durante 2 horas a 37°C con tampón de bloqueo (PBS, 0,05% de Tween-20, 1% de BSA). Después del bloqueo, 100 µl por pocillo de suero de ratón diluidos en serie se añadieron e incubaron durante 4 horas a temperatura ambiente. Las placas se lavaron después tres veces con tampón de lavado. Después, 100 µl por pocillo de anticuerpo de detección IgG2a[b] anti-ratón de rata biotinilado (BD Biosciences) diluido 1:600 en tampón de bloqueo se dejó unir durante 1 hora a temperatura ambiente. Las placas se lavaron de nuevo tres veces con tampón de lavado, seguidas por incubación durante 30 minutos a temperatura ambiente con 100 µl por pocillo de estreptavidina acoplada a peroxidasa de rábano. Después de cuatro lavados con tampón de lavado, se añadieron 100 µl por pocillo de 3,3',5,5'-tetrametilbencidina (Thermo

*		La base puede estar presente o no
---	--	-----------------------------------

11.2 La combinación de poli(A) e histonaSL incrementa la expresión de proteína a partir de ARNm de una manera sinérgica

5 Para investigar el efecto de la combinación de poli(A) e histonaSL en la expresión de proteínas a partir de ARNm, moléculas de ARNm con diferentes secuencias 3' de la 3'-UTR de alfa-globina fueron sintetizadas: moléculas de ARNm ya sea que concluían justo a 3' de la 3'-UTR, careciendo entonces de la secuencia poli(A) e histonaSL, o contenía ya sea una secuencia A64 poli(A) o una histonaSL en su lugar, pero ambas A64 poli(A) e histonaSL 3' de la 3'-UTR. Moléculas de ARNm que codificaban para luciferasa o ARNm de control se electroporaron en células HeLa. Los niveles de luciferasa se midieron 6, 24 y 48 horas después de la transfección (véase la siguiente Tabla 6 y Figura 18).

Tabla 6

ARNm	TLU a 6 horas	RLU a 24 h	RLU a 48 h
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 ARNm que no tenía ni secuencia poli(A) ni histonaSL. Tanto una secuencia poli(A) como la histonaSL incrementaron el nivel de luciferasa a un grado similar. Cada ARNm dio origen a un nivel de luciferasa más alto que el ARNm que carecía tanto de poli(A) como de histonaSL. Sorprendentemente, sin embargo, la combinación de poli(A) e histonaSL incrementó fuertemente más el nivel de luciferasa, varias veces por encima del nivel observado con cualquiera de los elementos individuales. La magnitud de la elevación en el nivel de luciferasa debido a la combinación de poli(A) e histonaSL en el mismo ARNm demuestra que actúan sinérgicamente.

La sinergia entre poli(A) e histonaSL se cuantificó al dividir la señal de poli(A)-histonaSL ARNm (+/+) entre la suma de las señales de histonaSL ARNm (-/+) más poli(A) ARNm (+/-) (véase la siguiente Tabla 7).

Tabla 7

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

25 El factor así calculado especifica que tan alto el nivel de luciferasa de ARNm que combina poli(A) e histonaSL sería el esperado si los efectos de poli(A) e histonaSL fueran puramente aditivos. El nivel de luciferasa a partir de ARNm que combina poli(A) e histonaSL fue hasta 16,6 veces más alto que si sus efectos fueran puramente aditivos. Este resultado confirma que la combinación de poli(A) e histonaSL efectúa un incremento marcadamente sinérgico en la expresión de proteínas.

11.3 La combinación de poli(A) e histonaSL incrementa la expresión de proteínas a partir de ARNm no obstante de su orden

35 El efecto de la combinación de poli(A) e histonaSL podría depender de la longitud de la secuencia poli(A) y el orden de la poli(A) e histonaSL. Así, se sintetizaron moléculas de ARNm con longitud de secuencia poli(A) cada vez más alta y ARNm con poli(A) e histonaSL en orden inverso. Dos moléculas de ARNm contenían 3' de la 3'-UTR una secuencia poli(A) ya sea A120 o A300. Una ARNm adicional contenía 3' del 3'-UTR primero una histonaSL seguida por una secuencia A250 poli(A). Las moléculas de ARNm que codificaban para luciferasa o ARNm de control fueron lipofectadas en células HeLa. Los niveles de luciferasa se midieron 6, 24 y 48 horas después del inicio de la transfección (véase la siguiente tabla 8 y figura 19).

Tabla 8

ARNm	RLU a 6 h	RLU a 24 h	RLU a 48 h
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

5 Tanto una secuencia A64 poli(A) como la histonaSL dieron origen a niveles de luciferasa comparables. De acuerdo con el experimento anterior, la combinación de A64 e histonaSL sí incrementó fuertemente el nivel de luciferasa, varias veces por encima del nivel observado con cualquiera de los elementos individuales. La magnitud de la elevación en el nivel de luciferasa debido a combinar poli(A) e histonaSL en el mismo ARNm demuestra que están actuando sinérgicamente. La sinergia entre A64 e histonaSL se cuantificó como antes con base en los niveles de luciferasa de ARNm de A64-histonaSL, A64 e histonaSL (véase la siguiente Tabla 9). El nivel de luciferasa de ARNm que combinaba A64 e histonaSL fue hasta 61,7 veces más alto que si los efectos de poli(A) e histonaSL fueran puramente aditivos.

10

Tabla 9

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

15 En contraste, incrementar la longitud de la secuencia poli(A) de A64 a A120 o a A300 incrementó el nivel de luciferasa sólo moderadamente (véase Tabla 8 y Figura 19). También se comparó ARNm con la secuencia poli(A) más larga, A300, con ARNm en el cual una secuencia poli(A) de longitud similar se combinó con la histonaSL, histonaSL-A250. Además de tener una larga secuencia poli(A), el orden de histonaSL y poli(A) se invierte en este ARNm con relación al ARNm de A64-histonaSL. La combinación de A250 e histonaSL fuertemente incrementó el nivel de luciferasa, varias veces por encima del nivel observado ya sea con histonaSL o A300. De nuevo, la sinergia entre A250 e histonaSL fue cuantificada como antes comparando RLU de ARNm de histonaSL-A250 con RLU de ARNm de A300 más ARNm de histonaSL (véase la siguiente Tabla 10). El nivel de luciferasa de ARNm que combinaba A250 e histonaSL fue hasta 17,0 veces más alto que si los efectos de poli(A) e histonaSL fueron puramente aditivos.

25

Tabla 10

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

30 En resumen, se ha demostrado un efecto altamente sinérgico de la combinación de histonaSL y poli(A) en la expresión de proteínas de ARNm para longitudes sustancialmente diferentes de poli(A) y no obstante el orden de poli(A) e histonaSL.

11.4 La elevación en la expresión de proteínas por la combinación de poli(A) e histonaSL es específica

35 Para investigar si el efecto de la combinación de poli(A) e histonaSL en la expresión de proteínas de ARNm es específico, se sintetizaron moléculas de ARNm con secuencias alternativas en combinación con poli(A): Estas moléculas de ARNm contenían 3' de A64 una de siete secuencias distintas, respectivamente. Moléculas de ARNm que codificaban para luciferasa o ARNm de control fueron electroporadas en células HeLa. Los niveles de luciferasa se midieron a 6, 24 y 48 horas después de la transfección (véase la siguiente Tabla 11 y Figura 20).

Tabla 11

ARNm	RLU a 6 h	RLU a 24 h	RLU a 48 h
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
poLuc(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

5 Tanto una secuencia poli(A) como la histonaSL dieron origen a niveles de luciferasa comparables. De nuevo, la combinación de poli(A) e histonaSL incrementó fuertemente el nivel de luciferasa, varias veces por encima del nivel observado con cualquiera de los elementos individuales, actuando así sinérgicamente. En contraste, combinar poli(A) con cualquiera de las secuencias alternativas no tuvo efecto en el nivel de luciferasa en comparación con ARNm que contenía sólo una secuencia poli(A). Así, la combinación de poli(A) e histona SL incrementan la expresión de proteínas de ARNm de una manera sinérgica, y su efecto es específico.

10 **11.5 La combinación de poli(A) e histonaSL incrementa la expresión de proteínas a partir de ARNm de una manera sinérgica *in vivo***

15 Para investigar el efecto de la combinación de poli(A) e histonaSL en la expresión de proteínas de ARNm *in vivo*, moléculas de ARNm que codificaban para luciferasa con diferentes secuencias 3' de la 3'-UTR de alfa-globina o ARNm de control se inyectaron intradérmicamente en ratones: las moléculas de ARNm contenían ya sea una secuencia A64 poli(A) o una histonaSL en su lugar, o tanto una A64 poli(A) como histonaSL 3' de la 3'-UTR. Los niveles de luciferasa se midieron 16 horas después de la inyección (véase la siguiente Tabla 12 y Figura 21).

Tabla 12

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

20 La luciferasa fue expresada a partir de ARNm que tenía ya sea una histonaSL o una secuencia poli(A). Sorprendentemente, sin embargo, la combinación de poli(A) e histonaSL incrementó más fuertemente el nivel de luciferasa, varias veces por encima del nivel observado con cualquiera de los elementos individuales. La magnitud de la elevación en el nivel de luciferasa debido a la combinación de poli(A) e histona SL en el mismo ARNm demuestra que están actuando sinérgicamente.

25 La sinergia entre poli(A) e histonaSL se cuantificó al dividir la señal de poli(A)-histonaSL ARNm (+/+) entre la suma de las señales de ARNm de histonaSL (-/+) más ARNm de poli(A) (+/-) (véase la siguiente Tabla 13).

Tabla 13

	A64	histonaSL	RLU a 16 h
	+	+	38081
	-	+	137
	+	-	4607
Sinergia			8,0

30 El factor así calculado especifica cuán alto es el nivel de luciferasa a partir de ARNm que combina poli(A) e histonaSL de lo que se esperaría si los efectos de poli(A) e histonaSL fueran puramente aditivos. El nivel de luciferasa de ARNm que combinaba poli(A) e histonaSL fue ocho veces más alto que si sus efectos fueran puramente aditivos. Este resultado confirma que la combinación de poli(A) e histonaSL efectúa un incremento marcadamente sinérgico en la expresión de proteínas *in vivo*.

35 **11.6 La combinación de poli(A) e histonaSL incrementan la expresión de proteína NY-ESO-1 a partir de ARNm**

40 Para investigar el efecto de la combinación de poli(A) e histonaSL en la expresión de proteínas a partir de ARNm, moléculas de ARNm que codificaban para NY-ESO-1 con diferentes secuencias 3' de la 3'-UTR de alfa-globina fueron sintetizadas: moléculas de ARNm que contenían ya sea una secuencia A64 poli(A) o tanto A64 poli(A) como histonaSL 3' de la 3'-UTR. Moléculas de ARNm que codificaban para NY-ESO-1 fueron electroporadas en células HeLa. Los niveles de NY-ESO-1 se midieron 24 horas después de la transfección mediante citometría de flujo (véase la siguiente Tabla 14 y Figura 22).

Tabla 14

ARNm	MFI a 24 horas	
	Anti-NY-ESO-1	Control de isotipo
NY-ESO-1(GC)-ag-A64-histonaSL	15600	1831
NY-ESO-1(GC)-ag-A64	1294	849

NY-ESO-1 fue expresada a partir de ARNm que tenía sólo una secuencia poli(A). Sorprendentemente, sin embargo, la combinación de poli(A) e histonaSL incrementó fuertemente el nivel de NY-ESO-1, varias veces por encima del nivel observado sólo con una secuencia poli(A).

5 **11.7 La combinación de poli(A) e histonaSL incrementa el nivel de anticuerpos desarrollados por vacunación con ARNm**

10 Para investigar el efecto de la combinación de poli(A) e histonaSL en la inducción de anticuerpos desarrollados por vacunación con ARNm, ratones C57BL/6 fueron vacunados intradérmicamente con moléculas de ARNm que codificaban para NY-ESO-1 complejadas a protamina con diferentes secuencias 3' de la 3'-UTR de alfa-globina. Las moléculas de ARNm contenían ya sea una secuencia A64 poli(A) o tanto A64 poli(A) como histonaSL 3' de la 3'-UTR. El nivel de anticuerpos específicos para NY-ESO-1 en ratones vacunados y de control se analizó por ELISA con diluciones en sueros en serie (véase la siguiente Tabla 15 y Figura 23).

Tabla 15

ARNm	Título de criterio de IgG2a[b] medio
NY-ESO-1(GC)-ag-A64-histonaSL	763
NY-ESO-1(GC)-ag-A64	20

15

IgG2a[b] anti NY-ESO-1 se indujo por ARNm que sólo tenía una secuencia poli(A): Sorprendentemente, sin embargo, la combinación de poli(A) e histonaSL incrementó fuertemente el nivel de IgG2a[b] anti-NY-SO-1, muchas veces por encima del nivel observado sólo con una secuencia poli(A).

LISTADO DE SECUENCIAS

- <110> CureVac GmbH
- 5 <120> Ácido nucleico que comprende un tallo-lazo de histona y opcionalmente una secuencia de poliadenilación para aumentar la expresión de una proteína codificada
- <130> CU01P088WO1
- 10 <160> 95
- <170> PatentIn version 3.5
- <210> 1
- 15 <211> 16
- <212> DNA
- <213> Artificial Sequence
- <220>
- 20 <223>estructura tallo-lazo to fórmula (Ic)
- <220>
- <221> variación
- 25 <222> (1)..(1)
- <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine" /replace=""
- <220>
- 30 <221> misc_feature
- <222> (1)..(1)
- <223> n es a, c, g, t or u
- <220>
- 35 <221> variación
- <222> (3)..(8)
- <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine" /replace=""
- <220>
- 40 <221> misc_feature
- <222> (3)..(8)
- <223> n es a, c, g, t or u
- <220>
- 45 <221> variación
- <222> (10)..(14)
- <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine" /replace=""
- 50 <220>
- <221> misc_feature
- <222> (10)..(14)
- <223> n es a, c, g, t or u
- 55 <220>
- <221> variación
- <222> (16)..(16)
- <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine" /replace=""
- 60 <220>
- <221> misc_feature
- <222> (16)..(16)
- 65 <223> n es a, c, g, t or u

ES 2 525 556 T3

<400> 1
 ngnnnnnnunun nnnncn 16

5 <210> 2
 <211> 26
 <212> DNA
 <213>Secuencia Artificial

10 <220>
 <223>estructura tallo-lazo según la fórmula(IIc)

15 <220>
 <221> variación
 <222> (1)..(6)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

20 <220>
 <221> misc_feature
 <222> (1)..(6)
 <223> n es a, c, g, t or u

25 <220>
 <221> variación
 <222> (8)..(13)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

30 <220>
 <221> misc_feature
 <222> (8)..(13)
 <223> n es a, c, g, t or u

35 <220>
 <221> variación
 <222> (15)..(19)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

40 <220>
 <221> misc_feature
 <222> (15)..(19)
 <223> n es a, c, g, t o u

45 <220>
 <221> variación
 <222> (21)..(26)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

50 <220>
 <221> misc_feature
 <222> (21)..(26)
 <223> n es a, c, g, t o u

55 <400> 2
 nnnnnngnnn nnnunnnnc nnnnnn 26

60 <210> 3
 <211> 16
 <212> DNA
 <213>Secuencia Artificial

65 <220>
 <223>estructura tallo-lazo según la fórmula (Id)

ES 2 525 556 T3

<220>
 <221> variación
 <222> (1)..(1)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 5 /replace=""

<220>
 <221> misc_feature
 <222> (1)..(1)
 10 <223> n es a, c, g, t o u

<220>
 <221> variación
 <222> (3)..(8)
 15 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

<220>
 <221> misc_feature
 20 <222> (3)..(8)
 <223> n es a, c, g, t o u

<220>
 <221>variación
 25 <222> (10)..(14)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

<220>
 <221> misc_feature
 30 <222> (10)..(14)
 <223> n es a, c, g, t o u

<220>
 <221> variación
 35 <222> (16)..(16)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

<220>
 <221> misc_feature
 40 <222> (16)..(16)
 <223> n es a, c, g, t o u

45 <400> 3
 ncnnnnnn nnnngn 16

<210> 4
 <211> 26
 50 <212> DNA
 <213>Secuencia Artificial

<220>
 <223>estructura tallo-lazo según la fórmula (IId)
 55

<220>
 <221> variación
 <222> (1)..(6)
 60 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

<220>
 <221> misc_feature
 65 <222> (1)..(6)
 <223> n es a, c, g, t o u

ES 2 525 556 T3

- 5 <220>
<221> variación
<222> (8)..(13)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 10 <220>
<221> misc_feature
<222> (8)..(13)
<223> n es a, c, g, t o u
- 15 <220>
<221> variación
<222> (15)..(19)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 20 <220>
<221> misc_feature
<222> (15)..(19)
<223> n es a, c, g, t o u
- 25 <220>
<221> variación
<222> (21)..(26)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 30 <220>
<221> misc_feature
<222> (21)..(26)
<223> n es a, c, g, t o u
- 35 <400> 4
nnnnnncnnn nnnunnnng nnnnnn 26
- 40 <210> 5
<211> 16
<212> DNA
<213>Secuencia Artificial
- 45 <220>
<223>estructura tallo-lazo según la fórmula (Ie)
- 50 <220>
<221>variación
<222> (3)..(8)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 55 <220>
<221> misc_feature
<222> (3)..(8)
<223> n es a, c, g, t o u
- 60 <220>
<221>variación
<222> (10)..(14)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 65 <220>
<221> misc_feature
<222> (10)..(14)
<223> n es a, c, g, t o u

<400> 5
dgnnnnnnun nnnch 16

5 <210> 6
<211> 26
<212> DNA
<213>Secuencia Artificial

10 <220>
<223>estructura tallo-lazo según la fórmula(Ile)

<220>
<221>variación

15 <222> (1)..(5)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""

<220>
20 <221> misc_feature
<222> (1)..(5)
<223> n es a, c, g, t o u

<220>
25 <221>variación
<222> (8)..(13)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""

30 <220>
<221> misc_feature
<222> (8)..(13)
<223> n es a, c, g, t o u

35 <220>
<221>variación
<222> (15)..(19)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""

40 <220>
<221> misc_feature
<222> (15)..(19)
<223> n es a, c, g, t o u

45 <220>
<221>variación
<222> (22)..(26)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""

50 <220>
<221> misc_feature
<222> (22)..(26)
<223> n es a, c, g, t o u

<400> 6
nnnnndggnnn nnnunnnnnc hnnnnn 26

60 <210> 7
<211> 16
<212> DNA
<213>Secuencia Artificial

65 <220>
<223>estructura tallo-lazo según la fórmula(If)

ES 2 525 556 T3

- 5 <220>
<221>variación
<222> (1)..(1)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 10 <220>
<221> misc_feature
<222> (1)..(1)
<223> n es a, c, g, t o u
- 15 <220>
<221>variación
<222> (3)..(3)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 20 <220>
<221> misc_feature
<222> (3)..(3)
<223> n es a, c, g, t o u
- 25 <220>
<221>variación
<222> (7)..(8)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 30 <220>
<221> misc_feature
<222> (7)..(8)
<223> n es a, c, g, t o u
- 35 <220>
<221>variación
<222> (10)..(10)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 40 <220>
<221> misc_feature
<222> (10)..(10)
<223> n es a, c, g, t o u
- 45 <220>
<221>variación
<222> (12)..(12)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 50 <220>
<221> misc_feature
<222> (12)..(12)
<223> n es a, c, g, t o u
- 55 <220>
<221>variación
<222> (14)..(14)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 60 <220>
<221> misc_feature
<222> (14)..(14)
<223> n es a, c, g, t o u
- 65 <220>
<221>variación
<222> (14)..(14)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""

- 5 <220>
<221>variación
<222> (16)..(16)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 10 <220>
<221> misc_feature
<222> (16)..(16)
<223> n es a, c, g, t o u
- 15 <400> 7
ngnbyynnun rndncn 16
- <210> 8
<211> 26
<212> DNA
<213>Secuencia Artificial
- 20 <220>
<223>estructura tallo-lazo según la fórmula(If)
- 25 <220>
<221>variación
<222> (1)..(6)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 30 <220>
<221> misc_feature
<222> (1)..(6)
<223> n es a, c, g, t o u
- 35 <220>
<221>variación
<222> (8)..(8)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 40 <220>
<221> misc_feature
<222> (8)..(8)
<223> n es a, c, g, t o u
- 45 <220>
<221>variación
<222> (12)..(13)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 50 <220>
<221> misc_feature
<222> (12)..(13)
<223> n es a, c, g, t o u
- 55 <220>
<221>variación
<222> (15)..(15)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 60 <220>
<221> misc_feature
<222> (15)..(15)
<223> n es a, c, g, t o u
- 65 <220>
<221>variación
<222> (15)..(15)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 65 <220>
<221> misc_feature
<222> (15)..(15)
<223> n es a, c, g, t o u

- 5 <220>
<221>variación
<222> (17)..(17)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 10 <220>
<221> misc_feature
<222> (17)..(17)
<223> n es a, c, g, t o u
- 15 <220>
<221>variación
<222> (19)..(19)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 20 <220>
<221> misc_feature
<222> (19)..(19)
<223> n es a, c, g, t o u
- 25 <220>
<221>variación
<222> (21)..(26)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 30 <220>
<221> misc_feature
<222> (21)..(26)
<223> n es a, c, g, t o u
- 35 <400> 8
nnnnnngnby ynnunrmdnc nnnnnn 26
- 40 <210> 9
<211> 16
<212> DNA
<213>Secuencia Artificial
- 45 <220>
<223>estructura tallo-lazo según la fórmula (lg)
- 50 <220>
<221>variación
<222> (1)..(1)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 55 <220>
<221> misc_feature
<222> (1)..(1)
<223> n es a, c, g, t o u
- 60 <220>
<221>variación
<222> (8)..(8)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 65 <220>
<221> misc_feature
<222> (8)..(8)
<223> n es a, c, g, t o u

- 5 <220>
<221>variación
<222> (16)..(16)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 10 <220>
<221> misc_feature
<222> (16)..(16)
<223> n es a, c, g, t o u
- 15 <400> 9
nghyydntn abrdcn 16
- <210> 10
<211> 26
<212> DNA
<213>Secuencia Artificial
- 20 <220>
<223> estructura tallo-lazo según la fórmula(Ilg)
- 25 <220>
<221>variación
<222> (1)..(2)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 30 <220>
<221> misc_feature
<222> (1)..(2)
<223> n es a, c, g, t o u
- 35 <220>
<221>variación
<222> (4)..(6)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 40 <220>
<221> misc_feature
<222> (4)..(6)
<223> n es a, c, g, t o u
- 45 <220>
<221>variación
<222> (13)..(13)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 50 <220>
<221> misc_feature
<222> (13)..(13)
<223> n es a, c, g, t o u
- 55 <220>
<221>variación
<222> (21)..(25)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 60 <220>
<221> misc_feature
<222> (21)..(25)
<223> n es a, c, g, t o u
- 65 <220>
<221>variación
<222> (21)..(25)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 65 <220>
<221> misc_feature
<222> (21)..(25)
<223> n es a, c, g, t o u

- <400> 10
nnhnnnghyy ydnthabrdc nnnnnh 26
- 5 <210> 11
<211> 16
<212> DNA
<213>Secuencia Artificial
- 10 <220>
<223>estructura tallo-lazo según la fórmula (Ih)
- <400> 11
dghyctdyuh asrrcc 16
- 15 <210> 12
<211> 26
<212> DNA
<213>Secuencia Artificial
- 20 <220>
<223>estructura tallo-lazo según la fórmula (IIh)
- <220>
- 25 <221>variación
<222> (1)..(1)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 30 <220>
<221> misc_feature
<222> (1)..(1)
<223> n es a, c, g, t o u
- 35 <220>
<221>variación
<222> (25)..(25)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- 40 <220>
<221> misc_feature
<222> (25)..(25)
<223> n es a, c, g, t o u
- 45 <400> 12
nhaahdghyc tdyhasrrc cvhbnh 26
- 50 <210> 13
<211> 16
<212> DNA
<213>Secuencia Artificial
- 55 <220>
<223>secuencia tallo-lazo de histona según la fórmula (Ic)
- <220>
- 60 <221>variación
<222> (4)..(4)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""
- <220>
- 65 <221> misc_feature
<222> (4)..(4)

- <223> n es a, c, g, t o u
- <220>
 <221>variación
 5 <222> (13)..(13)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
 10 <221> misc_feature
 <222> (13)..(13)
 <223> n es a, c, g, t o u
- <400> 13
 15 ggcncctttc agngcc 16
- <210> 14
 <211> 16
 <212> DNA
 20 <213>Secuencia Artificial
- <220>
 <223>secuencia tallo-lazo de histona según la fórmula(1e)
- <220>
 25 <221>variación
 <222> (5)..(5)
 <223> n = g o c o t/u o a (cualquier base o no presente)
- <220>
 30 <221> misc_feature
 <222> (5)..(5)
 <223> n es a, c, g, t o u
- <220>
 35 <221>variación
 <222> (7)..(8)
 <223> n = g o c o t/u o a (cualquier base o no presente)
- <220>
 40 <221> misc_feature
 <222> (7)..(8)
 <223> n es a, c, g, t o u
- <220>
 45 <221>variación
 <222> (10)..(10)
 <223> n = g o c o t/u o a (cualquier base o no presente)
- <220>
 50 <221> misc_feature
 <222> (10)..(10)
 <223> n es a, c, g, t o u
- <220>
 55 <221>variación
 <222> (12)..(13)
 <223> n = g o c o t/u o a (cualquier base o no presente)
- <220>
 60 <221> misc_feature
 <222> (12)..(13)
 <223> n es a, c, g, t o u
- <400> 14
 65 ggctntnntn anngcc 16

- <210> 15
 <211> 16
 <212> DNA
 5 <213>Secuencia Artificial

 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lf)
 10 <400> 15
 ggcbcttttc agdgcc 16

 <210> 16
 <211> 16
 15 <212> DNA
 <213>Secuencia Artificial

 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lg)
 20 <400> 16
 ggctcttth agagcc 16

 <210> 17
 <211> 16
 25 <212> DNA
 <213>Secuencia Artificial

 <220>
 30 <223>secuencia tallo-lazo de histona según la fórmula(lh)

 <400> 17
 ggcycyttth agrgcc 16

 35 <210> 18
 <211> 26
 <212> DNA
 <213>Secuencia Artificial

 40 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(IIc)

 <220>
 <221>variación
 45 <222> (1)..(2)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

 <220>
 50 <221> misc_feature
 <222> (1)..(2)
 <223> n es a, c, g, t o u

 <220>
 55 <221>variación
 <222> (4)..(5)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

 60 <220>
 <221> misc_feature
 <222> (4)..(5)
 <223> n es a, c, g, t o u

 65 <220>
 <221>variación

ES 2 525 556 T3

- <222> (9)..(9)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
 <220>
- 5 <221> misc_feature
 <222> (9)..(9)
 <223> n es a, c, g, t o u
- <220>
- 10 <221>variación
 <222> (18)..(18)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
- 15 <221> misc_feature
 <222> (18)..(18)
 <223> n es a, c, g, t o u
- <220>
- 20 <221>variación
 <222> (24)..(26)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
- 25 <221> misc_feature
 <222> (24)..(26)
 <223> n es a, c, g, t o u
- 30 <400> 18
 nnannggcnc tttcagngc cacnnn 26
- <210> 19
 <211> 26
 <212> DNA
 <213>Secuencia Artificial
- <220>
- 40 <223>secuencia tallo-lazo de histona según la fórmula(Ile)
- <220>
- 45 <221>variación
 <222> (1)..(3)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
- 50 <221> misc_feature
 <222> (1)..(3)
 <223> n es a, c, g, t o u
- <220>
- 55 <221>variación
 <222> (5)..(5)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
- 60 <221> misc_feature
 <222> (5)..(5)
 <223> n es a, c, g, t o u
- <220>
- 65 <221>variación

<222> (10)..(10)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

5 <220>
 <221> misc_feature
 <222> (10)..(10)
 <223> n es a, c, g, t o u

10 <220>
 <221>variación
 <222> (12)..(13)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

15 <220>
 <221> misc_feature
 <222> (12)..(13)
 <223> n es a, c, g, t o u

20 <220>
 <221>variación
 <222> (15)..(15)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"

25 /replace=""

<220>
 <221> misc_feature
 <222> (15)..(15)
 <223> n es a, c, g, t o u

30 <220>
 <221>variación
 <222> (17)..(18)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

35 <220>
 <221> misc_feature
 <222> (17)..(18)
 <223> n es a, c, g, t o u

40 <220>
 <221>variación
 <222> (24)..(26)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

45 <220>
 <221> misc_feature
 <222> (24)..(26)
 <223> n es a, c, g, t o u

50 <400> 19

55 nnnanggctn tntnanngc cacnnn 26

<210> 20
 <211> 26
 <212> DNA
 <213>Secuencia Artificial

60 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(IIf)

65 <220>
 <221>variación

- <222> (1)..(2)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- 5 <220>
 <221> misc_feature
 <222> (1)..(2)
 <223> n es a, c, g, t o u
- 10 <220>
 <221>variación
 <222> (4)..(5)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- 15 <220>
 <221> misc_feature
 <222> (4)..(5)
 <223> n es a, c, g, t o u
- 20 <220>
 <221>variación
 <222> (24)..(26)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- 25 <220>
 <221> misc_feature
 <222> (24)..(26)
 <223> n es a, c, g, t o u
- <400> 20
 nnannggcbc tttcagdc cacnnc 26
- 35 <210> 21
 <211> 26
 <212> DNA
 <213>Secuencia Artificial
- 40 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(IIg)
- <220>
 <221>variación
- 45 <222> (1)..(1)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
 <221> misc_feature
- 50 <222> (1)..(1)
 <223> n es a, c, g, t o u
- <220>
 <221>variación
- 55 <222> (5)..(5)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- 60 <220>
 <221> misc_feature
 <222> (5)..(5)
 <223> n es a, c, g, t o u
- 65 <220>
 <221>variación

ES 2 525 556 T3

<222> (25)..(25)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

5 <220>
 <221> misc_feature
 <222> (25)..(25)
 <223> n es a, c, g, t o u

10 <400> 21
 ncaanggctc tttthagagc caccnh 26

<210> 22
 <211> 26

15 <212> DNA
 <213>Secuencia Artificial

<220>
 <223>secuencia tallo-lazo de histona según la fórmula(IIh)

20 <220>
 <221>variación
 <222> (1)..(1)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"

25 /replace=""

<220>
 <221> misc_feature
 <222> (1)..(1)

30 <223> n es a, c, g, t o u

<220>
 <221>variación
 <222> (25)..(25)

35 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

<220>
 <221> misc_feature
 <222> (25)..(25)

40 <223> n es a, c, g, t o u

<400> 22
 nhaahggcyc tttthagrgc cvgbnh 26

45 <210> 23
 <211> 16
 <212> DNA
 <213>Secuencia Artificial

50 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ic)

<400> 23
 vgyyyyhhth rvvrcb 16

55 <210> 24
 <211> 16
 <212> DNA

60 <213>Secuencia Artificial

<220>
 <223>secuencia tallo-lazo de histona según la fórmula (Ic)

65 <400> 24
 rgyyytttm agrrcs 16

<210> 25
 <211> 16
 <212> DNA
 5 <213>Secuencia Artificial

 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lc)
 10 <400> 25
 rgyyytttm agrrcs 16

 <210> 26
 <211> 16
 15 <212> DNA
 <213>Secuencia Artificial

 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lc)
 20 <400> 26
 rgyyyyyytm rrrrcs 16

 <210> 27
 <211> 16
 25 <212> DNA
 <213>Secuencia Artificial

 <220>
 30 <223>secuencia tallo-lazo de histona según la fórmula(lc)

 <400> 27
 ggcycytttc agrgcc 16

 35 <210> 28
 <211> 16
 <212> DNA
 <213>Secuencia Artificial

 40 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lc)

 <400> 28
 45 ggccttttc agggcc 16

 <210> 29
 <211> 16
 <212> DNA
 <213>Secuencia Artificial
 50
 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(le)

 <220>
 55 <221>variación
 <222> (3)..(5)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

 60 <220>
 <221> misc_feature
 <222> (3)..(5)
 <223> n es a, c, g, t o u

 65 <220>
 <221>variación

- <222> (7)..(8)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- 5 <220>
 <221> misc_feature
 <222> (7)..(8)
 <223> n es a, c, g, t o u
- 10 <220>
 <221>variación
 <222> (12)..(14)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- 15 <220>
 <221> misc_feature
 <222> (12)..(14)
 <223> n es a, c, g, t o u
- 20 <400> 29
 dgnnnbnnth vnnch 16
- <210> 30
 <211> 16
 <212> DNA
 <213>Secuencia Artificial
- 30 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(1e)
- <220>
 <221>variación
 <222> (3)..(5)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- 35 <220>
 <221> misc_feature
 <222> (3)..(5)
 <223> n es a, c, g, t o u
- 40 <220>
 <221>variación
 <222> (13)..(14)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- 45 <220>
 <221> misc_feature
 <222> (13)..(14)
 <223> n es a, c, g, t o u
- 50 <400> 30
 rgnnnyhbth rdncy 16
- 55 <210> 31
 <211> 16
 <212> DNA
 <213>Secuencia Artificial
- 60 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(1e)
- 65 <220>
 <221>variación

<222> (3)..(3)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

5 <220>
 <221> misc_feature
 <222> (3)..(3)
 <223> n es a, c, g, t o u

10 <220>
 <221>variación
 <222> (14)..(14)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

15 <220>
 <221> misc_feature
 <222> (14)..(14)
 <223> n es a, c, g, t o u

20 <400> 31
 rgndbyhyth rdhncy 16

25 <210> 32
 <211> 16
 <212> DNA
 <213>Secuencia Artificial

30 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(le)

<400> 32
 rgykywytw rrmrcy 16

35 <210> 33
 <211> 16
 <212> DNA
 <213>Secuencia Artificial

40 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(le)

<400> 33
 45 ggctytwytw armgcc 16

<210> 34
 <211> 16
 <212> DNA

50 <213>Secuencia Artificial

<220>
 <223>secuencia tallo-lazo de histona según la fórmula(le)

55 <400> 34
 ggcttttta agagcc 16

<210> 35
 <211> 16
 <212> DNA

60 <213>Secuencia Artificial

<220>
 <223>secuencia tallo-lazo de histona según la fórmula(lf)

65 <400> 35

	vgyyytyhth ryrccb	16
	<210> 36	
	<211> 16	
5	<212> DNA	
	<213>Secuencia Artificial	
	<220>	
10	<223>secuencia tallo-lazo de histona según la fórmula(lf)	
	<400> 36	
	sgyycttytm agrrcs	16
	<210> 37	
15	<211> 16	
	<212> DNA	
	<213>Secuencia Artificial	
	<220>	
20	<223>secuencia tallo-lazo de histona según la fórmula(lf)	
	<400> 37	
	sgyyctttm agrrcs	16
25	<210> 38	
	<211> 16	
	<212> DNA	
	<213>Secuencia Artificial	
30	<220>	
	<223>secuencia tallo-lazo de histona según la fórmula(lf)	
	<400> 38	
35	sgyyyyyytm rrrrcs	16
	<210> 39	
	<211> 16	
	<212> DNA	
	<213>Secuencia Artificial	
40	<220>	
	<223>secuencia tallo-lazo de histona según la fórmula(lf)	
	<400> 39	
45	ggcycytttc agrgcc	16
	<210> 40	
	<211> 16	
	<212> DNA	
	<213>Secuencia Artificial	
50	<220>	
	<223>secuencia tallo-lazo de histona según la fórmula(lf)	
	<400> 40	
55	ggcccttttc agggcc	16
	<210> 41	
	<211> 16	
	<212> DNA	
60	<213>Secuencia Artificial	
	<220>	
	<223>secuencia tallo-lazo de histona según la fórmula(lg)	
65	<400> 41	
	ggyycttyth agrcc	16

<210> 42
 <211> 16
 <212> DNA
 5 <213>Secuencia Artificial

 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ig)

 10 <400> 42
 ggcyccttytm agrgcc 16

 <210> 43
 <211> 16
 15 <212> DNA
 <213>Secuencia Artificial

 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ig)
 20 <400> 43
 ggctctttm agrgcc 16

 <210> 44
 <211> 16
 25 <212> DNA
 <213>Secuencia Artificial

 <220>
 30 <223>secuencia tallo-lazo de histona según la fórmula(Ig)

 <400> 44
 rgyyyykytm asrrcb 16

 35 <210> 45
 <211> 16
 <212> DNA
 <213>Secuencia Artificial

 40 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ig)

 <400> 45
 ggctctttm agagcc 16
 45 <210> 46
 <211> 16
 <212> DNA
 <213>Secuencia Artificial

 50 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ig)

 <400> 46
 55 ggctctttc agagcc 16

 <210> 47
 <211> 16
 <212> DNA
 <213>Secuencia Artificial
 60 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ih)

 <400> 47
 65 kgcyctryth agrgcc 16

<210> 48
 <211> 16
 <212> DNA
 <213>Secuencia Artificial
 5 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lh)
 <400> 48
 10 ggcyccttth agrgcc 16
 <210> 49
 <211> 16
 <212> DNA
 15 <213>Secuencia Artificial
 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lh)
 20 <400> 49
 ggcyccttth agrgcc 16
 <210> 50
 <211> 16
 25 <212> DNA
 <213>Secuencia Artificial
 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lh)
 30 <400> 50
 kghyctkytm asrrcc 16
 <210> 51
 <211> 16
 35 <212> DNA
 <213>Secuencia Artificial
 <220>
 40 <223>secuencia tallo-lazo de histona según la fórmula(lh)
 <400> 51
 ggcycctttm agrgcc 16
 45 <210> 52
 <211> 16
 <212> DNA
 <213>Secuencia Artificial
 50 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lh)
 <400> 52
 55 ggctctttc agagcc 16
 <210> 53
 <211> 26
 <212> DNA
 <213>Secuencia Artificial
 60 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(lhc)
 <220>
 <221>variación
 65 <222> (26)..(26)

ES 2 525 556 T3

```

<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
/replace=""

<220>
5 <221> misc_feature
<222> (26)..(26)
<223> n e s a , c , g , t o u

<400> 53
10 hhhhhvgyyy yhhthrvrc bvhhhn 26

<210> 54
<211> 26
<212> DNA
15 <213>Secuencia Artificial

<220>
<223>secuencia tallo-lazo de histona según la fórmula(IIc)

20 <400> 54
hhmmrgyyy tttmagrrc sachhh 26

<210> 55
<211> 26
25 <212> DNA
<213>Secuencia Artificial

<220>
<223>secuencia tallo-lazo de histona según la fórmula(IIc)
30 <400> 55
hmmhrgyyy tttmagrrc sachhh 26

<210> 56
<211> 26
35 <212> DNA
<213>Secuencia Artificial

<220>
40 <223>secuencia tallo-lazo de histona según la fórmula(IIc)

<400> 56
mmmmrgyyy yyymrrrc smymmw 26

45 <210> 57
<211> 26
<212> DNA
<213>Secuencia Artificial

50 <220>
<223>secuencia tallo-lazo de histona según la fórmula(IIc)

<400> 57
55 mmammggcyc tttcagrc cacmmw 26

<210> 58
<211> 26
<212> DNA
<213>Secuencia Artificial

60 <220>
<223>secuencia tallo-lazo de histona según la fórmula(IIc)

<400> 58
65 acaacggccc tttcagggc caccaa 26
<210> 59

```

ES 2 525 556 T3

- <211> 26
 <212> DNA
 <213>Secuencia Artificial
- 5 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ile)
- <220>
 <221>variación
 10 <222> (1)..(5)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
 <221> misc_feature
 <222> (1)..(5)
 <223> n es a, c, g, t o u
- <220>
 <221>variación
 <222> (8)..(10)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
 <221> misc_feature
 <222> (8)..(10)
 <223> n es a, c, g, t o u
- <220>
 <221>variación
 <222> (12)..(13)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
 <221> misc_feature
 <222> (12)..(13)
 <223> n es a, c, g, t o u
- <220>
 <221>variación
 <222> (17)..(19)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
 <221> misc_feature
 <222> (17)..(19)
 <223> n es a, c, g, t o u
- <220>
 <221>variación
 <222> (22)..(22)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
- <220>
 <221> misc_feature
 <222> (22)..(22)
 <223> n es a, c, g, t o u
- <220>
 <221>variación
 <222> (24)..(26)

ES 2 525 556 T3

<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="" /replace="uracile" /replace="cytosine"

<220>

5 <221> misc_feature
<222> (24)..(26)
<223> n es a, c, g, t o u

<400> 59

10 nnnnndggnn bnthvnnnc hnhnnn 26
<210> 60
<211> 26
<212> DNA
<213>Secuencia Artificial

15 <220>
<223>secuencia tallo-lazo de histona según la fórmula(Ile)

20 <220>
<221>variación
<222> (1)..(2)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="" /replace="uracile" /replace="cytosine"

25 <220>
<221> misc_feature
<222> (1)..(2)
<223> n es a, c, g, t o u

30 <220>
<221>variación
<222> (8)..(10)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="" /replace="uracile" /replace="cytosine"

35 <220>
<221> misc_feature
<222> (8)..(10)
<223> n es a, c, g, t o u

40 <220>
<221>variación
<222> (18)..(19)
<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="" /replace="uracile" /replace="cytosine"

45 <220>
<221> misc_feature
<222> (18)..(19)
<223> n es a, c, g, t o u

<400> 60

55 nnhhhrgnnyhbthrdnnc ydhhhh 26
<210> 61
<211> 26
<212> DNA
<213>Secuencia Artificial

60 <220>
<223>secuencia tallo-lazo de histona según la fórmula(Ile)

65 <220>
<221>variación
<222> (1)..(1)

ES 2 525 556 T3

<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="" /replace="uracile" /replace="cytosine"

5 <220>
 <221> misc_feature
 <222> (1)..(1)
 <223> n es a, c, g, t o u

10 <220>
 <221>variación
 <222> (8)..(8)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="" /replace="uracile" /replace="cytosine"

15 <220>
 <221> misc_feature
 <222> (8)..(8)
 <223> n es a, c, g, t o u

20 <220>
 <221>variación
 <222> (19)..(19)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="" /replace="uracile" /replace="cytosine"

25 <220>
 <221> misc_feature
 <222> (19)..(19)
 <223> n es a, c, g, t o u

30 <400> 61
 nhhhhrgrndb yhythrdhnc yrhhhh 26

35 <210> 62
 <211> 26
 <212> DNA
 <213>Secuencia Artificial

40 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ile)

<400> 62
 mhwmrgyky ywytwrrmrc yrymmm 26

45 <210> 63
 <211> 26
 <212> DNA
 <213>Secuencia Artificial

50 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ile)

<400> 63
 mhwamggcty twytwarmgc cacmmm 26

55 <210> 64
 <211> 26
 <212> DNA
 <213>Secuencia Artificial

60 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(Ile)

<400> 64
 acaaaggctt ttttaagagc caccaa 26

65

ES 2 525 556 T3

<210> 65
 <211> 26
 <212> DNA
 <213>Secuencia Artificial
 5
 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(If)

 <220>
 10 <221>variación
 <222> (26)..(26)
 <223>/replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

 15 <220>
 <221> misc_feature
 <222> (26)..(26)
 <223> n es a, c, g, t o u

 20 <400> 65
 hhhhmvgyyy tyhthryrc bvmhhn 26

 <210> 66
 <211> 26
 25 <212> DNA
 <213>Secuencia Artificial

 <220>
 <223>secuencia tallo-lazo de histona según la fórmula(If)
 30
 <400> 66
 mmmmmmsgyyc ttytmagrrc smchhh 26

 <210> 67
 <211> 26
 35 <212> DNA
 <213>Secuencia Artificial

 <220>
 40 <223>secuencia tallo-lazo de histona según la fórmula (If)

 <400> 67
 mmmmmmsgyyc tttmagrrc sachmh 26

 45 <210> 68
 <211> 26
 <212> DNA
 <213>Secuencia Artificial

 50 <220>
 <223>secuencia tallo-lazo de histona según la fórmula (If)

 <400> 68
 55 mmmmmmsgyyy yytmmrrrc smmmmw 26

 <210> 69
 <211> 26
 <212> DNA
 <213>Secuencia Artificial
 60
 <220>
 <223>secuencia tallo-lazo de histona según la fórmula (If)

 <400> 69
 65 mmmmmmsgcyc tttcagrgc cacmmw 26

<210> 70
 <211> 26
 <212> DNA
 <213>Secuencia Artificial
 5
 <220>
 <223>secuencia tallo-lazo de histona según la fórmula (IIf)
 <400> 70
 10 acaacggccc tttcagggc caccaa 26
 <210> 71
 <211> 26
 <212> DNA
 15 <213>Secuencia Artificial
 <220>
 <223>secuencia tallo-lazo de histona según la fórmula (IIg)
 20 <220>
 <221>variación
 <222> (24)..(25)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""
 25 <220>
 <221> misc_feature
 <222> (24)..(25)
 <223> n es a, c, g, t o u
 30 <400> 71
 hhvamggyyc ttythagrrc cvhnnm 26
 <210> 72
 35 <211> 26
 <212> DNA
 <213>Secuencia Artificial
 <220>
 40 <223>secuencia tallo-lazo de histona según la fórmula (IIg)
 <400> 72
 hhaamggcyc ttytmagrgc cvchhm 26
 <210> 73
 45 <211> 26
 <212> DNA
 <213>Secuencia Artificial
 <220>
 50 <223>secuencia tallo-lazo de histona según la fórmula (IIg)
 <400> 73
 mhaamggctc tttmagrgc cmcymm 26
 <210> 74
 55 <211> 26
 <212> DNA
 <213>Secuencia Artificial
 <220>
 60 <223>secuencia tallo-lazo de histona según la fórmula (IIg)
 <400> 74
 mmmmmrgyyy ykytmasrrc bmmymm 26
 65 <210> 75

<211> 26
 <212> DNA
 <213>Secuencia Artificial

5 <220>
 <223>secuencia tallo-lazo de histona según la fórmula (IIg)

<400> 75
 mcaamggctc tttmagagc caccmm 26

10 <210> 76
 <211> 26
 <212> DNA
 <213>Secuencia Artificial

15 <220>
 <223>secuencia tallo-lazo de histona según la fórmula (IIg)

<400> 76
 20 acaaaggctc tttcagagc caccca 26

<210> 77
 <211> 26
 <212> DNA
 25 <213>Secuencia Artificial

<220>
 <223>secuencia tallo-lazo de histona según la fórmula (IIh)

30 <220>
 <221>variación
 <222> (1)..(1)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"
 /replace=""

35 <220>
 <221> misc_feature
 <222> (1)..(1)
 <223> n es a, c, g, t o u

40 <220>
 <221>variación
 <222> (25)..(25)
 <223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="uracile" /replace="cytosine"

45 /replace=""

<220>
 <221> misc_feature
 <222> (25)..(25)
 50 <223> n es a, c, g, t o u

<400> 77
 nhaahkgcyc trythagrrc cvhbnh 26

55 <210> 78
 <211> 26
 <212> DNA
 <213>Secuencia Artificial

60 <220>
 <223>secuencia tallo-lazo de histona según la fórmula (IIh)

<220>
 <221>variación
 65 <222> (25)..(25)

ES 2 525 556 T3

<223> /replace="guanine" /replace="adenine" /replace="thymine" /replace="" /replace="uracile" /replace="cytosine"

<220>

5 <221> misc_feature
<222> (25)..(25)
<223> n e s a , c , g , t o u

<400> 78

10 hhaamggcyc tttthagrc cvmynm 26

<210> 79
<211> 26
<212> DNA

15 <213>Secuencia Artificial

<220>
<223>secuencia tallo-lazo de histona según la fórmula (IIh)

20 <400> 79
hmaaaggcyc tttthagrc crmyhm 26

<210> 80
<211> 26
<212> DNA

25 <213>Secuencia Artificial

<220>

30 <223>secuencia tallo-lazo de histona según la fórmula (IIh)

<400> 80
mmaamkghyc tkytmasrc crmyym 26

35 <210> 81
<211> 26
<212> DNA

<213>Secuencia Artificial

40 <220>
<223>secuencia tallo-lazo de histona según la fórmula (IIh)

<400> 81
mmaamggcyc tttmagrc crmyym 26

45 <210> 82
<211> 26
<212> DNA

50 <213>Secuencia Artificial

<220>
<223>secuencia tallo-lazo de histona según la fórmula (IIh)

<400> 82

55 ccaaaggctc tttcagagc caccca 26

<210> 83
<211> 1747
<212> DNA

60 <213>Secuencia Artificial

<220>
<223> ppLuc(GC) - ag

65 <400> 83

ES 2 525 556 T3

gggagaaagc uugaggauvg aggacgcca gaacaucaag aagggcccgg cgcccuucua 60
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgcccu 120
 ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga 180
 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggpgc 300
 ccucuucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu 360
 gaacagcaug gggauacagc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480
 gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgcggg 540
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgccugau 600
 caugaacagc agcggcagca ccggccugcc gaagggggug gccucgccc accggaccgc 660
 cugcgugcgc uucucgcacg cccgggacc caucuucggc aaccagauca ucccggacac 720
 cgccauccug agcggugugc cguuccacca cggcuucggc auguucacga cccugggcu 780
 ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg 840
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugug ucagcuucuu 900
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960
 gggcgcctcg cugagcaagg agguuggcga ggccguggcc aagcgguucc accucccggg 1020
 cauccgccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca cccccgaggg 1080
 ggacgacaag cccgggcgcc ugggcaaggu ggucccguuc uucgaggcca agguuggugga 1140
 ccuggacacc ggcaagacc uggcgugaa ccagcggggc gagcugugcg ugcgggggccc 1200
 gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgcccua ucgacaagga 1260
 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgcgg cccgagcugga 1380
 gagcauccug cuccagcacc ccaacaucuu cgacgccggc guggccgggc ugccggacga 1440
 cgacgccggc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga 1500
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaagaagc ugcggggcgg 1560

 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620
 ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguuaa 1680
 agacugacua gcccgaugg ccucccaacg gcccuccuc cccuccuugc accgagauua 1740
 auagauc 1747

- <210> 84
- <211> 1806
- <212> DNA
- 5 <213>Secuencia Artificial
- <220>
- <223> ppLuc(GC) - ag - A64
- 10 <400> 84

ES 2 525 556 T3

gggagaaagc uugaggauagg aggacgccaa gaacaucaag aagggcccgg cgcccuucua 60
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgcccu 120
 ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga 180
 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc 300
 ccucuuauc ggcguggccg ucgccccggc gaacgcauc uacaacgagc gggagcugcu 360
 gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480
 gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgccggg 540
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600
 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660
 cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagavca ucccggacac 720
 cgccauccug agcguggugc cguuccacca cggcuucggc auguucacga cccugggcua 780
 ccucaucugc ggcuuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugc 840
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccuguu ucagcuucuu 900
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960
 gggcgccccg cugagcaagg agguuggcga ggccguggcc aagcgguucc accuccggg 1020
 cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgagg 1080
 ggacgacaag ccgggcgccc ugggcaaggu ggucccguuc uucgaggcca agguugguga 1140
 ccuggacacc ggcaagacc ugggcgugaa ccagcggggc gagcugugcg ugcgggggcc 1200
 gaugaucaug agcggcuacg ugaacaacc ccaggccacc aacgcccua ucgacaagga 1260
 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga 1380
 gagcauccug cuccagcacc ccaacaucuu cgacgcgggc guggccgggc ugccggacga 1440
 cgacgccggc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga 1500

gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcgg 1560
 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620
 ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguuaa 1680
 agacugacua gcccgauagg ccucccaacg ggccuccuc cccuccuugc accgagauua 1740
 auaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800
 aaaaaa 1806

- <210> 85
- <211> 1772
- <212> DNA
- 5 <213>Secuencia Artificial
- <220>
- <223> ppLuc(GC) - ag - histonaSL
- 10 <400> 85

ES 2 525 556 T3

gggagaaagc uugaggauwg aggacgcca gaacaucaag aagggcccg cgcccuucua 60
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgcccu 120
 ggugccgggc acgauwgccu ucaccgacgc ccacauwgag gucgacauc cuuacgwgga 180
 guacuucgag augagwgugc gccuggccga ggccaugaag cgguacwgcc ugaacaccaa 240
 ccaccggauw guggugugcu cggagaacag ccugcaguuc uucauwgccg ugcuggwggc 300
 ccucuucawc ggcguggccg ucgcccwggc gaacgacawc uacaacwgagc gggagwgugc 360
 gaacagcaug gggauwgagc agccgaccgu gguguucwgug agcaagaagc gccugcagaa 420
 gauccugaac gwgcagaaga agcwgcccau cauccagaag aucaucauca uggacagcaa 480
 gaccgacuac cagggcuuc agucgaugua cacguucwgug accagccacc ucccwgccgg 540
 cuucaacwgag uacgacuucg ucccwgagag cuucgaccwg gacaagacca ucgcccugau 600
 caugaacagc agcggcagca ccggccwgcc gaagggggwg gcccwgccgc accggaccgc 660
 cwgcgugwgc uucucgacg cccgggacc caucuucwgc aaccagauca ucccwgacac 720
 cgccaucwg agcwgugwg cguuccacca cggcuucwgc auguucacga ccwgggcua 780
 ccucaucwg gcguuccwg ugguccugau guaccgguuc gaggaggagc uguuccwgcg 840
 gagccwgag gacuacaaga uccagagwg gcwgucwgug ccgaccwgug ucagcuucuu 900
 cgccaagagc acccwgauwg acaaguacga ccwgucgaac cwgacwgaga ucgccagwg 960
 gggcgcccc cwgagcaag agguggwgga ggccwgwgcc aagcwggucc accuccwggg 1020
 cauccwgccg ggcuawgcc ugaccwgagc cacwgagwg gc auccwgauca ccccgaggwg 1080
 ggacgacaag cggggwgcc ugggcaaggu ggucccwguc uucgaggcca agguggwgga 1140
 ccwgacacc ggcaagacc ugggcgugaa ccagcggggc gagcwgugwg ucgggggcc 1200
 gaugaucaug agcggcuac ugaacaaccc ggaggccacc aacgcccua ucgacaagga 1260
 cggcwgwgug cacagwgwg acaucwgcu cugggacwgag gagwgacacu ucuucaucgu 1320
 cgaccwgwg aagucwguga ucaaguacaa gggcuaccag guggcwgccg ccgagcwgga 1380
 gagcaucwg cuccagcacc ccaacaucu cgacwgccwg guggccwggc ugccwgacga 1440
 cgacwgccwg gagcwgccg ccgwgugwg ggwgucwgag cacwgcaaga ccaugacwgga 1500
 gaaggagauw gucgacuac ugccagcca ggugaccacc gccagaagc ugcggggcg 1560
 cguggwguc guggacwgag ucccgaaggg ccwgaccwg aagcucwgac cccwggaagau 1620
 ccgwgagauw cugaucaag ccaagaaggg cggcaagauw gccwguaag acuauguaua 1680
 agacwgacu gcccwgug ccuccaawc ggccuccuc cccuccuugc accgagauua 1740
 auagauca aagwgcuuu ucagagccac ca 1772

- <210> 86
- 5 <211> 1835
- <212> DNA
- <213> Secuencia Artificial

- <220>
- 10 <223> ppLuc(GC) - ag - A64 - histonaSL

- <400> 86

ES 2 525 556 T3

gggagaaagc uugaggaugg aggacgccaa gaacaucaag aagggcccgg cgcccuucua 60
 cccgcuggag gacgggaccg cggcgagca gcuccacaag gccaugaagc gguacgcccu 120
 ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauc ccuacgcgga 180
 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc 300
 ccucuuauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu 360
 gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480
 gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgccggg 540
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600
 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660
 cugcgugcgc uucucgcacg cccgggacc caucuucggc aaccagauca ucccggacac 720
 cgccauccug agcguugguc cguuccacca cggcuucggc auguucacga ccugggcua 780
 ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugc 840
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugc ucagcuucuu 900
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960
 gggcgccccg cugagcaagg aggugggcca ggccguggcc aagcgguucc accucccggg 1020
 cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgaggg 1080
 ggacgacaag ccgggcgccc ugggcaaggu ggucccguuc uucgaggcca agguggugga 1140
 ccuggacacc ggcaagacc uggcgugaa ccagcggggc gagcugugc ugcgggggccc 1200
 gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgcccua ucgacaagga 1260
 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga 1380

 gagcauccug cuccagcacc ccaacaucuu cgacgccggc guggccgggc ugccggacga 1440
 cgacgccggc gagcugccgg ccgcgguugu ggugcuggag cacggcaaga ccaugacgga 1500
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaaagaagc ugcggggcgg 1560
 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620
 ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguua 1680
 agacugacua gcccgauagg ccuccaacg ggccuccuc cccuccuugc accgagauua 1740
 auaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1800
 aaaaaaugca ucaaaggcuc uuuucagagc cacca 1835

- <210> 87
- <211> 1869
- 5 <212> DNA
- <213> Secuencia Artificial

- <220>
- <223> ppLuc(GC) - ag - A120
- 10 <400> 87

ES 2 525 556 T3

```

gggagaaagc uugaggauwg aggacgcca gaacaucaag aagggcccg cgcccuucua      60
cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgcccu      120
ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga      180
guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa      240
ccaccggauw guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc      300
ccucuucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu      360
gaacagcaug gggauwagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa      420
gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa      480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgccggg      540
cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau      600
caugaacagc agcggcagca ccggccugcc gaagggggug gccucgccc accggaccgc      660
cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac      720
cgccauccug agcguwgugc cguuccacca cggcuucggc auguucacga cccugggcuw      780
ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg      840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugug ucagcuucuu      900
cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg      960
gggcgccccg cugagcaagg aggugggcga ggccguggcc aagcgguucc accuccggg      1020
cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgaggg      1080
ggacgacaag ccgggcgccc ugggcaaggu ggucccguuc uucgaggcca agguggugga      1140
ccuggacacc ggcaagacc ugggcgugaa ccagcggggc gagcugugcg ugcgggggccc      1200
gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccuca ucgacaagga      1260
cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu      1320
cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga      1380
gagcauccug cuccagcacc ccaacaucuu cgacgccggc guggccgggc ugccggacga      1440
cgacgccggc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga      1500
gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaugaagc ugcggggcg      1560
cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau      1620
ccgcgagauc cugaucaagg ccaagaagg cggaagauw gccguguaag acuaguuaa      1680
agacugacua gcccgauggg ccucccaacg ggcccuccuc cccuccuugc accgagauua      1740
auagaucuaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      1800
aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      1860
aaaaaaaaaa                                     1869

```

- <210> 88
- <211> 1858
- 5 <212> DNA
- <213> Secuencia Artificial
- <220>
- <223> ppLuc(GC) - ag - A64 - ag
- <400> 88

ES 2 525 556 T3

gggagaaagc uugaggaugg aggacgcca gaacaucaag aaggggcccgg cgcccuucua 60
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgcccu 120
 ggugccgggg acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga 180
 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggcg 300
 ccucuucauc ggcgugggcc ucgccccggc gaacgacauc uacaacgagc gggagcugcu 360
 gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480
 gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccggcggg 540
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600
 caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc 660
 cugcgugcgc uucucgcacg cccgggacct caucuucggc aaccagauca ucccggacac 720
 cgccauccug agcgguggugc cguuccacca cggcuucggc auguucacga cccugggcu 780
 ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg 840
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugug ucagcuucuu 900
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960
 gggcgccccg cugagcaagg aggugggcga ggccguggcc aagcgguucc accuccggg 1020
 cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgagg 1080
 ggacgacaag ccgggcgccc ugggcaaggu ggucccguuc uucgaggcca agguggugga 1140

 ccuggacacc ggcaagacc uggcgugaa ccagcggggc gagcugugcg ugcgggggcc 1200
 gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgcccua ucgacaagga 1260
 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320
 cgaccggcug aagucgcuga ucaagucaa gggcuaccag guggcgccgg ccgagcugga 1380
 gagcauccug cuccagcacc ccaataucuu cgacgccggc guggccgggc ugccggacga 1440
 cgacgccggc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga 1500
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcgg 1560
 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620
 ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguuaa 1680
 agacugacua gcccgauggg ccucccaacg ggccuccuc cccuccuugc accgagauua 1740
 auaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800
 aaaaaaugca uccugcccga ugggccuccc aacgggccc ucccccucc uugcaccg 1858

<210> 89

<211> 1894

<212> DNA

5 <213>Secuencia Artificial

<220>

<223> ppLuc(GC) - ag - A64 - aCPSL

10 <400> 89

ES 2 525 556 T3

gggagaaagc uugaggauagg aggacgcaa gaacaucaag aagggcccgg cgcccuucua 60
 cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgcccu 120
 ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga 180
 guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa 240
 ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc 300
 ccucuucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu 360
 gaacagcaug gggauagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa 420
 gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa 480
 gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgccggg 540
 cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau 600
 caugaacagc agcggcagca ccggccugcc gaagggggug gccucgccc accggaccgc 660
 cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac 720
 cgccauccug agcguuggugc cguuccacca cggcuucggc auguucacga cccugggcua 780
 ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg 840
 gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugug ucagcuucuu 900
 cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg 960
 gggcgccccg cugagcaagg agguggggcga ggccguggcc aagcggguucc accuccggg 1020
 cauccgccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca cccccgagg 1080
 ggacgacaag ccgggcgccc ugggcaaggu ggucccguuc uucgaggcca agguggugga 1140
 ccuggacacc ggcaagacc ugggcgugaa ccagcggggc gagcugugcg ugcgggggccc 1200
 gaugaucaug agcggcuacg ugaacaaccc ggaggccacc aacgcccua ucgacaagga 1260
 cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu 1320
 cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccc cggagcugga 1380
 gagcauccug cuccagcacc ccaacauuu cgacgccggc guggccgggc ugccggacga 1440
 cgacgccggc gagcugcccg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga 1500
 gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaaagaagc ugcggggcgg 1560
 cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau 1620
 ccgcgagauc cugaucaagg ccaagaagg cggaagauc gccguguaag acuaguuaa 1680
 agacugacua gcccgauggg ccucccaacg ggcuccuc cccuccuugc accgagauua 1740
 aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa 1800
 aaaaaaugca ucaauuccua cacgugaggc gcugugauuc ccuaucggc uucauuccu 1860
 auacauuagc acagcgccau ugcauguagg aauu 1894

- <210> 90
- 5 <211> 1909
- <212> DNA
- <213> Secuencia Artificial
- <220>
- 10 <223> ppLuc(GC) - ag - A64 - PolioCL

ES 2 525 556 T3

<400> 90

```

gggagaaagc uugaggauwg aggacgccaa gaacaucaag aagggcccg cgcccuucua      60
cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgcccu      120
ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga      180
guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa      240
ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc      300
ccucuucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu      360
gaacagcaug gggaucagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa      420
gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa      480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgccggg      540
cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau      600
caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc      660
cugcgugcgc uucucgcacg cccgggacc ccuucgugc aaccagauca ucccggacac      720
cgccauccug agcguuggug cguuccacca cggcuucggc auguucacga cccugggcu      780
ccucaucugc ggcuuuccggg ugguccugau guaccgguuc gaggaggagc uguuuccugc      840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccuugu ucagcuucuu      900

cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg      960
gggcgccccg cugagcaagg aggugggca ggcguggcc aagcgguucc accuccggg      1020
cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgagg      1080
ggacgacaag ccgggcgccg ugggcaaggu ggucccguuc uucgaggcca agguggugga      1140
ccuggacacc ggcaagacc uggcgugaa ccagcggggc gagcugugcg ugcggggcc      1200
gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgcccua ucgacaagga      1260
cggcuggcug cacagcggc acaucgccua cugggacgag gacgagcacu ucuucaucgu      1320
cgaccggcug aagucgcuga ucaaguacaa ggcuaaccag guggcgccgg ccgagcugga      1380
gagcauccug cuccagcacc ccaacauuu cgacgcggc guggccggc ugccggacga      1440
cgacgcggc gagcugccgg ccgcguggu ggugcuggag cacggcaaga ccaugacgga      1500
gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaagaagc ugcggggcg      1560
cgugguguuc guggacgag ucccgaagg ccugaccggg aagcucgacg cccggaagau      1620
ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguuaa      1680
agacugacua gcccgaugg ccucccaacg ggccuccuc cccuccuugc accgagaua      1740
aaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      1800
aaaaaauca ucaauucuaa aacagcucug gguuguuacc cccccagag gccacgugg      1860
cggcuaguac uccgguauug cgguaaccuu guacgccugu uuuagaauu      1909

```

<210> 91

<211> 1841

<212> DNA

<213>Secuencia Artificial

<220>

5

ES 2 525 556 T3

<223> ppLuc(GC) - ag - A64 - G30

<400> 91

```

gggagaaagc uugaggauagg aggacgcca gaacaucaag aagggcccgg cgcccuucua      60
cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgccc      120
ggugccgggc acgaucccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga      180
guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa      240
ccaccggauc guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcg      300
ccucuucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu      360
gaacagcaug gggauccagc agccgaccgu gguguucgug agcaagaagg gccugcagaa      420
gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa      480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgcggg      540
cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau      600
caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc      660
cugcgugcgc uucucgcacg cccgggacc caucuucggc aaccagauca ucccggacac      720
cgccauccug agcugggugc cguuccacca cggcuucggc auguucacga cccugggcua      780
ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg      840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugug ucagcuucuu      900
cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg      960
gggcgccccg cugagcaagg aggugggcga ggccguggcc aagcgggucc accuccggg      1020
cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgagg      1080
ggacgacaag ccgggcgccc ugggcaaggu ggucccguuc uucgaggcca agguggugga      1140
ccuggacacc ggcaagacc ugggcgugaa ccagcggggc gagcugugcg ugcgggggccc      1200
gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgcccua ucgacaagga      1260
cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu      1320
cgaccggcug aagucgcuga ucaaguacaa ggcuaaccag guggcggcg ccgagcugga      1380
gagcauccug cuccagcacc ccaacauuu cgacgccggc guggccggg ugcgggacga      1440
cgacgccggc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga      1500
gaaggagauc gucgacuacg uggccagcca ggugaccacc gccagaagc ugcggggcgg      1560
cgugguguuc guggacgagg ucccgaagg ccugaccggg aagcucgacg cccggaagau      1620
ccgcgagauc cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguuaa      1680
agacugacua gcccgaugg ccucccaacg ggcccuccuc cccuccuugc accgagaua      1740
auaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa      1800
5 aaaaaaugca uggggggggg gggggggggg gggggggggg g      1841

```

<210> 92

<211> 1841

<212> DNA

10 <213>Secuencia Artificial

<220>

<223> ppLuc(GC) - ag - A64 - U30

ES 2 525 556 T3

	<400> 92	
	gggagaaagc uugaggauagg aggacgcca gaacaucaag aagggcccgg cgcccuucua	60
	cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaagc gguacgcccu	120
	ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga	180
	guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa	240
	ccaccggau c guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcugggcgc	300
	ccucuucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu	360
	gaacagcaug gggaucaagg agccgaccgu gguguucgug agcaagaagg gccugcagaa	420
	gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa	480
	gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccgccggg	540
	cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau	600
	caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc	660
	cugcgugcgc uucucgcacg cccgggaccc caucuucggc aaccagauca ucccggacac	720
	cgccauccug agcguugguc cguuccacca cggcuucggc auguucacga cccugggcua	780
	ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg	840
	gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccugu ucagcuucuu	900
	cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg	960
	gggcgccccg cugagcaagg aggugggcga ggccguggcc aagcgguucc accuccggg	1020
	cauccgccag ggcuacggcc ugaccgagac cacgagcgcg auccugauca cccccgagg	1080
	ggacgacaag ccgggcgccc ugggcaaggu ggucccguuc uucgaggcca agguggugga	1140
	ccuggacacc ggcaagacc ugggcgugaa ccagcggggc gagcugugcg ugcggggcc	1200
	gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgcccua ucgacaagga	1260
	cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu	1320
	cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcggcg ccgagcugga	1380
	gagcauccug cuccagcacc ccaacauuu cgacgccggc guggccgggc ugccggacga	1440
	cgacgccggc gagcugccgg ccgcgguggu ggugcuggag cacggcaaga ccaugacgga	1500
	gaaggagauc gucgacuacg uggccagcca ggugaccacc gccaagaagc ugcggggcgg	1560
	cgugguguuc guggacgagg ucccgaaggg ccugaccggg aagcucgacg cccggaagau	1620
	ccgcgagauc cugaucaagg ccaagaaggg cggcaagauc gccguguaag acuaguuaa	1680
	agacugacua gcccgauggg ccucccaacg ggcccuccuc ccuccuugc accgagauua	1740
	auaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa	1800
5	aaaaaaaaugca uuuuuuuuuuu uuuuuuuuuuu uuuuuuuuuuu u	1841
	<210> 93	
	<211> 1857	
	<212> DNA	
10	<213>Secuencia Artificial	
	<220>	
	<223> ppLuc(GC) - ag - A64 - SL	

ES 2 525 556 T3

<400> 93
 gggagaaagc uugaggau^gg aggacgcca^a gaacauca^ag aagggccc^gg cgcccuuc^uca 60
 cccgcugg^ag gacgggacc^g ccggcgag^{ca} gcuccaca^{ag} gccauga^{ag}c gguacgcc^{cu} 120
 ggugccgg^gc acgau^cgccu ucaccgac^{gc} ccacaucg^{ag} gucgaca^{uca} ccuacgcg^{ga} 180
 guacuucg^{ag} augagc^gugc gccuggcc^{ga} ggccauga^{ag} cgguacgg^{cc} ugaacacca^a 240
 ccaccgga^{uc} guggugug^{cu} cggagaac^{ag} ccugcagu^{uc} uucaugcc^{gg} ugcugggc^{gc} 300
 ccucuuca^{uc} ggcguggc^g ucgcccc^{gg}c gaacgaca^{uc} uacaacg^{ag}c gggagcug^{cu} 360
 gaacagca^{ug} gggauca^gcc agccgacc^{gu} gguguucg^{ug} agcaaga^{ag} gccugcag^{aa} 420
 gauccuga^{ac} gugcaga^{aga} agcugccc^{au} cauccaga^{ag} aucauca^{uca} uggacag^{caa} 480
 gaccgacu^{ac} cagggcu^{ucc} agucgaug^{ua} cacguucg^{ug} accagcc^{acc} ucccgcgg^g 540
 cuucaacg^{ag} uacgacu^{ucg} ucccggag^{ag} cuucgacc^{gg} gacaagac^{ca} ugcgccug^{au} 600
 caugaac^{ag}c agcggcag^{ca} ccggccug^{cc} gaaggggg^{ug} gccucg^{cc}g accggacc^{gc} 660
 cugcgu^cgcg uucucgc^{acg} cccgggac^{ccc} caucuucg^{gc} aaccagau^{ca} ucccggac^{ac} 720
 cgccaucc^{ug} agcgu^{gg}ugc cguuccac^{ca} cggcuucg^{gc} auguucac^{ga} cccugggc^{ua} 780
 ccucauc^{ug}c ggcuu^{cc}ggg ugguccug^{au} guaccggu^{uc} gaggagg^{ag}c uguuccug^{cg} 840
 gagccug^cag gacuaca^{aga} uccagag^cgc gcugcucg^{ug} ccgacc^{cu}gu ucagcuu^{cu} 900
 cgccaag^{ag}c acccugau^c acaaguac^{ga} ccugucga^{ac} cugcacg^{aga} ugc^{cc}agcgg 960
 gggcgccc^{cg} cugagca^{ag}g agguggg^cga ggccgugg^{cc} aagcgg^{uu}cc accucc^{gg}g 1020
 cauccgcc^{ag} ggcua^cggcc ugaccgag^{ac} cacgagc^{gc}g auccugau^{ca} cccccg^{agg}g 1080
 ggacgaca^{ag} ccgggcg^{cc}g ugggcaag^{gu} ggucccgu^{uc} uucgagg^{cca} agguggu^{gga} 1140
 ccuggac^{acc} ggcaag^{acc} ugggcgug^{aa} ccagcggg^{gc} gagcugug^{cg} ugcgggg^gcc 1200
 gaugauca^{ug} agcggcu^{acg} ugaaca^{acc} ggaggcc^{acc} aacgccc^{uca} ucgaca^{agga} 1260
 cggcug^{gc}ug cacagcgg^{cg} acaucg^{cc}ua cugggacg^{ag} gacgagc^{acu} ucuucau^{cgu} 1320
 cgaccgg^cug aagucg^cuga ucaagu^{aca}a gggcuacc^{ag} guggcgc^{cc}g ccgagcug^{ga} 1380
 gagcaucc^{ug} cuccagc^{acc} ccaaca^{uc}uu cgacgc^{cc}ggc guggccg^{gg}c ugccggac^{ga} 1440
 cgacgc^{cc}ggc gagcug^{cc}gg ccgcgg^{ug}gu ggugcugg^{ag} cacggca^{aga} ccaugac^{gga} 1500
 gaaggaga^{uc} gucgacu^{acg} uggccag^{cca} ggugacc^{acc} gccaa^{ga}agc ugcgggg^cg 1560
 cguggugu^{uc} guggacg^{agg} ucccga^{agg}g ccugacc^{gg}g aagcucg^{acg} cccgga^{agau} 1620
 ccgcgaga^{uc} cugauca^{agg} ccaaga^{agg}g cggcaag^{auc} gccguga^{aag} acuagu^{uaua} 1680
 agacugac^{ua} gcccga^{ugg}g ccuccca^{acg} ggcccuc^{cc}uc cccuccu^{ug}c accgaga^{uua} 1740
 auaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa aaaaaaaaa 1800
 aaaaaaug^{ca} uuauggc^{gg}c cgugucc^{acc} acggau^{auca} ccguggu^{gga} cgcg^{gcc} 1857

5

<210> 94
 <211> 1838
 <212> DNA
 <213>Secuencia Artificial

10

<220>
 <223> ppLuc(GC) - ag - A64 - N32

<400> 94

ES 2 525 556 T3

```

gggagaaagc uugaggauwg aggacgcca gaacaucaag aagggcccg cgcccuucua      60
cccgcuggag gacgggaccg ccggcgagca gcuccacaag gccaugaac gguacgcccc      120
ggugccgggc acgaucgccu ucaccgacgc ccacaucgag gucgacauca ccuacgcgga      180
guacuucgag augagcgugc gccuggccga ggccaugaag cgguacggcc ugaacaccaa      240
ccaccggauw guggugugcu cggagaacag ccugcaguuc uucaugccgg ugcuggggcg      300
ccucuucauc ggcguggccg ucgccccggc gaacgacauc uacaacgagc gggagcugcu      360
gaacagcaug gggauwagcc agccgaccgu gguguucgug agcaagaagg gccugcagaa      420
gauccugaac gugcagaaga agcugcccau cauccagaag aucaucauca uggacagcaa      480
gaccgacuac cagggcuucc agucgaugua cacguucgug accagccacc ucccggccggg      540
cuucaacgag uacgacuucg ucccggagag cuucgaccgg gacaagacca ucgcccugau      600
caugaacagc agcggcagca ccggccugcc gaagggggug gccugccgc accggaccgc      660
cugcgugcgc uucucgcacg cccgggacc caucuucggc aaccagauca ucccggacac      720
cgccauccug agcguwgugc cguuccacca cggcuucggc auguucacga cccugggcua      780
ccucaucugc ggcuuccggg ugguccugau guaccgguuc gaggaggagc uguuccugcg      840
gagccugcag gacuacaaga uccagagcgc gcugcucgug ccgaccuugu ucagcuucuu      900
cgccaagagc acccugaucg acaaguacga ccugucgaac cugcacgaga ucgccagcgg      960
gggcgccccg cugagcaagg agguwggcga ggccguggcc aagcgguucc accuccggg      1020
cauccgccag ggcuaaggcc ugaccgagac cacgagcgcg auccugauca cccccgaggg      1080
ggacgacaag ccgggcgccc ugggcaaggu ggucccuuc uucgaggcca agguwgugga      1140
ccuggacacc ggcaagacc ugggcgugaa ccagcggggc gagcugugcg ugcggggggc      1200
gaugaucaug agcggcuacg ugaacaacc ggaggccacc aacgcccua ucgacaagga      1260
cggcuggcug cacagcggcg acaucgccua cugggacgag gacgagcacu ucuucaucgu      1320
cgaccggcug aagucgcuga ucaaguacaa gggcuaccag guggcgccgg ccgagcugga      1380
gagcauccug cuccagcacc ccaacauuu cgacgccggc guggccgggc ugccggacga      1440
cgacgccggc gagcugccgg ccgcgguwggu ggugcuggag cacggcaaga ccaugacgga      1500
gaaggagauw gucgacuacg uggccagcca ggugaccacc gccaaagaagc ugcggggcg      1560
cgugguguuc guggacgag ucccgaagg ccugaccggg aagcucgacg cccggaagau      1620
ccgcgagauw cugaucaagg ccaagaagg cggcaagauc gccguguaag acuaguuaa      1680
agacugacua gcccgaugg ccucccaacg ggccuccuc cccuccuugc accgagauua      1740
auaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa      1800
aaaaaaaaugca uccccucua gacaauwgga auuccaua      1838

```

- 5 <210> 95
- <211> 15
- <212> RNA
- <213> Artificial

10 <220>
 <223> Descripción de secuencia: secuencia estabilizante genérica de fórmula
 (C/U)CCANxCCC(U/A)PyxUC(C/U)CC

<220>
 <221>variación
 <222> (1)..(1)
 <223> /replace="cytosine" /replace="uracile"
 5

<220>
 <221> misc_feature
 <222> (1)..(1)
 <223>ácidoacid nucleico = citosina o uracilo
 10

<220>
 <221> misc_feature
 <222> (5)..(5)
 <223> Nx = a, g, c o u o cuanquier otro ácido nucleico
 15

<220>
 <221>variación
 <222> (5)..(5)
 <223> /replace="cytosine" /replace="uracile" /replace="guanosine" /replace="adenosine", o cualquier otro
 20 ácido nucleico

<220>
 <221> repeat_unit
 <222> (5)..(5)
 <223> x = cualquier número
 25

<220>
 <221> misc_feature
 <222> (9)..(9)
 <223>ácido nucleico = uracilo o adenosina
 30

<220>
 <221>variación
 <222> (9)..(9)
 <223> /replace="uracile" /replace="adenosine"
 35

<220>
 <221> misc_feature
 <222> (10)..(10)
 <223> Py = pyrimidine
 40

<220>
 <221> repeat_unit
 <222> (10)..(10)
 <223> x = cualquier número
 45

<220>
 <221>variación
 <222> (10)..(10)
 <223> /replace="pyrimidine"
 50

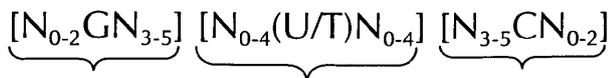
<220>
 <221> misc_feature
 <222> (13)..(13)
 <223>ácido nucleico = citosina o uracilo
 55

<220>
 <221>variación
 <222> (13)..(13)
 <223> /replace="cytosine" /replace="uracile"
 60

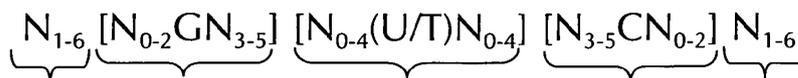
<400> 95
 nccanccnn ucnc
 65

REIVINDICACIONES

1. Secuencia de ácido nucleico que comprende o codifica para
- 5 a) una región de codificación que codifica para un péptido o proteína terapéuticamente activos, una proteína adyuvante, un antígeno, un antígeno tumoral, un antígeno patogénico, un antígeno animal, un antígeno viral, un antígeno protozoal, un antígeno bacteriano, un antígeno alérgico, un antígeno autoinmune, un alérgeno, un anticuerpo, un péptido o proteína inmunoestimulador o un antígeno específico del receptor de células T, siempre que la región de codificación no codifique para proteínas de histona, proteínas reporteras seleccionadas de EGFP y luciferasa y proteínas marcadoras o de selección seleccionadas de alfa-globina, galactoquinasa y xantina:guanina fosforribosil transferasa,
- 10 b) al menos un tallo-lazo de histona, y
- c) una secuencia poli(A) o una señal de poliadenilación,
- 15 para su uso con el fin de incrementar el nivel de expresión de dicha proteína codificada tal como se define en a) en el tratamiento de enfermedades cancerosas, enfermedades infecciosas, enfermedades autoinmunes, alergias o enfermedades alérgicas, enfermedades monogénicas, enfermedades genéticas, enfermedades cardiovasculares, enfermedades neuronales, enfermedades del sistema respiratorio, del sistema digestivo, de la piel, trastornos musculoesqueléticos, trastornos del tejido conectivo, inmunodeficiencias, enfermedades endocrinas, nutricionales y metabólicas, enfermedades oculares o del oído
- 20 2. Ácido nucleico para su uso según la reivindicación 1 en el tratamiento de enfermedades infecciosas virales, bacterianas o protozoológicas.
3. Ácido nucleico para su uso según la reivindicación 1, caracterizado porque el ácido nucleico no contiene uno o dos o al menos uno o todos pero uno o todos de los componentes del grupo consistente en: una secuencia que codifica para una ribozima o una ribozima de autoempalme, una secuencia de ácido nucleico viral, una señal de procesamiento de tallo-lazo de histona, en particular una secuencia de procesamiento de tallo-lazo de histona derivada del gen H2A614 de histona de ratón, un gen Neo, una secuencia promotora inactivada y una secuencia potenciadora inactivada.
- 25 4. Ácido nucleico para su uso según la reivindicación 1 ó 3, caracterizado porque el ácido nucleico no contiene una ribozima, preferentemente una ribozima de autoempalme, y uno del grupo consistente en: un gen Neo, una secuencia promotora inactivada, una secuencia potenciadora inactivada, una señal de procesamiento de tallo-lazo de histona, en particular una secuencia de procesamiento de tallo-lazo de histona derivada del gen H2A14 de histona de ratón.
- 30 5. Ácido nucleico para su uso según las reivindicaciones 1 a4, caracterizado porque el ácido nucleico es un ARN, preferentemente un ARNm.
- 35 6. Ácido nucleico para su uso según cualquiera de las reivindicaciones 1 a5, caracterizado porque el al menos un tallo-lazo de histona se selecciona de las siguientes fórmulas (I) o (II):



- 40 fórmula (II) (secuencia de tallo-lazo con elementos de borde de tallo):



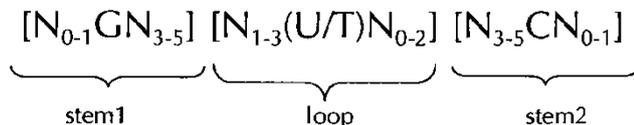
stem1 stem1 loop stem2 stem2
elemento de borde elemento de borde

- 45 donde:
los elementos de borde stem1 o stem2 N₁₋₆ es una secuencia consecutiva de 1 a 6, de preferencia de 2 a 6, muy preferiblemente de 2 a 5, aún más preferiblemente de 3 a 5, con total preferencia de 4 a 5 o 5 N, donde cada N se selecciona independientemente de entre un nucleótido seleccionado de A, U, T, G y C, o un análogo de nucleótido del mismo;
- 50 stem1 [N₀₋₂GN₃₋₅] es complementario inverso o complementario inverso parcialmente del elemento stem2, y es una secuencia consecutiva de entre 5 a 7 nucleótidos; donde N₀₋₂ es una secuencia

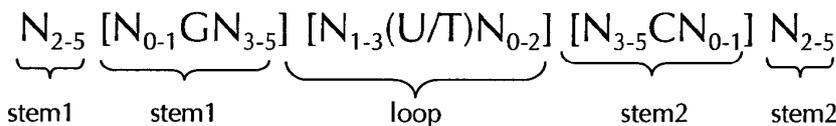
consecutiva de 0 a 2, de preferencia de 0 a 1, muy preferiblemente de 1 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C o un nucleótido análogo de los mismos; donde N₃₋₅ es una secuencia consecutiva de 3 a 5, de preferencia de 4 a 5, muy preferiblemente de 4 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C o un análogo de nucleótido del mismo, y donde G es guanosina o un análogo de la misma, y puede ser reemplazada opcionalmente por una citidina o un análogo de la misma siempre que su nucleótido citidina complementario en stem2 se reemplace por guanosina; la secuencia de asa [N₀₋₄(U/T)N₀₋₄] se ubica entre los elementos stem1 y stem2, y es una secuencia consecutiva de 3 a 5 nucleótidos, muy preferiblemente de 4 nucleótidos; donde cada N₀₋₄ es independientemente una secuencia consecutiva de 0 a 4, de preferencia de 1 a 3, muy preferiblemente de 1 a 2 N, donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C o un análogo de nucleótido del mismo; y donde U/T representa uridina, u opcionalmente timidina; stem2[N₃₋₅CN₀₋₂] es complementario inverso o complementario inverso parcialmente del elemento stem1, y es una secuencia consecutiva de entre 5 a 7 nucleótidos; donde N₃₋₅ es una secuencia consecutiva de 3 a 5, de preferencia de 4 a 5, muy preferiblemente de 4 N, en donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C o un análogo de nucleótido del mismo; donde N₀₋₂ es una secuencia consecutiva de 0 a 2, de preferencia de 0 a 1, muy preferiblemente de 1 N, en donde cada N se selecciona independientemente de un nucleótido seleccionado de A, U, T, G y C o un análogo de nucleótido del mismo; y donde C es citidina o un análogo de la misma, y puede ser reemplazada opcionalmente por una guanosina o un análogo de la misma siempre y cuando su nucleótido guanosina complementario en stem1 se reemplace por citidina;

donde stem1 y stem2 son capaces de apareamiento de bases entre sí formando una secuencia complementaria inversa, donde el apareamiento de bases puede ocurrir entre stem1 y stem2, o formando una secuencia complementaria parcialmente inversa, donde el apareo de bases incompleto puede ocurrir entre stem1 y stem2.

7. Ácido nucleico para su uso según la reivindicación 6, caracterizado porque el al menos un tallo-lazo de histona se selecciona de al menos una de las siguientes fórmulas (Ia) o (IIa):



Fórmula (Ia) (secuencia de tallo-lazo sin elementos de borde de tallo)



elemento de borde

elemento de borde

35 Fórmula (IIa) (secuencia de tallo-lazo con elementos de borde de tallo).

8. Ácido nucleico para su uso según cualquiera de las reivindicaciones 1 a7, caracterizado porque la secuencia poli(A) comprende una secuencia de aproximadamente 25 a alrededor de 400 nucleótidos de adenosina, preferentemente una secuencia de alrededor de 50 a aproximadamente 400 nucleótidos de adenosina, en especial una secuencia de alrededor de 50 a aproximadamente 300 nucleótidos de adenosina, en particular una secuencia de alrededor de 50 a aproximadamente 250 nucleótidos de adenosina, con especial preferencia una secuencia de alrededor de 60 a aproximadamente 250 nucleótidos de adenosina.

9. Ácido nucleico para su uso según cualquiera de las reivindicaciones 1 a7, caracterizado porque la señal de poliadenilación comprende la secuencia de consenso NNUANA, preferentemente AAUAAA o AUUAAA.

10. Ácido nucleico para su uso según cualquiera de las reivindicaciones 1 a 9, caracterizado porque el ácido nucleico es un ácido nucleico monocistrónico, bicistrónico o incluso multicistrónico.
- 5 11. Composición farmacéutica que comprende un ácido nucleico tal como se ha definido según cualquiera de las reivindicaciones 1 a 10 para su uso para incrementar la expresión de la proteína codificada por la región codificante de dicho ácido nucleico como se define en cualquiera de las reivindicaciones 1 a 10, en el tratamiento de enfermedades cancerosas, enfermedades infecciosas, enfermedades autoinmunes, alergias o enfermedades alérgicas, enfermedades monogénicas, enfermedades genéticas, enfermedades cardiovasculares, enfermedades neuronales, enfermedades del sistema respiratorio, enfermedades del sistema digestivo, enfermedades de la piel, trastornos musculoesqueléticos, trastornos del tejido conectivo, deficiencias inmunológicas, enfermedades endocrinas, nutricionales y metabólicas, enfermedades oculares o enfermedades del oído.
- 10 12. Composición farmacéutica para su uso según la reivindicación 11, que comprende un vehículo farmacéuticamente aceptable.
- 15 13. Composición farmacéutica para su uso según las reivindicaciones 11 o 12, en el tratamiento de enfermedades infecciosas virales, bacterianas o protozoológicas.
- 20 14. Vacuna que comprende un ácido nucleico según cualquiera de las reivindicaciones 1 a 13 que codifica para una proteína antigénica o que comprende un antígeno de proteína para uso para aumentar la expresión de dicho antígeno de proteína en el tratamiento de enfermedades cancerosas, enfermedades infecciosas, alergias o enfermedades alérgicas o enfermedades autoinmunes.
15. Vacuna para su uso según la reivindicación 14 en el tratamiento de enfermedades infecciosas virales, bacterianas o protozoológicas.
- 25 16. Vacuna para su uso según las reivindicaciones 14 o 15, caracterizada porque el antígeno se selecciona de entre antígenos tumorales, antígenos alérgicos, auto-antígenos autoinmunes, antígenos patogénicos, antígenos virales, antígenos bacterianos, antígenos fúngicos, antígenos protozoológicos, antígenos animales o antígenos alérgicos.

ppLuc(GC) – ag

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA
 CCCGCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCAUGCCGGUGCUGGGCGC
 CCUCUUCAUCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUCGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
 GAGCCUGCAGGACUACAAGAUCCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCCAGCGG
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUC AUGAGCGGCUACGUGAACAAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA
 CGGUGGUGCAGCAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGU
 CGACCGGCUGAAGUCGCUGAUC AAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
 GAGCAUCCUGCUC CAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCGGCCGGACGA
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCC GAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCCUGAUC AAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA *Agacuaguaua*
*agacugacua***GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCCUCCUUGCACCG***agauua*
auagauc-3'

FIGURA 6

ppLuc(GC) – ag – A64

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA
 CCCGCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCAUGCCGGUGCUGGGCGC
 CCUCUUCAUCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUAUCAUCAUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUCGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGUGGCCUCGCCGACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUAUCCCGGACAC
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCGCCAGCGG
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUCAUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGU
 CGACCGGCUGAAGUCGCUGAUAAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCUGAUAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA *Agacuaguaua*
agacugacua **GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCCUCCUUGCACCG** *agauua*
 auAA
 AAAAAA-3'

FIGURA 7

ppLuc(GC) – ag – histoneSL

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCAUGCCGGUGCUGGGCGC
 CCUCUUCAUCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUAUCAUCAUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCCGCCGGG
 CUUCAACGAGUACGACUUCGUUCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUCGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC
 CGCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
 GAGCCUGCAGGACUACAAGAUAUCCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCGCCAGCGG
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCCCGGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGGCAAGACCCUGGGCGUGAACCAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUCAUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGU
 CGACCGGCUGAAGUCGUGAUCGAUCAAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCUGAUCGAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAagacuaguua
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCUCCUCCCUCCUUGCACC**Gagauua
 auagaucuCAAAGGCUCUUUCAGAGCCACCA-3'

FIGURA 8

ppLuc(GC) – ag – A64 – histoneSL

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCCUUCUA
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCU
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCAUGCCGGUGCUGGGCGC
 CCUCUUCAUCGGCGUGGCCGUCGCCCGGCCAAGCACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
 GAGCCUGCAGGACUACAAGAUCCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCGCCAGCGG
 GGGCGCCCGCUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGGCAAGACCCUGGGCGUGAACCAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUC AUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCUCAUCGACAAGGA
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGU
 CGACCGGCUGAAGUCGCUGAUCAAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACCGCGGUGGCCGGGCGCCGGACGA
 CGACCGCGGCGAGCUGCCGGCCCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCCUGAUCAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAgacuaguaua
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCUCCUCCCUCCUUGCACCG**agauua
 auAA
 AAAAAAugcauCAAAGGCUCUUUCAGAGCCACCA-3'

FIGURA 9

ppluc(GC) – ag – A120

gggagaaagcuugaggaUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA
CCCGCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCCU
GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCUUCGCGGUGCUGGGCGC
CCUCUUCUUCGCGGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA
GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCGCGGG
CUUCAACGAGUACGACUUCGUCCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCCUGAU
CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC
CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUAUCCCGGACAC
CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUGGCCGACCCUGUUCAGCUUCUU
CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCCAGCGG
GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCGG
CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG
GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
CCUGGACACCGGCAAGACCCUGGGCGUGAACCAGCGGGGCGAGCUGUGCGUGCGGGGGCC
GAUGAUAUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA
CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCUUCGU
CGACCGGCUGAAGUCGCUGAUAAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA
CGACCGCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA
GAAGGAGAUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
CCGCGAGAUCUGAUAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAgacuaguaua
agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCCUCCUUGCACCG**agauua
auagaucuAAA
AAA
AAAAAAAAA-3'

FIGURA 10

ppLuc(GC) – ag – A64 – ag

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA
 CCCGUCGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUC AUGCCGGUGCUGGGCGC
 CCUCUUCAUCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG
 CUUCAACGAGUACGACUUCGUCCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUAUCCCGGACAC
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
 GAGCCUGCAGGACUACAAGAUCCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCGCCAGCGG
 GGGCGCCCCGCUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGGCAAGACCCUGGGCGUGAACCAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUAUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGU
 CGACCGGCUGAAGUCGUGAUCAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA
 CGACCGCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCCUGAUCAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA
 agacuaguaua
 agacugacua**GCCC GAUGGGCCUCCCAACGGGCCCUCUCCCUCCUUGCACCG**agauua
 auAA
 AAAAAAugcau**CCUGCCCGAUGGGCCUCCCAACGGGCCCUCUCCCUCCUUGCACCG**3'

FIGURA 11

ppLuc(GC) – ag – A64 – aCPSL

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA
 CCCGCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCUUCGCGGUGCUGGGCGC
 CCUCUUCUUCGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGUGGCCUGCCGCACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
 GAGCCUGCAGGACUACAAGAUCCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCGCCAGCGG
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUC AUGAGCGGCUACGUGAACAAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCUUCG
 CGACCGGCUGAAGUCGUGAUC AAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCUGAUC AAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAgacuaguaua
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCCUCUUGCACCG**agauua
 auAAA
 AAAAAAugcauCAAUUCUACACGUGAGGCGCUGUGAUUCCCUAUCCCCCUUCAUUCUU
 AUACAUUAGCACAGCGCCAUUGCAUGUAGGAAUU-3'

FIGURA 12

ppLuc(GC) – ag – A64 – PolioCL

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCCGGCGCCCUUCUA
 CCCGCUUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCAUGCCGGUGCUGGGCGC
 CCUCUUCAUCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCCGCCGGG
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUCGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
 CCUCAUCUGCGGCUUCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUCGUGCCGACCCUGUUCAGCUUCU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUUGUCGAACCUGCACGAGAUCCGACGCG
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCCGGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUC AUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGU
 CGACCGGCUGAAGUCGUGAUC AAGUACAAGGGCUACCAGGUGGGCGCCGCGGAGCUGGA
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCUGAUC AAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA *Agacuaguaua*
agacugacua **GCCCGAUGGGCCUCCCAACGGGCCCUCCUCCCUCCUUGCACCG** *agauua*
*au*AA
 AAAAA *Augcau* **CAAUUCUAAAACAGCUCUGGGGUUGUACCCACCCAGAGGCCACGUGG**
CGGCUAGUACUCCGGUAUUGCGGUACCCUUGUACGCCUGUUUAGAAUU-3'

FIGURA 13

ppLuc(GC) – ag – A64 – U30

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCAUGCCGGUGCUGGGCGC
 CCUCUUCAUCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCCGCGGG
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUAUCCCGGACAC
 CGCCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCGCCAGCGG
 GGGCGCCCCGCUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCCGGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCUGAUCACCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUCAUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCCUUCAUCGACAAGGA
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGU
 CGACCGGCUGAAGUCGUGAUCAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCUGAUCAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAGacuaguaua
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCUCCUUGCACCG**agauua
 auAA
 AAAAAAugcaUU-3'

FIGURA 15

ppLuc(GC) – ag – A64 – SL

gggagaaagcuugaggaUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA
 CCCGUCUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCCU
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCAUGCCGGUGCUGGGCGC
 CCUCUUCAUCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG
 CUUCAACGAGUACGACUUCGUUCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUCGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUCGCCGACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUAUCCCGGACAC
 CGCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCUCGGGCUA
 CCUCAUCUGCGGCCUUCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUUCUGCG
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCUU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCACGAGAUCGCCAGCGG
 GGGCGCCCGCUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUC AUGAGCGGCUACGUGAACAAACCCGGAGGCCACCAACGCCCUCAUCGACAAGGA
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUCAUCGU
 CGACCGGCUGAAGUCGCUGAUCAAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA
 CGACGCCGGCGAGCUGCCGGCCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAU CGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCCUGAUCAAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUA *Agacuaguaua*
agacugacua **GCCCGAUGGGCCUCCCAACGGGCCUCCUCCCUCCUUGCACCG** *agauua*
*au*AA
 AAAAA *AugcauUAUGGCGGCCGUGUCCACCACGGUAUACCGUGGUGGACGCGGCC* – 3'

FIGURA 16

ppLuc(GC) – ag – A64 – N32

gggagaaagcuugaggAUGGAGGACGCCAAGAACAUCAAGAAGGGCCCGGCGCCCUUCUA
 CCCGUGGAGGACGGGACCGCCGGCGAGCAGCUCCACAAGGCCAUGAAGCGGUACGCCC
 GGUGCCGGGCACGAUCGCCUUCACCGACGCCACAUCGAGGUCGACAUCACCUACGCGGA
 GUACUUCGAGAUGAGCGUGCGCCUGGCCGAGGCCAUGAAGCGGUACGGCCUGAACACCAA
 CCACCGGAUCGUGGUGUGCUCGGAGAACAGCCUGCAGUUCUUCAUGCCGGUGCUGGGCGC
 CCUCUUCAUCGGCGUGGCCGUCGCCCCGGCGAACGACAUCUACAACGAGCGGGAGCUGCU
 GAACAGCAUGGGGAUCAGCCAGCCGACCGUGGUGUUCGUGAGCAAGAAGGGCCUGCAGAA
 GAUCCUGAACGUGCAGAAGAAGCUGCCCAUCAUCCAGAAGAUCAUCAUCAUGGACAGCAA
 GACCGACUACCAGGGCUUCCAGUCGAUGUACACGUUCGUGACCAGCCACCUCGCCGGG
 CUUCAACGAGUACGACUUCGUCCCGGAGAGCUUCGACCGGGACAAGACCAUCGCCUGAU
 CAUGAACAGCAGCGGCAGCACCGGCCUGCCGAAGGGGGUGGCCUGCCGCACCGGACCGC
 CUGCGUGCGCUUCUCGCACGCCCGGGACCCCAUCUUCGGCAACCAGAUCAUCCCGGACAC
 CGCAUCCUGAGCGUGGUGCCGUUCCACCACGGCUUCGGCAUGUUCACGACCCUGGGCUA
 CCUCAUCUGCGGCUUCCGGGUGGUCCUGAUGUACCGGUUCGAGGAGGAGCUGUCCUGCG
 GAGCCUGCAGGACUACAAGAUCAGAGCGCGCUGCUCGUGCCGACCCUGUUCAGCUUCU
 CGCCAAGAGCACCCUGAUCGACAAGUACGACCUGUCGAACCUGCAGAGAUCCAGCGG
 GGGCGCCCCGUGAGCAAGGAGGUGGGCGAGGCCGUGGCCAAGCGGUUCCACCUCGCCGG
 CAUCCGCCAGGGCUACGGCCUGACCGAGACCACGAGCGCGAUCCUGAUCACCCCCGAGGG
 GGACGACAAGCCGGGCGCCGUGGGCAAGGUGGUCCCGUUCUUCGAGGCCAAGGUGGUGGA
 CCUGGACACCGGCAAGACCCUGGGCGUGAACAGCGGGGCGAGCUGUGCGUGCGGGGGCC
 GAUGAUC AUGAGCGGCUACGUGAACAACCCGGAGGCCACCAACGCCUCAUCGACAAGGA
 CGGCUGGCUGCACAGCGGCGACAUCGCCUACUGGGACGAGGACGAGCACUUCUUAUCGU
 CGACCGGCGUAAGUCGUGAUC AAGUACAAGGGCUACCAGGUGGCGCCGGCCGAGCUGGA
 GAGCAUCCUGCUCCAGCACCCCAACAUCUUCGACGCCGGCGUGGCCGGGCUGCCGGACGA
 CGACGCCGGCGAGCUGCCGGCCGCGGUGGUGGUGCUGGAGCACGGCAAGACCAUGACGGA
 GAAGGAGAUUCGUCGACUACGUGGCCAGCCAGGUGACCACCGCCAAGAAGCUGCGGGGCGG
 CGUGGUGUUCGUGGACGAGGUCCGAAGGGCCUGACCGGGAAGCUCGACGCCCGGAAGAU
 CCGCGAGAUCCUGAUC AAGGCCAAGAAGGGCGGCAAGAUCGCCGUGUAAgacuaguuaua
 agacugacua**GCCCGAUGGGCCUCCCAACGGGCCCUCUCCCCUCCUUGCACCG**agauua
 auAA
 AAAAAAugcauCCCCUCUAGACAAUUGGAAUCCAUA-3'

FIGURA 17

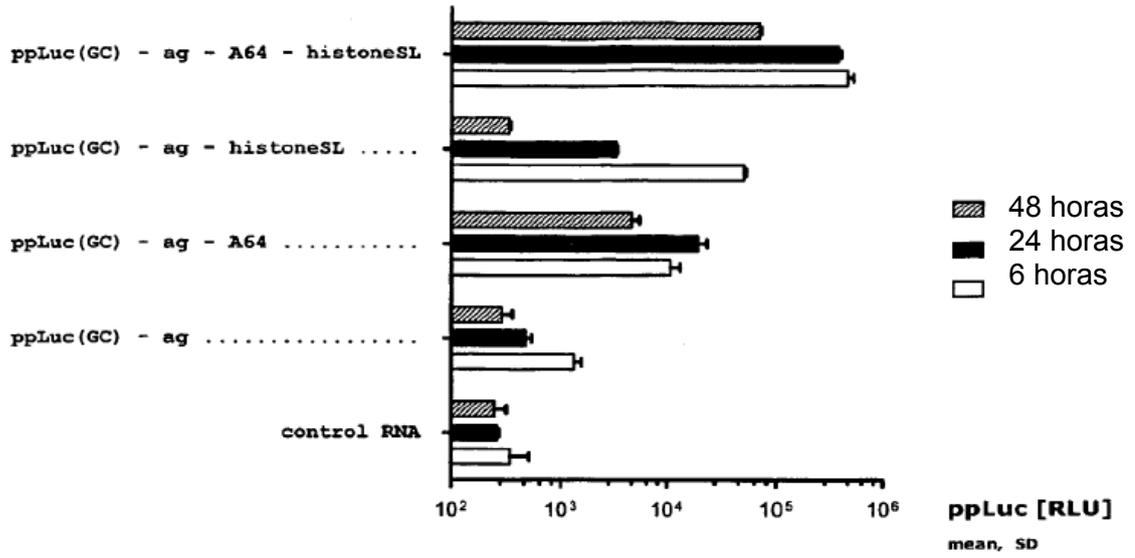


FIGURA 18

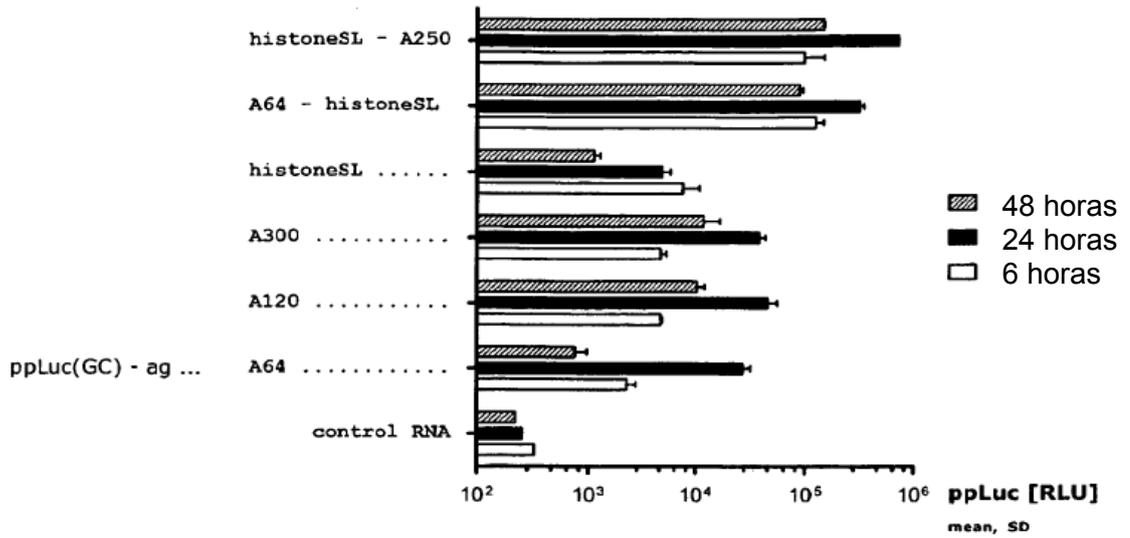


FIGURA 19

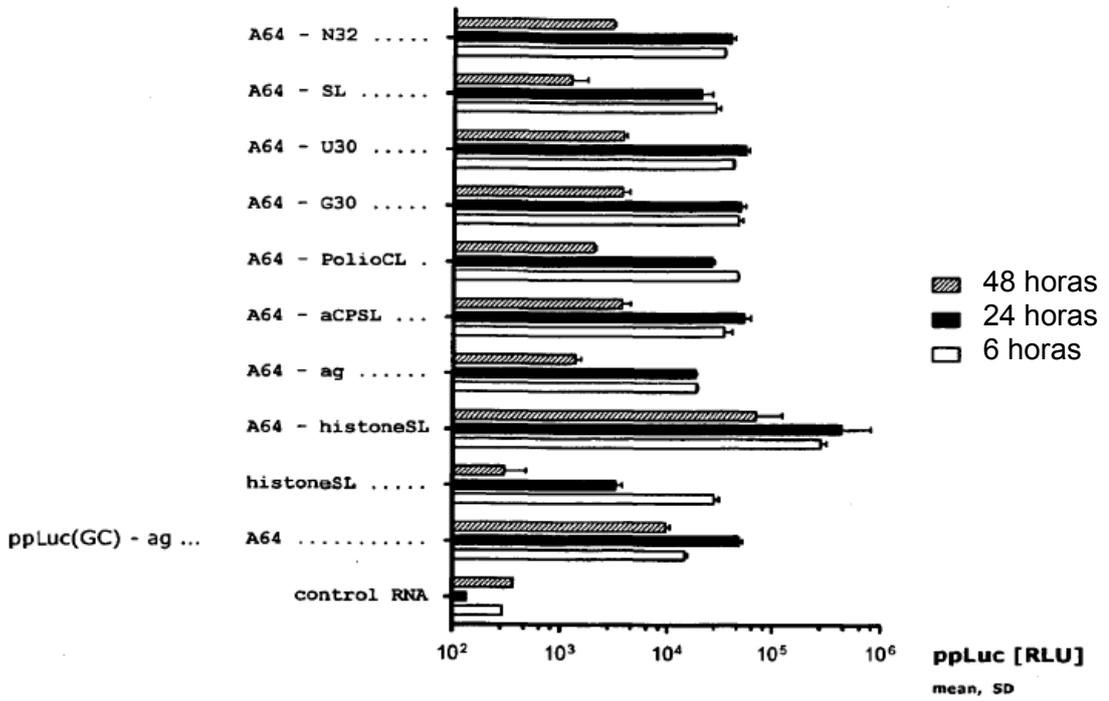


FIGURA 20

Expresión de luciferasa *in vivo*

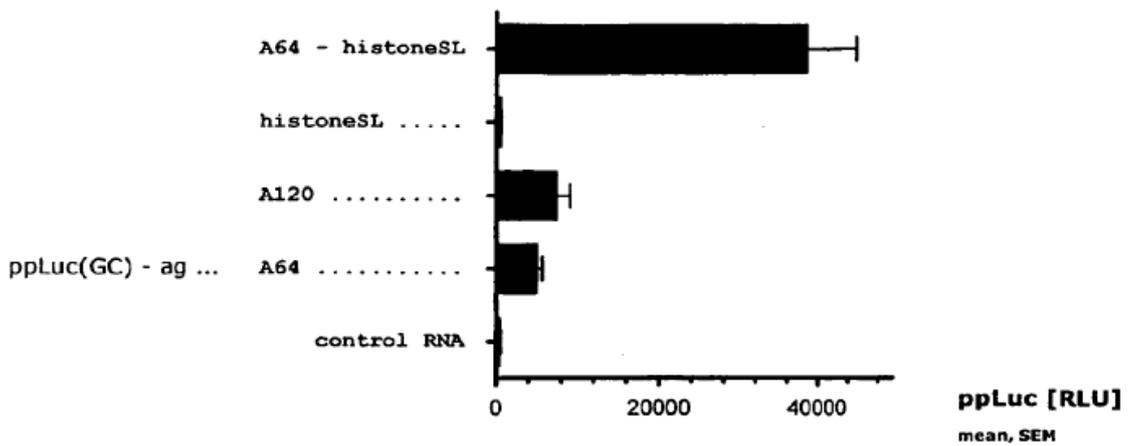


FIGURA 21

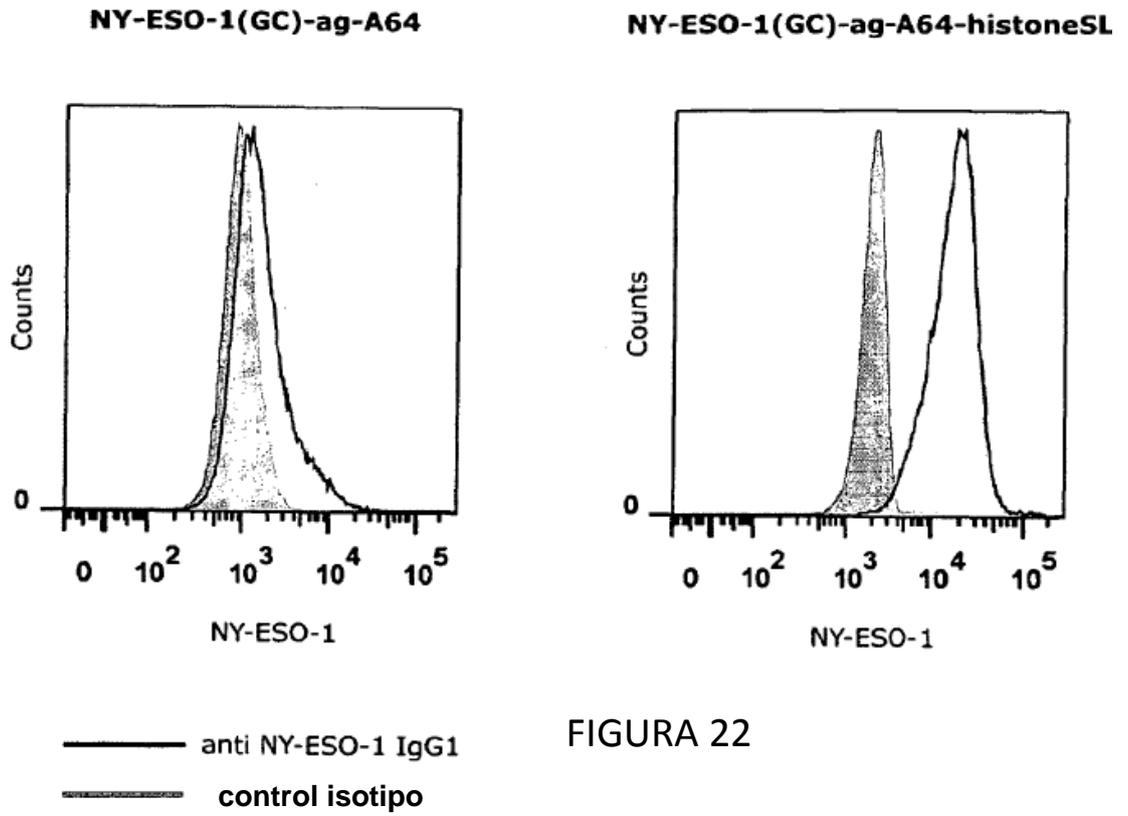


FIGURA 22

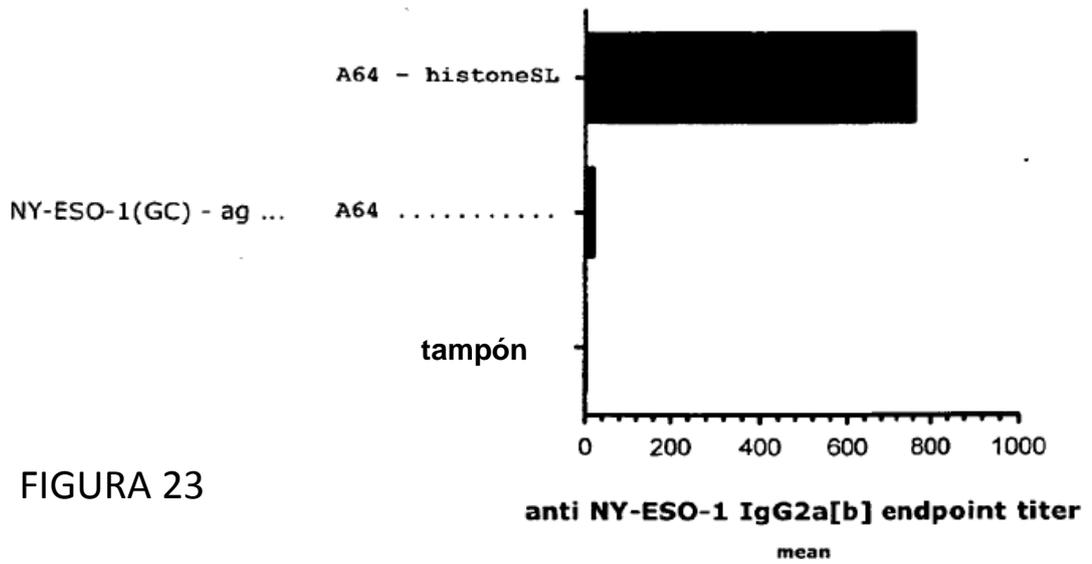


FIGURA 23