

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 993 107**

51 Int. Cl.:

C12N 15/86

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **03.10.2017** **PCT/US2017/054928**

87 Fecha y número de publicación internacional: **26.04.2018** **WO18075235**

96 Fecha de presentación y número de la solicitud europea: **03.10.2017** **E 17787724 (8)**

97 Fecha y número de publicación de la concesión europea: **26.06.2024** **EP 3526332**

54 Título: **Sistemas de replicón de virus recombinantes y usos de los mismos**

30 Prioridad:

17.10.2016 US 201662409228 P

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

23.12.2024

73 Titular/es:

JANSSEN PHARMACEUTICALS, INC. (100.00%)
1125 Trenton-Harbourton Road
Titusville, NJ 08560, US

72 Inventor/es:

KAMRUD, KURT, IVER

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

Observaciones:

Véase nota informativa (Remarks, Remarques o
Bemerkungen) en el folleto original publicado por la
Oficina Europea de Patentes

ES 2 993 107 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

Sistemas de replicón de virus recombinantes y usos de los mismos

5 **Aplicaciones relacionadas**

La presente solicitud reivindica prioridad según 35 USC § 119(e) a la solicitud provisional n.º US-62/409228, presentada el 17 de octubre de 2016.

10 **Incorporación del listado de secuencias**

Por la presente se incorpora a esta solicitud el material del listado de secuencias adjunto. El archivo de texto de lista de secuencias adjunto, denominado SGI011WO_SeqListing.txt, se creó el 3 de octubre de 2017 y tiene 34 KB.

15 **Campo**

La presente descripción se refiere al campo de la biología molecular, incluidas moléculas de ácido nucleico que comprenden replicones virales modificados y el uso de tales moléculas de ácido nucleico para la producción de productos deseados en células huésped adecuadas en cultivo celular o en un cuerpo vivo.

20 **Antecedentes**

En los últimos años, varios grupos diferentes de virus animales han sido sometidos a manipulación genética ya sea mediante recombinación homóloga o mediante ingeniería directa de sus genomas. La disponibilidad de sistemas de genética inversa para virus de ADN y ARN ha creado nuevas perspectivas para el uso de virus recombinantes, por ejemplo, como vacunas, vectores de expresión, agentes antitumorales, vectores de terapia génica y vehículos de administración de fármacos.

Por ejemplo, se han desplegado muchos vectores de expresión basados en virus para la expresión de proteínas heterólogas en células recombinantes cultivadas. En particular, la aplicación de vectores virales modificados para la expresión génica en células huésped continúa expandiéndose. Los avances recientes a este respecto incluyen un mayor desarrollo de técnicas y sistemas para la producción de complejos proteicos de subunidades múltiples y la coexpresión de enzimas modificadoras de proteínas para mejorar la producción de proteínas heterólogas. Otros avances recientes con respecto a las tecnologías de vectores de expresión viral incluyen muchas aplicaciones avanzadas de ingeniería genómica para controlar la expresión génica, la preparación de vectores virales, aplicaciones de terapia génica *in vivo* y la creación de vectores de administración de vacunas. La patente WO 2014/170.493 describe el ARN de replicón de alfavirus modificado con mutaciones en la UTR 5' en las posiciones 3, 24 y 30.

Sin embargo, todavía existe la necesidad de métodos y sistemas más eficientes para expresar genes de interés en sistemas de expresión recombinantes.

Resumen

Esta sección proporciona un resumen general de la presente solicitud y no abarca todo su alcance ni todas sus características.

En un aspecto, en la presente memoria se describe una molécula de ácido nucleico que incluye un ARN de replicón modificado, en el que el ARN de replicón modificado incluye una 5'-UTR modificada y está desprovisto de al menos una porción de una secuencia de ácido nucleico que codifica proteínas estructurales virales. En diversas realizaciones de este aspecto y otros aspectos de la presente descripción, la molécula de ácido nucleico como se describe en la presente memoria puede incluir una o más de las siguientes características. En algunas realizaciones, el ARN de replicón modificado es un ARN de replicón de alfavirus modificado. En algunas realizaciones, el ARN de replicón de alfavirus modificado incluye un genoma de alfavirus modificado. En las realizaciones, la UTR 5' modificada incluye una sustitución de nucleótidos en la posición 2, en donde las sustituciones de nucleótidos en la posición 2 de la UTR 5' modificada son una sustitución U->G.

En algunas realizaciones, la molécula de ácido nucleico incluye un ARN de replicón modificado que carece de una porción sustancial de la secuencia de ácido nucleico que codifica proteínas estructurales virales. En algunas realizaciones, el genoma de alfavirus modificado o el ARN de replicón como se describe en la presente memoria no incluye ninguna secuencia de ácido nucleico que codifique proteínas estructurales virales.

En diversas realizaciones de este aspecto y otros aspectos de la presente descripción, la molécula de ácido nucleico incluye además uno o más casetes de expresión, en donde cada uno de los casetes de expresión incluye un promotor unido operativamente a una secuencia de ácido nucleico heteróloga. En algunas realizaciones, la molécula de ácido nucleico incluye al menos dos, tres, cuatro, cinco o seis casetes de expresión. En algunas realizaciones, el promotor de al menos uno de los casetes de expresión es o comprende un promotor subgenómico 26S.

En algunas realizaciones, la secuencia de ácido nucleico heteróloga de al menos uno de los casetes de expresión como se describe en la presente memoria incluye una secuencia codificante de un gen de interés (GOI, por sus siglas en inglés). En algunas realizaciones, el GOI codifica un polipéptido seleccionado del grupo que consiste en un polipéptido terapéutico, un polipéptido profiláctico, un polipéptido de diagnóstico, un polipéptido neutracéutico, una enzima industrial y un polipéptido informador. En algunas realizaciones, el GOI codifica un polipéptido seleccionado del grupo que consiste en un anticuerpo, un antígeno, un modulador inmunológico y una citoquina. En algunas realizaciones particulares, la secuencia codificante del GOI se optimiza para la expresión a un nivel superior al nivel de expresión de una secuencia codificante de referencia.

En algunas realizaciones, la molécula de ácido nucleico incluye un ARN de replicón modificado que comprende un genoma o ARN de replicón modificado de un virus que pertenece al género *Alphavirus* de la familia *Togaviridae*. En algunas realizaciones, el genoma modificado o ARN de replicón es de un alfavirus que pertenece al grupo VEEV/EEEV, o al grupo SF, o al grupo SIN. En algunas realizaciones, el alfavirus se selecciona del grupo que consiste en virus de la encefalitis equina del este (EEEV), virus de la encefalitis equina venezolana (VEEV), virus de los Everglades (EVEV), virus Mucambo (MUCV), virus Pixuna (PIXV), virus de Middleburg (MIDV), virus Chikungunya (CHIKV), virus O'Nyong-Nyong (ONNV), virus del río Ross (RRV), virus del bosque de Barmah (BF), virus Getah (GET), virus Sagiyama (SAGV), virus Bebaru (BEBV), virus Mayaro (MAYV), virus Una (UNAV), virus Sindbis (SINV), virus Aura (AURAV), virus Whartaroa (WHAV), virus Babanki (BABV), virus Kyzylagach (KYZV), virus de la encefalitis equina occidental (WEEV), virus Highland J (HJV), virus Fort Morgan (FMV), Ndumu (NDUV) y virus Buggy Creek. En algunas realizaciones, el alfavirus es el virus de la encefalitis equina venezolana (VEEV).

Algunas realizaciones proporcionan una molécula de ácido nucleico que incluye un genoma de alfavirus modificado o un ARN de replicón que está unido operativamente a un elemento regulador heterólogo. En algunas realizaciones, el elemento regulador heterólogo incluye una secuencia promotora. En algunas realizaciones, la secuencia promotora incluye una secuencia promotora de T7. En algunas realizaciones, el elemento regulador heterólogo incluye una secuencia de terminación transcripcional. En algunas realizaciones, la secuencia de terminación transcripcional es o comprende una secuencia de terminación T7.

En algunas realizaciones, la molécula de ácido nucleico como se describe en la presente memoria incluye un genoma de alfavirus modificado o un ARN de replicón que incluye un genoma de alfavirus modificado o un ARN de replicón, en donde la molécula de ácido nucleico exhibe al menos 80 % de identidad de secuencia con la secuencia de ácido nucleico de Id. de sec. n.º: 1, en donde el genoma o ARN de replicón del alfavirus modificado comprende una sustitución U→G en la posición 2 de la región 5'-no traducida (UTR 5') y carece de al menos una porción de la secuencia que codifica la estructura viral. proteínas. En algunas realizaciones, la molécula de ácido nucleico exhibe al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % de identidad de secuencia con la secuencia de ácido nucleico de Id. de sec. n.º: 1.

En algunas realizaciones, la molécula de ácido nucleico como se describe en la presente memoria incluye un genoma de alfavirus modificado o ARN de replicón, en donde el genoma de alfavirus modificado o ARN de replicón comprende una UTR 5' que muestra al menos 80 % de identidad de secuencia con la secuencia de ácido nucleico de al menos uno de Id. de sec. n.º: 2-18 y una sustitución U→G en la posición 2 de la UTR 5', y en el que el genoma del alfavirus modificado o el ARN de replicón carece de al menos una porción de la secuencia que codifica proteínas estructurales virales. En algunas realizaciones, el genoma o ARN de replicón del alfavirus modificado comprende una UTR 5' que muestra al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % de identidad de secuencia con la secuencia de ácido nucleico de al menos una de las Id. de sec. n.º: 2-18. En determinadas realizaciones, el genoma de alfavirus modificado o el ARN de replicón carece de una porción sustancial de la secuencia de ácido nucleico que codifica proteínas estructurales virales. En determinadas realizaciones, el genoma de alfavirus modificado o el ARN de replicón no comprende ninguna secuencia de ácido nucleico que codifique proteínas estructurales virales.

En un aspecto, algunas realizaciones descritas en la presente memoria se relacionan con una célula recombinante que incluye una molécula de ácido nucleico descrita en la presente memoria. En algunas realizaciones, la célula recombinante es una célula procariótica o una célula eucariota. En algunas realizaciones, la célula recombinante es una célula animal. En algunas realizaciones, la célula recombinante es una célula de un animal vertebrado o una célula de invertebrado. En algunas realizaciones, la célula recombinante se selecciona del grupo que consiste en una célula endotelial de la arteria pulmonar equina, una célula de la dermis equina, una célula de riñón de hámster bebé (BHK), una célula de riñón de conejo, una célula muscular de ratón, una célula de tejido conectivo de ratón, una célula de cuello uterino humano, una célula de laringe epidermoide humana, una célula de ovario de hámster chino (CHO), una célula HEK-293 humana, una célula 3T3 de ratón, una célula Vero, una célula epitelial de riñón canino Madin-Darby (MDCK), primaria célula de fibroblasto de pollo, una célula HuT78, una célula de pulmón A549, una célula HeLa, una célula PER.C6®, una célula WI-38, una célula MRC-5, una FRhL-2 y una célula T CEM. Algunas realizaciones descritas en la presente memoria proporcionan un cultivo celular que incluye al menos una célula recombinante como se describe en la presente memoria.

En un aspecto, algunas realizaciones proporcionan un método para producir un polipéptido de interés que implica cultivar una célula huésped que comprende una molécula de ácido nucleico que comprende una 5'-UTR modificada y está desprovista de al menos una porción de una secuencia de ácido nucleico que codifica proteínas estructurales virales. En algunas realizaciones, la célula recombinante es una célula de vertebrado o una célula de invertebrado.

En un aspecto adicional, algunas realizaciones proporcionan un método para producir un polipéptido de interés en un sujeto que implica administrar al sujeto una molécula de ácido nucleico que comprende una 5'-UTR modificada y está desprovista de al menos una porción de una Secuencia de ácido nucleico que codifica proteínas estructurales virales. En algunas realizaciones, el sujeto es humano, caballo, cerdo, primate, ratón, ganado vacuno, cerdo, oveja, conejo, gato, perro, pájaro, pez, cabra, burro, hámster o búfalo.

Las implementaciones de realizaciones de los métodos según la presente descripción pueden incluir una o más de las siguientes características. En algunas realizaciones, el ARN de replicón modificado es un ARN de replicón de alfavirus modificado. En algunas realizaciones, el ARN de replicón de alfavirus modificado incluye un genoma de alfavirus modificado. En algunas realizaciones, la UTR 5' modificada incluye una o más sustituciones de nucleótidos en la posición 1, 2, 4 o una combinación de las mismas. En algunas realizaciones, al menos una de las sustituciones de nucleótidos es una sustitución de nucleótidos en la posición 2 de la UTR 5' modificada. En las realizaciones, las sustituciones de nucleótidos en la posición 2 de la UTR 5' modificada son una sustitución U->G. En determinadas realizaciones, el ARN de replicón modificado carece de una porción sustancial de la secuencia de ácido nucleico que codifica proteínas estructurales virales. En algunas realizaciones, el genoma de alfavirus modificado o el ARN de replicón no incluye ninguna secuencia de ácido nucleico que codifique proteínas estructurales virales.

En algunas realizaciones, la molécula de ácido nucleico como se describe en la presente memoria incluye además uno o más casetes de expresión, en donde cada uno de los casetes de expresión incluye un promotor unido operativamente a una secuencia de ácido nucleico heteróloga. En algunas realizaciones, la molécula de ácido nucleico incluye al menos dos, al menos tres, al menos cuatro, al menos cinco o al menos seis casetes de expresión. En algunas realizaciones, el promotor de al menos uno de los casetes de expresión incluye un promotor subgenómico 26S. En algunas realizaciones particulares, el promotor de al menos uno de los casetes de expresión incluye un promotor subgenómico 26S de alfavirus. Preferiblemente, el promotor comprende un promotor subgenómico 26S de la encefalitis equina venezolana (VEEV). En determinadas realizaciones, la secuencia de ácido nucleico heteróloga de al menos uno de los casetes de expresión incluye una secuencia codificante de un gen de interés (GOI). La secuencia codificante del GOI, en algunas realizaciones, está optimizada para la expresión a un nivel superior al nivel de expresión de una secuencia codificante de referencia. En algunas realizaciones, el promotor unido operativamente a la secuencia de ácido nucleico heteróloga comprende una secuencia promotora heteróloga. Los promotores heterólogos adecuados incluyen, entre otros, elementos reguladores del sitio de entrada al ribosoma interno (IRES) derivados de virus de la encefalomiocarditis (EMCV), virus de la diarrea viral bovina (BVDV), poliovirus, virus de la fiebre aftosa (FMD), enterovirus 71 o virus de la hepatitis C.

En algunas realizaciones, el ARN de replicón modificado incluye un genoma o ARN de replicón modificado de un virus que pertenece al género *Alphavirus* de la familia *Togaviridae*. En algunas realizaciones, el genoma modificado o ARN de replicón es de un alfavirus que pertenece al grupo VEEV/EEEV, o al grupo SF, o al grupo SIN. En algunas realizaciones, el alfavirus se selecciona del grupo que consiste en virus de la encefalitis equina del este (EEEV), virus de la encefalitis equina venezolana (VEEV), virus de los Everglades (EVEV), virus Mucambo (MUCV), virus Pixuna (PIXV), virus de Middleburg (MIDV), virus Chikungunya (CHIKV), virus O'Nyong-Nyong (ONNV), virus del río Ross (RRV), virus del bosque de Barmah (BF), virus Getah (GET), virus Sagiyama (SAGV), virus Bebaru (BEBV), virus Mayaro (MAYV), virus Una (UNAV), virus Sindbis (SINV), virus Aura (AURAV), virus Whataroa (WHAV), virus Babanki (BABV), virus Kyzylgach (KYZV), virus de la encefalitis equina occidental (WEEV), virus Highland J (HJV), virus Fort Morgan (FMV), Ndumu (NDUV) y virus Buggy Creek. En algunas realizaciones, el alfavirus es el virus de la encefalitis equina venezolana (VEEV).

En algunas realizaciones, el genoma de alfavirus modificado o el ARN de replicón está unido operativamente a un elemento regulador heterólogo. En algunas realizaciones, el elemento regulador heterólogo incluye una secuencia promotora. En algunas realizaciones, la secuencia promotora incluye una secuencia promotora de T7. En algunas realizaciones, el elemento regulador heterólogo comprende una secuencia de terminación transcripcional. En algunas realizaciones, la secuencia de terminación transcripcional es o comprende una secuencia de terminación T7.

En algunas realizaciones, la molécula de ácido nucleico incluye un genoma de alfavirus modificado o ARN de replicón, en donde el genoma de alfavirus modificado o ARN de replicón incluye una UTR 5' que muestra al menos 80 % de identidad de secuencia con la secuencia de ácido nucleico de Id. de sec. n.º: 1 y una sustitución U->G en la posición 2 de la UTR 5', y en el que el genoma del alfavirus modificado o el ARN de replicón carece de al menos una porción de la secuencia que codifica proteínas estructurales virales. En algunas realizaciones, la molécula de ácido nucleico exhibe al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % de identidad de secuencia con la secuencia de ácido nucleico de Id. de sec. n.º: 1. En algunas realizaciones, la molécula de ácido nucleico incluye un genoma de alfavirus modificado o ARN de replicón, en donde el genoma de alfavirus modificado o ARN de replicón incluye una UTR 5' que muestra al menos 80 % de identidad de secuencia con la secuencia de ácido nucleico de al menos una de las Id. de sec. n.º: 2-18 y una sustitución U->G en la posición 2 de la

UTR 5', y en el que el genoma del alfavirus modificado o el ARN de replicón carece de al menos una porción de la secuencia que codifica proteínas estructurales virales. En algunas realizaciones, el genoma o ARN de replicón del alfavirus modificado incluye una UTR 5' que muestra al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % de identidad de secuencia con la secuencia de ácido nucleico de al menos una de las Id. de sec. n.º: 2-18.

En un aspecto adicional, algunas realizaciones descritas en la presente memoria proporcionan polipéptidos recombinantes producidos mediante un método según una o más realizaciones descritas en la presente memoria.

En un aspecto, algunas realizaciones descritas en la presente memoria se relacionan con una composición que incluye un polipéptido recombinante como se describe en la presente memoria. En algunas realizaciones, la composición es una composición farmacéutica que comprende un portador farmacéuticamente aceptable. En algunas realizaciones, la composición es una composición profiláctica, una composición nutracéutica, una composición farmacéutica o una combinación de las mismas.

En un aspecto adicional, algunas realizaciones descritas en la presente memoria se relacionan con una composición que incluye una molécula de ácido nucleico como se describe en la presente memoria. En algunas realizaciones, la composición es una composición farmacéutica que comprende un portador farmacéuticamente aceptable. En algunas realizaciones, la composición es una composición profiláctica, una composición nutracéutica, una composición farmacéutica o una combinación de las mismas.

Aún en un aspecto adicional, algunas realizaciones descritas en la presente memoria se relacionan con una composición que incluye una célula recombinante como se describe en la presente memoria. En algunas realizaciones, la composición es una composición farmacéutica que comprende un portador farmacéuticamente aceptable. En algunas realizaciones, la composición es una composición profiláctica, una composición nutracéutica, una composición farmacéutica o una combinación de las mismas.

El resumen anterior es sólo ilustrativo y no pretende ser de ninguna forma limitante. Además de las realizaciones y características ilustrativas descritas en la presente memoria, otros aspectos, realizaciones, objetos y características de la aplicación resultarán completamente evidentes a partir de los dibujos y la descripción detallada y las reivindicaciones.

Breve descripción de las figuras

La figura 1A muestra una alineación de secuencia de las regiones no traducidas 5' (UTR 5') de alfavirus representativos: Aura virus (AURAV; Id. de sec. n.º: 2), virus Chikungunya (CHIKV, Id. de sec. n.º: 3), virus O'Nyong-Nyong (Id. de sec. n.º: 4, ONNV), virus Bebaru (Id. de sec. n.º: 5, BEBV), virus del bosque Semliki (Id. de sec. n.º: 6, SFV), virus Mayaro (Id. de sec. n.º: 7, MAYV), virus Getah (Id. de sec. n.º: 8, GETV), virus Sagiya (Id. de sec. n.º: 9, SAGV), virus Ndumu (Id. de sec. n.º: 10, NDUV), virus de Middleburg (Id. de sec. n.º: 11, MIDV), virus de la encefalitis equina del este (Id. de sec. n.º: 12, EEEV), virus de Fort Morgan (Id. de sec. n.º: 13, FMV), virus Buggy Creek (Id. de sec. n.º: 14, Buggy), virus de la encefalitis equina venezolana (Id. de sec. n.º: 15, VEEV), virus Whartaroa (Id. de sec. n.º: 16, WHAV), virus Sindbis (Id. de sec. n.º: 17, SINV), y Bebaru virus (Id. de sec. n.º: 18, BEBV). El alineamiento de secuencia de la figura 1A se generó usando el programa MUSCLE 3.6 con la configuración predeterminada. En la alineación de secuencias que se muestra en este documento, un guión en una secuencia alineada, que se crea mediante el programa MUSCLE 3.6 para una alineación óptima, representa una brecha, es decir, una falta de nucleótido en esa posición. Como se analiza en detalle a continuación, se han identificado varios residuos de nucleótidos conservados en este análisis de comparación de secuencias. Los asteriscos identifican residuos de nucleótidos idénticos entre las secuencias alineadas. La figura 1B muestra una representación gráfica de la secuencia consenso como consenso ponderada en la que el tamaño de la letra que designa un aminoácido determinado es proporcional a la conservación del residuo en las diferentes secuencias utilizadas para generar el motivo (el tamaño de la letra denota una frecuencia relativa del residuo en esa posición entre las secuencias alineadas). El tamaño del carácter refleja el contenido de información medido en bits.

La figura 2 muestra una ilustración esquemática de la estructura de un diseño de replicón de VEEV monovalente de base ejemplar no limitante, Rep-Alfa, que incluye una secuencia promotora de T7, una secuencia UTR 5' de VEEV que tiene una sustitución U2->G como se describe en la presente memoria. Secuencia codificante de los polipéptidos no estructurales nsp1, nsp2, nsp3 y nsp4 de un genoma de alfavirus. El replicón Rep-Alfa de VEEV monovalente básico también contiene una secuencia promotora subgenómica 26S, una secuencia UTR 3', una secuencia de terminación T7, una secuencia de poliadenilación PolyA y una serie de sitios de restricción únicos diseñados para facilitar la inserción de componentes adicionales en el replicón.

Las figuras 3 y 4 representan gráficamente las estructuras de dos diseños de replicón VEEV monovalentes ejemplares no limitantes, en los que el gen de interés (GOI) incorporado operativamente en el vector era un gen A Vietnam 1203 HA (figura 3) y un gen indicador de proteína de fluorescencia verde mejorada (eGFP, por sus siglas en inglés) (figura 4), respectivamente.

La figura 5 resume gráficamente los resultados de experimentos ejemplares que ilustran que una modificación U2->G en la posición 2 de la UTR 5' en un replicón VEEV-HA modificado no afecta la actividad biológica del replicón modificado. El análisis de citometría de flujo (FACS) se realizó en células sometidas a electroporación con un replicón U2->G VEEV-HA modificado que expresa un gen de hemaglutinina de influenza - HA (véase también **la Figura 3** para la organización estructural del replicón). Se utilizó como control un replicón VEEV-HA de tipo salvaje, *es decir*, que contenía un residuo U en la posición 2.

La figura 6 resume gráficamente los resultados de un análisis de citometría de flujo ejemplar realizado para demostrar que la expresión de un replicón de alfavirus que porta una modificación U2->G en la UTR 5' no está restringida a ningún gen de interés (GOI) en particular. En este experimento, el replicón de alfavirus modificado se diseñó para expresar un gen indicador de la proteína verde fluorescente (GFP) (consulte también **la Figura 4** para conocer la organización estructural del replicón). Se demostró que la modificación U2->G en la UTR 5' del replicón VEEV-GFP modificado mejora la expresión del gen GFP 3 veces en relación con la expresión detectada en un replicón de control de tipo salvaje.

La figura 7 resume gráficamente los resultados de otro análisis de citometría de flujo ejemplar que evalúa la expresión de un gen indicador de Firefly rojo a partir de un replicón VEEV modificado. En este experimento, se demostró que la modificación U2->G en la UTR 5' del replicón VEEV-rFF modificado mejora la expresión del gen de Firefly rojo 2 veces en relación con la expresión detectada en un replicón de control de tipo salvaje.

La figura 8 representa esquemáticamente una estructura genómica y expresión genómica de alfavirus ilustrativos no limitantes (adaptado de Strauss y col., Microbiological Reviews, págs. 491-562, septiembre de 1994). Se muestra la organización del genoma de un virus SIN. Se dan los nombres de los genes no estructurales y de los genes de proteínas estructurales. Se puede encontrar una referencia a la nomenclatura de los genes y proteínas en Strauss y col., supra, 1994. El ARN genómico 49S se ilustra esquemáticamente en el centro, con su ORF traducido mostrado como un cuadro abierto. Los pequeños cuadros negros son elementos de secuencia conservados; el diamante abierto denota el codón de terminación del ópalo con fugas. Las poliproteínas no estructurales y sus productos procesados se muestran arriba. La terminación en el codón de ópalo produce P123, cuya función principal en la replicación se cree que es la de una proteinasa que actúa en *trans* para procesar las poliproteínas en las replicasas de ARN activas; este dominio de proteinasa se encuentra en la región nsP2. La lectura completa del codón de parada del ópalo produce P1234, que puede formar una replicasa activa. El ARNm subgenómico 26S se amplía a continuación para mostrar el ORF estructural y sus productos de traducción. Los polipéptidos presentes en el virión están sombreados. vcRNA es el complemento de cadena negativa del ARN genómico.

Las características anteriores y otras de la presente descripción resultarán más evidentes a partir de la siguiente descripción y las reivindicaciones adjuntas, tomadas en conjunto con los dibujos adjuntos. Entendiendo que estos dibujos representan sólo varias realizaciones según la descripción y no deben considerarse limitativos de su alcance; la descripción se describirá con especificidad y detalle adicionales mediante el uso de los dibujos adjuntos.

Descripción detallada

La presente descripción se refiere generalmente a sistemas de expresión viral con potencial de expresión superior que son adecuados para expresar moléculas heterólogas tales como, por ejemplo, vacunas y polipéptidos terapéuticos, en células recombinantes. Por ejemplo, algunas realizaciones de la descripción se relacionan con moléculas de ácido nucleico tales como, *p. ej.*, construcciones y vectores de expresión, que contienen un ARN de replicón modificado que incluye una región 5'-no traducida modificada (UTR 5') y, opcionalmente, al menos algunas de sus secuencias virales originales que codifican proteínas estructurales que han sido eliminadas. También se incluyen según algunas realizaciones de la descripción vectores de expresión basados en virus que incluyen uno o más casetes de expresión que codifican polipéptidos heterólogos. En consecuencia, las células recombinantes que están modificadas genéticamente para incluir una o más de las moléculas de ácido nucleico descritas en la presente memoria, así como biomateriales y productos recombinantes derivados de dichas células también están dentro del alcance de la solicitud. Además se proporcionan en aspectos particulares de la descripción composiciones que incluyen una o más de las moléculas y/o células recombinantes descritas en la presente memoria.

Los ARN (replicones) autoamplificadores basados en virus de ARN (*p. ej.*, alfavirus) se pueden usar como sistemas de expresión robustos. Por ejemplo, las modificaciones en la región no traducida (UTR) 5' del virus de tipo salvaje (*p. ej.*, alfavirus) pueden permitir la disección de los nucleótidos de ARN clave que comprenden los elementos promotores implicados tanto en la replicación del ARN como en la transcripción del ARN. El desarrollo de sistemas de expresión viral mejorados (*p. ej.*, alfavirus) mediante la manipulación de la secuencia UTR 5' representa un avance importante en el desarrollo de plataformas de replicones. Sin estar limitado por ninguna teoría particular, se cree que una ventaja no limitante de usar alfavirus como vectores de expresión viral es que pueden dirigir la síntesis de grandes cantidades de proteínas heterólogas en células huésped recombinantes. En particular, entre otras ventajas, los sistemas de plataforma de replicón de alfavirus descritos en la presente memoria, en algunas realizaciones, son capaces de expresar hasta tres veces la cantidad de proteína normalmente expresada a partir de un replicón de alfavirus. Esta mejora es significativa dados los niveles de expresión naturalmente altos observados con los sistemas de replicones de alfavirus estándar y que la mutación UTR 5' que imparte esta capacidad se consideraba anteriormente una mutación

casi letal para la replicación y transcripción de alfavirus. Por ejemplo, polipéptidos tales como anticuerpos terapéuticos de cadena sencilla pueden ser más eficaces si se expresan en niveles elevados *in vivo*. Además, para producir anticuerpos recombinantes purificados a partir de células en cultivo (*ex vivo*), la expresión elevada de proteínas a partir de un ARN de replicón puede aumentar los rendimientos generales del producto de anticuerpo. Además, si la proteína que se expresa es un antígeno de vacuna, un alto nivel de expresión puede inducir la respuesta inmune más sólida *in vivo*.

En la siguiente descripción detallada, se hace referencia a los dibujos adjuntos, que forman parte de la misma. En los dibujos, símbolos similares normalmente identifican componentes similares, a menos que el contexto indique lo contrario. Las alternativas ilustrativas descritas en la descripción detallada, los dibujos y las reivindicaciones no pretenden ser limitantes. Se podrán utilizar otras alternativas y se podrán realizar otros cambios, sin salirse del alcance del tema aquí presentado. Se entenderá fácilmente que los aspectos, como se describen generalmente en la presente memoria y se ilustran en las Figuras, se pueden disponer, sustituir, combinar y diseñar en una amplia variedad de configuraciones diferentes, todas las cuales se contemplan explícitamente y forman parte de esta solicitud.

A menos que se defina lo contrario, todos los términos de la técnica, notaciones y otros términos o terminología científicos utilizados en la presente memoria pretenden tener los significados comúnmente entendidos por aquellos expertos en la técnica a la que pertenece esta solicitud. En algunos casos, los términos con significados comúnmente entendidos se definen en la presente memoria para mayor claridad y/o para una fácil referencia, y la inclusión de tales definiciones en la presente memoria no necesariamente debe interpretarse como que representa una diferencia sustancial sobre lo que generalmente se entiende en la técnica. Muchas de las técnicas y procedimientos descritos o a los que se hace referencia en la presente memoria se entienden bien y se emplean comúnmente utilizando metodología convencional por parte de los expertos en la técnica.

Algunas definiciones

La forma singular “un”, “una” y “el” incluye referencias en plural a menos que el contexto indique claramente lo contrario. Por ejemplo, el término “una célula” incluye una o más células, que comprenden mezclas de las mismas. “A y/o B” se utiliza en la presente memoria para incluir todas las siguientes alternativas: “A”, “B”, “A o B” y “A y B”.

El término “aproximadamente”, como se usa en la presente memoria, tiene su significado habitual de aproximadamente. Si el grado de aproximación no queda claro en el contexto, “aproximadamente” significa dentro de más o menos el 10 % del valor proporcionado, o redondeado a la cifra significativa más cercana, en todos los casos incluido el valor proporcionado. Cuando se proporcionan rangos, estos incluyen los valores límite.

Los términos “células”, “cultivos celulares”, “línea celular”, “células huésped recombinantes”, “células receptoras” y “células huésped” como se usan en la presente memoria, incluyen las células objeto primarias y cualquier progenie de las mismas, sin en cuanto al número de transferencias. Debe entenderse que no toda la progenie es exactamente idéntica a la célula parental (debido a mutaciones deliberadas o inadvertidas o diferencias en el entorno); sin embargo, dicha progenie alterada está incluida en estos términos, siempre que la progenie conserve la misma funcionalidad que la de la célula originalmente transformada.

Como se usa en la presente memoria, el término “construcción” pretende significar cualquier molécula de ácido nucleico recombinante tal como un casete de expresión, plásmido, cósmido, virus, molécula de polinucleótido que se replica de forma autónoma, fago o lineal o circular, monocatenario o doble, molécula polinucleotídica de ADN o ARN de cadena simple, derivada de cualquier fuente, capaz de integración genómica o replicación autónoma, que comprende una molécula de ácido nucleico donde una o más secuencias de ácido nucleico se han unido de una manera funcionalmente operativa, p. ej., operativamente unidas.

El término “gen” se usa ampliamente para referirse a cualquier segmento de molécula de ácido nucleico que codifica una proteína o que puede transcribirse en un ARN funcional. Los genes pueden incluir secuencias que se transcriben pero que no forman parte de un transcrito de ARN final, maduro y/o funcional, y los genes que codifican proteínas pueden comprender además secuencias que se transcriben pero no se traducen, por ejemplo, regiones no traducidas 5', regiones no traducidas 3', intrones, etc. Además, los genes pueden comprender opcionalmente secuencias reguladoras requeridas para su expresión, y tales secuencias pueden ser, por ejemplo, secuencias que no se transcriben ni se traducen. Los genes se pueden obtener de una variedad de fuentes, incluida la clonación de una fuente de interés o la síntesis a partir de información de secuencia conocida o predicha, y pueden incluir secuencias diseñadas para tener los parámetros deseados.

El término “heterólogo” cuando se usa en referencia a un polinucleótido, un gen o una molécula de ácido nucleico se refiere a un polinucleótido, gen o una molécula de ácido nucleico que no deriva de la especie huésped. Por ejemplo, “gen heterólogo” o “secuencia de ácido nucleico heterólogo”, como se usa en la presente memoria, se refiere a un gen o secuencia de ácido nucleico de una especie diferente a la especie del organismo huésped en el que se introduce. Cuando se hace referencia a una secuencia reguladora de gen o a una secuencia de ácido nucleico auxiliar usada para manipular la expresión de una secuencia de gen (p. ej., una región no traducida 5', una región no traducida 3', una secuencia de adición de poli A, etc.) o a una secuencia de ácido nucleico que codifica un dominio proteico o una

secuencia de localización de proteínas, “heterólogo” significa que la secuencia reguladora o auxiliar o la secuencia que codifica un dominio proteico o una secuencia de localización proviene de una fuente diferente que el gen con el que se codifica el gen regulador. o una secuencia de ácido nucleico auxiliar o una secuencia de ácido nucleico que codifica un dominio proteico o una secuencia de localización se yuxtaponen en un genoma. Por tanto, un promotor unido operativamente a un gen al que no está unido operativamente en su estado natural (por ejemplo, en el genoma de un organismo no modificado genéticamente) se denomina en la presente memoria “promotor heterólogo”, aunque el promotor puede derivar de la misma especie (o, en algunos casos, del mismo organismo) que el gen al que está vinculado. Por ejemplo, en algunas realizaciones descritas en la presente memoria, una secuencia codificante de un gen heterólogo de interés (GOI) no está unida a la secuencia del replicón de EAV en su estado natural. En algunas realizaciones, la secuencia codificante de GOI se deriva de otro organismo, tal como otro virus, bacteria, hongo, célula humana (*p. ej.*, Ag tumoral), parásito (*p. ej.*, malaria), *etc.*

Los términos “molécula de ácido nucleico” y “polinucleótido” se usan indistintamente en la presente memoria y se refieren tanto a moléculas de ADN como de ADN, incluidas moléculas de ácido nucleico que comprenden ADNc, ADN genómico, ADN sintético y moléculas de ADN o ARN que contienen análogos de ácidos nucleicos. Las moléculas de ácido nucleico pueden tener cualquier estructura tridimensional. Una molécula de ácido nucleico puede ser bicatenaria o monocatenaria (*p. ej.*, una cadena sentido o antisentido). Los ejemplos no limitantes de moléculas de ácido nucleico incluyen genes, fragmentos de genes, exones, intrones, ARN mensajero (ARNm), ARN de transferencia, ARN ribosómico, ARNip, microARN, ARNtracr, ARNcr, ARN guía, ribozimas, ADNc, polinucleótidos recombinantes, polinucleótidos ramificados, sondas de ácidos nucleicos y cebadores de ácidos nucleicos. Una molécula de ácido nucleico puede contener nucleótidos no convencionales o modificados. Los términos “secuencia de polinucleótidos” y “secuencia de ácido nucleico”, como se usan en la presente memoria, se refieren indistintamente a la secuencia de una molécula de polinucleótido. En la presente memoria se utiliza la nomenclatura para bases de nucleótidos establecida en 37 CFR §1,822.

Las moléculas de ácido nucleico pueden ser moléculas de ácido nucleico de cualquier longitud, incluidas, entre otras, moléculas de ácido nucleico que tienen entre aproximadamente 3 Kb y aproximadamente 50 Kb, por ejemplo entre aproximadamente 3 Kb y aproximadamente 40 Kb, entre aproximadamente 3 Kb y aproximadamente 40 Kb, entre aproximadamente 3 Kb y aproximadamente 30 Kb, entre aproximadamente 3 Kb y aproximadamente 20 Kb, entre aproximadamente 5 Kb y aproximadamente 40 Kb, entre aproximadamente 5 Kb y aproximadamente 40 Kb, entre aproximadamente 5 Kb y aproximadamente 30 Kb, entre aproximadamente 5 Kb y aproximadamente 20 Kb, o entre aproximadamente 10 Kb y aproximadamente 50 Kb, por ejemplo entre aproximadamente 15 Kb y 30Kb, entre aproximadamente 20 Kb y aproximadamente 50 Kb, entre aproximadamente 20 Kb y aproximadamente 40 Kb, entre aproximadamente 5 Kb y aproximadamente 25 Kb, o entre aproximadamente 30 Kb y aproximadamente 50 Kb. Las moléculas de ácido nucleico también pueden tener, por ejemplo, más de 50 kb.

Los polinucleótidos de la presente descripción pueden ser “biológicamente activos” con respecto a un atributo estructural, tal como la capacidad de un ácido nucleico para hibridarse con otro ácido nucleico, o la capacidad de una secuencia de polinucleótidos para ser reconocida y unida. por un factor de transcripción y/o una polimerasa de ácido nucleico.

El término molécula de ácido nucleico “recombinante” o “diseñada por ingeniería genética”, como se usa en la presente memoria, se refiere a una molécula de ácido nucleico que ha sido alterada mediante intervención humana. Como ejemplos no limitantes, un ADNc es una molécula de ADN recombinante, como lo es cualquier molécula de ácido nucleico que haya sido generada por reacción(es) de polimerasa *in vitro*, o a la que se hayan unido conectores, o que se haya integrado en un vector, tal como un vector de clonación o un vector de expresión. Como ejemplos no limitantes, una molécula de ácido nucleico recombinante: 1) ha sido sintetizado o modificado *ex vitro*, por ejemplo, utilizando técnicas químicas o enzimáticas (*p. ej.*, mediante el uso de síntesis química de ácidos nucleicos, o mediante el uso de enzimas para la replicación, polimerización, digestión exonucleolítica, digestión endonucleolítica, ligación, transcripción inversa, transcripción, modificación de bases (incluyendo, *p. ej.*, metilación) o recombinación (incluyendo recombinación homóloga y específica de sitio)) de moléculas de ácido nucleico; 2) incluye secuencias de nucleótidos unidas que no están unidas en la naturaleza, 3) ha sido diseñada utilizando técnicas de clonación molecular de modo que carece de uno o más nucleótidos con respecto a la secuencia de la molécula de ácido nucleico natural, y/o 4) ha sido manipulada utilizando técnicas de clonación molecular de tal modo que tenga uno o más cambios o reordenamientos de secuencia con respecto a la secuencia de ácido nucleico de origen natural. Como ejemplos no limitantes, un ADNc es una molécula de ADN recombinante, como lo es cualquier molécula de ácido nucleico que haya sido generada por reacción(es) de polimerasa *in vitro*, o a la que se hayan unido conectores, o que se haya integrado en un vector, tal como un vector de clonación o un vector de expresión.

Como se usa en la presente memoria, una “porción sustancial” de una secuencia de ácido nucleico que codifica un polipéptido estructural viral puede comprender suficiente de la secuencia de ácido nucleico que codifica el polipéptido estructural viral para permitir la identificación putativa de ese polipéptido, ya sea mediante evaluación manual de la secuencia por un experto en la técnica, o mediante comparación e identificación de secuencias automatizadas por ordenador usando algoritmos tales como BLAST (véase, por ejemplo, en “Herramienta de búsqueda de alineación local básica”; Altschul SF y col., J. Mol. Biol. 215:403-410, 1993). En general, un experto en la técnica reconocerá que es necesaria una secuencia de diez o más aminoácidos contiguos o treinta o más nucleótidos para identificar

supuestamente una secuencia de polipéptido o ácido nucleico como homóloga a una proteína o gen conocido. Además, con respecto a las secuencias de nucleótidos, se pueden usar sondas de oligonucleótidos específicas de genes que comprenden de 15 a 30 nucleótidos contiguos en métodos dependientes de la secuencia de identificación de genes (*p. ej.*, hibridación Southern) y aislamiento (*p. ej.*, hibridación in situ de colonias bacterianas o placas de bacteriófagos). Además, se pueden usar oligonucleótidos cortos de 10 a 15 bases como cebadores de amplificación en PCR para obtener un fragmento de ácido nucleico particular que comprende los cebadores. En consecuencia, una “porción sustancial” de una secuencia de nucleótidos comprende una cantidad suficiente de la secuencia para permitir la identificación y/o el aislamiento específicos de un fragmento de ácido nucleico que comprende la secuencia. La presente descripción proporciona moléculas de ácido nucleico que carecen de secuencias de ácido nucleico parciales o completas que codifican uno o más polipéptidos estructurales virales. El experto en la técnica, que tiene el beneficio de las secuencias descritas en la presente memoria, puede usar fácilmente todas o una parte sustancial de las secuencias descritas para fines conocidos por los expertos en esta técnica. En consecuencia, la presente solicitud comprende las secuencias completas tal como se describen en la presente memoria, *p. ej.*, las expuestas en el Listado de Secuencias adjunto, así como porciones sustanciales de esas secuencias tal como se definen anteriormente.

Como entenderá alguien con conocimientos habituales en la técnica, para todos y cada uno de los fines, tales como en términos de proporcionar una descripción escrita, todos los rangos descritos en la presente memoria también abarcan todos y cada uno de los subrangos y combinaciones posibles de subrangos de los mismos. Se puede reconocer fácilmente que cualquier rango enumerado describe suficientemente y permite dividir el mismo rango en al menos mitades, tercios, cuartos, quintos, décimos, *etc.*, iguales. Como ejemplo no limitante, cada rango discutido en la presente memoria se puede descomponer fácilmente en un tercio inferior, un tercio medio y un tercio superior, *etc.* Como también entenderá un experto en la técnica, todos los lenguajes tales como “hasta”, “al menos”, “mayor que”, “menor que” y el como incluir el número citado y hacer referencia a rangos que posteriormente pueden dividirse en subrangos como se analizó anteriormente. Finalmente, como entenderá un experto en la técnica, un rango incluye cada miembro individual. Así, por ejemplo, un grupo que tiene de 1 a 3 artículos se refiere a grupos que tienen 1, 2 o 3 artículos. De modo similar, un grupo que tiene de 1 a 5 artículos se refiere a grupos que tienen 1, 2, 3, 4 o 5 artículos, y así sucesivamente.

La discusión de los métodos generales proporcionados en la presente memoria tiene fines ilustrativos únicamente. Otros métodos y alternativas alternativas serán evidentes para los expertos en la técnica tras la revisión de esta solicitud, y deben incluirse dentro del alcance de esta solicitud.

Alfavirus

Alphavirus es un género de virus relacionados genética, estructural y serológicamente de la familia *Togaviridae* del grupo IV que incluye al menos 30 miembros, cada uno con genomas de ARN monocatenario de polaridad positiva encerrados en una nucleocápside rodeada por una envoltura que contiene proteínas de pico viral. Actualmente, el género alfavirus comprende entre otros el virus Sindbis (SIN), el virus del bosque Semliki (SFV), el virus del río Ross (RRV), el virus de la encefalitis equina venezolana (VEEV) y el virus de la encefalitis equina del este (EEEV), que son todos estrechamente relacionados y son capaces de infectar a diversos vertebrados como mamíferos, roedores, peces, especies de aves y mamíferos más grandes como humanos y caballos, así como a invertebrados como insectos. La transmisión entre especies e individuos se produce principalmente a través de mosquitos, lo que hace que los alfavirus contribuyan a la colección de arbovirus, o virus transmitidos por artrópodos. En particular, los virus Sindbis y Semliki Forest han sido ampliamente estudiados y el ciclo de vida, el modo de replicación, *etc.*, de estos virus están bien caracterizados. En particular, se ha demostrado que los alfavirus se replican muy eficientemente en células animales, lo que los hace valiosos como vectores para la producción de proteínas y ácidos nucleicos en dichas células.

Las partículas de alfavirus están envueltas, tienen un diámetro de 70 nm, tienden a ser esféricas (aunque ligeramente pleomórficas) y tienen una nucleocápside isométrica de aproximadamente 40 nm. La **figura 8** representa una estructura genómica y una expresión genómica típicas de alfavirus (adaptado de Strauss y col., *Microbiological Reviews*, págs. 491-562, septiembre de 1994). El genoma del alfavirus es un ARN monocatenario de polaridad positiva de aproximadamente 11-12 kb de longitud, que comprende una tapa en 5', una cola poli-A en 3' y dos marcos de lectura abiertos, con un primer marco que codifica las proteínas no estructurales con función enzimática y un segundo marco que codifica las proteínas estructurales virales (*p. ej.*, la proteína C de la cápside, la glicoproteína E1, la glicoproteína E2, la proteína E3 y la proteína 6K).

Los dos tercios 5' del genoma del alfavirus codifican una serie de proteínas no estructurales necesarias para la transcripción y replicación del ARN viral. Estas proteínas se traducen directamente del ARN y, junto con las proteínas celulares, forman la ARN polimerasa dependiente de ARN, esencial para la replicación del genoma viral y la transcripción del ARN subgenómico. Se producen cuatro proteínas no estructurales (nsP1-4) como una sola poliproteína que constituye la maquinaria de replicación del virus. El procesamiento de la poliproteína se produce de una manera altamente regulada, y la escisión en la unión P2/3 influye en el uso de la plantilla de ARN durante la replicación del genoma. Este sitio está ubicado en la base de una estrecha hendidura y no es de fácil acceso. Una vez escindido, nsP3 crea una estructura de anillo que rodea a nsP2. Estas dos proteínas tienen una interfaz extensa. Las mutaciones en nsP2 que producen virus no citopáticos o fenotipos sensibles a la temperatura se agrupan en la región

de interfaz P2/P3. Las mutaciones de P3 opuestas a la ubicación de las mutaciones no citopáticas de nsP2 impiden la escisión eficaz de P2/3. Esto a su vez puede afectar la infectividad del ARN alterando los niveles de producción de ARN viral.

El tercio 3' del genoma comprende ARN subgenómico que sirve como plantilla para la traducción de todas las proteínas estructurales necesarias para formar partículas virales: la proteína C de la nucleocápside central y las proteínas de la envoltura P62 y E1 que se asocian como heterodímero. Las glicoproteínas de superficie ancladas a la membrana viral son responsables del reconocimiento del receptor y la entrada a las células diana a través de la fusión de la membrana. El ARN subgenómico se transcribe a partir del promotor subgenómico p26S presente en el extremo 3' de la secuencia de ARN que codifica la proteína nsP4. La maduración proteolítica de P62 en E2 y E3 provoca un cambio en la superficie viral. Juntos, los "picos" de glicoproteína E1, E2 y, a veces, E3 forman un dímero E1/E2 o un trímero E1/E2/E3, donde E2 se extiende desde el centro hasta los vértices, E1 llena el espacio entre los vértices y E3, si está presente, está en el extremo distal de la espiga. Tras la exposición del virus a la acidez del endosoma, E1 se disocia de E2 para formar un homotrímero E1, que es necesario para que el paso de fusión una las membranas celular y viral. La glicoproteína alfaviral E1 es una proteína de fusión viral de clase II, que es estructuralmente diferente de las proteínas de fusión de clase I que se encuentran en el virus de la influenza y el VIH. La glicoproteína E2 funciona para interactuar con la nucleocápside a través de su dominio citoplasmático, mientras que su ectodominio es responsable de unirse a un receptor celular. La mayoría de los alfavirus pierden la proteína periférica E3, mientras que en los virus Semliki permanece asociada a la superficie viral.

Se ha informado que la replicación de alfavirus tiene lugar en el citoplasma de la célula. En el primer paso del ciclo infeccioso, el extremo 5' del ARN genómico se traduce en una poliproteína (nsP1-4) con actividad ARN polimerasa que produce una hebra negativa complementaria al ARN genómico. En un segundo paso, la cadena negativa se utiliza como plantilla para la producción de dos ARN, respectivamente: (1) un ARN genómico positivo correspondiente al genoma de los virus secundarios que produce, mediante traducción, otras proteínas nsP y actúa como genoma del virus; y (2) ARN subgenómico que codifica las proteínas estructurales del virus que forman las partículas infecciosas. La relación ARN genómico positivo/ARN subgenómico está regulada por autoescisión proteolítica de la poliproteína en nsP1, nsP2, nsP3 y nsP4. En la práctica, la expresión del gen viral se produce en dos fases. En una primera fase se produce la síntesis principal de cadenas genómicas positivas y de cadenas negativas. Durante la segunda fase, la síntesis de ARN subgenómico es prácticamente exclusiva, lo que da como resultado la producción de una gran cantidad de proteína estructural.

Los análisis detallados previos de las regiones no traducidas 5' (UTR 5') de alfavirus han revelado la importancia absoluta de los nucleótidos 5' extremos para soportar la replicación y transcripción del ARN. En particular, como se ilustra en la **figura 1**, la conservación de un dinucleótido AU en las posiciones de nucleótidos 1 y 2, respectivamente, de la secuencia UTR 5' se observa entre todos los alfavirus, lo que sugiere la importancia de estos nucleótidos. Como se usa en la presente memoria, "A1" se refiere al nucleótido A conservado en la posición 1 del nucleótido de la UTR 5' (p. ej., una UTR 5' de alfavirus), y "U2" se refiere al nucleótido U conservado en la posición 2 del nucleótido de la UTR 5' (p. ej., una UTR 5' de alfavirus). Además, para el virus de la encefalitis equina venezolana (VEEV), se ha realizado un análisis detallado de los tres nucleótidos más 5', así como de la región del bucle del tallo que se encuentra inmediatamente después de esta secuencia. En particular, la importancia de mantener el residuo U en la posición 2 de la UTR 5' se ha determinado previamente (Kulasegaran-Shylini y col., J. Virol. 83:17 págs. 8327-8339, 2009a; y Kulasegaran-Shylini y col. J. Virol. 83:17 págs. 8327-8339, 2009b). Específicamente, el ARN transcrito *in vitro* de un clon infeccioso de longitud completa denominado (G2)VEE/SINV que contiene un único cambio U2->G en la UTR 5' demostró una pérdida de casi tres registros de infectividad en comparación con el ARN transcrito *in vitro* de un clon salvaje. clon infeccioso tipo VEE/SINV. Este informe sugiere fuertemente que la U en la posición 2 es crítica para la replicación del ARN y no puede ser reemplazada por una G. Sin embargo, como se describe en detalle en este documento, un replicón VEEV con un cambio U2->G en la UTR 5' es, sorprendentemente y en directa contradicción con este informe anterior, no solo es completamente capaz de una replicación robusta sino que también da como resultado tres veces el potencial de expresión de un replicón VEEV en comparación con una UTR 5' de tipo salvaje que contiene el residuo U en la posición 2.

Las secuencias extremas 5' y 3' de la mayoría de los virus de ARN están muy limitadas y se tolera poca o ninguna variación; la mayoría de las modificaciones tienen como resultado resultados altamente paralizados o letales para la replicación del ARN. Kulasegaran-Shylini y col. completó un análisis en profundidad de las secuencias de nucleótidos 5' críticas para la replicación del ARN para un clon infeccioso quimérico VEEV/SINV, que es representativo de todos los alfavirus (Kulasegaran-Shylini y col. 2009a, supra). Este informe se basó en un análisis realizado a lo largo de 25 años por muchos investigadores que respalda claramente la restricción en la variación de la secuencia de ARN que puede ocurrir en el extremo 5' de cualquier alfavirus en particular. El estudio Kulasegaran-Shylini y col. Artículo 2009b (J. Virol. 83:17 p 8327-8339, 2009) establece/muestra específicamente que cambiar el nucleótido 2 en la UTR 5' de un residuo U a un residuo G (U2->G) reduce significativamente la viabilidad de ese ARN clon infeccioso. Es decir, ese cambio específico en la UTR 5' redujo la actividad biológica del ARN clon infeccioso en casi 3 órdenes de magnitud. Como se describe en la presente memoria, el cambio en la UTR 5' (p. ej., un cambio U2->G) incorporado en un ARN de replicón de VEEV (cepa TC-83) no sólo no paraliza la replicación del replicón sino que en realidad puede aumentar la capacidad biológica. actividad del replicón. Por ejemplo, el replicón que comprende la sustitución U2->G puede, en algunas realizaciones, conducir a la expresión de una proteína de interés hasta tres veces más que un replicón de tipo

salvaje que expresa la misma proteína. Este resultado es sorprendente y no se podría haber predicho el aumento de la actividad biológica del replicón que porta el cambio U2->G. Este replicón modificado tiene el potencial de ser una plataforma superior de expresión de ARN para respaldar aplicaciones terapéuticas y de vacunas.

Se ha observado la conservación de los 2 nucleótidos más 5' en todo el ARN genómico de los subtipos de alfavirus. También se ha demostrado que el dinucleótido AU conservado (A1 y U2) es críticamente necesario para la replicación del ARN (Kulasegaran-Shylini y col. 2009a y 2009b, supra). La demostración de que un ARN de replicón de alfavirus que lleva un dinucleótido AG en el extremo 5' no sólo es completamente funcional sino que demuestra una actividad biológica mejorada es sorprendente y es completamente contraria al dogma en este campo.

Como se describe en la presente memoria, se pueden generar sistemas de expresión de alfavirus monogénicos o multigénicos usando un ARN de replicón modificado que tiene actividad potenciadora de la expresión/traducción tal como, por ejemplo, un ARN de replicón que contiene una UTR 5' modificada. En algunas realizaciones, los sistemas de expresión virales (*p. ej.*, alfavirus) como se describen en la presente memoria carecen además de una parte o de la región codificante completa para una o más proteínas estructurales virales. Por ejemplo, el sistema de expresión de alfavirus puede estar desprovisto de una porción o de la secuencia codificante completa para una o más de la proteína C de la cápside viral, la glicoproteína E1, la glicoproteína E2, la proteína E3 y la proteína 6K. En algunas realizaciones, la modificación del nucleótido en la posición 2 en una copia de ADNc de la secuencia UTR 5' del virus de la encefalitis equina venezolana (VEEV) de un nucleótido de timina (T) a un nucleótido de guanina (G) (mutación T2->G), en el contexto de un ARN de replicón, confiere al replicón un potencial de expresión de proteínas significativamente mayor en comparación con un replicón de VEEV con una secuencia UTR 5' de tipo salvaje.

En algunas realizaciones, el nivel de actividad de mejora de la expresión y/o traducción de los ARN replicón modificados como se describe en la presente memoria es de al menos 1,1, 1,2, 1,3, 1,4, 1,5, 1,6, 1,7, 1,8, 1,9, 2 (2 - veces), 3, 4, 5, 6, 7, 8 o más veces, en relación con el nivel de expresión detectado a partir de un replicón no modificado correspondiente, *p. ej.*, replicón con una UTR 5' de tipo salvaje. Sin estar limitado por ninguna teoría particular, la traducción mejorada puede deberse a una mejora de la transcripción que da como resultado un mayor nivel de transcritos disponibles para la traducción y/o puede ser independiente de la transcripción y deberse, por ejemplo, a una unión ribosomal mejorada. El nivel de actividad de mejora se puede medir mediante cualquier método y técnica convenientes conocidos en la técnica, incluidos, entre otros, nivel de transcripción, cantidad de proteína, actividad proteica, etc. (*véase, p. ej.*, los Ejemplos 1, 3-5 a continuación).

Moléculas de ácido nucleico

En un aspecto, se describen en la presente memoria nuevas moléculas de ácido nucleico que incluyen un ARN de replicón modificado. Por ejemplo, un ARN de replicón modificado puede comprender mutación(es), delección(es), sustitución(es) y/o inserción(es) en una o más de las regiones genómicas originales (*p. ej.*, marcos de lectura abiertos (ORF) y /o regiones no codificantes (*p. ej.*, secuencias promotoras)) del ARN replicón original. En algunas realizaciones, el ARN de replicón modificado incluye una región no traducida 5' modificada (UTR 5'). En algunas realizaciones, la UTR 5' modificada incluye una o más sustituciones de nucleótidos en la posición 1, 2, 4 o una combinación de las mismas. En algunas realizaciones, al menos una de las sustituciones de nucleótidos es una sustitución de nucleótidos en la posición 1 de la UTR 5' modificada. En las realizaciones, al menos una de las sustituciones de nucleótidos es una sustitución de nucleótidos en la posición 2 de la UTR 5' modificada. En algunas realizaciones, al menos una de las sustituciones de nucleótidos es una sustitución de nucleótidos en la posición 4 de la UTR 5' modificada. En las realizaciones, las sustituciones de nucleótidos en la posición 2 de la UTR 5' modificada son una sustitución U->G. En algunos aspectos de la descripción, la sustitución de nucleótidos en la posición 2 de la UTR 5' modificada es una sustitución U->A. En algunos aspectos de la descripción, la sustitución de nucleótidos en la posición 2 de la UTR 5' modificada es una sustitución U->C.

Como se usa en la presente memoria, los términos “ARN replicón” se refieren a ARN que contiene toda la información genética necesaria para dirigir su propia amplificación o autorreplicación dentro de una célula permisiva. Para dirigir su propia replicación, la molécula de ARN 1) codifica polimerasa, replicasa u otras proteínas que pueden interactuar con proteínas, ácidos nucleicos o ribonucleoproteínas virales o derivadas de la célula huésped para catalizar el proceso de amplificación del ARN; y 2) contienen secuencias de ARN que actúan *en cis* necesarias para la replicación y transcripción del ARN codificado por el replicón subgenómico. Estas secuencias pueden unirse durante el proceso de replicación a sus proteínas autocodificadas, o proteínas derivadas de células no autocodificadas, ácidos nucleicos o ribonucleoproteínas, o complejos entre cualquiera de estos componentes. Para los fines de la presente descripción, una molécula de ARN replicón derivada de alfavirus normalmente contiene los siguientes elementos ordenados: Secuencia(s) de ARN viral o de interferencia defectuosa en 5' requeridas *en cis* para la replicación, secuencias que codifican proteínas no estructurales de alfavirus biológicamente activas (*p. ej.*, nsP1, nsP2, nsP3 y nsP4), promotor del ARN subgenómico, secuencias virales en 3' requerido *en cis* para la replicación y un tracto de poliadenilato. Además, el término ARN replicón generalmente se refiere a una molécula de polaridad positiva, o sentido de “mensaje”, y el ARN replicón puede tener una longitud diferente a la de cualquier alfavirus conocido de origen natural. En algunas realizaciones de la presente descripción, el ARN de replicón no contiene las secuencias de al menos una de las proteínas virales estructurales; las secuencias que codifican genes estructurales pueden sustituirse por secuencias heterólogas. En aquellos casos en los que el ARN de replicón se va a empaquetar en una partícula de

alfavirus recombinante, debe contener una o más secuencias, las llamadas señales de empaquetado, que sirven para iniciar interacciones con proteínas estructurales de alfavirus que conducen a la formación de partículas.

Como se usa en la presente memoria, “ARN subgenómico” se refiere a una molécula de ARN de una longitud o tamaño que es más pequeño que el ARN genómico del que se deriva. El ARN subgenómico del alfavirus debe transcribirse a partir de un promotor interno, cuyas secuencias residen dentro del ARN genómico o su complemento. La transcripción de un ARN subgenómico de alfavirus puede estar mediada por polimerasas codificadas por virus asociadas con proteínas codificadas por células huésped, ribonucleoproteínas o una combinación de las mismas. En algunas realizaciones de la presente descripción, el ARN subgenómico se produce a partir de un ARN de replicón modificado como se describe en la presente memoria y codifica o expresa uno o más genes de interés (GOI). En lugar del promotor subgenómico nativo, el ARN subgenómico puede colocarse bajo el control del sitio de entrada al ribosoma interno (IRES) derivado de los virus de la encefalomiocarditis (EMCV), los virus de la diarrea viral bovina (BVDV), los poliovirus y los virus de la fiebre aftosa (FMD), enterovirus 71 o virus de la hepatitis C.

En consecuencia, en algunas realizaciones, una parte o la secuencia codificante completa para una o más proteínas estructurales virales están ausentes y/o modificadas en las moléculas de ácido nucleico descritas en la presente memoria. Por lo tanto, en algunas realizaciones particulares, el ARN de replicón modificado como se describe en la presente memoria incluye una 5'-UTR modificada y está desprovisto de al menos una porción de una secuencia de ácido nucleico que codifica una o más proteínas estructurales virales, por ejemplo, desprovista de la primera, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez o más nucleótidos de la secuencia de ácido nucleico que codifica las proteínas estructurales virales. En algunas realizaciones, el genoma o ARN de replicón del alfavirus modificado puede estar desprovisto de aproximadamente el 10 %, 20 %, 30 %, 40 %, 50 %, 60 %, 70 %, 80 %, 90 %, 95 % o más del secuencia que codifica uno o más de los polipéptidos estructurales E1, E2, E3, 6K y la proteína C de la cápside. En algunas realizaciones, el genoma de alfavirus modificado o ARN de replicón carece de una porción sustancial o de la secuencia completa que codifica uno o más de los polipéptidos estructurales E1, E2, E3, 6K y la proteína C de la cápside. Como se usa en la presente memoria, una “porción sustancial” de una secuencia de ácido nucleico que codifica una proteína estructural viral comprende suficiente de la secuencia de ácido nucleico que codifica la proteína estructural viral para permitir la identificación putativa de esa proteína, ya sea mediante evaluación manual de la secuencia por parte de un experto en la técnica, o mediante comparación e identificación de secuencias automatizadas por computadora usando algoritmos tales como BLAST (véase, por ejemplo, Altschul SF y *col.* 1993, *supra*). En algunas realizaciones, el genoma de alfavirus modificado o ARN de replicón carece de la secuencia completa que codifica uno o más de los polipéptidos estructurales E1, E2, E3, 6K y la proteína C de la cápside.

En algunas realizaciones particulares de la solicitud, la molécula de ácido nucleico como se describe en la presente memoria incluye un genoma de alfavirus modificado o un ARN de replicón que incluye un genoma de alfavirus modificado o un ARN de replicón, en donde la molécula de ácido nucleico exhibe al menos 80 % de identidad de secuencia con el genoma de alfavirus modificado o un ARN de replicón. secuencia ácida de Id. de sec. n.º: 1, en donde el genoma de alfavirus modificado o ARN de replicón comprende una sustitución U->G en la posición 2 de la región 5'-no traducida (UTR 5') y está desprovisto de al menos una porción de la Secuencia que codifica proteínas estructurales virales. En algunas realizaciones, la molécula de ácido nucleico exhibe al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % de identidad de secuencia con la secuencia de ácido nucleico de Id. de sec. n.º: 1. En algunas realizaciones, la molécula de ácido nucleico muestra una identidad de secuencia del 100 % con la secuencia de ácido nucleico de Id. de sec. n.º: 1.

En algunas realizaciones, la molécula de ácido nucleico como se describe en la presente memoria incluye un genoma de alfavirus modificado o ARN de replicón, en donde el genoma de alfavirus modificado o ARN de replicón comprende una UTR 5' que muestra al menos 80 % de identidad de secuencia con la secuencia de ácido nucleico de al menos una UTR 5' descrita en la presente memoria y una sustitución U->G en la posición 2 de la UTR 5', y en el que el genoma del alfavirus modificado o el ARN de replicón carece de al menos una porción de la secuencia que codifica proteínas estructurales virales. En algunas realizaciones, el genoma de alfavirus modificado o ARN de replicón comprende una UTR 5' que muestra al menos un 80 % de identidad de secuencia con al menos una de las secuencias expuestas en las Id. de sec. n.º: 2-18. En algunas realizaciones, el genoma o ARN de replicón del alfavirus modificado comprende una UTR 5' que muestra al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 % o 100 % de identidad de secuencia con al menos una de las secuencias establecidas en Id. de sec. n.º: 2-18. En algunas realizaciones, el genoma de alfavirus modificado o ARN de replicón comprende una UTR 5' que muestra una identidad de secuencia del 100 % con al menos una de las secuencias expuestas en las Id. de sec. n.º: 2-18 del Listado de Secuencias.

Moléculas de ácido nucleico que tienen un alto grado de identidad de secuencia (*p. ej.*, al menos 85 %, al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o el 100 % de la secuencia de ácido nucleico de una UTR 5' descrita en la presente memoria se puede identificar y/o aislar usando las secuencias identificadas en la presente memoria (*p. ej.*, Id. de sec. n.º: 1-18) o cualquier otra UTR 5' de alfavirus tal como se conoce en la técnica, por ejemplo, las secuencias que tienen los números de acceso de GenBank/NCBI J02363, NC_001547, U38305, L04653, NC_001449, U38304, X04129, NC_003215 y la genómica TR339. secuencia (Klimstra y *col.*, J. Virol. 72:7357, 1988; McKnight y *col.*, J. Virol. 70:1981, 1996), mediante análisis de secuencia del genoma, hibridación y/o PCR con cebadores degenerados o cebadores específicos de genes de secuencias identificadas en el

respectivo genoma de alfavirus. Como se usa en la presente memoria, “identidad de secuencia” se refiere al grado en que dos polinucleótidos óptimamente alineados son invariantes a lo largo de una ventana de alineación de componentes, *p. ej.*, nucleótidos. Una “fracción de identidad” para segmentos alineados de una secuencia de prueba y una secuencia de referencia es el número de componentes idénticos que comparten las dos secuencias alineadas dividido por el número total de componentes en el segmento de secuencia de referencia, *p. ej.*, la secuencia de referencia completa o una parte definida más pequeña de la secuencia de referencia.

Algunas realizaciones descritas en la presente memoria se refieren a moléculas de ácido nucleico que comprenden un ARN de replicón modificado, en donde el ARN de replicón modificado comprende una secuencia de bases de alfavirus modificada tal como, *p. ej.*, una UTR 5', que tiene actividad potenciadora de la traducción. Dichos ARN replicón modificados se pueden usar para lograr niveles mejorados de expresión de una secuencia de ácido nucleico heteróloga (*p. ej.*, ADN o ADNc) que codifica un producto deseado. En algunas realizaciones, los ARN de replicón modificados se usan para lograr niveles mejorados de expresión de una secuencia de ácido nucleico heterólogo (*p. ej.*, ADN o ADNc) que codifica un producto deseado después de la introducción de los replicones modificados en una célula que puede ser, por ejemplo, una célula en un cultivo celular o puede ser una célula en un cuerpo vivo.

Además, en algunas realizaciones, las moléculas de ácido nucleico pueden incluir un genoma de alfavirus modificado o un ARN de replicón que contiene una o más mutaciones atenuantes para aumentar la seguridad de la manipulación y/o administración del virus. La frase “mutación atenuante” como se usa en la presente memoria significa una mutación de nucleótido o un aminoácido codificado en vista de dicha mutación que da como resultado una probabilidad disminuida de causar enfermedad en su huésped (*es decir*, una pérdida de virulencia), según la terminología estándar en la técnica, ya sea que la mutación sea una mutación de sustitución o una mutación de inserción o delección en el marco. Las mutaciones atenuantes pueden estar en las regiones codificantes o no codificantes (*p. ej.*, UTR 5') del genoma del alfavirus. Como saben los expertos en la técnica, la frase “mutación atenuante” excluye mutaciones o combinaciones de mutaciones que serían letales para el virus. Además, los expertos en la técnica apreciarán que algunas mutaciones atenuantes pueden ser letales en ausencia de una mutación supresora del segundo sitio.

Las técnicas y métodos moleculares mediante los cuales se construyeron y caracterizaron estas nuevas moléculas de ácido nucleico se describen más completamente en los Ejemplos de la presente solicitud. En la sección de Ejemplos, se ha utilizado el virus de la encefalitis equina venezolana (VEEV) para ilustrar las composiciones y métodos descritos en la presente memoria.

En algunas realizaciones, las moléculas de ácido nucleico son moléculas de ácido nucleico recombinantes. Como se usa en la presente memoria, el término recombinante significa cualquier molécula (*p. ej.*, ADN, ARN, polipéptido), es decir, o resulta, aunque sea indirectamente, de la manipulación humana. Como ejemplos no limitantes, un ADNc es una molécula de ADN recombinante, como lo es cualquier molécula de ácido nucleico que haya sido generada por reacción(es) de polimerasa *in vitro*, o a la que se hayan unido conectores, o que se haya integrado en un vector, tal como un vector de clonación o un vector de expresión. Como ejemplos no limitantes, una molécula de ácido nucleico recombinante: 1) ha sido sintetizado o modificado *in vitro*, por ejemplo, utilizando técnicas químicas o enzimáticas (por ejemplo, mediante el uso de síntesis química de ácidos nucleicos, o mediante el uso de enzimas para la replicación, polimerización, digestión exonucleolítica, digestión endonucleolítica, ligación, transcripción inversa, transcripción, modificación de bases (incluyendo, *p. ej.*, metilación) o recombinación (incluyendo recombinación homóloga y específica de sitio) de moléculas de ácido nucleico; 2) incluye secuencias de nucleótidos unidas que no están unidas por naturaleza; 3) ha sido diseñado utilizando técnicas de clonación molecular de modo que carece de uno o más nucleótidos con respecto a la secuencia de nucleótidos natural; y/o 4) ha sido manipulado usando técnicas de clonación molecular de modo que tenga uno o más cambios o reordenamientos de secuencia con respecto a la secuencia de nucleótidos de origen natural.

En algunas realizaciones, las moléculas de ácido nucleico descritas en la presente memoria se producen usando tecnología de ADN recombinante (*p. ej.*, amplificación por reacción en cadena de la polimerasa (PCR), clonación, *etc.*) o síntesis química. Las moléculas de ácido nucleico como se describen en la presente memoria incluyen moléculas de ácido nucleico naturales y homólogos de las mismas, incluidas, entre otras, variantes alélicas naturales y moléculas de ácido nucleico modificadas en las que se han insertado, eliminado y/o sustituido uno o más residuos de nucleótidos, de un modo tal que dichas modificaciones proporcionen la propiedad deseada para efectuar una actividad biológica como se describe en la presente memoria.

Una molécula de ácido nucleico, que incluye una variante de una secuencia de ácido nucleico natural, se puede producir usando varios métodos conocidos por los expertos en la técnica (véase, por ejemplo, Sambrook y col., en: Molecular Cloning, A Laboratory Manual, segunda edición, Cold Spring Harbor Press, Cold Spring Harbor, N. Y. (1989)). La secuencia de una molécula de ácido nucleico se puede modificar con respecto a una secuencia natural de la que se deriva usando una variedad de técnicas que incluyen, entre otras, técnicas de mutagénesis clásicas y técnicas de ADN recombinante, tales como, entre otras, técnicas de mutagénesis de sitio. -mutagénesis dirigida, tratamiento químico de una molécula de ácido nucleico para inducir mutaciones, escisión con enzimas de restricción de un fragmento de ácido nucleico, ligación de fragmentos de ácido nucleico, amplificación por PCR y/o mutagénesis de regiones seleccionadas de una secuencia de ácido nucleico, clonación recombinacional y síntesis, incluyendo síntesis química de mezclas de oligonucleótidos y ligación de grupos de mezclas para “construir” una mezcla de

moléculas de ácido nucleico, y combinaciones de las mismas. Los homólogos de moléculas de ácido nucleico se pueden seleccionar de una mezcla de moléculas de ácido nucleico modificadas mediante cribado de la función de la proteína o del replicón codificado por la molécula de ácido nucleico y/o mediante hibridación con un gen de tipo salvaje o fragmento del mismo, o mediante PCR, usando cebadores que tienen homología con una molécula o secuencia de ácido nucleico diana o de tipo salvaje.

En diversas realizaciones descritas en la presente memoria, la molécula de ácido nucleico descrita en la presente memoria puede incluir una o más de las siguientes características. En algunas realizaciones, el ARN de replicón modificado es un ARN de replicón de alfavirus modificado. En algunas realizaciones, el ARN de replicón de alfavirus modificado incluye un genoma de alfavirus modificado. En algunas realizaciones, la UTR 5' modificada incluye una o más sustituciones de nucleótidos en la posición 1, 2, 4 o una combinación de las mismas. En determinadas realizaciones, al menos una de las sustituciones de nucleótidos es una sustitución de nucleótidos en la posición 2 de la UTR 5' modificada. En todas las realizaciones, las sustituciones de nucleótidos en la posición 2 de la UTR 5' modificada son una sustitución U->G.

En algunas realizaciones descritas en la presente memoria, el genoma de alfavirus modificado o el ARN de replicón está unido operativamente a un elemento regulador heterólogo. Como se usa en la presente memoria, "elemento regulador", "secuencia reguladora" o "secuencia del elemento regulador" se refiere a una secuencia de nucleótidos ubicada aguas arriba (5'), dentro o aguas abajo (3') de una secuencia codificante tal como, por ejemplo, una secuencia que codifica un polipéptido o una secuencia que codifica un ARN funcional. La transcripción de la secuencia codificante y/o la traducción de una molécula de ARN resultante de la transcripción de la secuencia codificante normalmente se ven afectadas por la presencia o ausencia del elemento regulador. Estos elementos reguladores pueden comprender promotores, elementos *cis*, potenciadores, terminadores o intrones. Un experto en la técnica apreciará que los elementos reguladores descritos en la presente memoria pueden estar presentes en un elemento de expresión regulador quimérico o híbrido. En algunas realizaciones, el elemento regulador heterólogo es, o comprende, una secuencia promotora. La secuencia promotora heteróloga puede ser cualquier secuencia promotora heteróloga, por ejemplo, un promotor SP6, un promotor T3 o un promotor T7, o una combinación de los mismos. En algunas realizaciones particulares, la secuencia promotora incluye una secuencia promotora de T7.

Además, en algunas realizaciones, el genoma de alfavirus modificado o el ARN de replicón pueden incluir una o más secuencias señal de terminación transcripcional heterólogas. El término "señal de terminación de la transcripción", "terminador" o "secuencia terminadora" o "terminador de la transcripción", como se usa indistintamente en la presente memoria, se refiere a una sección reguladora de la secuencia genética que hace que la ARN polimerasa cese la transcripción. Las secuencias señal de terminación transcripcional heterólogas pueden ser generalmente cualquier secuencia señal de terminación transcripcional heteróloga y, por ejemplo, una secuencia señal de terminación SP6, una secuencia señal de terminación T3, una secuencia señal de terminación T7 o una variante de las mismas. En consecuencia, las moléculas de ácido nucleico según algunas realizaciones de la descripción pueden incluir al menos una de una o más secuencias señal de terminación transcripcional heterólogas seleccionadas del grupo que consiste en una secuencia señal de terminación SP6, una secuencia señal de terminación T3, una secuencia señal de terminación T7, secuencia, o una variante de la misma. En algunas realizaciones particulares, la secuencia de terminación transcripcional incluye una secuencia señal de terminación de T7.

En algunas realizaciones, las moléculas de ácido nucleico descritas en la presente memoria pueden incluir uno o más casetes de expresión. En principio, las moléculas de ácido nucleico descritas en la presente memoria generalmente pueden incluir cualquier número de casetes de expresión. En algunas realizaciones particulares, las moléculas de ácido nucleico descritas en la presente memoria pueden incluir al menos dos, al menos tres, al menos cuatro, al menos cinco o al menos seis casetes de expresión. Como se usa en la presente memoria, el término "casete de expresión" se refiere a un constructo de material genético que contiene secuencias codificantes y suficiente información reguladora para dirigir la transcripción y/o traducción adecuada de las secuencias codificantes en una célula receptora, *in vivo* y/o *ex vivo*. El casete de expresión se puede insertar en un vector para dirigirlo a una célula huésped deseada y/o a un sujeto. Además, el término casete de expresión puede usarse indistintamente con el término "construcción de expresión". En algunas realizaciones, el término "casete de expresión" se refiere a una construcción de ácido nucleico que incluye un gen que codifica una proteína o ARN funcional unido operativamente a elementos reguladores tales como, por ejemplo, un promotor y/o una señal de terminación, y opcionalmente, cualquier o una combinación de otras secuencias de ácidos nucleicos que afectan la transcripción o traducción del gen.

El término "unido operativamente", como se usa en la presente memoria, denota un enlace funcional entre dos o más secuencias. Por ejemplo, un enlace operativamente entre un polinucleótido de interés y una secuencia reguladora (por ejemplo, un promotor) es un enlace funcional que permite la expresión del polinucleótido de interés. En este sentido, el término "unido operativamente" se refiere al posicionamiento de una región reguladora y una secuencia codificante a transcribir de modo que la región reguladora sea eficaz para regular la transcripción o traducción de la secuencia codificante de interés. En algunas realizaciones descritas en la presente memoria, el término "unido operativamente" denota una configuración en la que una secuencia reguladora se coloca en una posición apropiada con respecto a una secuencia que codifica un polipéptido o ARN funcional de modo que la secuencia de control dirige o regula la expresión o localización celular. del ARNm que codifica el polipéptido, el polipéptido y/o el ARN funcional. Por tanto,

un promotor está en enlace operable con una secuencia de ácido nucleico si puede mediar en la transcripción de la secuencia de ácido nucleico. Los elementos vinculados operativamente pueden ser contiguos o no contiguos.

Las técnicas básicas para unir operativamente dos o más secuencias de ADN entre sí son familiares para el experto, y dichos métodos se han descrito en varios textos para la manipulación biológica molecular estándar (véase, por ejemplo, Maniatis y col., "Molecular Cloning: A Laboratory Manual", 2ª ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, Nueva York; y Gibson y col., Nature Methods 6:343-45, 2009).

En consecuencia, las moléculas de ácido nucleico proporcionadas en la presente memoria pueden encontrar uso, por ejemplo, como un vector de expresión que, cuando incluye un elemento regulador unido operativamente a una secuencia de ácido nucleico heteróloga, puede afectar la expresión de la secuencia de ácido nucleico heteróloga. En algunas realizaciones, la secuencia de nucleótidos heteróloga incluye una secuencia codificante de un gen de interés (GOI). En algunas realizaciones, la secuencia codificante del GOI se optimiza para la expresión a un nivel superior al nivel de expresión de una secuencia codificante de referencia. En algunas realizaciones, la secuencia codificante de referencia es una secuencia que no ha sido optimizada. En algunas realizaciones, la optimización de la secuencia codificante de GOI puede incluir optimización de secuencia. Con respecto a la optimización de secuencias de nucleótidos, la degeneración del código genético proporciona la posibilidad de sustituir al menos una base de la secuencia codificante de la proteína de un gen con una base diferente sin causar que la secuencia de aminoácidos del polipéptido producido a partir del gen cambie. ser cambiado. Por lo tanto, las moléculas de ácido nucleico de la presente solicitud también pueden tener cualquier secuencia de bases que haya sido cambiada con respecto a cualquier secuencia de polinucleótidos descrita en la presente memoria mediante sustitución según la degeneración del código genético. Las referencias que describen el uso de codones están disponibles públicamente. En algunas realizaciones, se pueden producir variantes de secuencia de polinucleótidos por una variedad de razones, *p. ej.*, para optimizar la expresión para un huésped particular (*p. ej.*, cambiar el uso de codones en el ARNm de alfavirus a los preferidos por otros organismos tales como humanos, hamsters, ratones o mono).

El polipéptido codificado por un GOI generalmente puede ser cualquier polipéptido y puede ser, por ejemplo, un polipéptido terapéutico, un polipéptido profiláctico, un polipéptido de diagnóstico, un polipéptido nutracéutico o una enzima industrial. En algunas realizaciones, el GOI codifica un polipéptido seleccionado del grupo que consiste en un anticuerpo, un antígeno, un modulador inmunológico y una citoquina.

En algunas realizaciones, la secuencia codificante del GOI se optimiza para una propiedad deseada. En algunas realizaciones, la secuencia codificante del GOI se optimiza para la expresión a un nivel superior al nivel de expresión de una secuencia codificante de referencia.

En algunas realizaciones, el genoma o ARN de replicón modificado descrito en la presente memoria es un genoma o ARN de replicón de un alfavirus, tal como un genoma o ARN de replicón de una especie viral que pertenece al género *Alphavirus* de la familia *Togaviridae*. En algunas realizaciones, el genoma modificado o ARN de replicón es de un alfavirus que pertenece al grupo VEEV/EEEV, o al grupo SF, o al grupo SIN (para una revisión, véase, *p. ej.*, Strauss y Strauss. Microbiol. Rev. 58:3 págs. 492-562, 1994). Los ejemplos no limitantes de alfavirus del grupo SF incluyen el virus del bosque Semliki, el virus O'Nyong-Nyong, el virus del río Ross, el virus Middelburg, el virus Chikungunya, el virus del bosque Barmah, el virus Getah, el virus Mayaro, el virus Sagiyama, el virus Bebaru y el virus Una. Los ejemplos no limitantes de alfavirus del grupo SIN incluyen el virus Sindbis, el virus Girdwood SA, el arbovirus sudafricano n.º 86, el virus Ockelbo, el virus Aura, el virus Babanki, el virus Whataroa y el virus Kyzylagach. Los ejemplos no limitantes de alfavirus del grupo VEEV/EEEV incluyen el virus de la encefalitis equina del este (EEEV), el virus de la encefalitis equina venezolana (VEEV), el virus de los Everglades (EVEV), el virus Mucambo (MUCV), el virus Pixuna (PIXV), el virus Middleburg (MIDV), virus chikungunya (CHIKV), virus O'Nyong-Nyong (ONNV), virus del río Ross (RRV), virus del bosque de Barmah (BF), virus Getah (GET), virus Sagiyama (SAGV), virus Bebaru (BEBV), Mayaro virus (MAYV) y virus Una (UNAV).

Los ejemplos no limitantes de especies de alfavirus incluyen el virus de la encefalitis equina del este (EEEV), el virus de la encefalitis equina venezolana (VEEV), el virus de los Everglades (EVEV), el virus Mucambo (MUCV), el virus Pixuna (PIXV), el virus Middleburg (MIDV), virus Chikungunya (CHIKV), virus O'Nyong-Nyong (ONNV), virus del río Ross (RRV), virus del bosque de Barmah (BF), virus Getah (GET), virus Sagiyama (SAGV), virus Bebaru (BEBV), Mayaro virus (MAYV), virus Una (UNAV), virus Sindbis (SINV), virus Aura (AURAV), virus Whataroa (WHAV), virus Babanki (BABV), virus Kyzylagach (KYZV), virus de la encefalitis equina occidental (WEEV), virus de las Tierras Altas Virus J (HJV), virus Fort Morgan (FMV), virus Ndumu (NDUV) y virus Buggy Creek. Son adecuadas tanto cepas de alfavirus virulentas como avirulentas. En algunas realizaciones particulares, el genoma o ARN de replicón modificado es de un virus Sindbis (SIN), un virus del bosque Semliki (SFV), un virus del río Ross (RRV), un virus de la encefalitis equina venezolana (VEEV) o un virus de la encefalitis equina del este. virus (EEEV). En algunas realizaciones, el genoma modificado o ARN de replicón es de un virus de la encefalitis equina venezolana (VEEV).

Células recombinantes

En un aspecto, algunas realizaciones descritas en la presente memoria se relacionan con un método para transformar una célula que incluye introducir en una célula huésped, tal como una célula animal, una molécula de ácido nucleico

como se proporciona en la presente memoria, y seleccionar o seleccionar una célula transformada. Los términos “célula huésped” y “célula huésped recombinante” se usan indistintamente en la presente memoria. Se entiende que dichos términos se refieren no sólo a la célula en cuestión en particular sino también a la progenie o progenie potencial de dicha célula. Como pueden producirse determinadas modificaciones en las generaciones sucesivas debido a mutación o influencias ambientales, tal descendencia puede no ser, de hecho, idéntica a la célula progenitora, pero aun se incluye dentro del alcance de la expresión como se usa en la presente memoria. En algunas realizaciones, la molécula de ácido nucleico se introduce en una célula huésped mediante un procedimiento de electroporación o un procedimiento biolístico.

En un aspecto relacionado, algunas realizaciones se refieren a células huésped recombinantes, por ejemplo, células animales recombinantes que incluyen una molécula de ácido nucleico descrita en la presente memoria. La molécula de ácido nucleico puede integrarse de modo estable en el genoma del huésped, o puede replicarse episomalmente, o estar presente en la célula huésped recombinante como un vector de expresión en minicírculo para una expresión estable o transitoria. En consecuencia, en algunas realizaciones descritas en la presente memoria, la molécula de ácido nucleico se mantiene y replica en la célula huésped recombinante como una unidad episomal. En algunas realizaciones, la molécula de ácido nucleico se integra de modo estable en el genoma de la célula recombinante. La integración estable se puede completar utilizando técnicas clásicas de recombinación genómica aleatoria o con técnicas de edición del genoma más precisas, como el uso de CRISPR/Cas9 dirigido por ARN guía, o la edición del genoma con endonucleasa guiada por ADN NgAgo (*Natronobacterium gregoryi* Argonaute), o la edición del genoma TALEN (activador de transcripción similar a las nucleasas efectoras). En algunas realizaciones, la molécula de ácido nucleico está presente en la célula huésped recombinante como un vector de expresión en minicírculo para una expresión estable o transitoria.

En algunas realizaciones, las células huésped pueden modificarse genéticamente (*p. ej.*, transducirse o transformarse o transfectarse) con, por ejemplo, una construcción de vector de la presente solicitud que puede ser, por ejemplo, un vector para recombinación homóloga que incluye ácido nucleico. secuencias homólogas a una porción del genoma de la célula huésped, o puede ser un vector de expresión para la expresión de cualquiera o una combinación de los genes de interés. El vector puede estar, por ejemplo, en forma de un plásmido, una partícula viral, un fago, etc. En algunas realizaciones, también se puede diseñar un vector para la expresión de un polipéptido de interés para su integración en el huésped, *p. ej.*, mediante recombinación homóloga. El vector que contiene una secuencia de polinucleótidos como se describe en la presente memoria, *p. ej.*, se puede emplear una molécula de ácido nucleico que comprende un genoma de alfavirus modificado o un ARN de replicón, así como, opcionalmente, un marcador seleccionable o un gen indicador, para transformar una célula huésped apropiada.

Los métodos y composiciones descritos en la presente memoria pueden implementarse para ingeniería genética de cualquier especie, incluidas, entre otras, especies procarióticas y eucariotas. Las células huésped adecuadas para modificar usando las composiciones y métodos según la presente descripción pueden incluir, entre otras, células de algas, células bacterianas, heterocontas, células fúngicas, células quitridio, microhongos, microalgas y células animales. En algunas realizaciones, las células animales son células animales invertebradas. En algunas realizaciones, las células de animales vertebrados son células de mamíferos. Las células huésped pueden ser células no transformadas o células que ya han sido transfectadas con al menos una molécula de ácido nucleico.

Los métodos y composiciones descritos en la presente memoria se pueden usar, por ejemplo, con células objeto y/o huésped que son importantes o interesantes para la acuicultura, la agricultura, la cría de animales y/o para aplicaciones terapéuticas y medicinales, incluida la producción de polipéptidos utilizados en la fabricación de vacunas, productos farmacéuticos, productos industriales, productos químicos y similares. En algunas realizaciones, las composiciones y métodos descritos en la presente memoria se pueden usar con células huésped de especies que son huéspedes naturales de alfavirus, como roedores, ratones, peces, aves y mamíferos más grandes como humanos, caballos, cerdos, monos y simios. así como invertebrados. Especies particularmente preferidas, en algunas realizaciones de la solicitud, son especies de animales vertebrados y especies de animales invertebrados. En principio, se puede utilizar generalmente cualquier especie animal y puede ser, por ejemplo, humanos, perros, pájaros, peces, caballos, cerdos, primates, ratones, vacas, cerdos, ovejas, conejos, gatos, cabras, burros, hámsteres o búfalo. Ejemplos no limitantes de especies de aves adecuadas incluyen pollo, pato, ganso, pavo, avestruz, emú, cisne, pavo real, faisán, perdiz y pintada. En algunas realizaciones particulares, la especie de pez es una especie de salmón. También son adecuadas las células primarias de mamífero y los tipos de células continuas/inmortalizadas. Los ejemplos no limitantes de células huésped animales adecuadas incluyen, entre otros, células endoteliales de la arteria pulmonar equina, células de la dermis equina, células de riñón de hámster bebé (BHK), células de riñón de conejo, células musculares de ratón, células de tejido conectivo de ratón, células de cuello uterino humano, célula epidermoide de laringe humana, célula de ovario de hámster chino (CHO), célula HEK-293 humana, célula 3T3 de ratón, célula Vero, célula epitelial del riñón canino Madin-Darby (MDCK), célula primaria de fibroblastos de pollo, una célula HuT78, célula pulmonar A549, célula HeLa, célula PER.C6®, célula WI-38, célula MRC-5, FRhL-2 y célula T CEM. En algunas realizaciones, la célula huésped es una célula de riñón de cría de hámster. En algunas realizaciones, la célula de riñón de cría de hámster es una célula BHK-21.

Las técnicas para transformar una amplia variedad de las células y especies huésped mencionadas anteriormente se conocen en la técnica y se describen en la literatura técnica y científica. En consecuencia, los cultivos celulares que

incluyen al menos una célula recombinante como se describe en la presente memoria también están dentro del alcance de esta solicitud. En la técnica se conocen métodos y sistemas adecuados para generar y mantener cultivos celulares.

Métodos para producir polipéptidos

Las células huésped de la presente descripción, tales como una célula huésped procariótica o eucariota, se pueden usar para producir (*es decir*, expresar) una molécula de interés tal como, *p. ej.*, un polipéptido, codificado en un marco de lectura abierto de un gen de interés (GOI) como se describe en la presente memoria. Por lo tanto, la presente solicitud proporciona además métodos para producir una molécula de interés tal como, *p. ej.*, un polipéptido, utilizando las células huésped de la descripción, que pueden ser, por ejemplo, células en cultivo celular o pueden ser células en un cuerpo vivo.

En consecuencia, algunas realizaciones descritas en la presente memoria proporcionan métodos para producir un polipéptido de interés en una célula huésped. Dicho método incluye el cultivo de una célula huésped recombinante, que incluye una molécula de ácido nucleico según uno cualquiera de los aspectos y realizaciones anteriores. En algunas realizaciones, los métodos incluyen cultivar la célula huésped de la invención (en la que se ha introducido un vector de expresión recombinante que codifica la molécula de interés) en un medio adecuado de modo que se produzca la molécula de interés. En algunas realizaciones, los métodos incluyen además aislar la molécula de interés del medio o de la célula huésped.

En otro aspecto, la descripción se refiere a métodos para producir un polipéptido de interés en un sujeto, incluyendo la administración al sujeto de una molécula de ácido nucleico según uno cualquiera de los aspectos y realizaciones anteriores.

Las células huésped y/o sujetos adecuados para su uso en los métodos y composiciones descritos en la presente memoria incluyen, entre otras, especies procarióticas y eucariotas. Las células huésped adecuadas para modificar usando las composiciones y métodos según la presente descripción pueden incluir, entre otras, células de algas, células bacterianas, heterocontas, células fúngicas, células quitridio, microhongos, microalgas y células animales. En algunas realizaciones, las células animales son células animales invertebradas. En algunas realizaciones, las células de animales vertebrados son células de mamíferos. Las células huésped pueden ser células no transformadas o células que ya han sido transfectadas con al menos una molécula de ácido nucleico. En consecuencia, las muestras biológicas, la biomasa y la progenie de una célula recombinante según uno cualquiera de los aspectos y realizaciones anteriores también están dentro del alcance de la presente solicitud. Por tanto, como se analiza con más detalle a continuación, los polipéptidos producidos mediante un método según este aspecto de la solicitud también están dentro del alcance de esta solicitud.

En algunas realizaciones, la célula recombinante es una célula animal. La producción de proteínas terapéuticas a pequeña y gran escala es un campo importante de desarrollo en la industria farmacéutica, porque se cree que las proteínas producidas en células animales generalmente tienen un procesamiento adecuado y una modificación postraduccional y, por lo tanto, tienen una actividad adecuada para el tratamiento de la condición fisiológica. En principio, se puede utilizar generalmente cualquier especie animal y puede ser, por ejemplo, humanos, perros, pájaros, peces, caballos, cerdos, primates, ratones, vacas, cerdos, ovejas, conejos, gatos, cabras, burros, hámsteres o búfalo. Ejemplos no limitantes de especies de aves adecuadas incluyen pollo, pato, ganso, pavo, avestruz, emú, cisne, pavo real, faisán, perdiz y pintada. En algunas realizaciones particulares, la especie de pez es una especie de salmón. También son adecuadas las células primarias de mamífero y los tipos de células continuas/inmortalizadas. Los ejemplos no limitantes de células huésped animales adecuadas incluyen, entre otros, células endoteliales de la arteria pulmonar equina, células de la dermis equina, células de riñón de hámster bebé (BHK), células de riñón de conejo, células musculares de ratón, células de tejido conectivo de ratón, células de cuello uterino humano, célula epidermoide de laringe humana, célula de ovario de hámster chino (CHO), célula HEK-293 humana, célula 3T3 de ratón, célula Vero, célula epitelial del riñón canino Madin-Darby (MDCK), célula primaria de fibroblastos de pollo, una célula HuT78, célula pulmonar A549, célula HeLa, célula PER.C6®, célula WI-38, célula MRC-5, FRhL-2 y célula T CEM. En algunas realizaciones, la célula huésped es una célula de riñón de cría de hámster. En algunas realizaciones, la célula de riñón de cría de hámster es una célula BHK-21.

Polipéptidos recombinantes

Algunos aspectos descritos en la presente memoria se relacionan con polipéptidos recombinantes producidos mediante un método según una o más realizaciones descritas en la presente memoria. Los polipéptidos recombinantes de la presente solicitud generalmente pueden ser cualquier polipéptido recombinante y pueden ser, por ejemplo, uno o más de polipéptidos terapéuticos, polipéptidos profilácticos, polipéptidos de diagnóstico, polipéptidos nutraceuticos, enzimas industriales y polipéptidos indicadores. En algunos aspectos descritos en la presente memoria, los polipéptidos recombinantes pueden ser uno o más de anticuerpos, antígenos, inmunomoduladores y citocinas. En algunos aspectos descritos en la presente memoria, el polipéptido de interés puede tener actividad terapéutica o profiláctica.

Composiciones

Algunos aspectos descritos en la presente memoria se relacionan con una composición que comprende cualquiera de los polipéptidos recombinantes descritos en la presente memoria. La composición puede ser, por ejemplo, una composición nutracéutica, una composición profiláctica, una composición farmacéutica que comprende un portador farmacéuticamente aceptable o una mezcla de los mismos. En algunos aspectos, las composiciones de la presente solicitud se pueden utilizar como vacuna.

Algunos aspectos descritos en la presente memoria se relacionan con una composición que incluye cualquiera de las moléculas de ácido nucleico descritas en la presente memoria. La composición puede ser, por ejemplo, una composición nutracéutica, una composición profiláctica, una composición farmacéutica que comprende un portador farmacéuticamente aceptable o una mezcla de los mismos. En algunos aspectos, las composiciones de la presente solicitud se pueden utilizar como vacuna.

Algunos aspectos descritos en la presente memoria se relacionan con una composición que incluye cualquiera de las células recombinantes descritas en la presente memoria. La composición puede ser, por ejemplo, una composición nutracéutica, una composición profiláctica, una composición farmacéutica que comprende un portador farmacéuticamente aceptable o una mezcla de los mismos. En algunos aspectos, las composiciones de la presente solicitud se pueden utilizar como vacuna.

Como se usa en la presente memoria, el término “portador farmacéuticamente aceptable” significa un portador que es útil para preparar una composición o formulación farmacéutica que es generalmente segura, no tóxica y ni biológicamente ni de otro modo indeseable, e incluye un portador que es aceptable para uso veterinario así como para uso farmacéutico humano. En algunas realizaciones, un portador farmacéuticamente aceptable es tan simple como agua, pero también puede incluir, por ejemplo, una solución de concentración de sal fisiológica. En algunas realizaciones, un portador farmacéuticamente aceptable puede ser, o puede incluir, estabilizadores, diluyentes y tampones. Como estabilizadores entran en consideración, por ejemplo, SPGA, hidratos de carbono (como leche en polvo, albúmina sérica o caseína) o sus productos de degradación. Los tampones adecuados son, por ejemplo, fosfatos de metales alcalinos. Los diluyentes incluyen agua, tampones acuosos (como solución salina tamponada), alcoholes y polioles (como glicerol). Para la administración a animales o seres humanos, la composición según la presente solicitud se puede administrar, *entre otras cosas*, por vía intranasal, mediante pulverización, por vía intradérmica, subcutánea, oral, mediante aerosol o por vía intramuscular.

No se admite que cualquier referencia citada en la presente memoria constituya técnica anterior. La discusión de las referencias expresa lo que afirman sus autores, y los inventores se reservan el derecho de cuestionar la exactitud y pertinencia de los documentos citados. Se entenderá claramente que, aunque en la presente memoria se hace referencia a varias fuentes de información, incluidos artículos de revistas científicas, documentos de patente y libros de texto; esta referencia no constituye una admisión de que cualquiera de estos documentos forme parte del conocimiento general común en el arte.

La discusión de los métodos generales proporcionados en la presente memoria tiene como objetivo únicamente fines ilustrativos. Otros métodos y alternativas alternativos serán evidentes para los expertos en la técnica tras la revisión de esta descripción, y deben incluirse dentro del alcance de esta solicitud.

Ejemplos

Se describen alternativas adicionales se describen con más detalle en los siguientes ejemplos, que no pretenden limitar el alcance de las reivindicaciones.

Ejemplo 1

Procedimientos experimentales generales

Transcripción *in vitro*

Las reacciones de transcripción *in vitro* (IVT, por sus siglas en inglés) se realizaron utilizando 1 µg de plantilla de ADN preparada como se describe anteriormente, en una reacción de 20 µl durante una incubación de una hora a 37 °C (cat. n.º NEB E2065S). Luego se añadió 1 unidad de DNasa I, proporcionada por el proveedor, directamente a la reacción IVT y se incubó a 37 °C durante 15 minutos más. Luego las reacciones se colocaron en hielo y se purificaron usando el método sugerido por el fabricante (Qiagen Cat. n.º 74104). Luego se cuantificó el ARN purificado utilizando un espectrofotómetro UV-Vis NanoDrop 2000c. El ARN se visualizó mediante electroforesis a través de geles de agarosa al 0,8 % (Life Technologies, cat. n.º G5018-08) y se comparó con Millennium RNA Marker (Ambion, cat. n.º AM7150), antes de proceder con la electroporación.

Se purificaron plantillas de ADN plasmídico (Qiagen, cat. n.º 12163) a partir de 300 ml de cultivos saturados de *E. coli* TransforMax Epi300 (Epicentre, cat. n.º EC300105) cultivados en medio caldo LB (Teknova, cat. n.º L8000 06) suplementado con 50 ng/ ml de carbacilina (Teknova, cat. n.º NC9730116). El ADN plasmídico se linealizó mediante

digestión con *Not*-I (New England Biolabs NEB, cat. n.º R3189S) durante una hora a 37 °C. Luego se volvió a purificar el ADN molde linealizado (Zymo, cat. n.º D4003) y se analizó mediante gel de agarosa al 0,8 % (Life Technologies, cat. n.º G5018-08) frente a una escalera de ADN comercial de 2 log (New England Biolabs, NEB, cat. n.º N3200S). Se confirmó la presencia de una sola banda en cada muestra, correspondiente al tamaño de fragmento esperado de la plantilla de ADN lineal, antes de proceder con la transcripción *in vitro*.

Transfección y análisis

En un experimento típico de transfección celular, se introdujo ARN replicón en células BHK-21 mediante electroporación utilizando el kit SF Cell Line Nucleofector™ para el sistema 4D-Nucleofector™ (Lonza). Se recogieron células BHK-21 usando tripsina al 0,25 % y se lavaron una vez con PBS frío. Las células se resuspendieron en tampón SF a una densidad celular de 1×10^6 células por 20 µl de reacción de electroporación. Se electroporaron tres microgramos de ARN en células por triplicado en una cubeta-tira de 16 pocillos y se incubaron a temperatura ambiente durante 10 minutos. Las células electroporadas se recuperaron en placas que contenían medio Eagle modificado por Dulbecco que contenía suero bovino fetal al 10 %, seguido de una incubación durante 16 a 18 h en condiciones estándar de cultivo celular.

Se realizaron análisis intracelulares de la eficacia de la transfección de replicones y la producción de proteínas mediante citometría de flujo. Las células BHK-21 transfectadas se fijaron y permeabilizaron utilizando concentrado de fijación/permanente y tampón de permeabilización (eBioscience). Las células se incubaron con anticuerpos para la producción de ARN bicatenario (anticuerpo monoclonal J2 anti-ARNbc IgG2A, English & Scientific Company) conjugado con R-ficoeritrina (Innova Biosciences). La producción de antígeno se evaluó mediante incubación adicional con anticuerpos específicos de antígeno conjugados con PE-Cy5 (Innova Biosciences) (*p. ej.*, anticuerpos para Renilla verde, Firefly roja, HA o RSV-F0 (Abcam)). Luego, las células se lavaron una vez y se analizaron utilizando un clasificador de células Fusion FACS Aria™ (BD Biosciences) o un clasificador de células FACS Aria™ II (BD Biosciences). Antes de la recolección de la muestra, se procesaron células BHK-21 transfectadas teñidas con colores únicos para controles de compensación. Los datos se recopilaron utilizando FACSDiva (BD Biosciences) y se analizaron adicionalmente utilizando el software FlowJo. La selección inicial se realizó para excluir células muertas y desechos mediante gráficos de dispersión frontal y lateral. Se realizaron controles adicionales para identificar poblaciones de células que eran positivas tanto para ARNbc (R-PE-positivo) como para expresión de proteínas (PE-Cy5-positivo o FITC-positivo para expresión de GFP). Se recogieron frecuencias e intensidades medias de fluorescencia y se utilizaron para comparar y optimizar construcciones.

Ejemplo 2

Modificaciones de la secuencia UTR 5'

Este ejemplo describe los resultados de experimentos en los que se modificaron secuencias UTR 5' para mejorar la expresión de los genes de interés codificados en un vector de ARN replicón de VEEV. La mutagénesis dirigida al sitio (SDM) se llevó a cabo en un plásmido que contenía una copia de ADNc de un vector replicón VEEV. Los cebadores SDM se diseñaron para cambiar el residuo de timina (T) en la posición 2 de la UTR 5' a un residuo G.

En estos experimentos, se usaron replicones de VEEV, cada uno de los cuales expresa el gen de luciferasa rFF, el gen de hemaglutinina (HA) de la influenza A Vietnam 1203 o el gen de la proteína fluorescente verde (GFP), como plantillas para llevar a cabo mutagénesis dirigida. El nucleótido T en la posición 2 en la UTR 5' de la secuencia VEEV de tipo salvaje (cepa TC-83) se cambió a G. Cebadores "VEE 5' T->G nt 2 F" y "VEE 5' T-> G nt 2 R" se utilizaron para introducir el cambio de nucleótido 2 mediante un kit de mutagénesis dirigida de Agilent. Los cebadores se diseñaron a través del sitio web de Agilent. Se identificaron clones positivos y se confirmó que la secuencia de clones representativos que expresaban rFF, HA y GFP era completamente correcta.

Los cebadores SDM para producir replicones Alpha-R-T2G son los siguientes.

Cebador directo: VEE 5' T->G nt 2 F (Id. de sec. n.º: 23): cgactcactatagaGaggcgccgcatgag.

Cebador inverso: VEE 5' T->G nt 2 R (Id. de sec. n.º: 24): ctcatgcgccgctCtctatagttagtgcg.

Después de la confirmación de la secuencia del cambio T2->G en el ADNc del replicón de VEEV, se generó ARN mediante transcripción *in vitro* usando ARN polimerasa T7 en ADN plasmídico linealizado. El ARN transcrito *in vitro* se purificó y se utilizó para someter a electroporación células BHK-21. Tanto la replicación como la expresión de GOI fueron monitoreadas por FACS utilizando anticuerpos específicos anti-ARNbc y específicos de GOI, respectivamente. La eficiencia de replicación y la expresión de GOI de los replicones U2->G VEEV se compararon directamente con los replicones VEEV de tipo salvaje que expresan el mismo GOI. La secuencia del replicón Alpha-R-rFF-T2G que comprende un gen indicador de Firefly rojo se proporciona como Id. de sec. n.º: 19 en el Listado de secuencias con el promotor T7 y una cola polyA con 40 residuos A. También se indica el nucleótido mutado en la posición 2 después de la secuencia del promotor T7.

Ejemplo 3

La sustitución U2->G en UTR 5' no afecta la actividad biológica del replicón VEEV-HA modificado

5 Este ejemplo describe los resultados de experimentos que evalúan el impacto de una sustitución U2->G en la UTR 5' de un replicón de alfavirus modificado sobre la expresión de un gen indicador del gen de hemaglutinina (HA) de la influenza A Vietnam 1203.

10 Para demostrar que un replicón de alfavirus que contiene un cambio U2->G en la UTR 5' puede efectivamente expresar proteína, se transcribió ARN del replicón *in vitro* a partir de un vector que porta un cambio U2->G en la UTR 5' (U2->G VEEV-HA). Como control positivo para la expresión, se transcribió ARN *in vitro* a partir de un vector que porta una UTR 5' de tipo salvaje (WT VEEV-HA). Se electroporaron células de riñón de hámster bebé (BHK-21) con 3 µg de ARN de VEEV-HA U2->G o de ARN de VEEV-HA de tipo salvaje. En la **figura 1** se muestra un ejemplo de análisis de citometría de flujo para esta comparación. Las células se analizaron mediante FACS con un anticuerpo específico de HA para demostrar tanto la presencia de proteína HA expresada como la cantidad relativa de HA expresada por célula (intensidad de fluorescencia media - MFI). No hay pérdida de actividad biológica de un ARN de replicón que porta el cambio U2->G UTR 5' en relación con la actividad detectada en un replicón con una UTR 5' de tipo salvaje.

20 Los datos experimentales presentados en este Ejemplo indican que no sólo el replicón U2->G VEEV-HA puede expresar la proteína HA sino que el nivel de expresión es equivalente al del replicón VEEV-HA de tipo salvaje. Este resultado es inesperado considerando que el clon infeccioso VEEV/SINV que porta el mismo cambio U2->G UTR 5' tuvo casi una pérdida de tres registros en actividad biológica; aquí no mostramos ninguna reducción en la actividad biológica para un replicón que porta el cambio U2->G UTR 5'.

Ejemplo 4

La sustitución U2->G en UTR 5' mejora la expresión del reportero GFP en 3 veces

30 Este ejemplo describe los resultados de experimentos que evalúan el impacto de una sustitución U2->G en la UTR 5' de un replicón de alfavirus modificado sobre la expresión de un gen indicador de la proteína de fluorescencia verde (GFP).

35 Para demostrar que la expresión de un replicón de alfavirus que contiene un cambio U2->G en la UTR 5' no está restringida a ningún GOI en particular, se compararon vectores de replicón que expresan el gen GFP de manera similar. El ARN replicón se transcribió *in vitro* a partir de un vector que porta un cambio U2->G en la UTR 5' (U2->G VEEV-GFP). Como control positivo para la expresión, se transcribió ARN *in vitro* a partir de un vector que porta una UTR 5' de tipo salvaje (WT VEEV-GFP). Las células BHK se sometieron a electroporación con 3 µg de ARN U2->G VEEV-GFP o ARN VEEV-GFP de tipo salvaje. En la **figura 2** se muestra un ejemplo de análisis de citometría de flujo para esta comparación. Las células se analizaron para determinar la expresión de GFP mediante FACS para demostrar tanto la presencia de la proteína GFP expresada como la cantidad relativa de GFP expresada por célula (intensidad media de fluorescencia - MFI). No sólo no hay pérdida de actividad biológica de un ARN de replicón que porta la UTR U2->G 5' sino que el cambio U2->G UTR 5' en realidad mejoró la expresión 3 veces en relación con la expresión detectada en un replicón con una naturaleza salvaje. -tipo UTR 5'.

45 Una vez más, los datos experimentales presentados en este ejemplo indican que un vector replicón que porta el cambio U2->G UTR 5' puede expresar proteína (esta vez GFP). Quizás incluso más inesperado que la simple expresión de proteínas es que el replicón U2->G VEEV-GFP expresó tres veces más GFP que el replicón VEEV-GFP de tipo salvaje. Una vez más, la pérdida esperada en la actividad biológica anticipada debido al cambio U2->G en la UTR 5' no se realizó y este resultado demostró que el cambio U2->G en la UTR 5' en realidad puede mejorar significativamente la expresión del replicón GOI.

Ejemplo 5

La sustitución U2->G en UTR 5' mejora la expresión del reportero rFF en 2 veces

55 Este ejemplo describe los resultados de experimentos que evalúan el impacto de una sustitución U2->G en la UTR 5' de un replicón de alfavirus modificado sobre la expresión de un gen indicador de luciérnaga roja (rFF).

60 En estos experimentos, se compararon de manera similar otro ejemplo de expresión de un replicón de alfavirus que contiene un cambio U2->G en la UTR 5', vectores de replicón que expresan el gen rFF. El ARN replicón se transcribió *in vitro* a partir de un vector que porta un cambio U2->G en la UTR 5' (U2->G VEEV-rFF). Como control positivo para la expresión, se transcribió ARN *in vitro* a partir de un vector que porta una UTR 5' de tipo salvaje (WT VEEV-rFF). Las células BHK se sometieron a electroporación con 3 µg de ARN U2->G VEEV-rFF o ARN VEEV-rFF de tipo salvaje. En la **figura 3** se muestra un ejemplo de expresión de la proteína luciferasa. En este experimento, se comparó la capacidad de un replicón de alfavirus modificado para llevar un cambio U2->G en la UTR 5' (U2->G alfa rFF) con un replicón de alfavirus que tenía una UTR 5' de tipo salvaje (Alfa rFF). Las células BHK se sometieron a electroporación

con una cantidad equivalente de ARN transcrito *in vitro* de cualquiera de los replicones y luego las células se analizaron para determinar la expresión de luciferasa rFF. Se presenta la cantidad de luciferasa (expresada como unidades relativas de luz (RLU)) expresada por célula. No sólo no hay pérdida de actividad biológica de un ARN de replicón que porta la UTR U2->G 5' sino que el cambio U2->G UTR 5' en realidad mejoró la expresión en aproximadamente 2 veces en relación con la expresión detectada en un replicón con una UTR 5' de tipo salvaje.

Una vez más, los datos experimentales presentados en este ejemplo indican que un vector replicón que porta el cambio U2->G UTR 5' puede expresar proteína (esta vez rFF). Quizás incluso más inesperado que la simple expresión de proteínas es que el replicón U2->G VEEV-rFF expresó ~ dos veces más rFF que el del replicón VEEV-rFF de tipo salvaje. Una vez más, la pérdida esperada en la actividad biológica anticipada debido al cambio U2->G en la UTR 5' no se realizó y este resultado demostró que el cambio U2->G en la UTR 5' en realidad puede mejorar significativamente la expresión del replicón GOI.

Ejemplo 6

Diseños de replicones VEEV multivalentes

Este ejemplo describe experimentos realizados para construir y evaluar replicones de VEEV multivalentes, que posteriormente se despliegan para la expresión de al menos dos polipéptidos diferentes en células recombinantes. En algunos experimentos, el replicón VEEV multivalente incluye en orden 5' a 3' (i) una secuencia 5' requerida para la amplificación mediada por proteínas no estructurales, (ii) una secuencia de nucleótidos que codifica las proteínas no estructurales nsP1, nsP2, nsP3 y nsP4 de VEEV, (iii) al menos dos promotores, cada uno de los cuales está operativamente unido a una secuencia de ácido nucleico heteróloga, en donde la secuencia de ácido nucleico heteróloga reemplaza uno o todos los genes de proteínas estructurales de VEEV, (iv) una secuencia 3' requerida para la transmisión mediada por proteínas no estructurales amplificación, y (v) un tracto de poliadenilato.

Si bien se han descrito alternativas particulares de la presente descripción, debe entenderse que son posibles diversas modificaciones y combinaciones y se contemplan dentro del alcance de las reivindicaciones adjuntas. Por lo tanto, no existe ninguna intención de limitar el resumen y la descripción exactos que se presentan en este documento.

REIVINDICACIONES

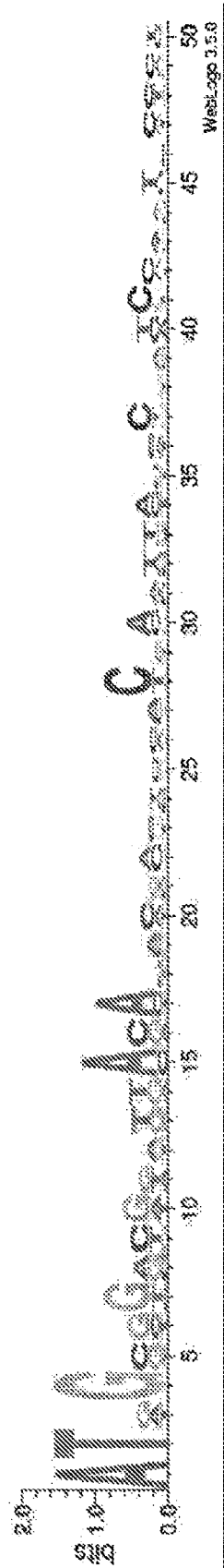
1. Una molécula de ácido nucleico que comprende un ARN de replicón de alfavirus modificado, en donde el ARN de replicón modificado comprende una región no traducida 5' (UTR 5') modificada y secuencias que codifican proteínas no estructurales de alfavirus biológicamente activas, y está desprovista de al menos una porción de una secuencia de ácido nucleico que codifica proteínas estructurales virales, en donde la UTR 5' modificada comprende una sustitución de nucleótidos en la posición 2, en donde las sustituciones de nucleótidos en la posición 2 de la UTR 5' modificada son una sustitución U->G.
2. La molécula de ácido nucleico de la reivindicación 1, en donde el ARN de replicón de alfavirus modificado comprende además un promotor para el ARN subgenómico, secuencias virales 3' requeridas en cis para la replicación, y un tracto de poliadenilato.
3. La molécula de ácido nucleico de la reivindicación 1 o 2, en donde el ARN de replicón de alfavirus modificado carece de una porción sustancial de la secuencia de ácido nucleico que codifica proteínas estructurales virales, o en donde el ARN de replicón de alfavirus modificado no comprende ninguna secuencia de ácido nucleico que codifique proteínas estructurales virales, en donde la UTR 5' modificada presenta al menos 90 %, al menos 95 %, al menos 96 %, al menos 97 %, al menos 98 %, al menos 99 %, o un 100 % de identidad de secuencia con la secuencia de ácido nucleico de al menos uno de Id. de sec. n.º: 2-18.
4. La molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 3, que comprende además uno o más casetes de expresión, en donde cada uno de los casetes de expresión comprende un promotor unido operativamente a una secuencia de ácido nucleico heteróloga.
5. La molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 4, en donde el alfavirus pertenece al género Alphavirus de la familia Togaviridae.
6. La molécula de ácido nucleico de la reivindicación 5, en donde el alfavirus pertenece al grupo VEEV/EEEV, o al grupo SF, o al grupo SIN.
7. La molécula de ácido nucleico de la reivindicación 6, en donde el alfavirus es el virus de la encefalitis equina venezolana (VEEV).
8. La molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 7, en donde el genoma de alfavirus modificado o el ARN de replicón está unido operativamente a un elemento regulador heterólogo.
9. La molécula de ácido nucleico de la reivindicación 8, en donde el genoma o ARN de replicón del alfavirus modificado carece de una porción sustancial de la secuencia de ácido nucleico que codifica proteínas estructurales virales, o en donde el genoma o ARN de replicón de alfavirus modificado no comprende ninguna secuencia de ácido nucleico que codifique proteínas estructurales virales.
10. Una célula recombinante que comprende la molécula de ácido nucleico de una cualquiera de las reivindicaciones 1 a 9, en donde la célula recombinante es una célula procariótica o una célula eucariota.
11. Un método para producir un polipéptido de interés, que comprende cultivar una célula huésped que comprende la molécula de ácido nucleico de una cualquiera de las reivindicaciones 9.

Figura 1A

SEQ ID NO: 2 (AURAV)	ATAGCGGACGGACTAGTACTTGTACTACAGATTAAAC-TGCCGTGTGCCGC-----
SEQ ID NO: 3 (CHIKV)	ATGGCTG-CGTGAGA-----CACACGTAGCCTACCAAGTTTCTTACTGCTCTACTCT-
SEQ ID NO: 4 (ONNV)	ATAGCTG-CGTGATA-----CACACAGCGAGCTTACGGGTTTCATATGCTCTACT
SEQ ID NO: 5 (BEBV)	ATGGCGGCTGTGTGA-CACACGAGCCGTCGATTTCAA-CCTTCTTGTCTCCCT-
SEQ ID NO: 6 (SFV)	ATGGCGGATGTGTGA-CATACACAGCCCAAAAGATT-TTGTTCAGCTCCT-
SEQ ID NO: 7 (MAYV)	ATGGCGGCGACGTGA-CACTTGTTCGCGCGGTCTCT-----CTAAGCTCTTCTCTC
SEQ ID NO: 8 (GETV)	ATGGCGGACGTGTGACATCAGCTTCGCTCTCTCTCTAG-GATCCTTTGTCTAC-----
SEQ ID NO: 9 (SAGV)	ATGGCGGACGTGTGACATCAGCTTCGCTCTCTCTCTAG-GATCCTTTGTCTAC-----
SEQ ID NO: 10 (NDUV)	ATGG--TGGGGAGTT-GA---GAGACGA-AGCACCAA-ACAACCTACGGCGCTCACC-AT-
SEQ ID NO: 11 (MIDV)	ATTGCTGTTACGTA---CACGTCCACCAACCCCCCA-CCCTCCAAAGCATCCA-----
SEQ ID NO: 12 (BEBV)	ATAGGGTACGGTGA-----GAGGCAACCAACCTAT----TTCCACCTATCCAAATGG--
SEQ ID NO: 13 (FMV)	ATAGGGTATGGTTA-----GAGGCGCTACCCCTAC----TTAACCGATCCAAACATGG--
SEQ ID NO: 14 (Buggy)	ATAGGGTATGGTTA-----GAGGCGCTACCCCTAC----TTAACCGATCCAAACATGG--
SEQ ID NO: 15 (VEEV)	ATGSGCGGCGCAAGA-----GAGAGGCCCAAAACCAA----TTACCTACCCAAATGGAG--
SEQ ID NO: 16 (WHAV)	ATGSGCGGCATAGTA-----CATACTATATAAAGAAACAGCGGACCAATTGCAC--
SEQ ID NO: 17 (SINV)	ATTGACGCGCGTAGTA-----CACACTAT-TGAATCAACAGCGGACCAATTGCACCT
SEQ ID NO: 18 (BABV)	ATTGSGCGGCGTAGTA-----CACACTAT-TGAATCAACAGCGGACCAATTGCACCT

★★ *

Figura 1B



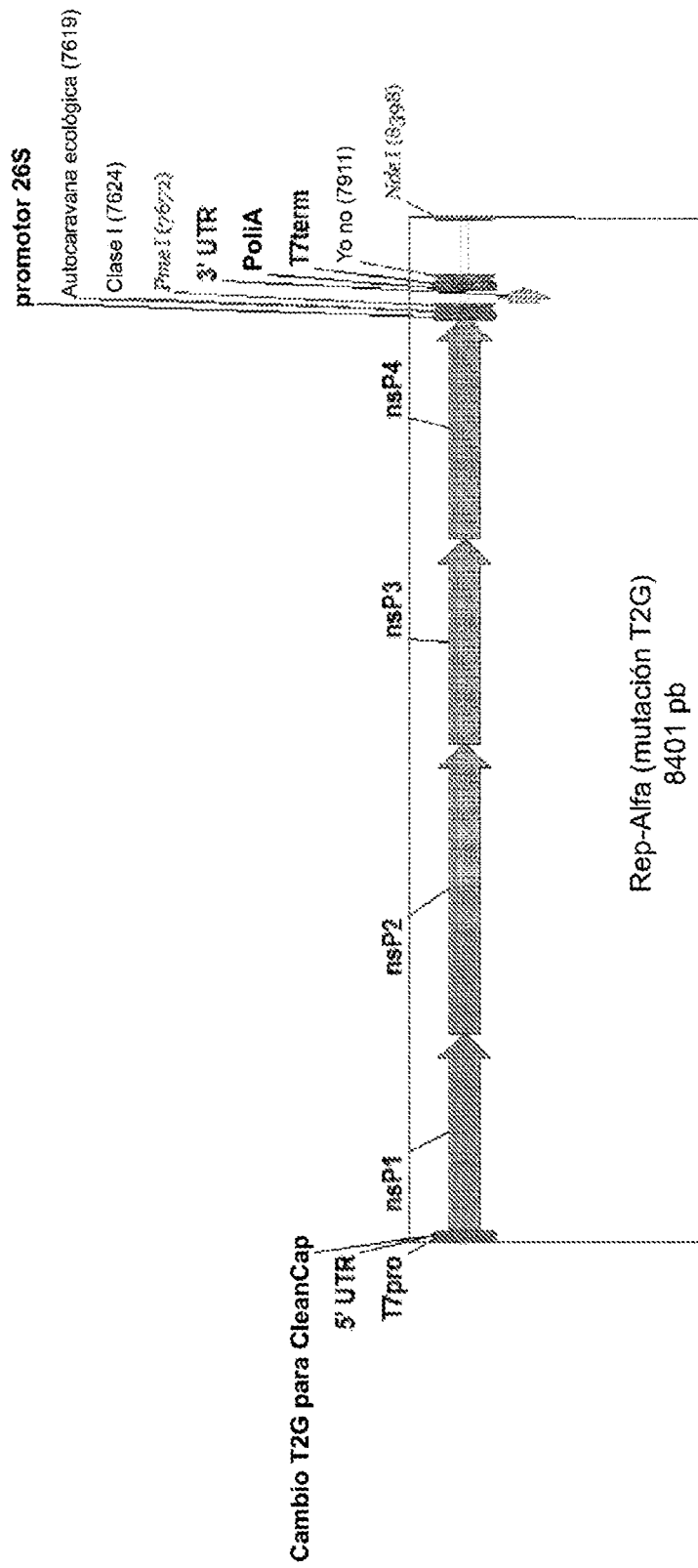


Figura 2

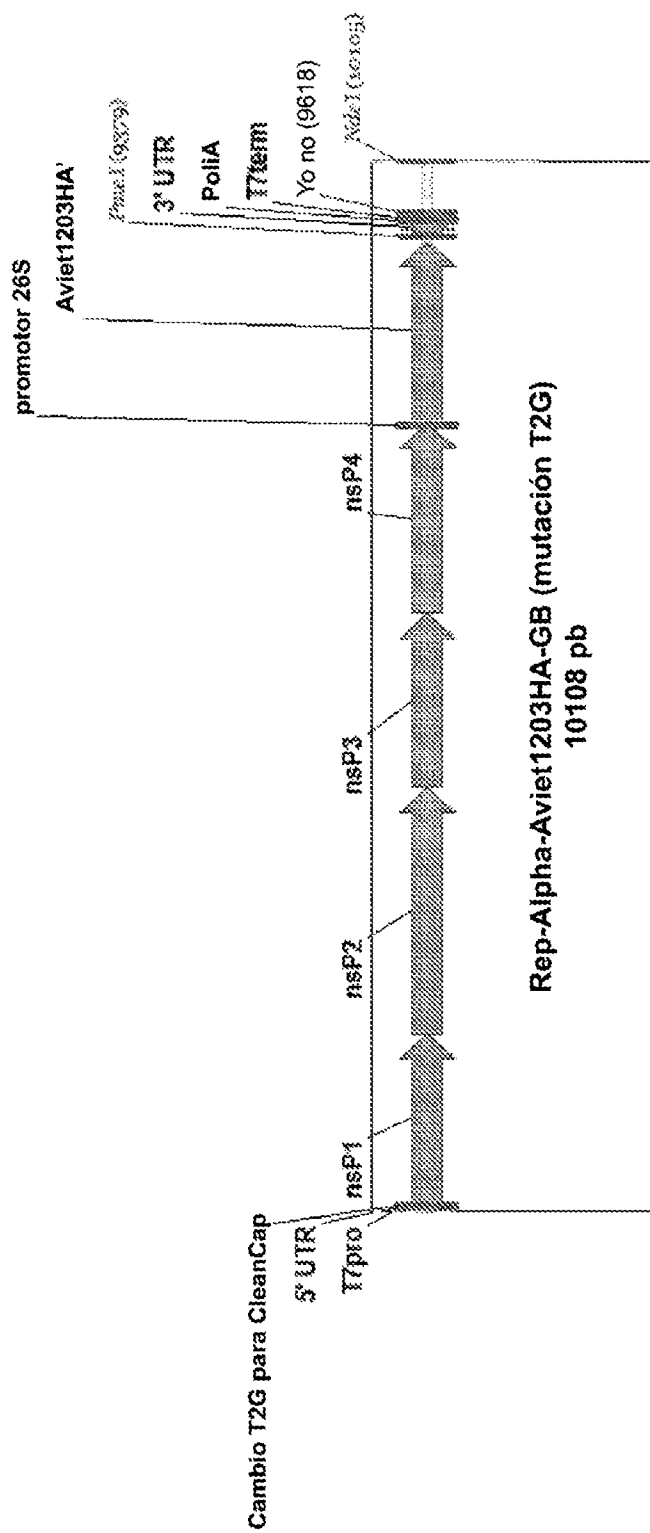


Figura 3

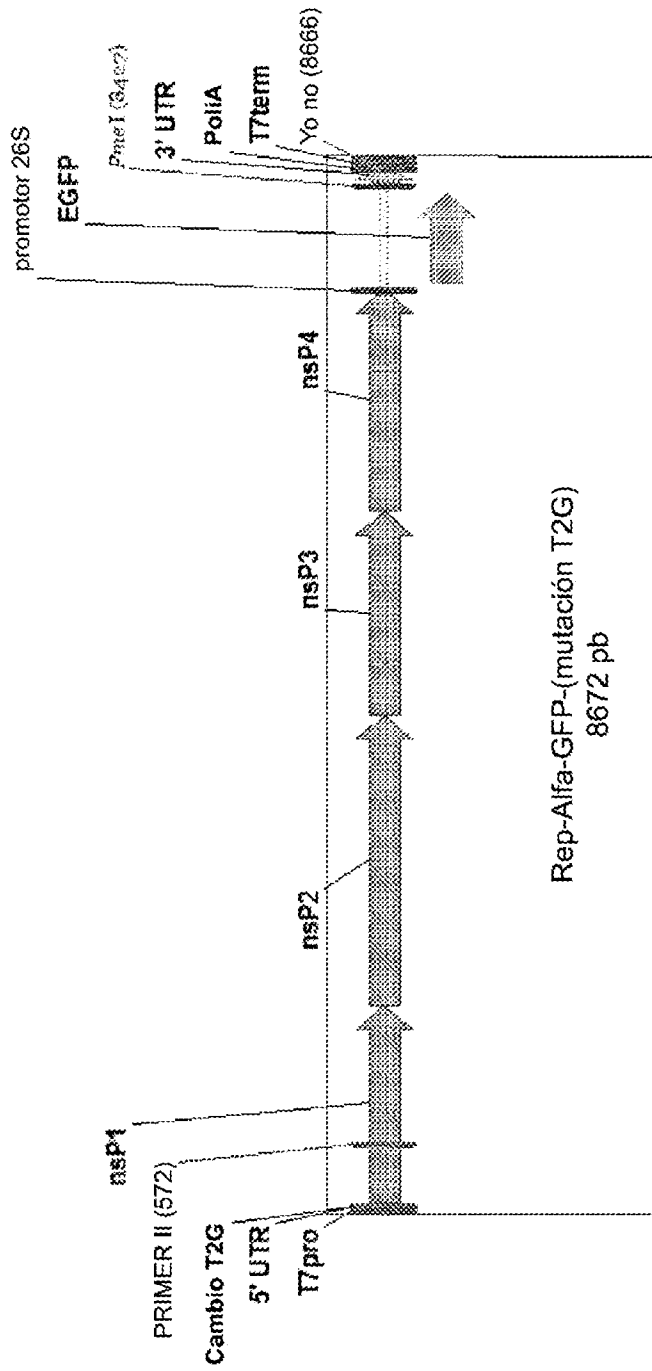


Figura 4

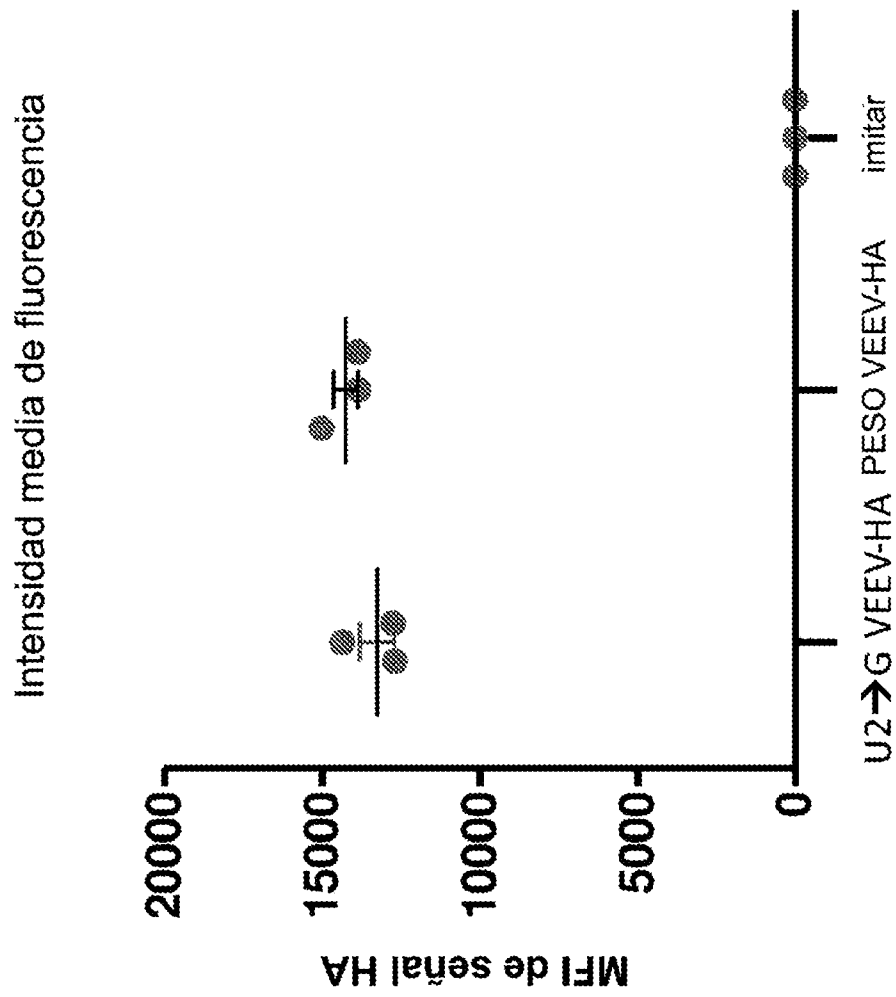


Figura 5

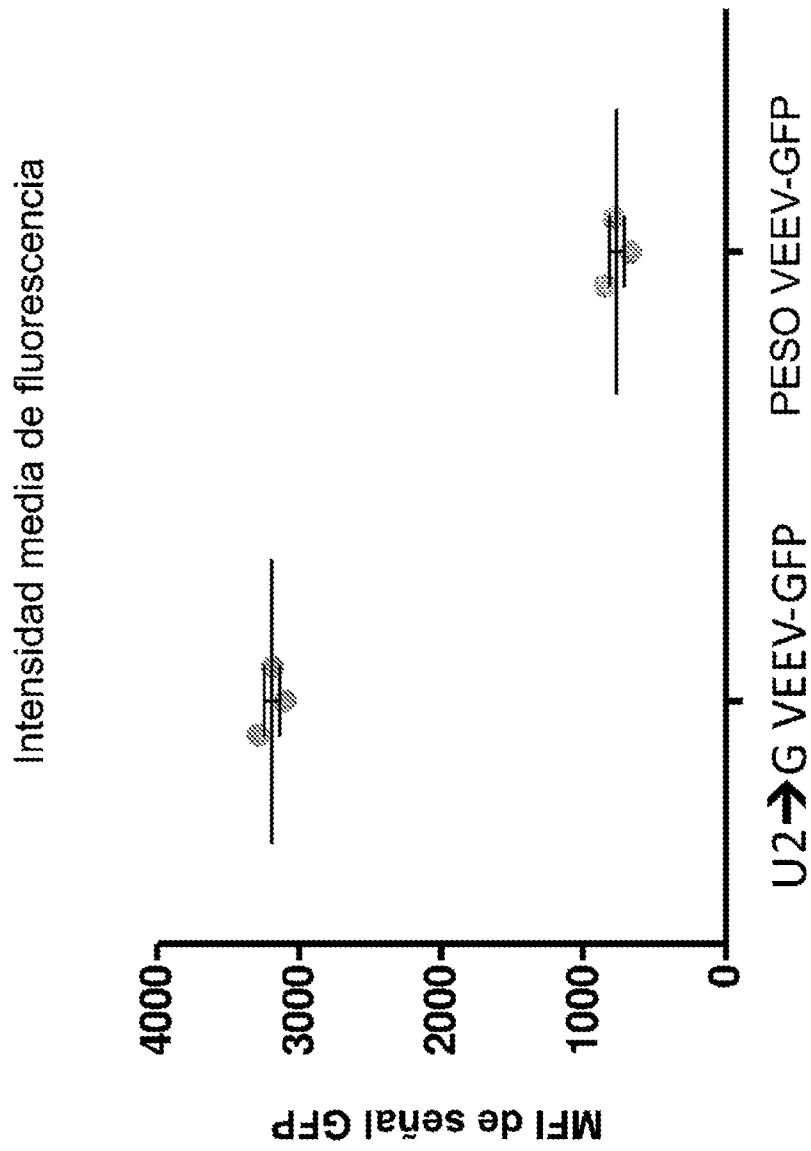


Figura 6

Expresión de luciferasa por célula

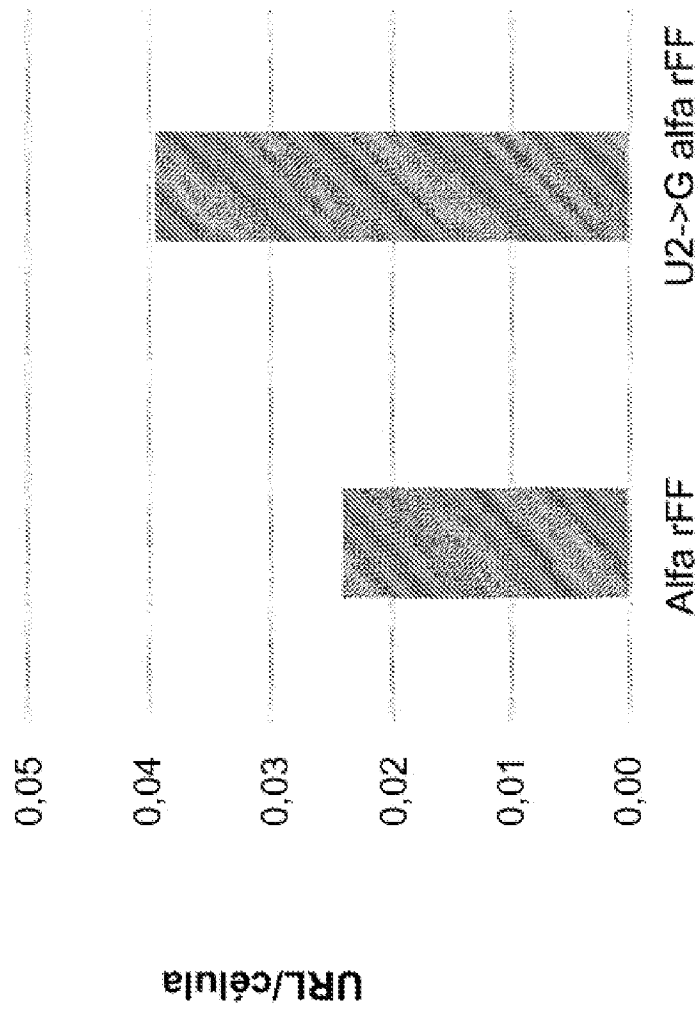


Figura 7

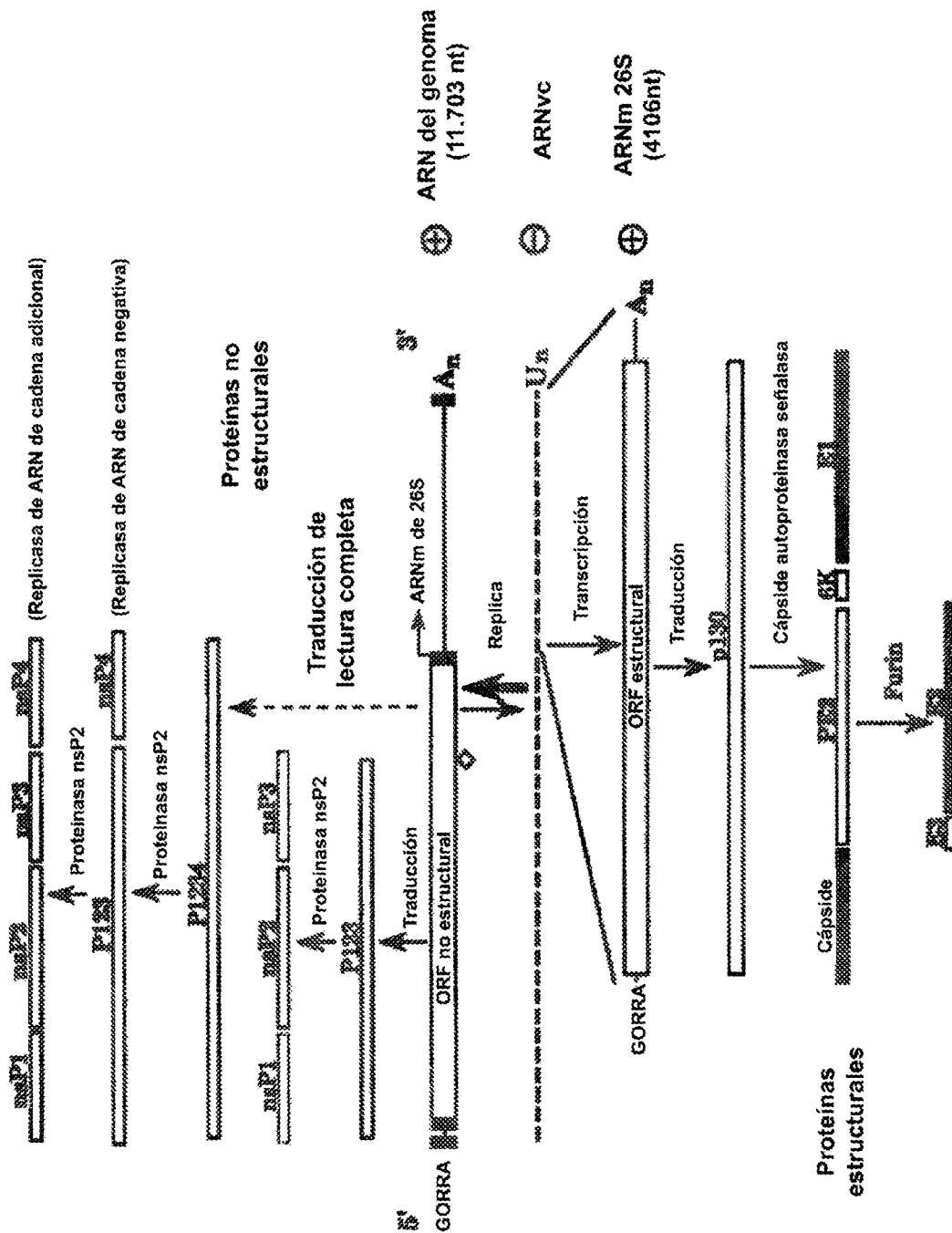


Figura 8