

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 735 728**

51 Int. Cl.:

C12N 15/09 (2006.01)

C12N 15/67 (2006.01)

C12N 15/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **06.07.2015 PCT/EP2015/065357**

87 Fecha y número de publicación internacional: **14.01.2016 WO16005324**

96 Fecha de presentación y número de la solicitud europea: **06.07.2015 E 15738014 (8)**

97 Fecha y número de publicación de la concesión europea: **26.06.2019 EP 3167059**

54 Título: **Estabilización de secuencias poli(a) que codifican secuencias de ADN**

30 Prioridad:

11.07.2014 WO PCT/EP2014/064924

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
20.12.2019

73 Titular/es:

**BIONTECH RNA PHARMACEUTICALS GMBH
(50.0%)**

**An der Goldgrube 12
55131 Mainz, DE y**

**TRON - TRANSLATIONALE ONKOLOGIE AN DER
UNIVERSITÄTSMEDIZIN DER JOHANNES
GUTENBERG- UNIVERSITÄT MAINZ
GEMEINNÜTZIGE GMBH (50.0%)**

72 Inventor/es:

**EBERLE, FLORIAN;
SAHIN, UGUR;
KUHN, ANDREAS;
VALLAZZA, BRITTA y
DIKEN, MUSTAFA**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 735 728 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

Estabilización de secuencias poli(a) que codifican secuencias de ADN

5 El uso del ARN ofrece una alternativa atractiva al ADN para evadir los riesgos potenciales de seguridad relacionados con el uso terapéutico del ADN. El ARN transcrito in vitro (ARN IVT) es de particular interés en los enfoques terapéuticos. Las ventajas de un uso terapéutico del ARN incluyen la expresión transitoria y un carácter no transformante. El ARN no necesita ingresar al núcleo para que se exprese y además, no puede integrarse en el genoma del huésped, lo que elimina de ese modo el riesgo de oncogénesis. Cuando se usa para la vacunación, la inyección de ARN puede inducir respuestas inmunes tanto celulares como humorales in vivo. Sin embargo, el uso de ARN para las aplicaciones clínicas está muy restringido especialmente por la corta vida media del ARN.

15 Los vectores IVT pueden usarse de una manera estandarizada como molde para la transcripción in vitro. Tales vectores IVT pueden tener la siguiente estructura: un promotor de ARN polimerasa 5' que permite la transcripción de ARN, seguido de un gen de interés que se flanquea en 3' y/o 5' por regiones no traducidas (UTR), y un casete de poliadenilo 3' que contiene nucleótidos A. Antes de la transcripción in vitro, el plásmido circular se linealiza aguas abajo del casete de poliadenilo mediante enzimas de restricción de tipo II (la secuencia de reconocimiento corresponde al sitio de escisión). El casete de poliadenilo por lo tanto, corresponde a la secuencia posterior de poli(A) en el transcrito.

20 La secuencia poli(A) 3' del ARN es importante para la exportación nuclear, la estabilidad del ARN y la eficacia de la traducción del ARN mensajero eucariótico (ARNm). La secuencia poli(A) 3' se acorta con el tiempo y si es lo suficientemente corta, el ARN se degrada enzimáticamente.

25 Hemos demostrado previamente que una secuencia poli(A) 3' con una longitud de 120 nucleótidos (A120) tiene un efecto predominante sobre la estabilidad del ARN y la eficiencia de la traducción y por lo tanto, es beneficiosa para la eficacia del ARN en todo momento.

30 Sin embargo, se ha observado que la secuencia de ADN que codifica la secuencia poli(A) 3' (casete de poliadenilo 3'), es decir, un tramo de pares de bases dA:dT consecutivos, se somete a acortamiento en algunos subclones bacterianos, cuando se propaga en *E. coli*. En consecuencia, antes de producir el ADN plasmídico como material de partida para la transcripción in vitro, debe probarse un gran número de clones bacterianos, por ejemplo, determinando la longitud del casete de poliadenilo 3' mediante un análisis de restricción adecuado, para obtener un solo clon con un casete de poliadenilo 3' de la longitud correcta que codifica una secuencia poli(A) 3' de la longitud correcta.

35 El objetivo de la presente invención fue encontrar un casete de poliadenilo 3' que muestre una propagación constante con el ADN plasmídico codificante en *E. coli* y que codifica una secuencia poli(A) 3' que mantiene los efectos con respecto al soporte de la estabilidad del ARN y la eficiencia de la traducción.

40 Este objetivo se logra de acuerdo con la invención mediante la materia de las reivindicaciones.

45 De acuerdo con la invención, se encontró que una alteración del casete de poliadenilo 3' (región poli(dA:dT)) por una secuencia aleatoria de 10 nucleótidos, con una distribución igual de los 4 nucleótidos (enlazador), tiene solo una influencia menor en la funcionalidad del ARN codificado, pero aumenta la estabilidad del casete de poliadenilo 3' en *E. coli*. Además, ni la secuencia ni la posición del enlazador dentro de la secuencia poli(A) 3' resultaron en una reducción de la eficiencia de la traducción y la estabilidad del ARN transcrito in vitro (ARN IVT).

50 Para las pruebas de estabilidad de la región del vector de IVT que codifica la secuencia poli(A) 3', el péptido SIINFEKL se clonó aguas arriba de la región poli(dA:dT). Este constructo mostró una inestabilidad de poli(dA:dT) (es decir, porcentaje de clones en la propagación con poli(dA:dT)) acortada de 50-60 %. El análisis detallado, con el uso del método de análisis de restricción descrito, identificó que la región en la posición 30-50 es particularmente sensible al acortamiento del estiramiento de poli(dA:dT). La introducción de una secuencia aleatoria de 10 nucleótidos en esta región sensible condujo a un aumento de la estabilidad de poli(dA:dT). Los constructos con 30 o 40 nucleótidos de adenosina, seguidas de la secuencia del enlazador y otras 70 o 60 adenosinas (A30L70 y A40L60) respectivamente, resultaron en una inestabilidad de poli(dA:dT) de solo 3-4 % en *E. coli*. Los resultados se confirmaron probando los constructos en varias cepas diferentes de *E. coli*.

60 La funcionalidad de ARN IVT codificado por el ADN que porta las colas poli(dA:dT) estabilizadas se probó en diferentes ensayos. La electroporación del ARN IVT en líneas celulares somáticas, pero además en células inmunes tales como las células dendríticas inmaduras, no mostró diferencias en la capacidad de traducción en comparación con el A120 durante un período de tiempo de 72 horas. La inyección de luciferasa que codifica el ARN IVT en ratones confirmó una traducción de proteína igual independiente del tipo insertado de la secuencia poli(A) 3'.

65 Se analizó un impacto en la respuesta inmunológica de las diferentes secuencias poli(A) 3' mediante la comparación de la cantidad de CD8 específico al antígeno+ células T tras la inyección de ARN IVT SIINFEKL. Los experimentos no revelaron diferencias entre el A120 y sus versiones estabilizadas A30L70 y A40L60.

En conjunto, mostramos que la inserción de una secuencia aleatoria de 10 nucleótidos entre las posiciones 30 y 50 de una región poli(dA:dT) resulta en una estabilización de la secuencia de más de 10 veces en *E. coli*. La correspondiente secuencia 3' poli(A) modificada del ARN transcrito a partir del molde de ADN tiene la misma funcionalidad, es decir, estabilidad y eficiencia de traducción in vivo e in vitro que la A120 clásica. Además, la respuesta inmunológica no se altera por el uso de una secuencia poli(A) modificada.

Resumen de la invención

El alcance de la presente invención se define en las reivindicaciones anexas.

En un aspecto, la invención se refiere a una molécula de ácido nucleico que comprende en la dirección 5' → 3' de la transcripción:

(a) un promotor;

(b) una secuencia de ácido nucleico transcribible o una secuencia de ácido nucleico para introducir una secuencia de ácido nucleico transcribible; y

(c) una secuencia de ácido nucleico que, cuando se transcribe bajo el control del promotor (a), codifica una secuencia de nucleótidos de al menos 80 nucleótidos consecutivos en el transcrito, en donde dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos en el transcrito es una secuencia de poliadenilo que comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A.

En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A es una secuencia, preferentemente una secuencia arbitraria, de 2 o más nucleótidos consecutivos, en donde el primer y el último nucleótido de dicha secuencia de 2 o más nucleótidos consecutivos es un nucleótido distinto de un nucleótido A.

En otras palabras, la molécula de ácido nucleico de la invención contiene el casete de poliadenilo 3' (región poli(dA:dT)) que contiene al menos una interrupción por una secuencia que no codifica una secuencia compuesta exclusivamente por residuos A, es decir, la región poli(dA:dT) se interrumpe por uno o más tramos de pares de bases que comprenden pares de bases distintos de (dA:dT). Así, la secuencia de ácido nucleico (c), cuando se transcribe bajo el control del promotor (a), codifica una secuencia de nucleótidos de al menos 80 nucleótidos consecutivos en el transcrito, en donde dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos en el transcrito es una secuencia de poliadenilo en donde al menos una parte de la secuencia de poliadenilo se reemplaza por una secuencia que contiene nucleótidos distintos de los nucleótidos A tales como una secuencia de 2 o más nucleótidos consecutivos, en donde el primer y el último nucleótido de dicha secuencia de 2 o más nucleótidos es un nucleótido distinto de un nucleótido A. En otras palabras, la secuencia de ácido nucleico (c), cuando se transcribe bajo el control del promotor (a), codifica para una secuencia de poliadenilo que contiene intercalada dentro de dicha secuencia de poliadenilo uno o más tramos de secuencia de uno o más nucleótidos, en donde cada uno de dichos tramos de secuencia no son un nucleótido A o un tramo de nucleótidos A, es decir, una secuencia oligo-A o una secuencia poli-A.

En una modalidad, dicha molécula de ácido nucleico es una molécula de ADN. En una modalidad, dicha molécula de ácido nucleico es un vector de expresión o plásmido tal como un vector IVT.

En una modalidad, dicha secuencia de ácido nucleico (c) exhibe una mayor estabilidad tras la propagación de dicha molécula de ácido nucleico en *E. coli* comparada con una molécula de ácido nucleico que comprende en lugar de dicha secuencia de ácido nucleico (c) una secuencia de ácido nucleico (c) que, cuando se transcribe bajo el control del promotor (a), codifica una secuencia de poliadenilo de la misma longitud que dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos en la transcripción.

En una modalidad, dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos comprende al menos 90 nucleótidos, preferentemente al menos 100 nucleótidos, preferentemente al menos 110 nucleótidos. En una modalidad, dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos comprende aproximadamente 120 nucleótidos. En modalidades particulares, dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos comprende hasta 200, preferentemente hasta 150, y, particularmente, hasta 130 nucleótidos. En una modalidad, al menos 90%, preferentemente al menos 92%, preferentemente al menos 95%, 97% o 98% de los nucleótidos de dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos son nucleótidos A en dicha secuencia de poliadenilo (sin incluir los nucleótidos A en dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A).

En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A se ubica dentro de una región desde la posición 21 a la posición 80, preferentemente desde la posición 21 a la posición 60, con mayor preferencia desde la posición 31 a la posición 50 de dicha secuencia de poliadenilo.

De acuerdo con la presente invención, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A se precede por al menos 20 residuos A, preferentemente al menos 30, 40 o 50 residuos A en dicha secuencia de poliadenilo. En modalidades particulares, dicha secuencia de uno o más nucleótidos consecutivos

que contienen nucleótidos distintos de los nucleótidos A se precede por hasta 80 residuos de A, preferentemente hasta 70 o 60 residuos A en dicha secuencia de poliadenilo.

5 De acuerdo con la presente invención, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A es seguida por al menos 20 residuos A, preferentemente al menos 30, 40, 50, 60 o 70 residuos A en dicha secuencia de poliadenilo. En modalidades particulares, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A es seguida por hasta 100 residuos A, preferentemente hasta 80 residuos A en dicha secuencia de poliadenilo.

10 En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A se precede de 20 a 50, preferentemente de 30 a 40 residuos A en dicha secuencia de poliadenilo y va seguida de 30 a 80, preferentemente de 40 a 70 residuos A en dicha secuencia de poliadenilo.

15 En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A tiene una longitud de al menos 3, al menos 4, al menos 5, al menos 6, al menos 8, preferentemente al menos 10, con mayor preferencia al menos 15 nucleótidos.

20 En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A tiene una longitud de no más de 50, preferentemente no más de 30, con mayor preferencia no más de 20 nucleótidos.

En una modalidad, dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A no comprende más de 3, preferentemente no más de 2, preferentemente no hay residuos de A consecutivos.

25 En una modalidad, las secuencias de ácido nucleico (b) y (c) bajo el control del promotor (a) pueden transcribirse para dar un transcrito común.

30 En una modalidad, en el transcrito dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos se ubica en el extremo 3'.

En una modalidad, la molécula de ácido nucleico de la invención es una molécula circular cerrada o una molécula lineal.

35 En una modalidad, la secuencia de ácido nucleico transcribible comprende una secuencia de ácido nucleico que codifica para un péptido o proteína y la secuencia de ácido nucleico para introducir una secuencia de ácido nucleico transcribible es un sitio de clonación múltiple.

En una modalidad, la molécula de ácido nucleico de la invención comprende además uno o más miembros seleccionados del grupo que consiste en: (i) un gen reportero; (ii) un marcador seleccionable; y (iii) un origen de replicación.

40 En una modalidad, la molécula de ácido nucleico de la invención es adecuada, particularmente después de la linealización, para la transcripción in vitro de ARN, particularmente ARNm.

45 Preferentemente, la secuencia de ácido nucleico transcrita de la secuencia de ácido nucleico (c), es decir, dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos, es preferentemente activa para aumentar la eficiencia de traducción y/o la estabilidad de la secuencia de ácido nucleico transcrita de la secuencia de ácido nucleico transcribible (b).

50 Antes de la transcripción in vitro, los vectores circulares de IVT generalmente se linealizan aguas abajo del casete de poliadenilo por enzimas de restricción de tipo II (la secuencia de reconocimiento corresponde al sitio de escisión). El casete de poliadenilo por lo tanto, corresponde a la secuencia posterior de poli(A) en el transcrito. Como resultado de este procedimiento, algunos nucleótidos permanecen como parte del sitio de escisión de la enzima después de la linealización y extienden o enmascaran la secuencia poli(A) en el extremo 3'. Sin embargo, se encontró que el ARN que tiene una secuencia poli(A) abierta se traduce de manera más eficiente que el ARN que tiene una secuencia poli(A) con un término enmascarado.

55 Como consecuencia, las moléculas de ácido nucleico de la invención, cuando se usan como vectores de expresión, permiten preferentemente la transcripción del ARN con una secuencia poli(A) que tiene preferentemente un extremo abierto en dicho ARN, es decir, no existen nucleótidos distintos de los nucleótidos A que flanquean dicha secuencia poli(A) en su extremo 3'. Puede lograrse una secuencia poli(A) de extremo abierto en el ARN introduciendo un sitio de escisión de restricción de tipo IIS en un vector de expresión que permita que el ARN se transcriba bajo el control de un promotor de la polimerasa de ARN 5' y que contenga un casete de poliadenilo, en donde la secuencia de reconocimiento se ubica aguas abajo del casete de poliadenilo, mientras que el sitio de escisión se ubica aguas arriba y así dentro del casete de poliadenilo. La escisión de restricción en el sitio de escisión de restricción de tipo IIS permite que un plásmido se linealice dentro del casete de poliadenilo. El plásmido linealizado puede usarse después como molde para la transcripción in vitro, y la transcripción resultante termina en una secuencia poli(A) sin enmascarar.

- 5 Como consecuencia, en una modalidad, se prefiere que la molécula de ácido nucleico de la invención pueda escindirse, preferentemente enzimáticamente o de otra manera bioquímica, dentro de la secuencia de ácido nucleico (c) de tal manera que dicha escisión resulte en una molécula de ácido nucleico que comprende, en la dirección 5'→ 3' de la transcripción, el promotor (a), la secuencia de ácido nucleico (b), y al menos una parte de la secuencia de ácido nucleico (c), en donde al menos una parte de la secuencia de ácido nucleico (c), cuando se transcribe bajo el control del promotor (a), codifica dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos en el transcrito y en donde en el transcrito el nucleótido 3' terminal es un nucleótido A de dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos.
- 10 Preferentemente, después de la escisión, la molécula de ácido nucleico, al final de la cadena que sirve como molde para la secuencia de nucleótidos de al menos 80 nucleótidos consecutivos, tiene un nucleótido T que es parte de la secuencia de ácido nucleico que sirve como molde para la secuencia de nucleótido de al menos 80 nucleótidos consecutivos en el transcrito.
- 15 La molécula de ácido nucleico de la invención es preferentemente una molécula circular cerrada antes de la escisión y una molécula lineal después de la escisión.
- Preferentemente, la escisión se lleva a cabo con la ayuda de un sitio de escisión de restricción que es preferentemente un sitio de escisión de restricción para una endonucleasa de restricción de tipo IIS.
- 20 En una modalidad, la secuencia de reconocimiento de la endonucleasa de restricción de tipo IIS se ubica 5-26 pares de bases, preferentemente 24-26 pares de bases, aguas abajo del extremo 3' de la secuencia de ácido nucleico (c).
- 25 En una modalidad, una molécula de ácido nucleico de acuerdo con la invención está en una conformación circular cerrada y es preferentemente adecuada para la transcripción in vitro de ARN, particularmente ARNm, particularmente después de la linealización. En aspectos adicionales, la invención se refiere a una molécula de ácido nucleico obtenible por linealización de una molécula de ácido nucleico descrita anteriormente, preferentemente por escisión dentro de la secuencia de ácido nucleico (c), y a ARN obtenible por transcripción, preferentemente transcripción in vitro, con las moléculas de ácido nucleico descritas anteriormente bajo el control del promotor (a).
- 30 En un aspecto adicional, la invención se refiere a un método de propagación de una molécula de ácido nucleico, que comprende:
 (i) proporcionar una molécula de ácido nucleico de la invención, y
 (ii) propagar dicha molécula de ácido nucleico en *E. coli*.
- 35 En una modalidad, propagar dicha molécula de ácido nucleico en *E. coli* comprende transformar *E. coli* con dicha molécula de ácido nucleico y cultivar dicha *E. coli* transformada.
- 40 En una modalidad, el método de la invención comprende además aislar dicha molécula de ácido nucleico a partir de *E. coli* siguiendo la propagación.
- 45 En un aspecto adicional, la invención se refiere a un método para obtener ARN, que comprende:
 (i) propagar una molécula de ácido nucleico de acuerdo con un método de la invención para propagar una molécula de ácido nucleico, y
 (ii) transcribir el ARN in vitro con el uso de la molécula de ácido nucleico como un molde.
- 50 En un aspecto adicional, la invención se refiere a un método para obtener un péptido o proteína, que comprende:
 (i) obtener un ARNm que codifica el péptido o la proteína de acuerdo con un método de la invención para obtener ARN, y
 (ii) traducir el ARNm.
- 55 En una modalidad, el método para obtener ARN o el método para obtener un péptido o proteína comprende además, antes de la transcripción de la molécula de ácido nucleico, la escisión de la molécula de ácido nucleico.
- En una modalidad, la escisión está dentro de la secuencia de ácido nucleico (c) de tal manera que la transcripción del ácido nucleico obtenida de esta manera genera un transcrito que tiene en su extremo 3' terminal dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos, en donde el nucleótido 3' terminal de dicho transcrito es un nucleótido A de la secuencia de nucleótidos de al menos 80 nucleótidos consecutivos.
- 60 En todos los aspectos de los métodos de acuerdo con la invención, la escisión se lleva a cabo preferentemente con la ayuda de un sitio de escisión de restricción que es preferentemente un sitio de escisión de restricción para una endonucleasa de restricción de tipo IIS.
- En una modalidad, la secuencia de reconocimiento para la endonucleasa de restricción tipo IIS es de 5-26 pares de bases, preferentemente de 24-26 pares de bases, aguas abajo del extremo 3' de la secuencia de ácido nucleico (c).
- 65 La invención además se refiere a ARN obtenible mediante los métodos de acuerdo con la invención para obtener ARN.

La invención puede utilizarse, por ejemplo, para aumentar la expresión de proteínas recombinantes en la transcripción y expresión celular. Más específicamente, es posible, cuando se producen proteínas recombinantes, usar vectores de expresión de la invención para la transcripción de ácidos nucleicos recombinantes y la expresión de proteínas recombinantes en sistemas basados en células. Esto incluye, por ejemplo, la preparación de anticuerpos recombinantes, hormonas, citoquinas, enzimas, y similares. Esto permite, entre otras cosas, reducir los costos de producción.

Además es posible usar las moléculas de ácido nucleico de la invención para aplicaciones de terapia génica. Como consecuencia, una molécula de ácido nucleico de la invención puede ser un vector de terapia génica y usarse para la expresión de un transgén. Para este fin, puede usarse cualquier sistema de vector basado en ácido nucleico (ADN/ARN) (por ejemplo, plásmidos, adenovirus, vectores de poxvirus, vectores de virus de la influenza, vectores de alfavirus y similares). Las células pueden transfectarse con estos vectores in vitro, por ejemplo en linfocitos o células dendríticas, o si no in vivo mediante administración directa.

El ARN de la invención (obtenido usando una molécula de ácido nucleico descrita en la presente descripción como un molde de transcripción) puede emplearse, por ejemplo, para la expresión transitoria de genes, con posibles campos de aplicación que son vacunas basadas en ARN que se transfectan a células in vitro o se administran directamente in vivo, expresión transitoria de proteínas recombinantes funcionales in vitro, por ejemplo, para iniciar procesos de diferenciación en células o estudiar funciones de proteínas, y expresión transitoria de proteínas recombinantes funcionales tales como eritropoyetina, hormonas, inhibidores de la coagulación, etc., en vivo, particularmente como productos farmacéuticos.

El ARN de la invención puede usarse particularmente para transfectar células presentadoras de antígeno y así como una herramienta para administrar el antígeno a presentar y para cargar las células presentadoras de antígeno, presentándose dicho antígeno correspondiente al péptido o proteína expresada a partir de dicho ARN o que se deriva de él, particularmente a través de un procesamiento intracelular tal como la escisión, es decir, el antígeno que se presenta es, por ejemplo, un fragmento del péptido o proteína expresado a partir del ARN. Tales células presentadoras de antígeno pueden usarse para estimular las células T, particularmente células T CD4⁺ y/o CD8⁺.

Como consecuencia, en un aspecto adicional, la invención se refiere a un uso del ARN de la invención para transfectar una célula huésped. En una modalidad preferida, la célula huésped es una célula presentadora de antígeno, particularmente una célula dendrítica, un monocito o un macrófago.

En un aspecto adicional, la invención se refiere a un uso del ARN de la invención para la vacunación.

Descripción detallada de la invención

Aunque la presente invención se describe en detalle más abajo, debe entenderse que esta invención no se limita a las metodologías, protocolos y reactivos particulares descritos en la presente descripción ya que estos pueden variar. Debe entenderse, también, que la terminología utilizada en la presente descripción es para el propósito de describir modalidades particulares solamente, y no pretende limitar el alcance de la presente invención la cual se limitará solamente por las reivindicaciones anexas. A menos que se defina de otra manera, todos los términos técnicos y científicos que se usan en la presente descripción tienen el mismo significado que el que entiende comúnmente un experto en la técnica.

A continuación, se describirán los elementos de la presente invención. Estos elementos se enumeran con modalidades específicas, sin embargo, debe entenderse que pueden combinarse de cualquier manera y en cualquier número para crear modalidades adicionales. Los ejemplos descritos de diversas maneras y las modalidades preferidas no deben interpretarse para limitar la presente invención a solamente las modalidades descritas explícitamente. Debe entenderse que esta descripción sustenta y abarca modalidades que combinan las modalidades descritas explícitamente con cualquier cantidad de los elementos descritos y/o preferidos. Además, cualquiera de las permutaciones y combinaciones de todos los elementos descritos en esta solicitud deben considerarse descritas por la descripción de la presente solicitud a menos que el contexto lo indique de cualquier otra manera.

Preferentemente, los términos que se usan en la presente descripción se definen como se describe en "A multilingual glossary of biotechnological terms: (IUPAC Recommendations)", H.G.W. Leuenberger, B. Nagel, y H. Kölbl, Eds., Helvética Chimica Acta, CH-4010 Basel, Suiza, (1995).

La práctica de la presente invención empleará, a menos que se indique de otra manera, métodos convencionales de química, bioquímica, biología celular, inmunología y técnicas de ADN recombinante que se explican en la literatura del campo (consultar, por ejemplo, Molecular Cloning: A Laboratory Manual, 2da Edición, J. Sambrook y otros eds., Cold Spring Harbor Laboratory Press, Cold Spring Harbor 1989).

A lo largo de esta descripción y de las reivindicaciones que siguen, a menos que el contexto lo requiera de cualquier otra manera, la palabra "comprender", y variaciones tales como "comprende" y "que comprende", se entenderán que implican la inclusión de un miembro, entero o etapa o grupo de miembros, enteros o etapas indicadas pero no la exclusión de cualquier otro miembro, entero o etapa o grupo de miembros, enteros o etapas. Los términos "un" y "una" y "el/la" y referentes similares usados en el contexto para describir la invención (especialmente en el contexto de las reivindicaciones) se deben interpretar para cubrir tanto el singular como el plural, a menos que se indique de cualquier

otra forma en la presente descripción o claramente se contradiga por el contexto. La enumeración de los intervalos de valores en la presente descripción pretende servir simplemente como un método abreviado para referirse individualmente a cada valor separado que se encuentra dentro del intervalo. A menos que se indique de otra manera en la presente descripción, cada valor individual se incorpora en la descripción como si se enumerara individualmente en la presente descripción. Todos los métodos que se describen en la presente descripción pueden llevarse a cabo en cualquier orden adecuado a menos que se indique de otra manera en la presente descripción o que el contexto lo contradiga claramente de otra manera. El uso de cualquiera y de todos los ejemplos, o lenguaje ilustrativo (por ejemplo, "tal como") que se proporciona en la presente descripción simplemente tiene el propósito de ilustrar mejor la invención y no representa una limitación en el alcance de la invención que se reivindica de otra manera. El lenguaje en la descripción no deberá interpretarse como indicación de cualquier elemento no reivindicado esencial para la práctica de la invención.

Se citan varios documentos a lo largo del texto de esta especificación. Nada debe interpretarse en la presente descripción como una admisión de que la invención no tiene derecho a preceder dicha descripción en virtud de una invención anterior.

La presente invención describe moléculas de ácido nucleico tales como plásmidos de ADN útiles como vectores de expresión de ARN que comprenden un casete de poli(dA:dT) modificado 3' (que codifica una secuencia de poli(A) 3' modificada) que muestra una propagación constante sin estar sujeto a acortamiento en *E. coli*.

E. coli es una bacteria gramnegativa, facultativamente anaeróbica, con forma de bastón del género *Escherichia* que se encuentra comúnmente en el intestino delgado de organismos de sangre caliente. La bacteria puede cultivarse de forma fácil y económica en un entorno de laboratorio, y se ha investigado de forma intensiva durante más de 60 años. *E. coli* es el organismo modelo procariótico más estudiado, y una especie importante en los campos de la biotecnología y la microbiología, donde ha servido como organismo huésped para la mayoría de los trabajos con ADN recombinante. Las cepas de *E. coli* de acuerdo con la invención incluyen: AG1, AB1157, B2155, BL21, BNN93, BNN97, BW26434, C600, CSH50, D1210, DB3.1, DH1, DH5 α , DH10B, DH12S, DM1, *E. clon(r)*, *E. coli* K12 ER2738, ER2566, ER2267, HB101, IJ1126, IJ1127, JM83, JM101, JM103, JM105, JM106, JM107, JM108, JM109, JM110, JM2.300, LE392, Mach1, MC1061, MC4100, MFDpir, MG1655, OmniMAX2, RR1, RV308, SOLR, SS320, STBL2, STBL3, STBL4, SURE, SURE2, TG1, TOP10, Top10F', W3110, WM3064, XL1-Blue, XL2-Blue, XL1-Red y XL10-Gold.

De acuerdo con la invención, una molécula de ácido nucleico o una secuencia de ácido nucleico se refiere a un ácido nucleico que es preferentemente ácido desoxirribonucleico (ADN) o ácido ribonucleico (ARN). De acuerdo con la invención, los ácidos nucleicos comprenden moléculas de ADN genómico, ADNc, ARNm preparadas recombinantemente y sintetizadas químicamente. De acuerdo con la invención, un ácido nucleico puede estar en forma de una molécula monocatenaria o bicatenaria y lineal o circular cerrada covalentemente.

En el contexto de la presente invención, el término "ARN" se refiere a una molécula que comprende residuos de ribonucleótidos y preferentemente está compuesta totalmente o esencialmente de residuos de ribonucleótidos. El término "ribonucleótido" se refiere a un nucleótido con un grupo hidroxilo en la posición 2' de un grupo β -D-ribofuranosilo. El término "ARN" comprende ARN bicatenario, ARN monocatenario, ARN aislado tales como ARN parcialmente purificado, completamente purificado, ARN esencialmente puro, ARN sintético, ARN generado de manera recombinante tal como ARN modificado que difiere del ARN presente de manera natural por la adición, delección, sustitución y/o alteración de uno o más nucleótidos. Dichas alteraciones pueden incluir la adición de material no nucleotídico, tal como al(los) extremo(s) de un ARN o internamente, por ejemplo, a uno o más nucleótidos del ARN. Los nucleótidos en las moléculas de ARN pueden comprender también nucleótidos no estándar, tales como nucleótidos que no están presentes de manera natural o nucleótidos o desoxinucleótidos sintetizados químicamente. Estos ARN alterados pueden referirse como análogos, particularmente análogos de ARN presentes de manera natural. De acuerdo con la invención, el ARN incluye ARNm.

El término "ARNm" significa "ARN mensajero" y se refiere a un transcrito que se genera con el uso de un molde de ADN y codifica un péptido o proteína. Típicamente, el ARNm comprende una 5'-UTR, una región codificante de proteínas, y una 3'-UTR. El ARNm puede generarse por transcripción *in vitro* a partir de un molde de ADN. La metodología de la transcripción *in vitro* se conoce por la persona experta. Por ejemplo, existe una variedad de kits de transcripción *in vitro* disponibles de manera comercial. De acuerdo con la invención, el ARNm puede modificarse mediante modificaciones y caperuzas estabilizantes, además de las modificaciones de acuerdo con la invención.

En una modalidad, el término "modificación" se refiere a proporcionar un ARN con una caperuza 5' o análogo de caperuza 5'. El término "caperuza 5'" se refiere a una estructura de caperuza que se encuentra en el extremo 5' de una molécula de ARNm y generalmente consiste en un nucleótido de guanosina que se conecta al ARNm a través de un enlace trifosfato inusual 5' a 5'. En una modalidad, esta guanosina se metila en la posición 7. El término "caperuza 5' convencional" se refiere a una caperuza 5' de ARN presente de manera natural, preferentemente a la caperuza 7-metilguanosa (m⁷G). En el contexto de la presente invención, el término "caperuza 5'" incluye un análogo de la caperuza 5' que se asemeja a la estructura de la caperuza de ARN y se modifica para que posea la capacidad de estabilizar el ARN si se une a este, preferentemente *in vivo* y/o en una célula. Proporcionar un ARN con una caperuza 5' o análogo de la caperuza 5' puede lograrse por transcripción *in vitro* de un molde de ADN en presencia de dicha caperuza 5' o análogo de la caperuza 5', en donde dicha caperuza 5' se incorpora cotranscripcionalmente a la cadena de ARN generada, o el ARN puede generarse, por ejemplo, por transcripción *in vitro*, y la caperuza 5' puede generarse de manera postranscripcional con el uso de enzimas de la caperuza, por ejemplo, las enzimas que añaden la caperuza del virus vaccinia.

El término "ácido nucleico" de acuerdo con la invención también comprende una derivación química de un ácido nucleico en una base de nucleótidos, en el azúcar o en el fosfato, y ácidos nucleicos que contienen nucleótidos no naturales y análogos de nucleótidos.

De acuerdo con la invención, una "secuencia de ácido nucleico que se deriva de una secuencia de ácido nucleico" se refiere a un ácido nucleico que contiene, en comparación con el ácido nucleico del que se deriva, sustituciones, deleciones y/o adiciones de nucleótidos simples o múltiples. Preferentemente, existe un cierto grado de homología entre dichos ácidos nucleicos y las secuencias de nucleótidos de dichos ácidos nucleicos se corresponden de una manera significativa directa o complementaria. De acuerdo con la invención, un ácido nucleico derivado de un ácido nucleico tiene una propiedad funcional del ácido nucleico del que se deriva. Tales propiedades funcionales incluyen particularmente la capacidad de aumentar, en un enlace funcional a un ácido nucleico que puede transcribirse en ARN (secuencia de ácido nucleico transcribible), la estabilidad y/o la eficiencia de traducción del ARN producido a partir de este ácido nucleico en la molécula de ARN completo.

De acuerdo con la invención, "enlace funcional" o "unido funcionalmente" se refiere a una conexión dentro de una relación funcional. Un ácido nucleico está "unido funcionalmente" si se relaciona funcionalmente con otra secuencia de ácido nucleico. Por ejemplo, un promotor se une funcionalmente a una secuencia codificante si esta influye en la transcripción de dicha secuencia codificante. Los ácidos nucleicos unidos funcionalmente son típicamente adyacentes entre sí, cuando se separan de manera apropiada por otras secuencias de ácido nucleico y, en modalidades particulares, se transcriben mediante la ARN polimerasa para dar una molécula de ARN única (transcrito común).

Los ácidos nucleicos descritos de acuerdo con la descripción se aíslan preferentemente. El término "ácido nucleico aislado" significa de acuerdo con la invención que el ácido nucleico se ha (i) amplificado in vitro, por ejemplo mediante reacción en cadena de la polimerasa (PCR), (ii) producido por vía recombinante mediante clonación, (iii) purificado, por ejemplo, mediante escisión y fraccionamiento electroforético en gel, o (iv) sintetizado, por ejemplo, mediante síntesis química. Un ácido nucleico aislado es un ácido nucleico disponible para la manipulación mediante técnicas de ADN recombinante.

Un ácido nucleico es "complementario" con otro ácido nucleico si las dos secuencias pueden hibridar y de formar un dúplex estable entre sí, dicha hibridación se lleva a cabo, preferentemente, en unas condiciones que permiten la hibridación específica entre polinucleótidos (condiciones rigurosas). Las condiciones rigurosas se describen, por ejemplo, en *Molecular Cloning: A Laboratory Manual*, J. Sambrook y otros, Editores, 2da Edición, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989 o *Current Protocols in Molecular Biology*, F.M. Ausubel y otros, Editores, John Wiley & Sons, Inc., New York y se refieren, por ejemplo, a una hibridización a 65°C en tampón de hibridación (3,5 x SSC, Ficoll al 0,02%, polivinilpirrolidona al 0,02%, albúmina de suero bovino al 0,02%, NaH₂PO₄ 2,5 mM (pH 7), SDS al 0,5%, EDTA 2 mM). SSC es cloruro de sodio 0,15 M/citrato de sodio 0,15 M, pH 7. Después de la hibridación, se lava la membrana a la que se ha transferido el ADN, por ejemplo, en 2 x SSC a temperatura ambiente y después en 0,1 - 0,5 x SSC/0,1 x SDS a temperaturas de hasta 68°C.

De acuerdo con la invención, los ácidos nucleicos homólogos tienen nucleótidos que son al menos 60%, al menos 70%, al menos 80%, al menos 90% y, preferentemente, al menos 95%, al menos 98% o al menos 99%, idénticos.

El término "% idéntico" pretende referirse a un porcentaje de nucleótidos que son idénticos en una alineación óptima entre dos secuencias que se comparan, siendo dicho porcentaje puramente estadístico, y las diferencias entre las dos secuencias pueden distribuirse aleatoriamente en toda la longitud de la secuencia y la secuencia que se compara pueden comprender adiciones o eliminaciones en comparación con la secuencia de referencia, para obtener una alineación óptima entre dos secuencias. Las comparaciones de dos secuencias generalmente se llevan a cabo comparando dichas secuencias, después del alineamiento óptimo, con respecto a un segmento o "ventana de comparación", para identificar las regiones locales de secuencias correspondientes. El alineamiento óptimo de una comparación puede llevarse a cabo manualmente o con ayuda del algoritmo de homología local de Smith y Waterman, 1981, *Ads App. Math.* 2, 482, con la ayuda del algoritmo de homología local de Needleman y Wunsch, 1970, *J. Mol. Biol.* 48, 443, y con la ayuda del algoritmo de búsqueda de similitud de Pearson y Lipman, 1988, *Proc. Natl Acad. Sci. USA* 85, 2444 o con la ayuda de programas de ordenador que usan dichos algoritmos (GAP, BESTFIT, FASTA, BLAST P, BLAST N y TFASTA en Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Drive, Madison, Wis.).

El porcentaje de identidad se obtiene mediante la determinación del número de posiciones idénticas en las que las secuencias que se comparan corresponden, dividiendo este número entre el número de posiciones comparadas y multiplicando este resultado por 100.

Por ejemplo, puede usarse el programa BLAST "Secuencias BLAST 2" que está disponible en el sitio web <http://www.ncbi.nlm.nih.gov/blast/bl2seq/wblast2.cgi>.

"Extremo 3'" de un ácido nucleico" se refiere, de acuerdo con la invención, a ese extremo que tiene un grupo hidroxilo libre. En una representación esquemática de ácidos nucleicos de doble cadena, particularmente el ADN, el extremo 3' está siempre en el lado derecho. "Extremo 5'" de un ácido nucleico" se refiere de acuerdo con la invención a ese extremo que

tiene un grupo fosfato libre. En una representación esquemática de ácidos nucleicos de doble cadena, particularmente el ADN, el extremo 5' está siempre en el lado izquierdo.

Extremo 5' 5'--P-NNNNNNN-OH-3' extremo 3' 3'-HO-NNNNNNN-P--5'

5 En modalidades particulares, un ácido nucleico está unido funcionalmente de acuerdo con la invención a secuencias de control de expresión que pueden ser homólogas o heterólogas con respecto al ácido nucleico.

10 Una secuencia de ácido nucleico transcribible, particularmente una secuencia de ácido nucleico que codifica un péptido o una proteína, y una secuencia de control de la expresión están unidas "funcionalmente" entre sí, si se unen covalentemente entre sí de tal manera que la transcripción o expresión de la secuencia de ácido nucleico transcribible y particularmente codificante está bajo el control o bajo la influencia de la secuencia de control de la expresión. Si la secuencia de ácido nucleico debe traducirse en una proteína o péptido funcional, la inducción de una secuencia de control de la expresión unida funcionalmente a la secuencia codificante resulta en la transcripción de dicha secuencia codificante, sin provocar un cambio de marco en la secuencia codificante o siendo la secuencia codificante incapaz de traducirse en la proteína o péptido deseado.

15 El término "secuencia de control de expresión" comprende de acuerdo con la invención promotores, secuencias de unión a ribosomas y otros elementos de control que controlan la transcripción de un gen o la traducción del ARN derivado. En modalidades particulares de la invención, pueden regularse las secuencias de control de la expresión. La estructura precisa de las secuencias control de la expresión puede variar en dependencia de la especie o tipo de célula, pero generalmente incluye las secuencias 5' no transcritas y 5' y 3' no traducidas que se involucran en el inicio de la transcripción y traducción, respectivamente, tales como caja TATA, secuencia de la caperuza, secuencia CAAT y similares. Más específicamente, las secuencias control de expresión 5' no transcritas incluyen una región promotora que abarca una secuencia promotora para el control de la transcripción del gen unido funcionalmente. Las secuencias control de la expresión pueden incluir además las secuencias potenciadoras o secuencias activadoras aguas arriba.

20 Las secuencias de ácido nucleico especificadas en la presente descripción, particularmente las secuencias de ácido nucleico transcribibles y codificante, pueden combinarse con cualquier secuencia de control de expresión, particularmente promotores, que pueden ser homólogos o heterólogos a dichas secuencias de ácido nucleico, con el término "homólogo" en referencia a el hecho de que una secuencia de ácido nucleico además se une funcionalmente de manera natural a la secuencia de control de la expresión, y el término "heterólogo" se refiere al hecho de que una secuencia de ácido nucleico no se une funcionalmente de manera natural a la secuencia de control de la expresión.

25 El término "promotor" o "región promotora" se refiere a una secuencia de ADN aguas arriba (5') de la secuencia codificante de un gen, que controla la expresión de dicha secuencia codificante al proporcionar un sitio de reconocimiento y unión para la ARN polimerasa. La región promotora puede incluir sitios de reconocimiento y unión adicionales para otros actores que se involucran en la regulación de la transcripción de dicho gen. Un promotor puede controlar la transcripción de un gen procariontico o eucariótico. Un promotor puede ser "inducible" e iniciar la transcripción en respuesta a un inductor, o puede ser "constitutivo" si la transcripción no se controla por un inductor. Un promotor inducible se expresa solo en un grado muy pequeño o en absoluto, si un inductor está ausente. En presencia del inductor, el gen se "enciende" o se aumenta el nivel de transcripción. Esto se media generalmente por la unión de un factor de transcripción específico.

Ejemplos de promotores preferidos de acuerdo con la invención son los promotores para SP6, T3 o T7 polimerasa.

30 De acuerdo con la invención, el término "expresión" se usa en su significado más general y comprende la producción de ARN o de ARN y proteína. Comprende también la expresión parcial de ácidos nucleicos. Además, la expresión puede ser transitoria o estable. Con respecto al ARN, el término "expresión" o "traducción" se refiere al proceso en los ribosomas de una célula por el cual una cadena de ARN mensajero dirige el ensamble de una secuencia de aminoácidos para hacer un péptido o proteína.

35 El término "secuencias de ácido nucleico que pueden transcribirse para dar un transcrito común" significa que dichas secuencias de ácido nucleico se unen funcionalmente entre sí de tal manera que, cuando sea apropiado después de la linealización, tal como la escisión por enzimas de restricción de la molécula de ácido nucleico que comprende dichas secuencias de ácido nucleico, particularmente de una molécula de ácido nucleico circular cerrada, la transcripción bajo el control de un promotor resulta una molécula de ARN que comprende los transcritos de dichas secuencias de ácido nucleico unidas covalentemente entre sí, cuando sea apropiado separadas por secuencias localizadas entre ellas.

40 En el contexto de la presente invención, el término "transcripción" se refiere a un proceso, en donde el código genético en una secuencia de ADN se transcribe a ARN. Posteriormente, el ARN puede traducirse en proteína. De acuerdo con la presente invención, el término "transcripción" comprende "transcripción in vitro", en donde el término "transcripción in vitro" se refiere a un proceso en donde el ARN, particularmente el ARNm, se sintetiza in vitro en un sistema libre de células. Preferentemente, los vectores de clonación se aplican para la generación de transcritos. Estos vectores de clonación se designan generalmente como vectores de transcripción y se abarcan de acuerdo con la presente invención por el término "vector". De acuerdo con la presente invención, el ARN es preferentemente un ARN transcrito in vitro (IVT-ARN) y puede obtenerse mediante la transcripción in vitro de un molde de ADN apropiada. El promotor para controlar la transcripción puede ser cualquier promotor para cualquier ARN polimerasa. Puede obtenerse un molde de ADN para la

transcripción in vitro mediante la clonación de un ácido nucleico, particularmente el ADNc, e introducir en un vector apropiado para la transcripción in vitro. El ADNc puede obtenerse por transcripción inversa de ARN.

5 El término "secuencia de ácido nucleico transcrita de una secuencia de ácido nucleico" se refiere a ARN, cuando sea apropiado como parte de una molécula de ARN completa, que es un producto de la transcripción de la última secuencia de ácido nucleico.

10 El término "secuencia de ácido nucleico que es activa para aumentar la eficiencia de traducción y/o la estabilidad de una secuencia de ácido nucleico" significa que la primera secuencia de ácido nucleico es capaz de modificar, en una transcripción común con la segunda secuencia de ácido nucleico, la eficiencia de traducción y/o estabilidad de dicha segunda secuencia de ácido nucleico de tal manera que dicha eficiencia de traducción y/o estabilidad aumenta en comparación con la eficiencia de traducción y/o estabilidad de dicha segunda secuencia de ácido nucleico sin dicha primera secuencia de ácido nucleico. En este contexto, el término "eficiencia de traducción" se refiere a la cantidad de producto de traducción provisto por una molécula de ARN dentro de un período de tiempo particular y el término "estabilidad" se relaciona con la vida media de una molécula de ARN.

20 Se ha demostrado que una doble región 3' no traducida (UTR), particularmente del gen de la beta globina humana, en una molécula de ARN mejora la eficiencia de traducción de una manera que claramente excede el efecto total que se espera usando dos UTR individuales.

25 La modificación y, por lo tanto, la estabilización y/o el aumento de la eficiencia de traducción del ARN pueden lograrse de acuerdo con la invención modificando genéticamente las moléculas de ácido nucleico de la expresión de la invención cuando se usan como vectores de expresión de tal manera que permitan la transcripción del ARN con dos o más regiones 3' no traducidas en su extremo 3', y preferentemente entre la secuencia que codifica un péptido o proteína (marco de lectura abierto) y la secuencia poli(A) o la secuencia poli(A) que comprende una secuencia de uno o más consecutivos nucleótidos que contienen nucleótidos distintos de los nucleótidos A.

30 En consecuencia, una molécula de ácido nucleico de la invención puede comprender, preferentemente entre la secuencia de ácido nucleico (b) y la secuencia de ácido nucleico (c) dos o más secuencias de ácido nucleico, cada una de dichas dos o más secuencias de ácido nucleico correspondientes a la región 3' no traducida de un gen o que se deriva de ella. Dichas dos o más secuencias de ácido nucleico pueden ser idénticas o diferentes. En modalidades preferidas, dichas dos o más secuencias de ácido nucleico son independientemente entre sí derivadas de un gen seleccionado del grupo que consiste en genes de globina tales como alfa2 globina, alfa1 globina, beta globina y hormona de crecimiento, preferentemente beta globina humana.

35 La región 3' no traducida se refiere a una región que se encuentra en el extremo 3' de un gen, aguas abajo del codón de terminación de una región que codifica una proteína, y que se transcribe pero no se traduce en una secuencia de aminoácidos.

40 De acuerdo con la invención, se considera que una primera región polinucleotídica se ubica aguas abajo de una segunda región polinucleotídica, si el extremo 5' de dicha primera región polinucleotídica es la parte de dicha primera región polinucleotídica más cercana al extremo 3' de dicha segunda región polinucleotídica.

45 La región 3' no traducida se extiende típicamente desde el codón de terminación para un producto de traducción a la secuencia poli(A) que generalmente se adjunta después del proceso de transcripción. Las regiones 3' no traducidas del ARNm de mamíferos tienen típicamente una región de homología conocida como la secuencia de hexanucleótidos AAUAAA. Esta secuencia es presumiblemente la señal de unión de poli(A) y con frecuencia se ubica desde 10 a 30 bases aguas arriba del sitio de unión de poli(A).

50 Las regiones 3' no traducidas pueden contener una o más repeticiones invertidas que pueden plegarse para dar estructuras de tallo-lazo que actúan como barreras para las exorribonucleasas o que interactúan con proteínas que se sabe que aumentan la estabilidad del ARN (por ejemplo, proteínas de unión al ARN).

55 Las regiones 5' no traducidas y/o 3' pueden, de acuerdo con la invención, estar unidas funcionalmente a un ácido nucleico codificable y particularmente codificante, para que estas regiones se asocien con el ácido nucleico de tal manera que la estabilidad y/o la eficiencia de traducción del ARN transcrito se aumenten a partir de dicho ácido nucleico transcribible.

60 Las regiones 3' no traducidas de los ARNm de inmunoglobulina son relativamente cortas (menos de aproximadamente 300 nucleótidos), mientras que las regiones 3' no traducidas de otros genes son relativamente largas. Por ejemplo, la región 3' no traducida de tPA tiene aproximadamente 800 nucleótidos de longitud, la del factor VIII tiene aproximadamente 1800 nucleótidos de longitud y la de la eritropoyetina tiene aproximadamente 560 nucleótidos de longitud.

65 De acuerdo con la invención, puede determinarse si una región 3' no traducida o una secuencia de ácido nucleico derivada de la misma aumenta la estabilidad y/o la eficiencia de traducción del ARN, por la incorporación de la región 3' no traducida o la secuencia de ácido nucleico derivada de la misma en la región 3' no traducida de un gen y por la medición de si dicha incorporación aumenta la cantidad de proteína sintetizada.

Lo anterior se aplica de acuerdo con el caso en el que, de acuerdo con la invención, un ácido nucleico comprende dos o más regiones no traducidas 3' que preferentemente se acoplan secuencialmente con o sin un enlazador intermedio, preferentemente en una "relación de cabeza a cola" (es decir, las regiones no traducidas 3' tienen la misma orientación, preferentemente la orientación que ocurre de manera natural en un ácido nucleico).

De acuerdo con la invención, el término "gen" se refiere a una secuencia particular de ácido nucleico que es responsable de producir uno o más productos celulares y/o de lograr una o más funciones intercelulares o intracelulares. Más específicamente, dicho término se refiere a una sección de ADN que comprende un ácido nucleico que codifica una proteína específica o una molécula de ARN funcional o estructural.

La poliadenilación es la adición de una secuencia poli(A) o cola a un ARN transcrito primario. La secuencia poli(A) consiste en múltiples monofosfatos de adenosina. En otras palabras, es un tramo de ARN que solo tiene bases de adenina. En los eucariotas, la poliadenilación es parte del proceso que produce el ARN mensajero maduro (ARNm) para la traducción. Esto, por lo tanto, forma parte del proceso de expresión génica. El proceso de poliadenilación comienza a medida que la transcripción de un gen acaba, o termina. El segmento más a lo largo de 3' del pre-ARNm recién creado se separa primero por un conjunto de proteínas; estas proteínas después sintetizan la secuencia poli(A) en el extremo 3' del ARN. La secuencia poli(A) es importante para la exportación nuclear, la traducción, y la estabilidad del ARNm. La secuencia se acorta con el tiempo y, cuando es lo suficientemente corta, el ARNm se degrada enzimáticamente.

Los términos "secuencia de poliadenilo", "secuencia de poli(A)" o "cola de poli(A)" se refieren a una secuencia de residuos de adenilo que se encuentra típicamente en el extremo 3' de una molécula de ARN. La invención proporciona tal secuencia para unirse durante la transcripción del ARN por medio de un molde de ADN sobre la base de residuos timidilo repetidos en la cadena complementaria a la cadena codificante, mientras que dicha secuencia normalmente no se codifica en el ADN pero se une a la cadena de ADN al extremo 3' libre del ARN por una ARN polimerasa independiente del molde después de la transcripción en el núcleo. El término "nucleótidos A" o "A" se refiere a residuos de adenilo.

En una modalidad preferida, una molécula de ácido nucleico de acuerdo con la invención es un vector. El término "vector" se usa en la presente en su significado más general y comprende cualquier vehículo intermediario para un ácido nucleico que, por ejemplo, permite que dicho ácido nucleico, por ejemplo, se introduzca en células procariotas y/o eucariotas y, cuando sea apropiado, se integre en un genoma. Tales vectores se replican preferentemente y/o se expresan en las células. Los vectores comprenden plásmidos, fagémidos o genomas virales. El término "plásmido", como se usa en la presente descripción, se refiere generalmente a un constructo de material genético extracromosómico, generalmente, un ADN dúplex circular, que puede replicarse independientemente del ADN cromosómico.

De acuerdo con la invención, el término "célula huésped" se refiere a cualquier célula que puede transformarse o transfectarse con un ácido nucleico exógeno. El término "célula huésped" comprende de acuerdo con la invención a células procariotas (por ejemplo *E. coli*) o eucariotas (por ejemplo células de levadura y células de insectos). Se da preferencia particular a células de mamífero tales como células de humanos, ratones, hámsteres, cerdos, cabras, primates. Las células pueden derivarse de una multiplicidad de tipos de tejidos y comprenden células primarias y líneas celulares. Los ejemplos específicos incluyen queratinocitos, leucocitos de sangre periférica, células madre de la médula ósea y células madre embrionarias. En otras modalidades, la célula huésped es una célula presentadora de antígeno, particularmente una célula dendrítica, un monocito o un macrófago. Un ácido nucleico puede estar presente en la célula huésped en una sola copia o en varias copias y, en una modalidad se expresa en la célula huésped.

De acuerdo con la presente invención, el término "péptido" comprende oligo y polipéptidos y se refiere a las sustancias que comprenden dos o más, preferentemente 3 o más, preferentemente 4 o más, preferentemente 6 o más, preferentemente 8 o más, preferentemente 10 o más, preferentemente 13 o más, preferentemente 16 o más, preferentemente 20 o más y hasta preferentemente 50, preferentemente 100 o preferentemente 150, aminoácidos consecutivos unidos entre sí por enlaces peptídicos.

El término "proteína" se refiere a péptidos grandes, preferentemente péptidos que tienen al menos 151 aminoácidos, pero los términos "péptido" y "proteína" se usan en la presente descripción usualmente como sinónimos. Los términos "péptido" y "proteína" comprenden de acuerdo con la invención sustancias que contienen no solo componentes de aminoácidos, sino también componentes que no son de aminoácidos tales como los azúcares y estructuras de fosfato, y también comprenden sustancias que contienen enlaces tales como enlaces éster, tioéter o disulfuro.

De acuerdo con la presente invención, un ácido nucleico tal como ARN puede codificar un péptido o proteína. En consecuencia, una secuencia de ácido nucleico transcribible o un transcrito de este pueden contener un marco de lectura abierto (ORF) que codifica un péptido o proteína. Dicho núcleo puede expresar el péptido o proteína codificado. Por ejemplo, dicho ácido nucleico puede ser un ácido nucleico que codifica y expresa un antígeno o un péptido o proteína farmacéuticamente activo tal como un compuesto inmunológicamente activo (que preferentemente no es un antígeno).

De acuerdo con la invención, el término "ácido nucleico que codifica un péptido o proteína" significa que el ácido nucleico, si está presente en el ambiente apropiado, preferentemente dentro de una célula, puede dirigir el ensamblaje de aminoácidos para producir el péptido o proteína durante el proceso de traducción. Preferentemente, el ARN de acuerdo

con la invención es capaz de interactuar con la maquinaria de traducción celular que permite la traducción del péptido o proteína.

5 De acuerdo con la invención, en una modalidad, el ARN comprende o consiste en ARN farmacéuticamente activo. Un "ARN farmacéuticamente activo" puede ser un ARN que codifica un péptido o proteína farmacéuticamente activo.

10 Un "péptido o proteína farmacéuticamente activo" tiene un efecto positivo o ventajoso sobre el estado de la afección o enfermedad de un sujeto cuando se administra al sujeto en una cantidad terapéuticamente eficaz. Preferentemente, un péptido o proteína farmacéuticamente activo tiene propiedades curativas o paliativas y puede administrarse para mejorar, atenuar, aliviar, revertir, retardar el inicio o disminuir la gravedad de uno o más síntomas de una enfermedad o trastorno. Un péptido o proteína farmacéuticamente activo puede tener propiedades profilácticas y puede usarse para retardar la aparición de una enfermedad o para disminuir la gravedad de dicha enfermedad o afección patológica. El término "péptido o proteína farmacéuticamente activo" incluye proteínas o polipéptidos completos, y puede referirse además a los fragmentos farmacéuticamente activos de estos. Puede incluir además los análogos farmacéuticamente activos de un péptido o proteína. El término "péptido o proteína farmacéuticamente activo" incluye a los péptidos y proteínas que son antígenos, es decir, la administración del péptido o proteína a un sujeto provoca una respuesta inmune en un sujeto que puede ser terapéutica o parcial o totalmente protectora.

20 Los ejemplos de proteínas farmacéuticamente activas incluyen, pero no se limitan a, citoquinas y proteínas del sistema inmunológico, tales como los compuestos inmunológicamente activos (por ejemplo, interleucinas, factor estimulante de colonias (CSF), factor estimulante de colonias de granulocitos (G-CSF), factor estimulante de colonias de granulocitos y macrófagos (GM-CSF), eritropoyetina, factor de necrosis tumoral (FNT), interferones, integrinas, adreínas, seletinas, receptores homing, receptores de células T, inmunoglobulinas, antígenos solubles del complejo mayor de histocompatibilidad, antígenos inmunológicamente activos tales como antígenos bacterianos, parasitarios o virales, alérgenos, autoantígenos, anticuerpos), hormonas (insulina, hormona tiroidea, catecolaminas, gonadotropinas, hormonas tróficas, prolactina, oxitocina, dopamina, somatotropina bovina, leptinas y similares), hormonas de crecimiento (por ejemplo, hormona de crecimiento humano), factores de crecimiento (por ejemplo, factor de crecimiento epidérmico, factor de crecimiento nervioso, factor de crecimiento de tipo insulina y similares), receptores de factor de crecimiento, enzimas (activador del plasminógeno de tejido, estreptoquinasa, colesterol biosintético o degradante, enzimas estereodogénicas, quinasas, fosfodiesterasas, metilasas, desmetilasas, deshidrogenasas, celulasas, proteasas, lipasas, fosfolipasas, aromatasas, citocromos, ciclasas de adenilato y guanilato, neuromidasas y similares), receptores (receptores de la hormona esteroidea, receptores de péptidos), proteínas de unión (proteínas de unión a hormona del crecimiento o proteínas de unión al factor de crecimiento y similares), factores de transcripción y traducción, proteínas supresoras de crecimiento tumoral (por ejemplo, proteínas que inhiben la angiogénesis), proteínas estructurales (tales como colágeno, fibroína, fibrinógeno, elastina, tubulina, actina y miosina), proteínas sanguíneas (trombina, albúmina sérica, Factor VII, Factor VIII, insulina, Factor IX, Factor X, activador del plasminógeno tisular, proteína C, factor de von Willebrand, antitrombina III, glucocerebrosidasa, factor estimulante de colonias de granulocitos eritropoyetina (G-CSF) o factor VIII modificado, anticoagulantes y similares.

40 En una modalidad, la proteína farmacéuticamente activa de acuerdo con la invención es una citoquina que participa en la regulación de la homeostasis linfoide, preferentemente una citoquina que participa y preferentemente induce o mejora el desarrollo, cebado, expansión, diferenciación y/o supervivencia de las células T. En una modalidad, la citoquina es una interleuquina. En una modalidad, la proteína farmacéuticamente activa de acuerdo con la invención es una interleuquina seleccionada del grupo que consiste en IL-2, IL-7, IL-12, IL-15, e IL-21.

45 El término "compuesto inmunológicamente activo" se refiere a cualquier compuesto que altera una respuesta inmune, preferentemente mediante la inducción y/o supresión de la maduración de las células inmunes, mediante la inducción y/o supresión de la biosíntesis de citoquinas, y/o mediante la alteración de la inmunidad humoral estimulando la producción de anticuerpos por las células B. Los compuestos inmunológicamente activos poseen una potente actividad inmunostimulante que incluye, pero no se limita a, la actividad antiviral y antitumoral, y pueden además regular negativamente otros aspectos de la respuesta inmunitaria, por ejemplo, desviar la respuesta inmunitaria de una respuesta inmunitaria TH2, que es útil para el tratamiento de un amplio rango de enfermedades mediadas por TH2. Los compuestos inmunológicamente activos pueden ser útiles como adyuvantes de vacunas.

55 Si, de acuerdo con la presente invención, se desea inducir o mejorar una respuesta inmune con el uso del ARN como se describe en la presente, la respuesta inmune puede activarse o mejorarse por el ARN. Por ejemplo, las proteínas o los péptidos codificados por los ARN o sus productos de procesamiento pueden presentarse por las proteínas del complejo mayor de histocompatibilidad (MHC) expresadas en las células presentadoras de antígenos. El complejo peptídico MHC puede reconocerse después por las células inmunes tales como las células T que conducen a su activación.

60 El término "enfermedad" se refiere a una condición anormal que afecta el cuerpo de un individuo. Una enfermedad a menudo se interpreta como una condición médica que se asocia con signos y síntomas específicos. Una enfermedad puede ser causada por factores originalmente de una fuente externa, tal como enfermedad infecciosa, o puede ser causada por disfunciones internas, tal como enfermedades autoinmunes.

65

De acuerdo con la invención, el término "enfermedad" se refiere además a las enfermedades del cáncer. Los términos "enfermedad del cáncer" o "cáncer" (término médico: neoplasia maligna) se refieren a una clase de enfermedades en las que un grupo de células muestra un crecimiento incontrolado (división más allá de los límites normales), invasión (intrusión y destrucción de tejidos adyacentes) y, a veces, metástasis (propagación a otras ubicaciones del cuerpo a través de la linfa o la sangre). Estas tres propiedades malignas de los cánceres los diferencian de los tumores benignos, que son autolimitados y no invaden ni metastatizan. La mayoría de los cánceres forman un tumor, es decir, una inflamación o lesión formada por un crecimiento anormal de las células (denominadas células neoplásicas o células tumorales), pero algunos, como la leucemia, no lo hacen. Los ejemplos de tales cánceres incluyen, pero no se limitan a, carcinoma, linfoma, blastoma, sarcoma, glioma y leucemia. Más particularmente, los ejemplos de tales cánceres incluyen el cáncer óseo, cáncer sanguíneo, cáncer de pulmón, cáncer de hígado, cáncer de páncreas, cáncer de piel, cáncer de la cabeza o cuello, melanoma maligno cutáneo o intraocular, cáncer uterino, cáncer de ovario, cáncer rectal, cáncer de la región anal, cáncer de estómago, cáncer de colon, cáncer de mama, cáncer de próstata, cáncer del útero, carcinoma de los órganos sexuales y reproductivos, enfermedad de Hodgkin, cáncer del esófago, cáncer del intestino delgado, cáncer del sistema endocrino, cáncer de la glándula tiroidea, cáncer de la glándula paratiroides, cáncer de la glándula adrenal, sarcoma de tejido blando, cáncer de vejiga, cáncer de riñón, carcinoma celular renal, carcinoma de la pelvis renal, neoplasias del sistema nervioso central (CNS), cáncer neuroectodermal, tumores del eje espinal, glioma, meningioma, y adenoma pituitario. El término "cáncer" de acuerdo con la invención también comprende, las metástasis del cáncer.

El término "enfermedad infecciosa" se refiere a cualquier enfermedad que puede transmitirse de un individuo a otro o de un organismo a otro, y es causada por un agente microbiano (por ejemplo, resfriado común). Los ejemplos de enfermedades infecciosas incluyen enfermedades infecciosas virales, tales como SIDA (VIH), hepatitis A, B o C, herpes, herpes zóster (varicela), sarampión alemán (virus de la rubéola), fiebre amarilla, dengue, etc. flavivirus, virus de la influenza, enfermedades infecciosas hemorrágicas (virus de Marburg o Ébola) y síndrome respiratorio agudo severo (SARS), enfermedades infecciosas bacterianas, tal como la enfermedad del legionario (*Legionella*), enfermedades de transmisión sexual (por ejemplo, clamidia o gonorrea), úlcera gástrica (*Helicobacter*), cólera (*Vibrio*), tuberculosis, difteria, infecciones por *E. coli*, *Estafilococos*, *Salmonela* o *Streptococos* (tétanos); infecciones por patógenos protozoarios tales como la malaria, la enfermedad del sueño, leishmaniasis; toxoplasmosis, es decir, infecciones por *Plasmodium*, *Trypanosoma*, *Leishmania* y *Toxoplasma*; o infecciones fúngicas, que son causadas por ejemplo por *Cryptococcus neoformans*, *Histoplasma capsulatum*, *Coccidioides immitis*, *Blastomyces dermatitidis* o *Candida albicans*.

El término "enfermedad autoinmune" se refiere a cualquier enfermedad en la que el cuerpo produce una respuesta inmunogénica (es decir, del sistema inmunitario) a algún componente de su propio tejido. En otras palabras, el sistema inmunológico pierde su capacidad de reconocer algunos tejidos o sistemas dentro del cuerpo como propios y objetivos y lo ataca como si fueran extraños. Las enfermedades autoinmunes pueden clasificarse en aquellas en las que predominantemente se afecta un órgano (por ejemplo, anemia hemolítica y tiroiditis antiinmune), y en aquellas en las que el proceso de la enfermedad autoinmune se difunde a través de muchos tejidos (por ejemplo, lupus eritematoso sistémico). Por ejemplo, se cree que la esclerosis múltiple es causada por las células T que atacan las vainas que rodean las fibras nerviosas del cerebro y la médula espinal. Esto resulta en la pérdida de coordinación, debilidad y visión borrosa. Las enfermedades autoinmunes son conocidas en la técnica e incluyen, por ejemplo, la tiroiditis de Hashimoto, la enfermedad de Grave, el lupus, la esclerosis múltiple, la artritis reumática, la anemia hemolítica, la tiroiditis antiinmune, el lupus eritematoso sistémico, la enfermedad celíaca, la enfermedad de Crohn, la colitis, la diabetes, la esclerodermia, la psoriasis, y similares.

De acuerdo con la invención, puede estimularse una respuesta inmunitaria al introducir en un sujeto un ARNm adecuado que codifique un antígeno o un fragmento de este, por ejemplo, un antígeno asociado a la enfermedad.

El término "antígeno" se refiere a un agente que comprende un epítipo contra el cual se genera y/o se direcciona una respuesta inmunitaria. El término "antígeno" incluye en particular proteínas, péptidos, polisacáridos, ácidos nucleicos, especialmente ARN y ADN, y nucleótidos. El término "antígeno" también incluye agentes, que se convierten en antigénicos - y sensibilizantes - solo a través de la transformación (por ejemplo de manera intermedia en la molécula o por terminación con la proteína del cuerpo). Un antígeno es preferentemente presentable por las células del sistema inmune, tales como células presentadoras de antígeno tipo células dendríticas o macrófagos. Un antígeno o un producto procesado de este se reconoce preferentemente por un receptor de células T o B, o por una molécula de inmunoglobulina tal como un anticuerpo. En una modalidad preferida, el antígeno es un antígeno asociado a la enfermedad, tales como un antígeno asociado a tumor, un antígeno viral, o un antígeno bacteriano.

El término "antígeno asociado a la enfermedad" se usa en su sentido más amplio para referirse a cualquier antígeno asociado con una enfermedad. Un antígeno asociado a la enfermedad es una molécula que contiene epítopos que estimularán el sistema inmunitario del huésped para que genere una respuesta inmunitaria específica al antígeno celular y/o una respuesta humoral de anticuerpos contra la enfermedad. Por tanto, el antígeno asociado a la enfermedad puede usarse con fines terapéuticos. Los antígenos asociados a las enfermedades se asocian preferentemente con la infección por microbios, típicamente antígenos microbianos, o se asocian con cáncer, típicamente tumores.

El término "enfermedad que involucra un antígeno" se refiere a cualquier enfermedad que implica un antígeno, por ejemplo, una enfermedad que se caracteriza por la presencia de un antígeno. La enfermedad que involucra un antígeno puede ser una enfermedad infecciosa, una enfermedad autoinmune, una enfermedad de cáncer o simplemente un cáncer.

Como se mencionó anteriormente, el antígeno puede ser un antígeno asociado a la enfermedad, tales como un antígeno asociado a tumor, un antígeno viral, o un antígeno bacteriano.

En una modalidad, un antígeno asociado a la enfermedad es un antígeno asociado a un tumor. En esta modalidad, la presente invención puede ser útil en el tratamiento del cáncer o metástasis del cáncer. Preferentemente, el órgano o tejido enfermo se caracteriza por las células enfermas tales como las células cancerosas que expresan un antígeno asociado a la enfermedad y/o se caracterizan por la asociación de un antígeno asociado a la enfermedad con su superficie. La inmunización con los antígenos asociados a tumores intactos o sustancialmente intactos o fragmentos de estos, tales como los péptidos o ácidos nucleicos del MHC de clase I y clase II, en particular el ARNm, que codifica dicho antígeno o fragmento hace posible inducir una respuesta tipo MHC de clase I y/o una clase II y, por lo tanto, estimulan las células T como los linfocitos T citotóxicos CD8+ que son capaces de lisar las células cancerosas y/o células T CD4+. Dicha inmunización puede inducir además una respuesta inmune humoral (respuesta de células B) que resulta en la producción de anticuerpos contra el antígeno asociado a tumor. Además, las células presentadoras de antígenos (APC), tal como las células dendríticas (DC), pueden cargarse con péptidos presentados por MHC de clase I mediante transfección con los ácidos nucleicos que codifican los antígenos tumorales *in vitro* y administrarse a un paciente. En una modalidad, el término "antígeno asociado a tumor" se refiere a un constituyente de células cancerosas que pueden derivarse del citoplasma, la superficie celular y el núcleo celular. En particular, se refiere a los antígenos que se producen, preferentemente en gran cantidad, intracelularmente o como antígenos de superficie en las células tumorales. Los ejemplos de antígenos tumorales incluyen HER2, EGFR, VEGF, antígeno-CAMPATH1, CD22, CA-125, HLA-DR, linfoma de Hodgkin o mucina-1, pero no se limitan a ellos.

De acuerdo con la presente invención, un antígeno asociado a tumor comprende preferentemente cualquier antígeno que sea característico para tumores o cánceres, así como para células tumorales o cancerosas con respecto al tipo y/o nivel de expresión. En una modalidad, el término "antígeno asociado a tumor" se relaciona con proteínas que en condiciones normales, es decir, en un sujeto sano, se expresan específicamente en un número limitado de tejidos y/u órganos o en etapas específicas del desarrollo, por ejemplo, el antígeno asociado a tumor puede expresarse en condiciones normales específicamente en tejido de estómago, preferentemente, en la mucosa gástrica, en órganos reproductores, por ejemplo, en testículos, en tejido trofoblástico, por ejemplo, en la placenta, o en células de la línea germinal, y se expresan o se expresan de manera aberrante en uno o más tejidos tumorales o de cáncer. En este contexto, "un número limitado" significa, preferentemente, no más de 3, con mayor preferencia no más de 2 o 1. Los antígenos asociados a tumores en el contexto de la presente invención incluyen, por ejemplo, antígenos de diferenciación, preferentemente, antígenos de diferenciación de tipo celular específico, es decir, proteínas que en condiciones normales se expresan específicamente en un cierto tipo de célula, en cierta etapa de diferenciación, antígenos de cáncer/testículos, es decir, proteínas que en condiciones normales se expresan específicamente en testículos y algunas veces en la placenta, y antígenos específicos de línea germinal. En el contexto de la presente invención, el antígeno asociado a tumor preferentemente no se expresa o raramente se expresa en los tejidos normales o se muta en las células tumorales. Preferentemente, el antígeno asociado a tumor o la expresión aberrante del antígeno asociado a tumor identifica las células cancerosas. En el contexto de la presente invención, el antígeno asociado a tumor que se expresa por una célula cancerosa en un sujeto, por ejemplo, un paciente que padece una enfermedad de cáncer, es preferentemente una autoproteína en dicho sujeto. En modalidades preferidas, el antígeno asociado a tumor en el contexto de la presente invención se expresa en condiciones normales específicamente en un tejido u órgano que no es esencial, es decir, los tejidos u órganos que cuando se dañan por el sistema inmunitario no conducen a la muerte del sujeto, o en órganos o estructuras del cuerpo que no son, o son difícilmente asequibles, al sistema inmunitario. Preferentemente, un antígeno asociado a tumor se presenta en el contexto de moléculas de MHC por una célula cancerosa en la que se expresa.

Los ejemplos de antígenos de diferenciación que cumplen idealmente los criterios para antígenos asociados a tumor como se contempla por la presente invención como estructuras objetivo en la inmunoterapia tumoral, en particular, en la vacunación tumoral son las proteínas de superficie celular de la familia claudina, tales como CLDN6 y CLDN18.2. Estos antígenos de diferenciación se expresan en tumores de diferentes orígenes, y son particularmente adecuados como estructuras objetivo en relación con la inmunoterapia del cáncer mediada por anticuerpos debido a su expresión selectiva (no se expresan en un tejido normal de toxicidad relevante) y localización en la membrana plasmática.

Otros ejemplos de antígenos que pueden ser útiles en la presente invención son p53, ART-4, BAGE, beta-catenina/m, BCR-abL CAMEL, CAP-1, CASP-8, CDC27/m, CDK4/m, CEA, CLAUDIN-12, c-MYC, CT, Cyp-B, DAM, ELF2M, ETV6-AML1, G250, GAGE, GnT-V, Gap100, HAGE, HER-2/neu, HPV-E7, HPV-E6, HAST-2, hTERT (o hTRT), LAGE, LDLR/FUT, MAGE-A, preferentemente MAGE-A1, MAGE-A2, MAGE-A3, MAGE-A4, MAGE-A5, MAGE-A6, MAGE-A7, MAGE-A8, MAGE-A9, MAGE-A10, MAGE-A11 o MAGE-A12, MAGE-B, MAGE-C, MART-1/Melan-A, MC1R, miosina/m, MUC1, MUM-1, -2, -3, NA88-A, NF1, NY-ESO-1, NY-BR-1, p190 BCR-abL menor, Pml/RARa, PRAME, proteinasa 3, PSA, PSM, RAGE, RU1 o RU2, SAGE, SART-1 o SART-3, SCGB3A2, SCP1, SCP2, SCP3, SSX, SURVIVIN, TEL/AML1, TPI/m, TRP-1, TRP-2, TRP-2/INT2, TPTE y WT, preferentemente WT-1.

El término "antígeno viral" se refiere a cualquier componente viral que tenga propiedades antigénicas, es decir, que sea capaz de provocar una respuesta inmune en un individuo. El antígeno viral puede ser una ribonucleoproteína viral o una proteína de la envoltura.

El término "antígeno bacteriano" se refiere a cualquier componente bacteriano que tenga propiedades antigénicas, es decir, que sea capaz de provocar una respuesta inmune en un individuo. El antígeno bacteriano puede derivarse de la pared celular o membrana del citoplasma de la bacteria.

5 El término "respuesta inmune", como se usa en la presente descripción, se refiere a una reacción del sistema inmune tal como a organismos inmunogénicos, tales como bacterias o virus, células o sustancias. El término "respuesta inmune" incluye la respuesta inmune innata y la respuesta inmune adaptativa. Preferentemente, la respuesta inmune se relaciona con una activación de las células inmunes, una inducción de biosíntesis de citoquinas y/o producción de anticuerpos. Se prefiere que la respuesta inmune comprenda las etapas de activación de las células presentadoras de antígeno, tales como células dendríticas y/o macrófagos, la presentación de un antígeno o fragmento de este por dichas células presentadoras de antígeno y la activación de células T citotóxicas debido a esta presentación.

15 El término "tratar" o "tratamiento" se refiere a cualquier tratamiento que mejora el estado de salud y/o prolonga (aumenta) la esperanza de vida de un individuo. Dicho tratamiento puede eliminar la enfermedad en un individuo, detener o enlentecer el desarrollo de una enfermedad en un individuo, inhibir o enlentecer el desarrollo de una enfermedad en un individuo, disminuir la frecuencia o gravedad de los síntomas en un individuo, y/o disminuir la recurrencia en un individuo quien actualmente tiene o quien previamente tuvo una enfermedad.

20 En particular, el término "tratamiento de una enfermedad" incluye curar, acortar la duración, mejorar, retrasar o inhibir la progresión o el empeoramiento, de una enfermedad o los síntomas de esta.

El término "inmunoterapia" se refiere a un tratamiento que involucra preferentemente una reacción inmune específica y/o función(es) inmune(s) efectora(s).

25 El término "inmunización" o "vacunación" describe el proceso de tratar a un sujeto por razones terapéuticas o profilácticas.

30 El término "sujeto" o "individuo", como se usa en la presente descripción, se refiere preferentemente a mamíferos. Por ejemplo, los mamíferos en el contexto de la presente invención son seres humanos, primates no humanos, animales domesticados tales como perros, gatos, ovejas, vacuno, cabras, cerdos, caballos, etc., animales de laboratorio tales como ratones, ratas, conejos, conejillos de indias, etc., así como animales en cautiverio, tal como animales de zoológicos. En una modalidad preferida, el sujeto es un humano.

35 El término "célula presentadora de antígeno" (APC) se relaciona con una célula de una variedad de células capaces de mostrar, adquirir y/o presentar al menos un antígeno o fragmento antigénico sobre (o en) su superficie celular. Las células presentadoras de antígeno pueden distinguirse en células presentadoras de antígeno profesionales y células presentadoras de antígeno no profesionales.

40 El término "células presentadoras de antígenos profesionales" se refiere a las células presentadoras de antígenos que expresan de forma constitutiva las moléculas del Complejo Mayor de Histocompatibilidad de clase II (MHC clase II) necesarias para la interacción con las células T vírgenes. Si una célula T interactúa con el complejo de moléculas MHC de clase II en la membrana de la célula presentadora de antígeno, la célula presentadora de antígeno produce una molécula coestimuladora que induce la activación de las células T. Las células presentadoras de antígenos profesionales comprenden las células dendríticas y los macrófagos.

45 El término "células presentadoras de antígenos no profesionales" se refiere a las células presentadoras de antígenos que no expresan constitutivamente moléculas MHC de clase II, pero que se estimulan después por ciertas citoquinas tal como interferón-gamma. A modo de ejemplo, las células presentadoras de antígenos no profesionales incluyen los fibroblastos, las células epiteliales tímicas, las células epiteliales tiroideas, las células gliales, las células beta pancreáticas o las células endoteliales vasculares.

50 En una modalidad de la invención, los ácidos nucleicos tal como ARN se administran al paciente mediante métodos ex vivo, es decir, al extraer las células de un paciente, modificar genéticamente dichas células y reintroducir las células modificadas en el paciente. Los métodos de transfección y transducción se conocen por el personal experto.

55 De acuerdo con la invención, el término "transfección" se refiere a la introducción de uno o más ácidos nucleicos en un organismo o en una célula huésped. Pueden emplearse diversos métodos in vitro o in vivo para introducir de acuerdo con la invención los ácidos nucleicos en las células. Tales métodos incluyen la transfección de precipitados de ácidos nucleicos-CaPO₄, la transfección de ácidos nucleicos asociados con DEAE, la transfección o infección con los virus que portan los ácidos nucleicos de interés, y la transfección mediada por liposomas, y similares.

60 De acuerdo con la invención, los ácidos nucleicos pueden dirigirse a las células particulares. En tales modalidades, un portador usado para la administración de un ácido nucleico en una célula (por ejemplo, un retrovirus o un liposoma) puede presentar una molécula unida al objetivo. Por ejemplo, puede incorporarse una molécula, tal como un anticuerpo específico a una proteína de membrana superficial en la célula objetivo o un ligando para un receptor en la célula objetivo en o unido al portador del ácido nucleico. Si se desea la administración de un ácido nucleico a través de liposomas, pueden incorporarse en la formulación del liposoma las proteínas que se unen a una proteína de membrana superficial

asociada con la endocitosis para permitir el direccionamiento y/o la adsorción. Tales proteínas incluyen las proteínas de la cápsida o fragmentos de estas que son específicos a un tipo particular de células, anticuerpos para proteínas que se internalizan, proteínas que se dirigen a un sitio intracelular, y similares.

5 "Reportero" se refiere a una molécula, típicamente un péptido o proteína, que está codificado por un gen reportero y se mide en un ensayo de reportero. Los sistemas convencionales usualmente emplean un reportero enzimático y miden la actividad de dicho reportero.

10 El término "sitio de clonación múltiple" se refiere a una región del ácido nucleico que contiene sitios de enzimas de restricción, cualquiera de los cuales puede usarse para la escisión de, por ejemplo, un vector y la inserción de un ácido nucleico.

15 De acuerdo con la invención, dos elementos tales como nucleótidos o aminoácidos son consecutivos, si están directamente adyacentes entre sí, sin ninguna interrupción. Por ejemplo, una secuencia de x nucleótidos consecutivos N se refiere a la secuencia (N)_x.

20 "Endonucleasa de restricción" o "enzima de restricción" se refiere a una clase de enzimas que escinden los enlaces fosfodiéster en ambas cadenas de una molécula de ADN dentro de las secuencias de bases específicas. Se reconocen los sitios de unión específicos, denominados como secuencias de reconocimiento, en una molécula de ADN de doble cadena. Los sitios en los cuales dichos enlaces fosfodiéster en el ADN se escinden por dichas enzimas se denominan como sitios de escisión. En el caso de las enzimas tipo IIS, el sitio de escisión se encuentra a una distancia definida del sitio de unión al ADN. De acuerdo con la invención, el término "endonucleasa de restricción" comprende, por ejemplo, las enzimas SapI, EciI, BpI, AarI, AclI, BaeI, BbvCI, PpI and PstI, BsrD1, BtsI, EarI, BmrI, BsaI, BsmBI, FauI, BbsI, BciVI, BfuAI, BspMI, BseRI, EciI, BtgZI, BpuEI, BsgI, MmeI, CspCI, BaeI, BsaMI, Mva1269I, PctI, Bse3DI, BseMI, Bst6I, Eam1104I, Ksp632I, BfiI, Bso31I, BspTNI, Eco31I, Esp3I, BfuI, Acc36I, AarI, Eco57I, Eco57MI, GsuI, AclI, Hin4I, PpI, y PstI.

30 El término "estabilidad" del ARN se refiere a la "vida media" del ARN. "Vida media" se refiere al período de tiempo que se necesita para eliminar la mitad de la actividad, cantidad, o número de moléculas. En el contexto de la presente invención, la vida media de un ARN es indicativo de la estabilidad de dicho ARN.

35 Los ácidos nucleicos, como el ARN descrito en la presente descripción, en particular cuando se usan para los tratamientos descritos en la presente descripción, pueden estar presentes en forma de una composición farmacéutica o estuche que comprende el ácido nucleico y opcionalmente uno o más portadores, diluyentes y/o excipientes farmacéuticamente aceptables.

Las composiciones farmacéuticas son preferentemente estériles y contienen una cantidad eficaz del ácido nucleico.

40 Las composiciones farmacéuticas se proporcionan usualmente en una forma de dosificación uniforme y pueden prepararse de una forma conocida en la técnica. La composición farmacéutica puede estar, por ejemplo, en la forma de una solución o una suspensión.

45 La composición farmacéutica de la invención puede comprender sales, sustancias tamponantes, conservantes, portadores, diluyentes y/o excipientes los que son preferentemente farmacéuticamente aceptables. El término "farmacéuticamente aceptable" se refiere a la no toxicidad de un material que no interactúa con la acción del componente activo de la composición farmacéutica.

50 Las sales que no son farmacéuticamente aceptables pueden usarse para preparar sales farmacéuticamente aceptables y se incluyen en la invención. Las sales farmacéuticamente aceptables de este tipo comprenden de manera no limitante, las preparadas a partir de los ácidos siguientes: ácidos clorhídrico, bromhídrico, sulfúrico, nítrico, fosfórico, maleico, acético, salicílico, cítrico, fórmico, malónico, succínico y similares. Las sales aceptables farmacéuticamente pueden prepararse también como sales de metales alcalinos o sales de metales alcalinotérreos, tales como sales de sodio, sales de potasio o sales de calcio.

55 Las sustancias tampón adecuadas para su uso en una composición farmacéutica de la invención incluyen el ácido acético en una sal, el ácido cítrico en una sal, el ácido bórico en una sal y el ácido fosfórico en una sal.

60 Los conservantes adecuados para su uso en una composición farmacéutica incluyen el cloruro de benzalconio, el clorobutanol, el parabeno y el timerosal.

65 El término "portador" se refiere a un componente orgánico o inorgánico, de una naturaleza natural o no-natural (sintética), en el que el componente activo se combina para facilitar, mejorar o habilitar la aplicación. De acuerdo con la invención, el término "portador" también incluye uno o más agentes de relleno sólidos o líquidos, diluyentes o sustancias de encapsulación, que son adecuadas para la administración a un paciente.

Las sustancias portadoras posibles para la administración parenteral son, por ejemplo, agua estéril, soluciones de glucosa, Ringer, lactato de Ringer, solución de cloruro de sodio estéril, polialquilenglicoles, naftalenos hidrogenados y, en particular, polímeros biocompatibles de lactida, copolímeros de lactida/glicolida o copolímeros de polioxietileno/polioxipropileno.

5 El término "excipiente" cuando se usa en la presente descripción pretende indicar todas las sustancias que pueden presentarse en una composición farmacéutica y que no son ingredientes activos tales como, por ejemplo, portadores, aglutinantes, lubricantes, espesantes, agentes activos de superficie, conservantes, emulsionantes, tampones, agentes saborizantes, o colorantes.

10 Los agentes y composiciones descritas en la presente descripción pueden administrarse por cualquier vía convencional, tal como por la administración parenteral que incluyen la inyección o la infusión. La administración es preferentemente por la vía parenteral, por ejemplo, intravenosa, intraarterial, subcutánea, intradérmica o intramuscular.

15 Las composiciones adecuadas para la administración parenteral comprenden usualmente una preparación acuosa o no acuosa estéril del componente activo, que es, preferentemente, isotónica para la sangre del receptor. Los ejemplos de portadores y disolventes compatibles son la solución de Ringer y la solución de cloruro de sodio isotónica. Adicionalmente, se usan aceites fijados, usualmente estériles, como solución o medio de suspensión.

20 Los agentes y composiciones descritas en la presente descripción se administran en cantidades efectivas. Una "cantidad efectiva" se refiere a la cantidad que logra una reacción deseada o un efecto deseado solo o junto con dosis adicionales. En el caso del tratamiento de una enfermedad particular o de una afección particular, la reacción deseada se relaciona, preferentemente, con la inhibición del curso de la enfermedad. Esto comprende ralentizar el progreso de la enfermedad y, particularmente, interrumpir o revertir el progreso de la enfermedad. La reacción deseada en un tratamiento de una enfermedad o de una afección puede ser también un retraso en la aparición o una prevención de la aparición de dicha enfermedad o de dicha afección.

25 Una cantidad efectiva de un agente o composición de la que se describe en la presente descripción dependerá de la afección a tratar, la gravedad de la enfermedad, los parámetros individuales del paciente, que incluyen la edad, la condición fisiológica, tamaño y el peso, la duración del tratamiento, el tipo de una terapia de acompañamiento (si está presente), la vía específica de administración y factores similares. En consecuencia, las dosis administradas de los agentes que se describen en la presente descripción pueden depender de varios de tales parámetros. En el caso que una reacción en un paciente sea insuficiente con una dosis inicial pueden usarse dosis más altas (o dosis efectivamente mayores alcanzadas por una vía de administración diferente, más localizada).

30 La presente invención se describe en detalle mediante las siguientes figuras y ejemplos, que deben interpretarse únicamente a modo de ilustración y no a modo de limitación. Sobre la base de la descripción y los ejemplos, modalidades adicionales son accesibles para la persona experta y se encuentran igualmente dentro del alcance de la invención. La materia descrita en las figuras y los ejemplos que no está cubierta con las reivindicaciones solo se describe con propósitos comparativos.

40 Figuras

Figura 1: Pantalla semiautomática sobre la estabilidad de poli(dA:dT)

45 96 clones de *E. coli* que portan un ADN plasmídico con una región poli(dA:dT) se seleccionaron y se inocularon en 1,4 mL en una placa de 96 pocillos durante 14-16 h (37 °C, 225 rpm). Las suspensiones de cultivo bacteriano se recogieron y el ADN plasmídico se purificó con el uso de un estuche de Nucleospin de 96 pocillos (Macherey y Nagel) de acuerdo con el protocolo del fabricante. La concentración de ADN plasmídico se determinó mediante espectroscopia UV (Nanodrop 2000, Thermo Scientific). La integridad de poli(dA:dT) se determinó por análisis de restricción con *SacI* (New England Biolabs). Los fragmentos resultantes se resolvieron en una electroforesis capilar en gel automatizada (Qiagen). **A)** Ejemplo del análisis poli(dA:dT) de 8 clones. Las bandas del marcador interno de tamaño a 25 pb y 500 pb están marcadas con asteriscos negros. Las bandas esperadas para una secuencia poli(dA:dT) de la longitud correcta a 142 pb y 270 pb están marcadas con flechas negras. El clon 1 y el clon 4 muestran una banda adicional resultante de una secuencia de poli(dA:dT) acortada, marcada con asteriscos rojos. **B)** Ejemplo de un mapa del vector que codifica un ARNm que consiste en una región 5' no traducida (5UTR), un gen de interés (GOI), la región 3' no traducida (3UTR) y la cola poli(A) (A120). Se representan los sitios de restricción de *SacI* y se dan las longitudes de los fragmentos tras la incubación con *SacI*.

Figura 2: Estabilidad de los diferentes constructos de poli(dA:dT)

60 ADN plásmido de 96 *E. coli* Los clones de cada construcción poli(dA:dT) se purificaron y se realizó el análisis de restricción *SacI*. Nombres de los constructos: A+números: número de adenosinas 5' de la secuencia del enlazador + L: secuencia del enlazador + número: número de adenosinas 3' de la secuencia del enlazador. Los clones con secuencia de poli(dA:dT) acortada se determinaron y se dan como porcentaje del número total de clones de *E. coli*.

65 Figura 3: Estabilidad de los constructos de poli(dA:dT) en diferentes cepas de *E. coli*

Cepas de *E. coli* TOP10, DH5 α y XL1-blue se usaron para las pruebas de integridad de poli(dA:dT) por análisis de restricción con *SacI*. Se probaron 96 clones para los constructos A120, A30L70 y A40L60. El número de clones con secuencia de poli(dA:dT) acertada se da en porcentaje con respecto al total.

5 Figura 4: Caracterización funcional *in vitro* de diferentes colas de poli(A)

Los plásmidos que codifican para el gen de la luciferasa de luciérnaga que contiene ya sea A120, A30L70 o A40L60 se linealizaron aguas abajo del poli(dA:dT) con una enzima de restricción ClassIIS, que genera por lo tanto un molde sin nucleótido adicional más allá del poli(dA:dT). El ADN plasmídico linealizado se purificó con el uso de perlas magnéticas carboxiladas (Invitrogen), se cuantificó espectrofotométricamente y se sometió a transcripciones *in vitro*. Para las transcripciones *in vitro* se usaron la ARN polimerasa T7 (Fermentas), el tampón de reacción respectivo y los NTP a 6 mM. Para una caperuza eficaz del ARN, la concentración de GTP se redujo a 1,5 mM y se agregaron 6 mM de β -S-ARCA(D2) a la reacción y se incubaron durante 2,5 horas a 37 °C. El ARN se purificó mediante perlas magnéticas carboxiladas (Invitrogen) y la concentración y la calidad del ARN se evaluaron mediante espectrofotometría y análisis en un Bionalyzer 2100 (Agilent). **A)-C) Se mezclaron** 1 x 10⁶ células dendríticas inmaduras humanas (iDC), fibroblastos humanos (CCD) o células de mioblastoma murino (C2C12) con 10 pmol de ARN respectivamente y se sometieron a electroporación. 5 x 10⁴ las células se sembraron en medio X-VIVO15 (Lonza) con los aditivos en las placas de 24 pocillos. A las 2, 4, 8, 24, 48 y 72 horas después de la siembra, se determinaron las actividades de luciferasa de luciérnaga mediante la adición de Luciferina (Promega) en un lector de fluorescencia (TECAN).

20 Figura 5: Caracterización funcional *in vivo* de diferentes colas de poli(A)

Los ratones BALB/c (n=5) se inyectaron por vía intravenosa con lipoplexos de ARN que contenían 20 μ g de ARN que codifica para luciferasa (Luc-ARN) y que portan las diferentes colas de poli(A) A120, A30L70 o A40L60. La captación y traducción de Luc-ARN se evaluaron mediante la formación de imágenes bioluminiscentes *in vivo* con el uso del sistema de imágenes IVIS Lumina (Caliper Life Sciences). En resumen, se inyectó por la vía i.p. una solución acuosa de D-luciferina (150 mg/kg de peso corporal) (BD Biosciences). 6 horas después de la administración de los lipoplexos de ARN. 5 minutos después, se cuantificaron los fotones emitidos (tiempo de integración de 1 minuto). La bioluminiscencia *in vivo* en las regiones de interés se cuantificaron como radiancia promedio (fotones/seg/cm).²/sr) con el uso del programa IVIS Living Image 4.0. La intensidad de la luz transmitida que se origina en las células que expresan luciferasa dentro del animal se representó como una imagen a escala de color, donde la señal de bioluminiscencia azul es la menos intensa y la rojo es la más intensa. Las imágenes de referencia en escala de grises de los ratones se obtuvieron con iluminación LED de baja intensidad. Las imágenes se superpusieron con el uso del programa Living Image 4.0. La señal de la luciferasa se monitorizó durante 48 h. **A)** Se muestra la actividad de la luciferasa en el bazo de los ratones. **B)** Cuantificación de la señal acumulada de la luciferasa monitorizada durante 48 horas.

35 Figura 6: Comparación de la respuesta inmunológica de diferentes colas poli(A)

Ratones C57BL/6 (n=5) se inmunizaron por vía intravenosa en duplicados con lipoplexos de ARN que contenían 20 μ g del ARN que codifica el péptido SIINFEKL que porta las diferentes colas de poli A A120, A30L70 o A40L60 en los días 0 y 3. Las frecuencias de células T CD8⁺ específicas a antígeno se determinaron en sangre periférica a través de la tinción con tetrámero de SIINFEKL-MHC 5 días después de la última inmunización (Día 8). En resumen, las muestras de sangre lisadas por hipotonicidad ase incubaron 4 °C con el anticuerpo anti-CD8 (Invitrogen) y el tetrámero H-2 K^b/SIINFEKL (Beckman-Coulter) y se lavaron para eliminar los anticuerpos no unidos antes del análisis por citometría de flujo. Los datos de citometría de flujo se adquirieron en un citómetro de flujo analítico FACS-Calibur y se analizaron con el uso del programa FlowJo (Tree Star). El perfil de ARN se obtuvo a partir del Bioanalyzer 2100 del ARN que codifica para la luciferasa que porta las colas de poli(A) A40L60, A30L70 y A120 respectivamente. **(A)** Estrategia de activación para las células T CD8⁺ específicas a antígeno. **(B)** Frecuencias de células T CD8⁺ específicas a antígeno en células T CD8⁺.

50 EJEMPLOS

Ejemplo 1: Pantalla semiautomática sobre la estabilidad de poli(dA:dT)

Se estableció un proceso semiautomático para detectar un gran número de clones de *E. coli* para la integridad de la región de secuencia poli(dA:dT) crítica codificada en el plásmido portado por los clones individuales de *E. coli*. Para el tamizaje de un constructo de poli(dA:dT) específico, 96 clones de *E. coli* se inocularon y se incubaron en una placa de 96 pocillos a 37 °C. Las células se recogieron por centrifugación y los plásmidos se purificaron en la plataforma de purificación basada en vacío de una placa de 96 pocillos. Los ADN plasmídicos ensayados contenían tres sitios de restricción *SacI*, escindiendo el vector dos veces en la 3-UTR (región 3' no traducida) y una vez aguas abajo de la secuencia poli(dA:dT). **La restricción *SacI*** dio como resultado 2 bandas específicas siempre en 142 pb y 270 pb de tamaño, lo que permitió el cálculo de la longitud del poli(dA:dT). La tercera banda representó el esqueleto del vector y el antígeno con un tamaño que depende del antígeno insertado (GOI = gen de interés). En la Figura 1B se representa un mapa vectorial ilustrativo con la posición de los sitios de restricción y las longitudes de los fragmentos.

65 Para monitorear un gran número de clones, las muestras de la digestión de restricción con *SacI* se aplicaron en un sistema de electroforesis capilar semiautomático y los patrones de bandas entre 25 pb y 500 pb se analizaron en alta resolución

(la Figura 1A muestra un ejemplo de un análisis de restricción de 8 clones). Las bandas del estándar interno de tamaño a 25 pb y 500 pb están marcadas con asteriscos negros (*). Las bandas que representan un poli(dA:dT) intacto a 142 pb y 270 pb están marcadas con flechas negras (->). El clon 1 y el clon 4 muestran una subpoblación con una región poli(dA:dT) acortada que resulta en una banda adicional entre 142 pb y 270 pb (marcada con asteriscos rojos (*)). La inestabilidad de la poli(dA:dT) se da como la proporción de clones con secuencia de poli(dA:dT) acortada para los clones con una secuencia de poli(dA:dT) intacta.

Ejemplo 2: Pruebas de estabilidad de diferentes constructos de poli(dA:dT)

Como antígeno modelo, se eligió el péptido SIINFEKL debido a que en experimentos previos la inestabilidad de poli(dA:dT) de este antígeno se determinó de manera reproducible entre 50-60% y, por lo tanto, proporciona una gran ventana experimental para las pruebas de estabilidad. Se diseñaron 10 constructos de poli(dA:dT) diferentes y se fusionaron directamente detrás del péptido SIINFEKL. Se insertó un enlazador de 10 nucleótidos (L) en el estiramiento poli(dA:dT) en diferentes posiciones de la secuencia poli(dA:dT). La secuencia del enlazador (GCATATGACT) se eligió de manera que contenga una contribución equilibrada de los 4 nucleótidos (2xG, 2xC, 3xT y 3xA). Se diseñaron 4 constructos con el enlazador en el medio del poli(dA:dT), comenzando con 45 residuos de adenosina (45 x A) en cada lado (A45L45) con un incremento gradual de 5 x A, ambos lados terminando con 60 x A en cada lado del enlazador (A50L50, A55L55 y A60L60, respectivamente).

Los 6 constructos restantes contenían el tipo A50L50 100 x A en total. Sin embargo, el enlazador se insertó después de 20 x A, seguido de la secuencia del enlazador y otro 80 x A (A20L80). En consecuencia, el enlazador se insertó después de 30 x A (A30L70), 40 x A (A40L60), 60 x A (A60L40), 70 x A (A70L30) y 80 x A (A80L20) respectivamente.

Se analizaron 96 clones de cada uno de los 10 constructos para determinar la integridad del poli(dA:dT) con el método de análisis de restricción descrito. Todos los 10 constructos que contienen enlazadores mostraron un efecto beneficioso sobre la estabilidad de poli(dA:dT) en comparación con el A120 (ver Figura 2). Los datos de estabilidad determinados se resumen en la Tabla 1. El constructo A45L45 mostró una estabilidad más de 6 veces mayor en comparación con el control A120, sin embargo, el aumento gradual de la longitud total del poli(dA:dT) condujo a una mayor inestabilidad, lo que se refleja en solo 1,66 veces la estabilización que permanece en A60L60. La estabilización de constructos con 100 x A y la secuencia del enlazador en posiciones variables de la secuencia poli(dA:dT) varió de 2,9 veces para A20L80 a 13 veces para A40L60. Sorprendentemente, A30L70 y A40L60 mostraron una alta estabilización particular de la región poli(dA:dT). En conjunto, nuestros resultados demuestran que la inserción de una secuencia aleatoria de 10 nucleótidos tiene un efecto estabilizador en la integridad de poli(dA:dT). Especialmente la región entre la posición 30 y la posición 50 de la región poli(dA:dT) es particularmente sensible al poli(dA:dT) acortado. La introducción de secuencias enlazadoras en esta área de secuencia condujo a un aumento adicional de la estabilidad de poli(dA:dT) en al menos 2 veces en comparación con otros constructos (ver la Tabla 1 y Figura 2).

Tabla 1: Resumen de las pruebas de estabilidad de poli(dA:dT).

Se muestra el porcentaje de clones con la secuencia poli(dA:dT) acortada y la estabilización resultante de la secuencia poli(dA:dT) en comparación con la poliA120.		
Constructo de poli(dA:dT)	Escisión [% de clones probados]	Estabilización [pliegue de A120]
A120	55,9	1
A45L45	8,8	6,4
A50L50	10,7	5,2
A55L55	21,1	2,7
A60L60	33,7	1,7
A60L40	8,9	6,3
A70L30	13,8	4,0
A80L20	13,6	4,1
A40L60	4,3	13,0
A30L70	4,4	12,7
A20L80	19,3	2,9

Ejemplo 3: Estabilidad de los constructos de poli(dA:dT) en diferentes cepas de *E. coli*

En experimentos adicionales se probó la especificidad y la funcionalidad de la estabilidad superior de los constructos A30L70 y A40L60. La posibilidad de que los resultados observados de las pruebas de estabilidad estén restringidos a la cepa de *E. coli* TOP10 probada se evaluó incluyendo otras dos cepas de *E. coli* en la prueba. Las pruebas para A30L70 y A40L60 se repitieron con DH5 α , XL1-blue y TOP10 como control respectivamente. Estas cepas se eligieron como i) que tienen una alta diversidad genética (ver Tabla 2) y ii) que representan las cepas de *E. coli* que se usan ampliamente en los laboratorios de biología molecular.

La inestabilidad de la A120 se midió para DH5 α al 42% y para XL1-blue al 61,8% y, por lo tanto, se consideró comparable a la inestabilidad detectada para la cepa de *E. coli* TOP10 (ver Figura 3). Ambos, A30L70 y A40L60 mostraron una inestabilidad entre 3-4%, solo para A40L60 en TOP10 la inestabilidad se elevó ligeramente a 6,8%. Las pruebas de las 3 cepas diferentes de laboratorio de *E. coli* confirmaron los resultados sobre la estabilización de poli(dA:dT). La introducción de una secuencia enlazadora de 10 nucleótidos en la región sensible a la escisión en la posición 30-50 se identificó como un principio general para la estabilización genética de las secuencias poli(dA:dT) en diferentes cepas de *E. coli*.

Tabla 2: Genotipos de las cepas de *E. coli* testadas

Cepa	Genotipo
TOP10	F-, mcrA, Δ (mrr-hsdRMS-mcrBC), ϕ 80lacZ Δ M15, Δ lacX74, nupG, recA1, araD139, Δ (ara-leu)7697, galK galU rpsL(Str ^R), endA1, λ -
DH5 α	F-, endA1, glnV44, thi-1, recA1, relA1, gyrA96, deoR, nupG, Φ 80, lacZ Δ M15, Δ (lacZYA-argF)U169, hsdR17 (r _K ⁻ m _K ⁺), λ -
XL1-blue	endA1, gyrA96(nal ^R), thi-1, recA1, relA1, lac, glnV44, F[: :Tn10, proAB ⁺ , lacI ^q , Δ (lacZ)M15], hsdR17 (r _K ⁻ m _K ⁺)

Ejemplo 4: Caracterización funcional *in vitro* de diferentes colas de poli(A)

Se realizaron experimentos basados en el reportero de la luciferasa para dilucidar el impacto de las colas poli(A) A30L70 y A40L60 estabilizadas e identificadas en la funcionalidad de las moléculas de ARN. Los constructos A30L70, A40L60 y A120 se fusionaron con un gen reportero de la luciferasa de luciérnaga y el ARN mensajero respectivo se generó por transcripción *in vitro*. Las moléculas de ARN mostraron una integridad comparable y se usaron para la electroporación celular (ver Tabla 3). Los ARN se sometieron a la electroporación en células dendríticas inmaduras humanas, aisladas de sangre humana que representan las células objetivo para el enfoque de la vacuna de ARNm contra tumores de la compañía. La traducción de la luciferasa se monitorizó durante un período de 72 horas. Las 3 moléculas de ARNm diferentes se expresaron igualmente con solo pequeñas diferencias (Figura 4A). Para probar que la funcionalidad de los ARNm en general no está influenciada por la naturaleza de las colas poli(A), el experimento se repitió en una línea celular de fibroblastos humanos (células CCD) y una línea celular de mioblastos murinos (C2C12) (Figura 4B+C). Aunque un patrón específico de tipo celular de la traducción del ARNm se monitorizó a lo largo del tiempo, ni las líneas celulares humanas ni las murinas mostraron diferencias en la expresión de proteínas por los ARNm que contenían diferentes colas poli(A).

Tabla 3: Integridad de luciferasa que codifica los ARN de IVT

ARN	Integridad [%]
hAg-Kozak-Luciferasa-2hBgUTR-A40L60	79
hAg-Kozak-Luciferasa-2hBgUTR-A30L70	81
hAg-Kozak-Luciferasa-2hBgUTR-A120	83

Estos resultados demuestran que las colas poli(A) elegidas tienen solo un impacto menor en la funcionalidad del ARNm total *in vitro*. Por lo tanto, las inserciones de secuencia enlazadora en la región poli(dA:dT) en la posición 30 y la posición 40, respectivamente, permiten una estabilización genética sustancial al mantener la funcionalidad completa de las respectivas moléculas de ARN.

Ejemplo 5: Caracterización funcional *in vivo* de diferentes colas de poli(A)

Para la aplicación sistémica *in vivo* del ARNm para la vacunación, los lipoplexos de ARN se generan mediante la formulación del ARN junto con los lípidos y se administran por vía intravenosa. Los lipoplexos de ARN están destinados a dirigirse al bazo y absorberse por las células dendríticas inmaduras que traducen el ARNm respectivo. El objetivo fue probar las dos colas poli(A) estabilizadas, es decir, A30L70 y A40L60 en un experimento con ratones para garantizar la expresión funcional de proteínas *in vivo*. El ARN con A120 sirvió como control de expresión. Tres grupos de ratones BALB/c con 5 animales cada uno se inyectaron por vía intravenosa con lipoplexos de ARN que contenían ARN que codifica

para la luciferasa de luciérnaga que se había usado para la prueba funcional *in vitro* (Tabla 3) con las diferentes colas poli(A) (A30L70, A40L60 y A120). La expresión de luciferasa de luciérnaga se monitorizó durante 48 horas con el uso de un sistema de imágenes de bioluminiscencia *in vivo* (Figura 5A). La cuantificación de las señales de luciferasa acumuladas se muestra en la Figura 5B. Ni la ubicación ni la intensidad de la señal de la luciferasa difirieron significativamente entre los ARN con diferentes colas de poli(A), lo que demuestra que ambas colas de poli(A) estabilizadas son adecuadas para las aplicaciones sistémicas *in vivo*.

Ejemplo 6: Respuesta inmunológica a diferentes colas poli(A)

En un último conjunto de experimentos, se evaluó si las colas estabilizadas de poli(A), A30L70 y A40L60 tienen una influencia en la respuesta inmune específica inducida por la vacuna de ARNm. Por lo tanto, las colas de poli(A) estabilizadas y el control A120 se fusionaron con el péptido SIINFEKL como en las anteriores pruebas de estabilidad. Los 3 ARN que contienen las colas poli(A) A30L70, A40L60 y A120 se generaron por transcripción *in vitro* y mostraron calidad e integridad comparables (Tabla 4). Los 3 grupos de ratones C57/BL6, 5 animales dobles cada uno, se inyectaron por vía intravenosa por duplicado el día 0 y el día 3 con lipoplexos de ARN que contenían los diferentes ARN de SIINFEKL. La respuesta inmune inducida por ARN se analizó determinando la frecuencia de células T CD8⁺ específica a antígeno, 5 días después de la última inmunización (día 8) mediante la tinción con el tetrámero SIINFEKL-MHC. La estrategia de activación respectiva por el análisis FACS se muestra en la Figura 6A. La comparación de la frecuencia de células T CD8⁺ específica a antígeno mostraron que los ARN con todas las colas poli(A) probadas indujeron una respuesta inmune. Por lo tanto, no se detectaron diferencias significativas dentro del mismo grupo (2 x 5 animales para cada ARN) ni entre los 3 grupos que recibieron los diferentes ARN IVT, lo que demuestra que las colas poli(A) estabilizadas no influyeron en la respuesta inmune específica inducida por el ARNm (Figura 6B).

Tabla 4: Integridad de los ARN de IVT que codifican SIINFEKL

ARN IVT	Integridad [%]
hAg-Kozak-sec-SIINFEKL-MITD-2hBgUTR-A40L60	82
hAg-Kozak-sec-SIINFEKL-MITD-2hBgUTR-A30L70	81
hAg-Kozak-sec-SIINFEKL-MITD-2hBgUTR-A120	83

Al establecer un método de análisis basado en la restricción, podemos mostrar en la presente que la región poli(dA:dT) que codifica la cola poli(A) de un ARNm es genéticamente inestable en cepas de *E. coli* comunes. Esta inestabilidad conduce a esfuerzos de tamizaje intensivos para obtener clones con una secuencia estable de poli(dA:dT). Se demuestra que la inserción de una secuencia enlazadora de 10 nucleótidos estabiliza esta extensión de secuencia. Por lo tanto, las posiciones 30 a 50 se identificaron como especialmente sensibles al acortamiento de poli(dA:dT). Las inserciones del enlazador en esta región particular aumentaron la estabilidad aún más en al menos el factor 2 en comparación con las inserciones en otras posiciones. Las pruebas de estabilidad se confirmaron en varias cepas de *E. coli* comúnmente usadas. Las inserciones de la secuencia no alteraron la funcionalidad *in vitro* de los respectivos ARN transcritos como se demuestra en varias líneas celulares y mediante la comparación de la actividad en ratones *in vivo*. Por último, la modificación de la cola poli(A) no influyó en la respuesta inmune inducida por el ARN. En conjunto, se identifica una herramienta para estabilizar genéticamente la región poli(dA:dT) que facilita el manejo con el ADN plasmídico respectivo y, por lo tanto, no influye en la funcionalidad *in vitroe in vivo* del ARN ni en la inducción de una respuesta inmune específica de ARN.

REIVINDICACIONES

1. Una molécula de ácido nucleico que comprende en la dirección 5'→ 3' de la transcripción:
 - (a) un promotor;
 - (b) una secuencia de ácido nucleico transcribible o una secuencia de ácido nucleico para introducir una secuencia de ácido nucleico transcribible; y
 - (c) una secuencia de ácido nucleico que, cuando se transcribe bajo el control del promotor (a), codifica una secuencia de nucleótidos de al menos 80 nucleótidos consecutivos en el transcrito, en donde dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos en el transcrito es una secuencia de poliadenilo que comprende dentro de la secuencia de poliadenilo una secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A, en donde dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A está precedida por al menos 20 residuos A en dicha secuencia de poliadenilo y está seguida por al menos 20 residuos A en dicha secuencia de poliadenilo.
2. La molécula de ácido nucleico de acuerdo con la reivindicación 1, en donde dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de A es una secuencia, preferentemente una secuencia arbitraria, de 2 o más nucleótidos consecutivos, en donde el primer y el último nucleótido de dicha secuencia de 2 o más nucleótidos consecutivos es un nucleótido distinto de un nucleótido A.
3. La molécula de ácido nucleico de acuerdo con la reivindicación 1 o 2, en donde dicha secuencia de ácido nucleico (c) muestra una mayor estabilidad tras la propagación de dicha molécula de ácido nucleico en Escherichia coli en comparación con una molécula de ácido nucleico que comprende en lugar de dicha secuencia de ácido nucleico (c) una secuencia de ácido nucleico (c)' que, cuando se transcribe bajo el control del promotor (a), codifica para una secuencia de poliadenilo de igual longitud que dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos en el transcrito.
4. La molécula de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1 a 3, en donde dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos comprende al menos 90 nucleótidos, preferentemente al menos 100 nucleótidos.
5. La molécula de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1 a 4, en donde dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A se localiza dentro de una región desde la posición 21 a la posición 80, preferentemente desde la posición 21 a la posición 60, con mayor preferencia desde la posición 31 a la posición 50 de dicha secuencia de poliadenilo.
6. La molécula de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1 a 5, en donde dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A tiene una longitud de al menos 3, al menos 4, al menos 5, al menos 6, al menos 8, preferentemente al menos 10, con mayor preferencia al menos 15 nucleótidos; y/o en donde dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A tiene una longitud de no más de 50, preferentemente no más de 30, con mayor preferencia no más de 20 nucleótidos.
7. La molécula de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1 a 6, en donde dicha secuencia de uno o más nucleótidos consecutivos que contienen nucleótidos distintos de los nucleótidos A no comprende más de 3, preferentemente no más de 2, preferentemente sin residuos de A consecutivos.
8. La molécula de ácido nucleico de acuerdo con cualquiera de las reivindicaciones 1 a 7, que tiene una o más de las siguientes características:
 - (i) las secuencias de ácido nucleico (b) y (c) bajo el control del promotor (a) pueden transcribirse para dar un transcrito común;
 - (ii) en el transcrito dicha secuencia de nucleótidos de al menos 80 nucleótidos consecutivos se ubica en el extremo 3';
 - (iii) la molécula de ácido nucleico es una molécula circular cerrada o una molécula lineal;
 - (iv) la secuencia de ácido nucleico transcribible comprende una secuencia de ácido nucleico que codifica un péptido o proteína y la secuencia de ácido nucleico para introducir una secuencia de ácido nucleico transcribible es un sitio de clonación múltiple;
 - (v) la molécula de ácido nucleico comprende además uno o más miembros seleccionados del grupo que consiste en: (i) un gen reportero; (ii) un marcador seleccionable; y (iii) un origen de replicación;
 - (vi) la molécula de ácido nucleico es adecuada, en particular después de la linealización, para la transcripción in vitro del ARN, en particular del ARNm.
9. Como se reivindica en cualquiera de las reivindicaciones 1 a 8, el ARN que puede obtenerse mediante la transcripción, preferentemente la transcripción in vitro, con el uso de una molécula de ácido nucleico como un molde.

10. Un método para propagar una molécula de ácido nucleico, que comprende:
(i) proporcionar una molécula de ácido nucleico según reivindica cualquiera de las reivindicaciones 1 a 8, y
(ii) propagar dicha molécula de ácido nucleico en *Escherichia coli*, en donde, opcionalmente, propagar dicha molécula de ácido nucleico en *Escherichia coli* comprende transformar *Escherichia coli* con dicha molécula de ácido nucleico y cultivar dicha *Escherichia coli* transformada.
- 5
11. El método de la reivindicación 10, que comprende además aislar dicha molécula de ácido nucleico de *Escherichia coli* después de la propagación.
- 10
12. Un método de obtención de ARN, que comprende:
(i) propagar una molécula de ácido nucleico de acuerdo con el método de la reivindicación 10 u 11, y
(ii) transcribir in vitro el ARN con el uso de la molécula de ácido nucleico como molde, en donde, opcionalmente, el método se caracteriza porque además comprende, antes de la transcripción de la molécula de ácido nucleico, la escisión de la molécula de ácido nucleico.
- 15
13. Un método para obtener un péptido o proteína, que comprende:
(i) obtener un ARNm que codifica el péptido o la proteína de acuerdo con el método de la reivindicación 12, y
(ii) traducir el ARNm.
- 20
14. El uso del ARN según reivindica la reivindicación 9 para transfectar una célula huésped, en donde, opcionalmente, la célula huésped es una célula presentadora de antígeno, en particular una célula dendrítica, un monocito o un macrófago.
- 25
15. El ARN según reivindica la reivindicación 9, para su uso en la vacunación.

Figura 1A

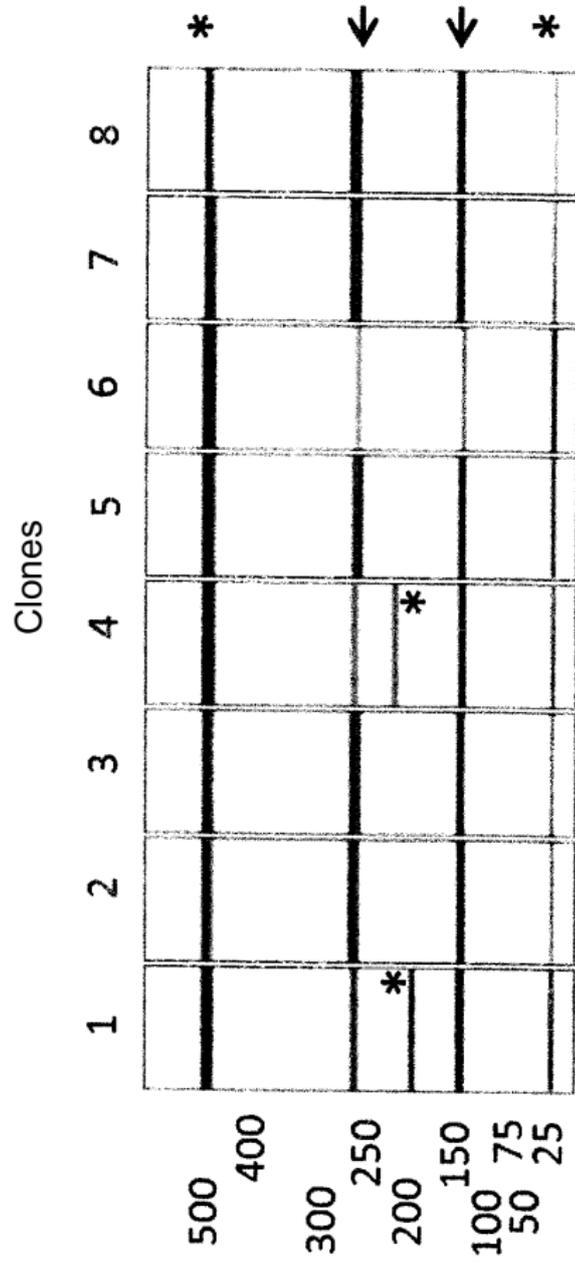
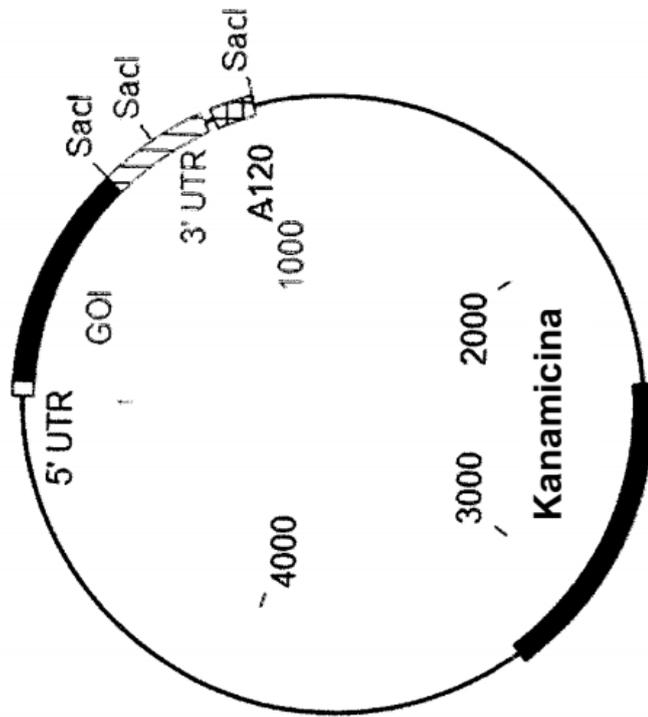


Figura 1B



Fragmentos SacI	Tamaño [bp]
1	142
2	270
3	3892

Figura 2

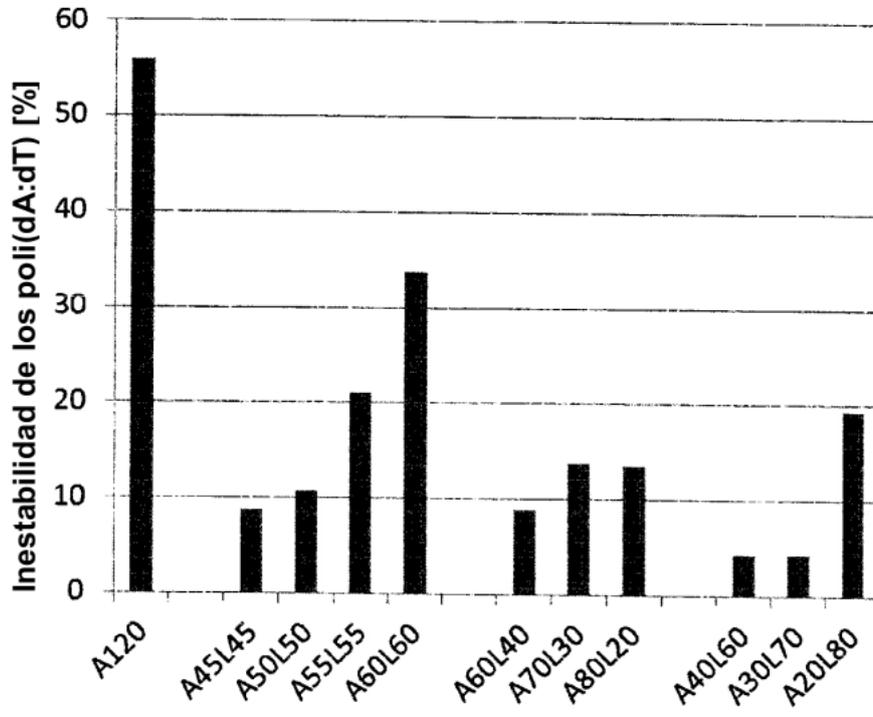
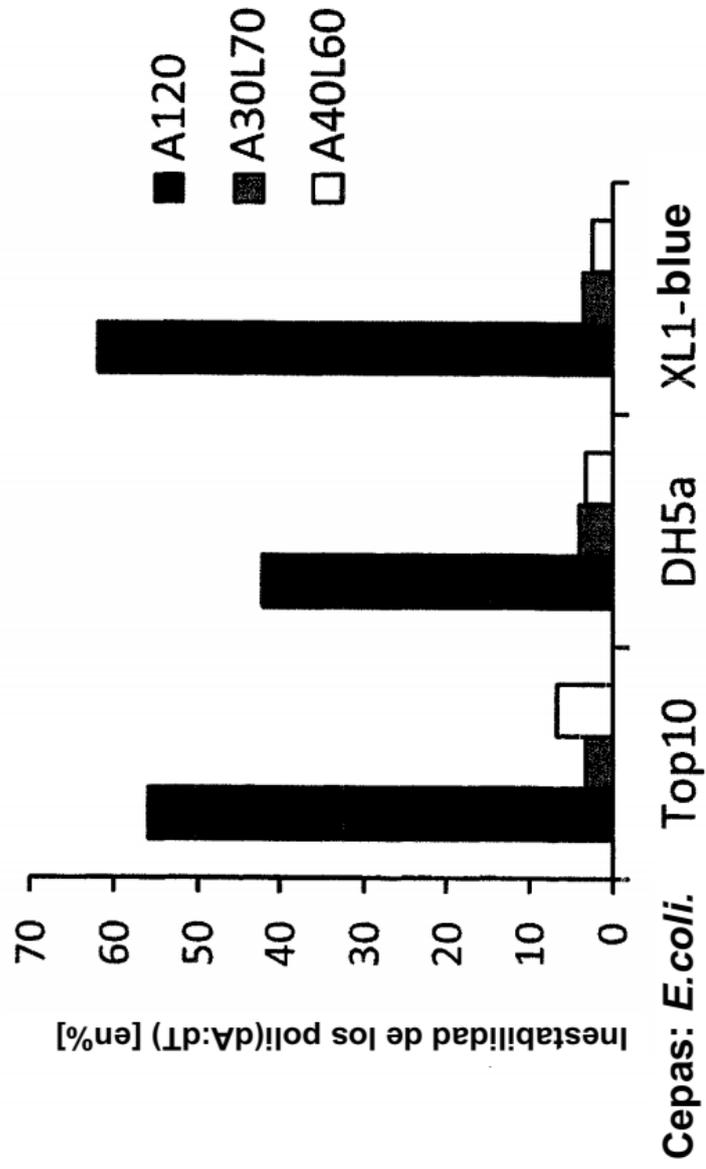
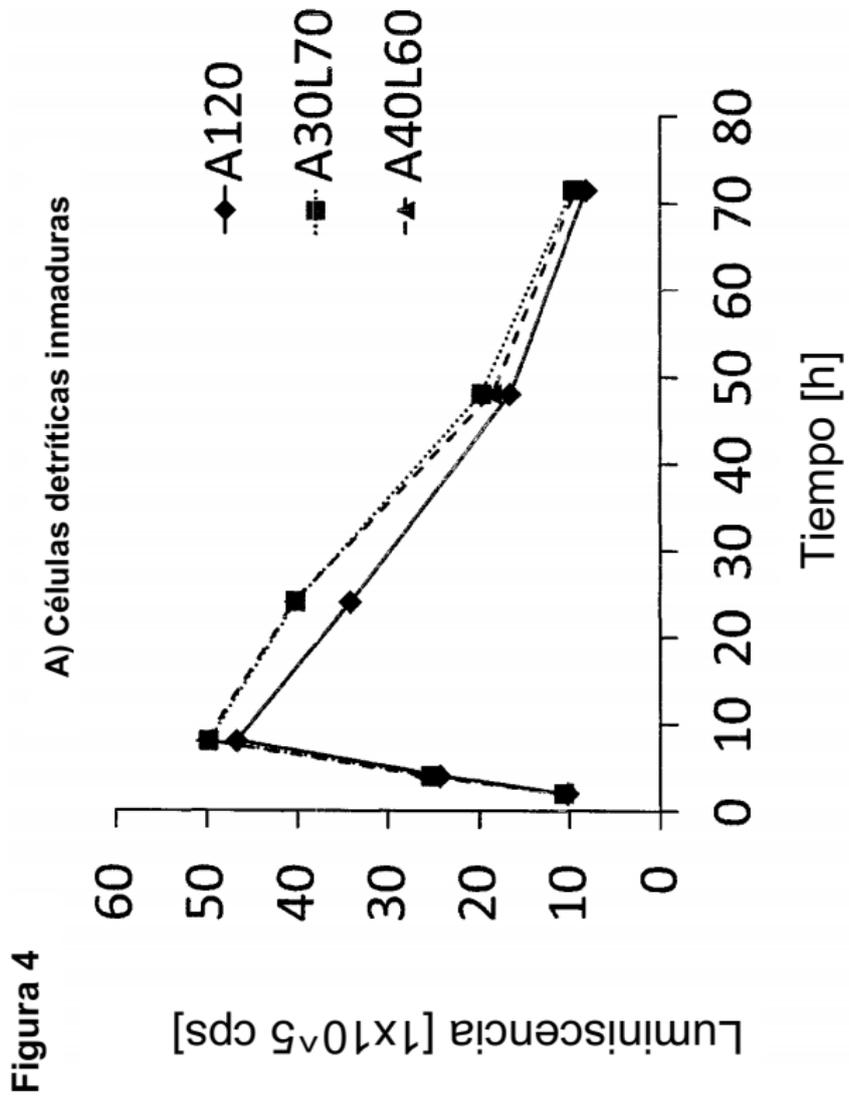
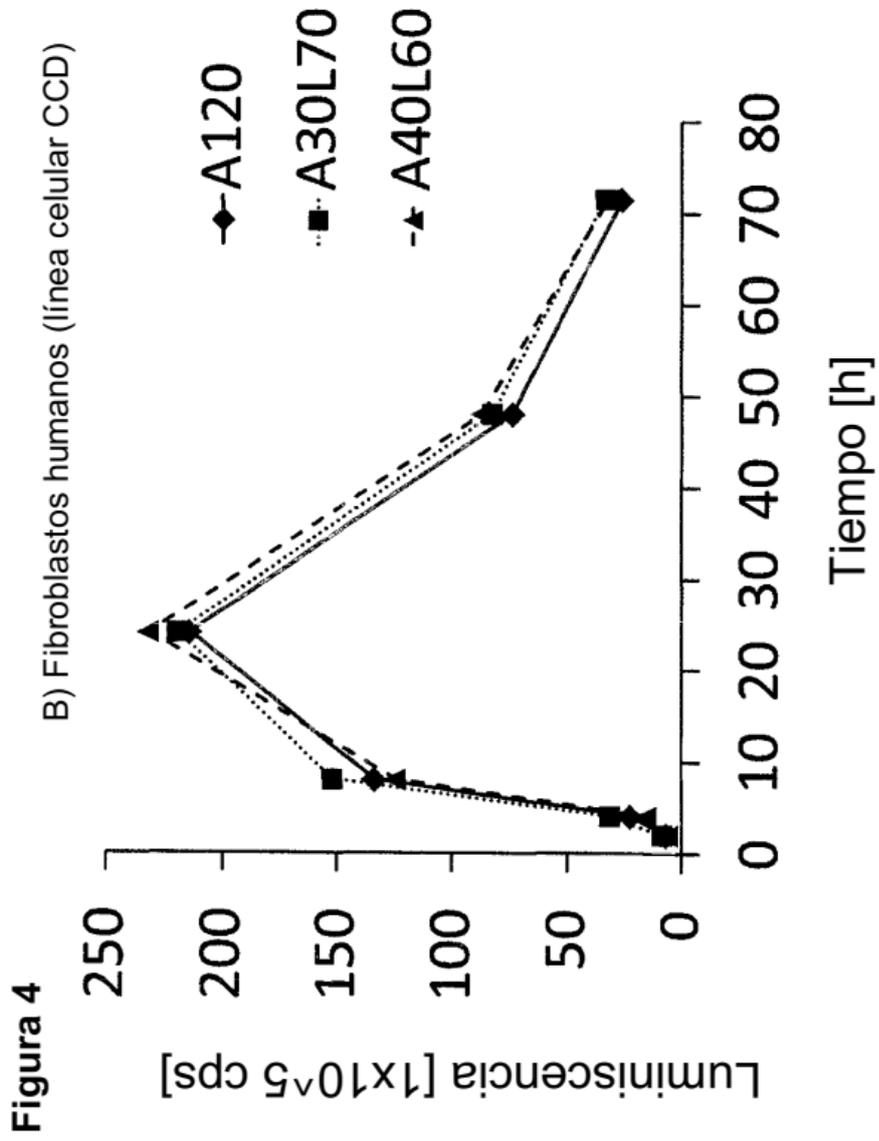


Figura 3







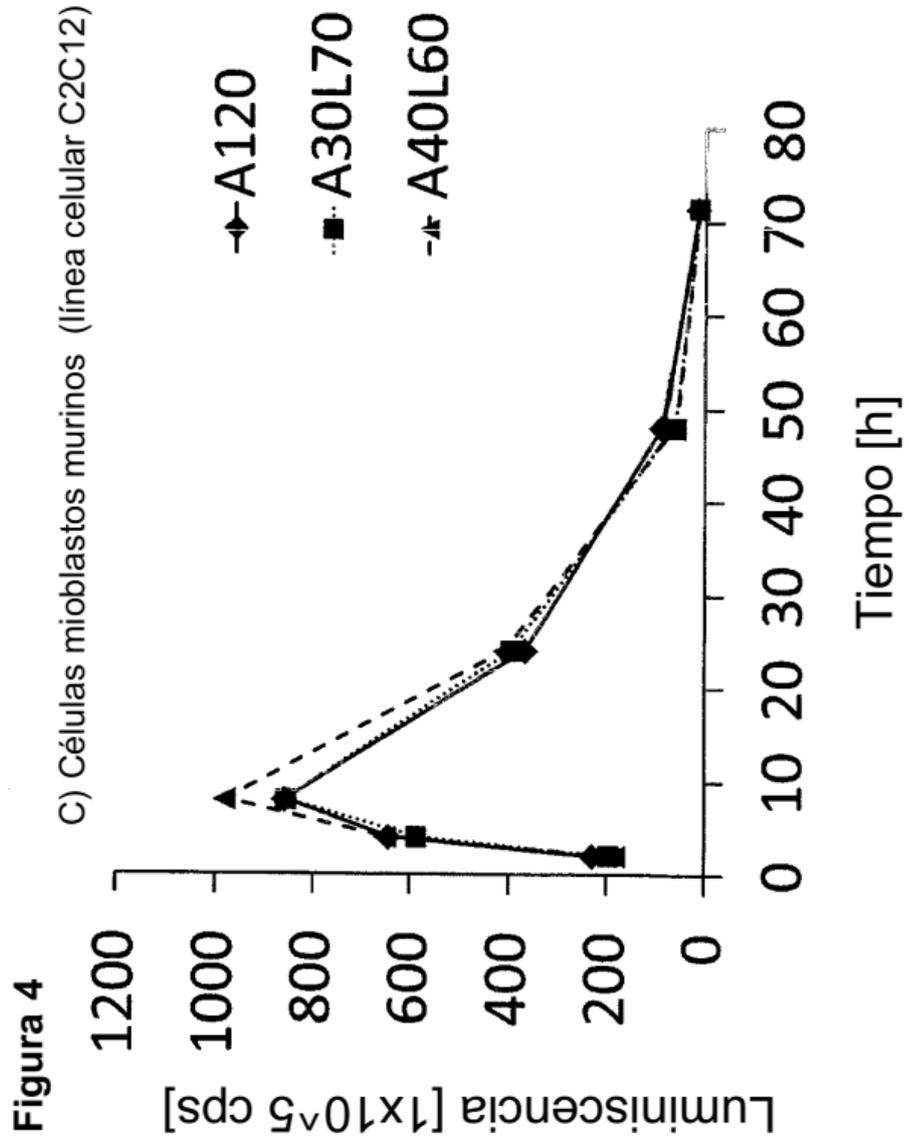


Figura 5 A

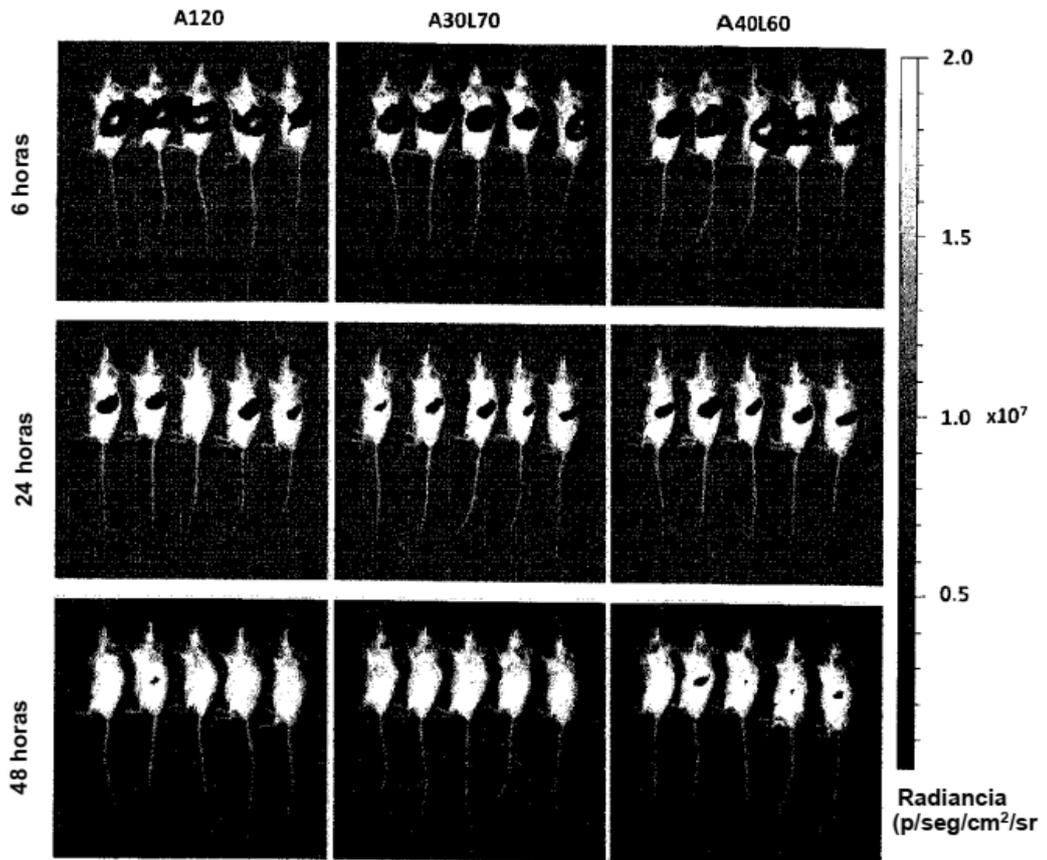


Figura 5B

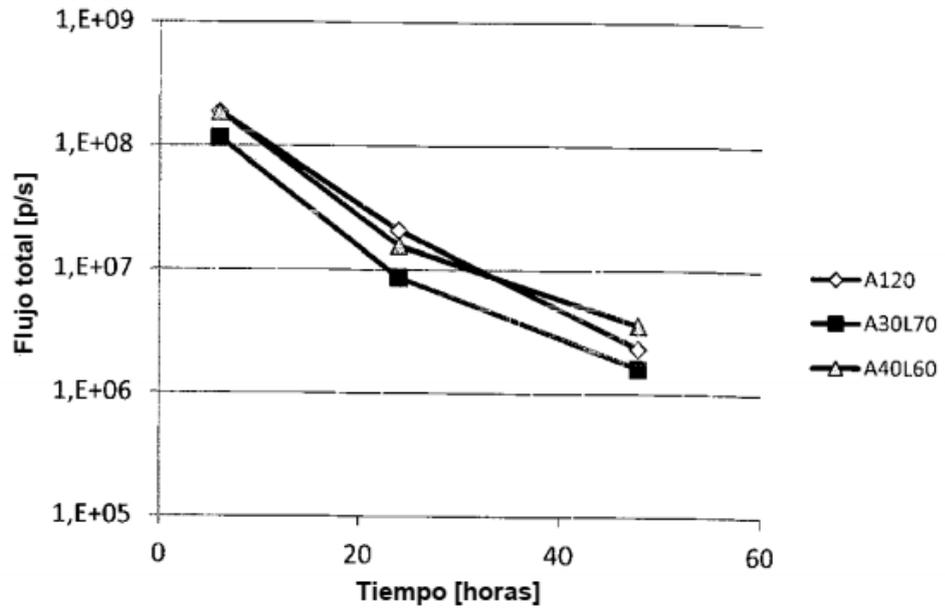
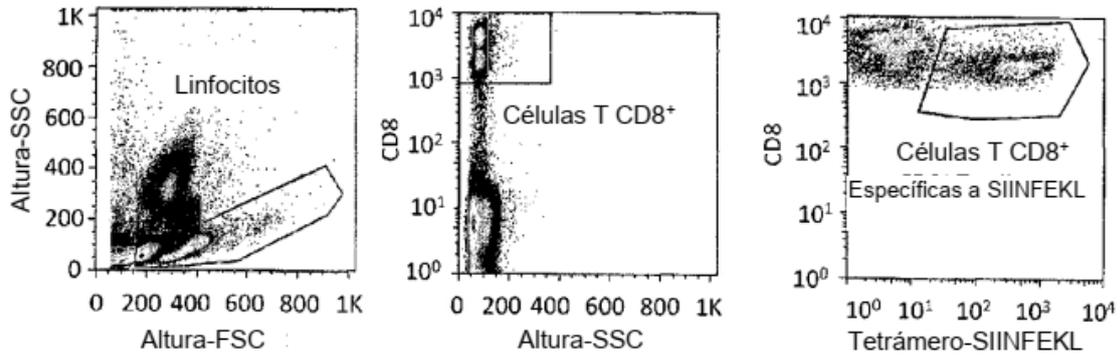


Figura 6

A



B

