

SECUENCIA MODIFICADA DEL GEN E2A DE ADENOVIRUS PARA LA PRODUCCIÓN DE VECTORES BASADOS EN VIRUS ADENOASOCIADOS (AAV) (VARIANTES)

Campo de la invención

La presente invención se relaciona con el campo de la ingeniería genética, la genética, y la biotecnología. Más específicamente, la presente invención se relaciona con las variantes de una secuencia modificada del gen E2a de adenovirus, a un plásmido ayudante que comprende las variantes de la secuencia modificada del gen E2a de adenovirus para producir vectores en función de virus adeno-asociados, y al uso de las variantes de la secuencia modificada del gen E2a de adenovirus y de los plásmidos ayudantes que comprenden estas secuencias para producir vectores en función de virus adeno-asociados.

Antecedentes de la invención

Al día de la fecha, se han utilizado ampliamente varios sistemas para la entrega de fármacos, mientras que, en el campo de la terapia de genes, la entrega de los constructos de terapia de genes que utilizan virus adeno-asociados (AAV) se encuentran en mayor demanda. AAV consiste en un virus pequeño no encapsulado de la familia de Parvoviridae caracterizado por un genoma de ADN lineal de una sola hebra. El genoma de AAV incluye secuencias, flanqueadas por repeticiones terminales invertidas (ITRs), del gen regulatorio *rep* cuyos productos están involucrados en el proceso de la replicación viral, y del gen *cap* que codifica las proteínas estructurales VP1, VP2, VP3 que forman la cápside de AAV. A pesar de que el virus adeno-asociado no puede replicarse a sí mismo, para vincularse efectivamente en su ciclo de vida lítico, requiere la co-infección simultánea de la célula con el denominado virus ayudante, en particular con adenovirus (Ad) o el virus herpes simplex (HSV) (Manuel AFV Gonçalves, Adeno-associated virus: from defective virus to effective vector, Virol J. 2005 May 6;2:43).

Los adenovirus son miembros de la familia de Adenoviridae y están caracterizados por la presencia de una cápside icosaedra sin una envoltura de lipoproteína, así como también por la presencia de un genoma de ADN de doble hebra. A diferencia del AAV, los adenovirus poseen un genoma significativamente más complejo, donde las secuencias del gen del virus y varios elementos regulatorios se ven comprendidos en un ácido nucleico limitado en tamaño debido a la superposición de secuencias. En general, el genoma de los miembros del gen Mastadenovirus incluye secuencias flanqueadas por repetición terminal invertida (ITR) de genes tempranos, como por ejemplo E1-E4, que codifican proteínas involucradas en los procesos de replicación y transcripción, y secuencias de genes tardíos, como por ejemplo L1-L5, que codifican predominantemente las proteínas estructurales del virus. El uso de adenovirus ayudante en la producción de vectores en función de virus adeno-asociados no ha visto una amplia aplicación para el desarrollo clínico de fármacos de terapia de genes, debido a que se obtuvo como resultado la presencia de toda clase de impurezas, en particular en forma de partículas adenovirales. Posteriormente, para superar las ventajas previamente mencionadas del procedimiento, se desarrolló un sistema de tres plásmidos, libre de partículas de adenovirus, para producir vectores en función de virus adeno-asociados, debido a que previamente se había descubierto que para la

producción de AAV es necesario y suficiente poseer secuencias funcionales de genes E1 de adenovirus, inclusive E1a y E1b, E2a, E4, y ARN VA (asociado al virus) (Juan Jose Aponte-Ubillus et al, Molecular design for recombinant adeno-associated virus (rAAV) vector production. Appl Microbiol Biotechnol, 2018; 102(3): 1045–1054). Por lo tanto, se considera a E1a un transactivador, en particular que facilita la expresión de los genes *rep* y *cap* de AAV; mientras que E1b interactúa con el gen E4 de adenovirus que facilita el transporte intracelular de mRNA viral en el momento apropiado. Este último también está involucrado en la replicación de ADN. Los genes E2a y ARN VA proveen estabilidad de mRNAs virales durante el proceso de traducción eficiente (Xiao Xiao et al, Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus. J Virol, 1998 Mar; 72(3): 2224–2232).

Al día de la fecha, el sistema más común libre de partículas adenovirales para producir vectores en función de virus adeno-asociados es un conjunto de plásmidos para transferir células animales que comprendan la secuencia exógena del gen E1 de adenovirus. Generalmente, el conjunto de plásmidos de transfección consiste en un plásmido que incluye un casete con el gen de interés flanqueado por secuencias ITR, un plásmido que comprende secuencias de los genes *rep* y *cap* de AAV, y un plásmido ayudante que incluye regiones del genoma de adenovirus que comprende a su vez secuencias de los genes E2, E4, y ARN VA de adenovirus y elementos regulatorios requeridos para la producción de AAV. Varias empresas comerciales modernas poseen estos sistemas de tres plásmidos para producir vectores con base en AAV.

Para el tratamiento de patologías genéticas *in vivo*, la ventaja obvia será el uso de fármacos con una alta concentración de partículas de entrega que, a su vez, permitirán lograr un efecto terapéutico. Para producir formulaciones altamente concentradas de fármacos de terapia de genes, es necesario poseer procedimientos de producción que provean un alto rendimiento del fármaco y/o sus componentes, inclusive las partículas de entrega. El uso de vectores de plásmido grandes en el trabajo experimental podría ser un factor que complique el trabajo del experimentador, y también afectaría directamente la productividad del proceso para producir vectores con base en partículas virales adeno-asociadas.

Por lo tanto, en la actualidad, existe una necesidad constante de desarrollar nuevos sistemas para producir vectores con base en virus adeno-asociados, por ejemplo, que comprendan un conjunto de plásmidos que difieran de los componentes existentes del sistema por el hecho de que comprenden secuencias de ácido nucleico más pequeñas que codifican proteínas de adenovirus, que permitirían, al mismo tiempo que conservan toda la funcionalidad de los componentes del sistema, aumentar la productividad en la producción de AAV en comparación con el uso de plásmidos más grandes convencionales de los sistemas disponibles en el mercado.

Divulgación de la invención

Uno de los objetivos urgentes de la investigación en el campo del desarrollo de terapia de genes efectiva es el de crear procedimientos novedosos y mejorar los métodos existentes para producir fármacos de terapia de genes para así aumentar el rendimiento del producto.

Los autores de la invención han desarrollado variantes de una secuencia modificada del gen E2a de adenovirus y plásmidos ayudantes para producir vectores con base en virus adeno-

asociados, que comprende secuencias modificadas del gen E2a de adenovirus. Los autores de la invención han descubierto que el uso de variantes de la secuencia modificada del gen E2a de adenovirus da como resultado, de manera sorprendente, un nivel mayor de productividad durante la generación de AAV.

Definiciones y procedimientos generales

A menos que se defina de otra manera en el presente documento, todos los términos técnicos y científicos utilizados en relación con la presente invención tendrán el mismo significado que se entiende comúnmente por parte de las personas versadas en la técnica.

Asimismo, a menos que el contexto requiera lo contrario, los términos singulares incluirán los términos plurales, y los términos plurales incluirán los términos singulares. Generalmente, la clasificación y procedimientos utilizados de cultivo celular, biología molecular, microbiología, genética descritos en el presente documento son reconocidos y utilizados ampliamente por las personas versadas en la técnica. Las reacciones de enzimas y procedimientos de purificación se llevan a cabo de acuerdo con los lineamientos del fabricante, tal como es común en la técnica, o tal como se describan en el presente documento.

Los términos “ocurre naturalmente”, “nativo”, o “de tipo salvaje” se utilizan para describir un objeto que puede encontrarse en la naturaleza al contrario de los que se producen artificialmente. Por ejemplo, una proteína o secuencia de nucleótidos presente en un organismo, inclusive en un virus, que puede aislarse de una fuente en la naturaleza y que no ha sido modificada intencionalmente por una persona en el laboratorio, ocurre naturalmente.

Tal como se utilizan en la presente descripción y reivindicaciones que siguen, a menos que el contexto sugiera lo contrario, los términos “incluir” y “comprender”, o sus variaciones como “incluye”, “inclusive”, “comprende”, o “que comprende”, serán entendidos como términos que supongan la inclusión de un número entero establecido o grupo de números enteros pero no la exclusión de cualquier otro número entero o grupo de números enteros.

Descripción detallada de la invención

Ácido nucleico

En un aspecto, la presente invención se relaciona con un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende una modificación seleccionada del siguiente grupo:

- i. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o
- ii. supresión de una región de secuencia correspondiente a 1696-3492 bp; o
- iii. supresión de una región de secuencia correspondiente a 1696-2321 bp; o
- iv. supresión de una región de secuencia correspondiente a 2400-2634 bp; o

- v. supresión de una región de secuencia correspondiente a 2839-3492 bp; o
 - vi. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp; o
 - vii. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2839-3492 bp; o
 - viii. supresiones de regiones de secuencia correspondientes a 2400-2634 bp, 2839-3492 bp
- en la secuencia de SEQ ID NO:1 o en una secuencia homóloga a la secuencia de SEQ ID NO:1 con un grado de identidad de al menos 63%.

Los términos “ácido nucleico”, “polinucleótido”, “oligonucleótido”, “secuencia de polinucleótido” utilizados de forma indistinta en la presente descripción se refieren a una secuencia precisa de nucleótidos, modificados o no, que determinan un fragmento o una región de un ácido nucleico, que contiene nucleótidos no naturales o no, y que es ADN o ARN de doble hebra, ADN o ARN de una sola hebra, o productos de transcripción de estos ADN.

Tal como se utiliza en la presente descripción, el término polinucleótidos puede incluir, a modo de ejemplos no limitantes, todas las secuencias de ácido nucleico que se obtienen mediante cualquier medio disponible en la técnica, entre los que se incluye, como ejemplos no limitantes, medios recombinantes, es decir, la clonación de secuencias de ácido nucleico de una librería recombinante o un genoma celular, que utiliza la tecnología de clonación normal y PCR y procedimientos similares, y mediante medios sintéticos.

Cabe mencionar que la presente invención no se relaciona con las secuencias de nucleótidos en su medioambiente cromosómico natural, es decir, en un estado natural. Por lo tanto, los ácidos nucleicos aislados producidos mediante genética recombinante, por ejemplo, con el uso de células anfitrionas, o producidos mediante la síntesis química, deberían incluirse en la presente invención.

A menos que se indique lo contrario, el término “secuencia de nucleótidos” abarca su complemento. De esta manera, un ácido nucleico con una secuencia en particular deberá entenderse como uno que abarca su hebra complementaria con su secuencia complementaria.

El término “secuencia homóloga” se refiere a dos o más secuencias de nucleótidos con un grado de identidad de al menos 61 por ciento, en particular 61 por ciento, 62 por ciento, 63 por ciento, 64 por ciento, 65 por ciento, 66 por ciento, 67 por ciento, 68 por ciento, 69 por ciento, 70 por ciento, 71 por ciento, 72 por ciento, 73 por ciento, 74 por ciento, 75 por ciento, 76 por ciento, 77 por ciento, 78 por ciento, 79 por ciento, 80 por ciento, 81 por ciento, 82 por ciento, 83 por ciento, 84 por ciento, 85 por ciento, 86 por ciento, 87 por ciento, 88 por ciento, 89 por ciento, 90 por ciento, 91 por ciento, 92 por ciento, 93 por ciento, 94 por ciento, 95 por ciento, 96 por ciento, 97 por ciento, 98 por ciento, 99 por ciento, 100 por ciento. De esta manera, de acuerdo con la presente invención, las secuencias de nucleótido homólogas pueden tener un origen evolutivo común o pueden producirse mediante optimización de codón en la secuencia inicial.

El término “identidad” se refiere a la comparación de una secuencia de nucleótidos de dos ácidos nucleicos. La identidad se determina en función de una secuencia de referencia. La secuencia del gen E2a de adenovirus se utiliza como secuencia de referencia de acuerdo con la presente invención. Los algoritmos para el análisis de secuencias son ampliamente conocidos en la técnica.

Una de las propiedades del código genético es la degeneración, es decir, la capacidad de diferentes codones (trinucleótidos) de codificar el mismo aminoácido. Estos codones que se traducen al mismo aminoácido se denominan codones sinónimos. En las secuencias naturales, uno de los codones sinónimos se selecciona aleatoriamente en el transcurso de la evolución, pero las frecuencias de uso de codones sinónimos son diferentes: cada aminoácido posee de mayor y menor preferencia. La optimización de codón es una técnica usada ampliamente para amplificar la producción de moléculas de proteína, lo que provee un mapeado racional de uno de los codones sinónimos apropiados a cada aminoácido en una secuencia de proteína. Uno de los principios más comunes de la optimización de codón incluye el uso de los codones más frecuentes, mientras que se introdujeron otros enfoques posteriormente, como por ejemplo la armonización (reproducción de distribución de frecuencias de uso de codón), pero no siempre aumentan la productividad. Además de las frecuencias de codón, el contenido de GC de secuencia (relación entre guanina y citosina y la longitud total de la secuencia) puede afectar la eficiencia de la producción, en particular, se ha demostrado que un alto contenido de GC se relaciona con niveles mayores de mRNA en células mamíferas (Grzegorz Kudla ET AL., High Guanine and Cytosine Content Increases mRNA Levels in Mammalian Cells, June 2006, Volumen 4, Edición 6, e180, pp. 933-942). Además, cabe mencionar que los elementos de estructura secundaria estables de mRNA, es decir, los que tienen baja energía de plegado libre, pueden reducir la eficiencia.

El término “adenovirus” se refiere a un virus de la familia de Adenoviridae. Los virus de esta familia son partículas virales que incluyen un genoma de ADN de doble hebra y una cápside icosaedra que no posee una envoltura de lipoproteína. La familia de Adenoviridae puede dividirse en 6 géneros y, en particular, comprende el gen de Mastadenovirus que infecta mamíferos. El criterio para la demarcación en la familia es el resultado del análisis filogenético, características de organización del genoma, así como también la infección de ciertos animales con el virus (Mária Benkő et al, ICTV Virus Taxonomy Profile: Adenoviridae 2022, J Gen Virol. 2022 Mar;103(3):001721). En el gen Mastadenovirus, se distingue un grupo de adenovirus de primates, que combina siete subgrupos o especies de adenovirus (HAdV-A-G). En vista de la presencia de diferentes características biológicas, también se han caracterizado serotipos de virus en estos subgrupos de adenovirus, como los serotipos 2, 5 en la especie Mastadenovirus C, serotipo 54 en la especie Mastadenovirus D y otros. Los adenovirus infectan el tracto respiratorio, el tracto intestinal, los riñones, y principalmente utilizan el receptor CAR (coxsackievirus y receptor AdV) para penetrar la célula (Urs F. Greber, Adenoviruses - Infection, pathogenesis and therapy, FEBS Lett. 2020 Jun;594(12):1818-1827). Se ha suministrado información más detallada sobre los varios miembros de la familia Adenoviridae en la literatura (Balázs Harrach et al, Adenoviruses across the animal kingdom: a walk in the zoo, FEBS Lett. 2019 Dec;593(24):3660-3673).

La organización del genoma de adenovirus se caracteriza por una variabilidad significativa entre géneros en términos del contenido de elementos funcionales y regulatorios y en términos del tamaño del genoma. En general, el genoma de adenovirus incluye secuencias de genes tempranos (E1-E4), intermedios (transcriptos IX y IVa2) y tardíos (L1-L5) flanqueados por repeticiones terminales invertidas (ITRs), que comprenden una secuencia conservadora y sirven como el origen de la replicación del virus. En función del tipo de adenovirus, los genes de ARN adeno-asociados (VA RNA) también pueden representarse en el genoma (Thomas Lion, Adenovirus Infections in Immunocompetent and Immunocompromised Patients, Clin Microbiol Rev. 2014 Jul; 27(3): 441–462).

La replicación de adenovirus incluye la proteína pre-terminal (pTP), polimerasa de adenovirus (AdV Pol), y proteína de enlace a ADN, codificada por el gen E2, así como también algunos factores de transcripción celular. En las etapas tempranas de replicación, estas proteínas se enlazan a varias ITRs de los virus, lo que determina la formación del complejo de pre-iniciación y posteriormente determina la formación de un nuevo genoma bidireccional (Rob C. Hoebe et al., Adenovirus DNA Replication, Cold Spring Harb Perspect Biol. 2013 Mar; 5(3): a013003). Las proteínas de adenovirus codificadas por el gen E1 son necesarias para la activación de promotores de genes tempranos durante el ciclo lítico del virus; asimismo, la transcripción se activa solamente mediante la interacción física de los productos de genes E1 y E4. A su vez, E4 codifica un complejo de productos de proteína que facilitan la transcripción y están involucradas en el empalme y transporte de ARN (M Bondesson, Adenovirus E4 open reading frame 4 protein autoregulates E4 transcription by inhibiting E1A transactivation of the E4 promoter, J Virol. 1996 Jun; 70(6):3844–51).

Los términos “adenovirus” o “Ad” se utilizan en el presente documento de forma indistinta.

La producción de vectores de rAAV al usar la transfección de células mamíferas como por ejemplo HEK293 con un sistema de tres plásmidos es el procedimiento más eficiente y más ampliamente utilizado. El primero en el sistema de estos plásmidos comprende un casete de expresión que comprende un gen de interés (GOI) conectado operativamente a las secuencias de control de expresión, en las que estos elementos se flanquean mediante repeticiones terminales invertidas (ITRs) del virus. Las ITRs son componentes esenciales para la replicación viral y empaquetado de ese casete de expresión de interés en la cápside del virus. El segundo plásmido es el que comprende las secuencias de nucleótidos de los genes *rep* y *cap* del virus adeno-asociado, que codifican las enzimas de replicación del virus y proteínas estructurales de su cápside de un serotipo deseado. El tercer plásmido, también denominado plásmido ayudante, comprende genes de adenovirus que facilitan la replicación y empaquetado del genoma recombinante de AAV (Parminder Singh Chahal et al., Production of adeno-associated virus (AAV) serotypes by transient transfection of HEK293 cell suspension cultures for gene delivery, J Virol Methods. 2014 Feb; 196: 163–173).

Las variantes resultantes de la secuencia modificada del gen E2a de adenovirus con elementos regulatorios que comprenden una modificación seleccionada del siguiente grupo:

i. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o

ii. supresión de una región de secuencia correspondiente a 1696-3492 bp; o

iii. supresión de una región de secuencia correspondiente a 1696-2321 bp; o

iv. supresión de una región de secuencia correspondiente a 2400-2634 bp; o

v. supresión de una región de secuencia correspondiente a 2839-3492 bp; o

vi. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp;
o

vii. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2839-3492 bp;
o

viii. supresiones de regiones de secuencia correspondientes a 2400-2634 bp, 2839-3492 bp
en la secuencia de SEQ ID NO:1 o en una secuencia homóloga a la secuencia de SEQ ID NO:1 con un grado de identidad de al menos 63% arrojó como resultado un nivel mayor de productividad durante la generación de vectores en función de AAV en comparación con la secuencia de SEQ ID NO: 1 o a una secuencia homóloga a SEQ ID NO: 1 con un grado de identidad de al menos 63%. Las supresiones de estas regiones no afectaron negativamente la expresión del gen E2a y la función de la proteína respectiva.

La SEQ ID NO 1 es la secuencia del gen de tipo salvaje con elementos regulatorios que codifica la proteína E2a de adenovirus y corresponde a la secuencia de genoma del serotipo de adenovirus 2 (N° de Ref. NCBI AC_000007.1) de 22393 bp a 27214 bp.

En algunas realizaciones de la invención, la secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 1, o una variante optimizada de codón de la secuencia SEQ ID NO: 1.

Las personas versadas en la técnica comprenderán fácilmente que las secuencias de genoma de adenovirus de diferentes serotipos pueden tener similitud significativa y caracterizarse por un alto porcentaje de identidad, especialmente, por ejemplo, las secuencias de genoma de diferentes serotipos de adenovirus del mismo gen. Por lo tanto, los serotipos de adenovirus 1, 5, 6, 7, 11, 54 y 89 se relacionan con el gen Mastadenovirus, mientras que la secuencia de SEQ ID NO: 1 es homóloga a la secuencia del gen E2a con elementos regulatorios del serotipo de adenovirus 1, que es SEQ ID NO: 32, con 98% de identidad; a la secuencia del gen E2a con elementos regulatorios de serotipo de adenovirus 5, que es SEQ ID NO: 33, con 97% de identidad; a la secuencia del gen E2a con elementos regulatorios del serotipo de adenovirus 6, que es SEQ ID NO: 34, con 99% de identidad; a la secuencia del gen E2a con elementos regulatorios del serotipo de adenovirus 7, que es SEQ ID NO: 35, con 64% de identidad; a la secuencia del gen E2a con elementos regulatorios del serotipo de adenovirus 11, que es SEQ ID NO: 36, con 64% de identidad; a la secuencia del gen E2a con elementos regulatorios del serotipo de adenovirus 54,

que es SEQ ID NO: 19, con 63% de identidad; a la secuencia del gen E2a con elementos regulatorios del serotipo de adenovirus 89, que es SEQ ID NO: 37, con 98% de identidad. La secuencia del serotipo de adenovirus humano 2 en el caso de marras sirve como secuencia de referencia con respecto a las posiciones de ácido nucleico para todas las otras secuencias de genoma de adenovirus mencionadas en el presente documento. Las respectivas posiciones de ácido nucleico de adenovirus pueden identificarse mediante el alineamiento de secuencia convencional con el uso del software disponible públicamente BlastP del Centro Nacional para la Información de la Biotecnología (NCBI).

Por lo tanto, el gen E2a de adenovirus que comprende una modificación seleccionada del siguiente grupo:

i. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o

ii. supresión de una región de secuencia correspondiente a 1696-3492 bp; o

iii. supresión de una región de secuencia correspondiente a 1696-2321 bp; o

iv. supresión de una región de secuencia correspondiente a 2400-2634 bp; o

v. supresión de una región de secuencia correspondiente a 2839-3492 bp; o

vi. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp; o

vii. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2839-3492 bp; o

viii. supresiones de regiones de secuencia correspondientes a 2400-2634 bp, 2839-3492 bp

en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% puede codificarse mediante una amplia gama de diferentes secuencias de nucleótido que representan la secuencia del gen E2a con elementos regulatorios de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 1. Estas variantes de secuencia de nucleótido se encuentran dentro del alcance de la presente invención.

En algunas realizaciones de la invención, una variante de la secuencia SEQ ID NO: 1, que es una secuencia que ocurre naturalmente del gen E2a con elementos regulatorios de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 1 es la secuencia de SEQ ID NO: 19 y corresponde a la secuencia del genoma del serotipo de adenovirus 54 (Secuencia de Referencia NCBI: NC_012959.1) de 21166 bp a 25369 bp. La variante que ocurre naturalmente de la secuencia del gen E2a de un serotipo de adenovirus, que no sea el serotipo con SEQ ID NO: 1, con la secuencia de SEQ ID NO: 19 se provee solamente a fines ilustrativos y no debe interpretarse como variantes que ocurren naturalmente limitantes de la secuencia del gen E2a del serotipo de adenovirus que no sea el serotipo de SEQ ID NO: 1. Además, la designación de

i. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o

- ii. supresión de una región de secuencia correspondiente a 1696-3492 bp; o
- iii. supresión de una región de secuencia correspondiente a 1696-2321 bp; o
- iv. supresión de una región de secuencia correspondiente a 2400-2634 bp; o
- v. supresión de una región de secuencia correspondiente a 2839-3492 bp; o
- vi. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp; o
- vii. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2839-3492 bp; o
- viii. supresiones de regiones de secuencia correspondientes a 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO: 1 corresponde a la designación de
- i. supresiones de regiones de secuencia correspondientes a 1534-2036 bp, 2115-2340 bp, 2545-3198 bp; o
- ii. supresión de una región de secuencia correspondiente a 1534-3198 bp; o
- iii. supresión de una región de secuencia correspondiente a 1534-2036 bp; o
- iv. supresión de una región de secuencia correspondiente a 2115-2340 bp; o
- v. supresión de una región de secuencia correspondiente a 2545-3198 bp; o
- vi. supresiones de regiones de secuencia correspondientes a 1534-2036 bp, 2115-2340 bp; o
- vii. supresiones de regiones de secuencia correspondientes a 1534-2036 bp, 2545-3198 bp; o
- viii. supresiones de regiones de secuencia correspondientes a 2115-2340 bp, 2545-3198 bp en la secuencia de SEQ ID NO: 19 (véase Figura 2).

El gen E2a de adenovirus que comprende una modificación seleccionada del siguiente grupo:

- i. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o
- ii. supresión de una región de secuencia correspondiente a 1696-3492 bp; o
- iii. supresión de una región de secuencia correspondiente a 1696-2321 bp; o
- iv. supresión de una región de secuencia correspondiente a 2400-2634 bp; o
- v. supresión de una región de secuencia correspondiente a 2839-3492 bp; o

o vi. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp;

o vii. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2839-3492 bp;

viii. supresiones de regiones de secuencia correspondientes a 2400-2634 bp, 2839-3492 bp en la variante optimizada de codón de la secuencia SEQ ID NO: 1 también puede codificarse mediante una amplia gama de secuencias de nucleótidos distintivas debido a la redundancia del código genético. Las personas versadas en la técnica tienen amplio conocimiento que los habilita a crear estas secuencias alternativas que codifican una y a las mismas secuencias de aminoácidos. Estas variantes de secuencia de nucleótidos están dentro del alcance de la presente invención.

En algunas realizaciones de la invención, la variante optimizada de codón de la secuencia de SEQ ID NO: 1 es la secuencia de SEQ ID NO: 10. La secuencia con SEQ ID NO: 1 y su variante optimizada de codón con la secuencia de SEQ ID NO: 10 poseen 92% de identidad. La variante optimizada de codón de la secuencia SEQ ID NO: 1 con la secuencia de SEQ ID NO: 10 se provee solamente a fines ilustrativos y no deberán interpretarse como variantes optimizadas de codón limitantes de la secuencia de SEQ ID NO: 1.

De esta manera, el ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende una modificación seleccionada del siguiente grupo:

i. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o

ii. supresión de una región de secuencia correspondiente a 1696-3492 bp; o

iii. supresión de una región de secuencia correspondiente a 1696-2321 bp; o

iv. supresión de una región de secuencia correspondiente a 2400-2634 bp; o

v. supresión de una región de secuencia correspondiente a 2839-3492 bp; o

o vi. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp;

o vii. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2839-3492 bp;

viii. supresiones de regiones de secuencia correspondientes a 2400-2634 bp, 2839-3492 bp en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% también es parte de la presente invención.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp, 2400-2634 bp y 2839-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de

identidad de al menos 63%, incluye la secuencia de SEQ ID NO: 2 o una secuencia homóloga a la secuencia de SEQ ID NO: 2 con identidad de al menos 62%.

De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 2 con un grado de identidad de al menos 62% en una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 2, o una variante optimizada de codón de la secuencia de SEQ ID NO: 2.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus que no sea el serotipo con SEQ ID NO: 2, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp, 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO: 1, es la secuencia de SEQ ID NO: 20. La secuencia de SEQ ID NO: 20 es un ácido nucleico que comprende una secuencia modificada del gen E2a del serotipo de adenovirus 54, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1534-2036 bp, 2115-2340 bp, 2545-3198 bp en la secuencia de SEQ ID NO: 19. Las secuencias del gen E2a del serotipo de adenovirus con SEQ ID NO: 2 y el serotipo con SEQ ID NO: 20 poseen 62% de identidad. Una variante de la secuencia modificada del gen E2a de un serotipo de adenovirus, que no sea el serotipo con SEQ ID NO: 2, con la secuencia de SEQ ID NO: 20 se provee solamente a fines ilustrativos y no deberá interpretarse como limitante de las variantes de la secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 2.

En algunas realizaciones de la invención, la variante optimizada de codón de la secuencia de SEQ ID NO: 2 es la secuencia de SEQ ID NO: 11. La secuencia con SEQ ID NO: 2 y su variante optimizada de codón con la secuencia de SEQ ID NO: 11 poseen 88% de identidad. La variante optimizada de codón de la secuencia SEQ ID NO: 2 con la secuencia de SEQ ID NO: 11 se provee solamente a fines ilustrativos y no deberá interpretarse como limitante de las variantes optimizadas de codón de la secuencia de SEQ ID NO: 2.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende una supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% incluye la secuencia de SEQ ID NO: 3 o una secuencia homóloga a la secuencia de SEQ ID NO: 3 con identidad de al menos 61%.

De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 3 con un grado de identidad de al menos 61% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 3, o una variante optimizada de cordón de la secuencia de SEQ ID NO: 3.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus que no sea el serotipo con SEQ ID NO: 3, que comprende una supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1, es la secuencia de SEQ ID NO: 21. La secuencia de SEQ ID NO: 21 es un ácido nucleico que comprende una secuencia modificada del gen E2a del serotipo de

adenovirus 54, que comprende una supresión de una región de secuencia correspondiente a la región de 1534-3198 bp en la secuencia de SEQ ID NO: 19. Las secuencias del gen E2a del serotipo de adenovirus con SEQ ID NO: 3 y el serotipo con SEQ ID NO: 21 tienen 61% de identidad. Una variante de la secuencia modificada del gen E2a de un serotipo de adenovirus, que no sea el serotipo con SEQ ID NO: 3, con la secuencia de SEQ ID NO: 21 se provee solamente a fines ilustrativos y no deberá interpretarse que limita las variantes de la secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 3.

En algunas realizaciones de la invención, la variante optimizada de codón de la secuencia de SEQ ID NO: 3 es la secuencia de SEQ ID NO: 12. La secuencia con SEQ ID NO: 3 y su variante optimizada de codón con la secuencia de SEQ ID NO: 12 poseen 87% de identidad. La variante optimizada de codón de la secuencia de SEQ ID NO: 3 con la secuencia de SEQ ID NO: 12 se provee solamente a fines ilustrativos y no deberá interpretarse como limitante de las variantes optimizadas en codón de la secuencia de SEQ ID NO: 3.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende una supresión de una región de secuencia correspondiente a la región de 1696-2321 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% incluye la secuencia de SEQ ID NO: 4 o una secuencia homóloga a la secuencia de SEQ ID NO: 4 con identidad de al menos 64%.

De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 4 con identidad de al menos 64% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 4, o una variante optimizada de codón de la secuencia de SEQ ID NO: 4.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus que no sea el serotipo con SEQ ID NO: 4, que comprende una supresión de una región de secuencia correspondiente a la región de 1696-2321 bp en la secuencia de SEQ ID NO: 1 es la secuencia de SEQ ID NO: 22. La secuencia de SEQ ID NO: 22 es un ácido nucleico que comprende una secuencia modificada del gen E2a del serotipo de adenovirus 54, que comprende una supresión de una región de secuencia correspondiente a la región de 1534-2036 bp en la secuencia de SEQ ID NO: 19. Las secuencias del gen E2a del serotipo de adenovirus con SEQ ID NO: 4 y serotipo con SEQ ID NO: 22 poseen 64% de identidad. Una variante de la secuencia modificada del gen E2a del serotipo de adenovirus, que no sea el serotipo con SEQ ID NO: 4, con la secuencia de SEQ ID NO: 22 se provee solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes de la secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 4.

En algunas realizaciones de la invención, la variante optimizada de codón de la secuencia de SEQ ID NO: 4 es la secuencia de SEQ ID NO: 13. La secuencia con SEQ ID NO: 4 y su variante optimizada de codón con la secuencia de SEQ ID NO: 13 poseen 91% de identidad. La variante optimizada de codón de la secuencia con SEQ ID NO: 4 con la secuencia de SEQ ID NO: 13 se

proporciona solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes optimizadas de codón de la secuencia de SEQ ID NO: 4.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende una supresión de una región de secuencia correspondiente a la región de 2400-2634 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% incluye la secuencia de SEQ ID NO: 5 o una secuencia homóloga a la secuencia de SEQ ID NO: 5 con identidad de al menos 64%.

De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 5 con identidad de al menos 64% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 5, o una variante optimizada de codón de la secuencia de SEQ ID NO: 5.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus que no sea el serotipo con SEQ ID NO: 5, que comprende una supresión de una región de secuencia correspondiente a la región de 2400-2634 bp en la secuencia de SEQ ID NO: 1 es la secuencia de SEQ ID NO: 23. La secuencia de SEQ ID NO: 23 es un ácido nucleico que comprende una secuencia modificada del gen E2a del serotipo de adenovirus 54, que comprende una supresión de una región de secuencia correspondiente a la región de 2115-2340 bp en la secuencia de SEQ ID NO: 19. Las secuencias del gen E2a del serotipo de adenovirus con SEQ ID NO: 5 y serotipo con SEQ ID NO: 23 tienen 64% de identidad. Una variante de la secuencia modificada del gen E2a de un serotipo de adenovirus, que no sea el serotipo con SEQ ID NO: 5, con la secuencia de SEQ ID NO: 23, se proporciona solamente a fines ilustrativo y no debería interpretarse como limitante de las variantes de la secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 5.

En algunas realizaciones de la invención, la variante optimizada de codón de la secuencia de SEQ ID NO: 5 es la secuencia de SEQ ID NO: 14. La secuencia con SEQ ID NO: 5 y su variante optimizada de codón con la secuencia de SEQ ID NO: 14 poseen 91,5% de identidad. La variante optimizada de codón de la secuencia de SEQ ID NO: 5 con la secuencia de SEQ ID NO: 14 se proporciona solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes optimizadas de codón de la secuencia de SEQ ID NO: 5.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende una supresión de una región de secuencia correspondiente a la región de 2839-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% incluye la secuencia de SEQ ID NO: 6 o una secuencia homóloga a la secuencia de SEQ ID NO: 6 con identidad de al menos 61%.

De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 6 con identidad de al menos 61% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 6, o una variante optimizada de codón de la secuencia de SEQ ID NO: 6.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus que no sea el serotipo con SEQ ID NO: 6, que comprende una supresión de una región de secuencia correspondiente a la región de 2839-3492 bp en la secuencia de SEQ ID NO: 1 es la secuencia de SEQ ID NO: 24. La secuencia de SEQ ID NO: 24 es un ácido nucleico que comprende una secuencia modificada del gen E2a del serotipo de adenovirus 54, que comprende una supresión de una región de secuencia correspondiente a la región de 2545-3198 bp en la secuencia de SEQ ID NO: 19. Las secuencias del gen E2a del serotipo de adenovirus con SEQ ID NO: 6 y serotipo con SEQ ID NO: 24 poseen 61% de identidad. Una variante de la secuencia modificada del gen E2a de un serotipo de adenovirus, que no sea el serotipo con SEQ ID NO: 6, con la secuencia de SEQ ID NO: 24, se provee solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes de la secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 6.

En algunas realizaciones de la invención, la variante optimizada de codón de la secuencia de SEQ ID NO: 6 es la secuencia de SEQ ID NO: 15. La secuencia con SEQ ID NO: 6 y su variante optimizada de codón con la secuencia de SEQ ID NO: 15 poseen 91% de identidad. La variante optimizada de codón de la secuencia de SEQ ID NO: 6 con la secuencia de SEQ ID NO: 15 se provee solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes optimizadas de codón de la secuencia de SEQ ID NO: 6.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende supresiones de regiones de secuencias correspondientes a las regiones de 1696-2321 bp y 2400-2634 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% incluye la secuencia de SEQ ID NO: 7 o una secuencia homóloga a la secuencia de SEQ ID NO: 7 con identidad de al menos 63%.

De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 7 con identidad de al menos 63% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 7, o una variante optimizada de codón de la secuencia de SEQ ID NO: 7.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus que no sea el serotipo con SEQ ID NO: 7, que comprende supresiones de las regiones de secuencia correspondientes a las regiones de 1696-2321 bp y 2400-2634 bp en la secuencia de SEQ ID NO: 1 es la secuencia de SEQ ID NO: 25. La secuencia de SEQ ID NO: 25 es un ácido nucleico que comprende una secuencia modificada del gen E2a del serotipo de adenovirus 54, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1534-2036 bp, 2115-2340 bp en la secuencia de SEQ ID NO: 19. Las secuencias del gen E2a del serotipo de adenovirus con SEQ ID NO: 7 y serotipo con SEQ ID NO: 25 poseen 63% de identidad. Una variante de la secuencia modificada del gen E2a de un serotipo de adenovirus, que no sea el serotipo de SEQ ID NO: 7, con la secuencia de SEQ ID NO: 25 se provee solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes de las secuencias del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 7.

En algunas realizaciones de la invención, la variante optimizada de codón de la secuencia de SEQ ID NO: 7 es la secuencia de SEQ ID NO: 16. La secuencia con SEQ ID NO: 7 y su variante optimizada de codón con la secuencia de SEQ ID NO: 16 poseen 90% de identidad. La variante optimizada de codón de la secuencia de SEQ ID NO: 7 con la secuencia de SEQ ID NO: 16 se provee solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes optimizadas de codón de la secuencia de SEQ ID NO: 7.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp y 2839-3492 en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% incluye la secuencia de SEQ ID NO: 8 o una secuencia homóloga a la secuencia de SEQ ID NO: 8 con identidad de al menos 62%.

De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 8 con identidad de al menos 62% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 8, o una variante optimizada de codón de la secuencia de SEQ ID NO: 8.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus que no sea el serotipo con SEQ ID NO: 8, que comprende supresiones de las regiones de secuencia correspondientes a las regiones de 1696-2321 bp y 2839-3492 bp en la secuencia de SEQ ID NO: 1 es la secuencia de SEQ ID NO: 26. La secuencia de SEQ ID NO: 26 es un ácido nucleico que comprende una secuencia modificada del gen E2a del serotipo de adenovirus 54, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1534-2036 bp y 2545-3198 en la secuencia de SEQ ID NO: 19. Las secuencias del gen E2a del serotipo de adenovirus con SEQ ID NO: 8 y serotipo con SEQ ID NO: 26 poseen 62% de identidad. Una variante de la secuencia modificada del gen E2a de un serotipo de adenovirus, que no sea el serotipo de SEQ ID NO: 8, con la secuencia de SEQ ID NO: 28 se provee solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes de la secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 8.

En algunas realizaciones de la invención, la variante optimizada de codón de la secuencia de SEQ ID NO: 8 es la secuencia de SEQ ID NO: 17. La secuencia con SEQ ID NO: 8 y su variante optimizada de codón con la secuencia de SEQ ID NO: 17 poseen 89% de identidad. La variante optimizada de codón de la secuencia de SEQ ID NO: 8 con la secuencia de SEQ ID NO: 17 se provee solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes optimizadas de codón de la secuencia de SEQ ID NO: 8.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 2400-2634 bp y 2839-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% incluye la secuencia de SEQ ID NO: 9 o una secuencia homóloga a la secuencia de SEQ ID NO: 9 con identidad de al menos 61%.

De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 9 con identidad de al menos 61% es una respectiva secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 9, o una variante optimizada de codón de la secuencia de SEQ ID NO: 9.

En algunas realizaciones de la invención, un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus que no sea el serotipo con SEQ ID NO: 9, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 2400-2634 bp y 2839-3492 en la secuencia de SEQ ID NO: 1 es la secuencia de SEQ ID NO: 27. La secuencia de SEQ ID NO: 27 es un ácido nucleico que comprende una secuencia modificada del gen E2a del serotipo de adenovirus 54, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 2115-2340 bp, 2545-3198 bp en la secuencia de SEQ ID NO: 19. Las secuencias del gen E2a del serotipo de adenovirus con SEQ ID NO: 9 y serotipo con SEQ ID NO: 27 poseen 61% de identidad. Una variante de la secuencia modificada del gen E2a de un serotipo de adenovirus, que no sea el serotipo con SEQ ID NO: 9, con la secuencia de SEQ ID NO: 27, se provee solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes de la secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 9.

En algunas realizaciones de la invención, la variante optimizada de codón de la secuencia de SEQ ID NO: 9 es la secuencia de SEQ ID NO: 18. La secuencia con SEQ ID NO: 9 y su variante optimizada de codón con la secuencia de SEQ ID NO: 18 poseen 90% de identidad. La variante optimizada de codón de la secuencia de SEQ ID NO: 9 con la secuencia de SEQ ID NO: 18 se provee solamente a fines ilustrativos y no debería interpretarse como limitante de las variantes optimizadas de codón de la secuencia de SEQ ID NO: 9.

En un aspecto, la presente invención se relaciona con un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende a su vez la secuencia de SEQ ID NO: 2 o una secuencia homóloga a la secuencia de SEQ ID NO: 2 con identidad de al menos 62%. De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 2 con identidad de al menos 62% es una respectiva secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 2, o una variante optimizada de codón de la secuencia de SEQ ID NO: 2.

En un aspecto, la presente invención se relaciona con un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende a su vez la secuencia de SEQ ID NO: 3 o una secuencia homóloga a la secuencia de SEQ ID NO: 3 con identidad de al menos 61%. De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 3 con identidad de al menos 61% es una respectiva secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 3, o una variante optimizada de codón de la secuencia de SEQ ID NO: 3.

En un aspecto, la presente invención se relaciona con un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende a su vez la secuencia de SEQ ID NO: 4 o una secuencia homóloga a la secuencia de SEQ ID NO: 4 con identidad de al menos 64%. De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO:

4 con identidad de al menos 64% es una respectiva secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 4, o una variante optimizada de codón de la secuencia de SEQ ID NO: 4.

En un aspecto, la presente invención se relaciona con un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende a su vez la secuencia de SEQ ID NO: 5 o una secuencia homóloga a la secuencia de SEQ ID NO: 5 con identidad de al menos 64%. De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 5 con identidad de al menos 64% es una respectiva secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 5, o una variante optimizada de codón de la secuencia de SEQ ID NO: 5.

En un aspecto, la presente invención se relaciona con un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende a su vez la secuencia de SEQ ID NO: 6 o una secuencia homóloga a la secuencia de SEQ ID NO: 6 con identidad de al menos 61%. De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 6 con identidad de al menos 61% es una respectiva secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 6, o una variante optimizada de codón de la secuencia de SEQ ID NO: 6.

En un aspecto, la presente invención se relaciona con un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende a su vez la secuencia de SEQ ID NO: 7 o una secuencia homóloga a la secuencia de SEQ ID NO: 7 con identidad de al menos 63%. De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 7 con identidad de al menos 63% es una respectiva secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 7, o una variante optimizada de codón de la secuencia de SEQ ID NO: 7.

En un aspecto, la presente invención se relaciona con un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende a su vez la secuencia de SEQ ID NO: 8 o una secuencia homóloga a la secuencia de SEQ ID NO: 8 con identidad de al menos 62%. De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 8 con identidad de al menos 62% es una respectiva secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 8, o una variante optimizada de codón de la secuencia de SEQ ID NO: 8.

En un aspecto, la presente invención se relaciona con un ácido nucleico que comprende una secuencia modificada del gen E2a de adenovirus, que comprende a su vez la secuencia de SEQ ID NO: 9 o una secuencia homóloga a la secuencia de SEQ ID NO: 9 con identidad de al menos 61%. De acuerdo con la presente invención, la secuencia homóloga a la secuencia de SEQ ID NO: 9 con identidad de al menos 61% es una respectiva secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 9, o una variante optimizada de codón de la secuencia de SEQ ID NO: 9.

En algunas realizaciones, la presente invención es un ácido nucleico aislado.

Una molécula de ácido nucleico “aislado” es la que se identifica y separa de al menos una impureza de molécula de ácido nucleico. Una molécula de ácido nucleico aislado es diferente de la forma o conjunto en que se encuentra en condiciones naturales. De esta manera, una molécula de ácido nucleico es diferente a una molécula de ácido nucleico que existe en las células en condiciones naturales.

Las secuencias de SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, que representan secuencias modificadas del gen E2a de adenovirus con supresiones de regiones de secuencia correspondientes a las regiones de

- i. 1696-2321 bp, 2400-2634 bp, 2839-3492 bp (SEQ ID NO: 2),
- ii. 1969-3492 bp (SEQ ID NO: 3),
- iii. 1696-2321 bp (SEQ ID NO: 4),
- iv. 2400-2634 bp (SEQ ID NO: 5),
- v. 2839-3492 bp (SEQ ID NO: 6),
- vi. 1696-2321 bp, 2400-2634 bp (SEQ ID NO: 7),
- vii. 1696-2321 bp, 2839-3492 bp (SEQ ID NO: 8),
- viii. 2400-2634 bp, 2839-3492 bp (SEQ ID NO: 9).

en la secuencia de SEQ ID NO: 1, respectivamente, se proveen en este caso y más adelante solamente a fines ilustrativos. Como se describiera previamente, las personas versadas en la técnica comprenderán que por secuencias con una modificación seleccionada del siguiente grupo:

- i. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o
- ii. supresión de una región de secuencia correspondiente a 1696-3492 bp; o
- iii. supresión de una región de secuencia correspondiente a 1696-2321 bp; o
- iv. supresión de una región de secuencia correspondiente a 2400-2634 bp; o
- v. supresión de una región de secuencia correspondiente a 2839-3492 bp; o
- vi. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp; o
- vii. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2839-3492 bp; o
- viii. supresiones de regiones de secuencia correspondientes a 2400-2634 bp, 2839-3492 bp

en la secuencia de SEQ ID NO: 1 se puede interpretar cualquier secuencia cercana a SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9.

Las secuencias optimizadas de codón de SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, que representan secuencias modificadas del gen E2a de adenovirus de regiones de secuencia correspondientes a las regiones de

- i. 1696-2321 bp, 2400-2634 bp, 2839-3492 bp (SEQ ID NO: 11),
- ii. 1969-3492 bp (SEQ ID NO: 12),
- iii. 1696-2321 bp (SEQ ID NO: 13),
- iv. 2400-2634 bp (SEQ ID NO: 14),
- v. 2839-3492 bp (SEQ ID NO: 15),
- vi. 1696-2321 bp, 2400-2634 bp (SEQ ID NO: 16),
- vii. 1696-2321 bp, 2839-3492 bp (SEQ ID NO: 17),
- viii. 2400-2634 bp, 2839-3492 bp (SEQ ID NO: 18).

en la secuencia optimizada de codón de SEQ ID NO: 10, respectivamente, se proveen en este caso y más adelante solamente a fines ilustrativos. Como se describiera previamente, las personas versadas en la técnica comprenderán que por secuencias con una modificación seleccionada del siguiente grupo:

- i. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o
- ii. supresión de una región de secuencia correspondiente a 1696-3492 bp; o
- iii. supresión de una región de secuencia correspondiente a 1696-2321 bp; o
- iv. supresión de una región de secuencia correspondiente a 2400-2634 bp; o
- v. supresión de una región de secuencia correspondiente a 2839-3492 bp; o
- vi. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp; o
- vii. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2839-3492 bp; o
- viii. supresiones de regiones de secuencia correspondientes a 2400-2634 bp, 2839-3492 bp

en la secuencia de SEQ ID NO: 10 se puede interpretar cualquier secuencia cercana a SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18.

Las secuencias de SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, que representan secuencias modificadas del serotipo 54 del gen E2a de adenovirus con supresiones de regiones de secuencia correspondientes a las regiones de

- i. 1534-2036 bp, 2115-2340 bp, 2545-3198 bp (SEQ ID NO: 20),
- ii. 1534-3198 bp (SEQ ID NO: 21),
- iii. 1534-2036 bp (SEQ ID NO: 22),
- iv. 2115-2340 bp (SEQ ID NO: 23),
- v. 2545-3198 bp (SEQ ID NO: 24),
- vi. 1534-2036 bp, 2115-2340 bp (SEQ ID NO: 25),
- vii. 1534-2036 bp, 2545-3198 bp (SEQ ID NO: 26),
- viii. 2115-2340 bp, 2545-3198 bp (SEQ ID NO: 27).

en la secuencia de SEQ ID NO: 19, respectivamente, se proveen en este caso y más adelante solamente a fines ilustrativos. Como se describiera previamente, las personas versadas en la técnica comprenderán que por secuencias con una modificación seleccionada del siguiente grupo:

- i. supresiones de regiones de secuencia correspondientes a 1534-2036 bp, 2115-2340 bp, 2545-3198 bp; o
- ii. supresión de una región de secuencia correspondiente a 1534-3198 bp; o
- iii. supresión de una región de secuencia correspondiente a 1534-2036 bp; o
- iv. supresión de una región de secuencia correspondiente a 2115-2340 bp; o
- v. supresión de una región de secuencia correspondiente a 2545-3198 bp; o
- vi. supresiones de regiones de secuencia correspondientes a 1534-2036 bp, 2115-2340 bp;
o
- vii. supresiones de regiones de secuencia correspondientes a 1534-2036 bp, 2545-3198 bp;
o
- viii. supresiones de regiones de secuencia correspondientes a 2115-2340 bp, 2545-3198 bp

en la secuencia de SEQ ID NO: 19 se puede interpretar cualquier secuencia cercana a SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27.

Todos los ácidos nucleicos, que comprenden una secuencia modificada del gen E2a de adenovirus, de acuerdo con la invención decodifican la proteína E2a de adenovirus.

En algunas realizaciones, los ácidos nucleicos, que comprenden una secuencia modificada del gen E2a de adenovirus, de acuerdo con la invención se utilizan para producir vectores con base en virus adeno-asociados.

Plásmido ayudante

En un aspecto, la presente invención se relaciona con un plásmido ayudante para producir vectores con base en virus adeno-asociados que comprenden un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende una modificación seleccionada del siguiente grupo:

- i. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o
- ii. supresión de una región de secuencia correspondiente a 1696-3492 bp; o
- iii. supresión de una región de secuencia correspondiente a 1696-2321 bp; o
- iv. supresión de una región de secuencia correspondiente a 2400-2634 bp; o
- v. supresión de una región de secuencia correspondiente a 2839-3492 bp; o
- vi. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2400-2634 bp; o
- vii. supresiones de regiones de secuencia correspondientes a 1696-2321 bp, 2839-3492 bp; o
- viii. supresiones de regiones de secuencia correspondientes a 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO:1 o en una secuencia homóloga a la secuencia de SEQ ID NO:1 con un grado de identidad de al menos 63%.

Los términos “plásmido ayudante” y “pAyudante” se utilizan de forma indistinta en el presente documento. En la presente invención, “plásmido ayudante”, “pAyudante” se refieren a un plásmido con funciones necesarias para producir vectores con base en AAV.

El término “virus adeno-asociado” se refiere a un virus de la familia de Parvoviridae. La organización genómica de todos los serotipos de AAV reconocidos es muy similar. El genoma de AAV es una molécula de ADN de una sola hebra y lineal que es menor a 5000 nucleótidos (nt) de longitud. Las repeticiones terminales invertidas (ITRs) flanquean las secuencias de nucleótidos de codificación única de replicación de proteínas no estructurales (Rep) y proteínas estructurales (Cap). El gen Cap codifica las proteínas VP (VP 1, VP2, y VP3) que forman la cápside. Los nucleótidos terminales 145 son auto-complementarios y se organizan de forma tal que puede formarse un dúplex intramolecular enérgicamente estable que cree una forma de T. Estas estructuras en forma de horquilla funcionan como un origen de la replicación del ADN del virus, y sirven como cebadores para el complejo de polimerasa de ADN celular. Luego de la infección del tipo salvaje de AAV (wtAAV) en las células mamíferas, los genes Rep (por ejemplo, Rep78 y Rep52) se expresan con el uso del promotor P5 y el promotor P19, respectivamente, y ambas

proteínas Rep poseen cierta función en la replicación del genoma viral. Un evento de empalme en el marco abierto de lectura (Rep ORF) da como resultado la expresión realmente de cuatro proteínas Rep (por ejemplo, Rep78, Rep68, Rep52, y Rep40). No obstante, se ha demostrado que el mRNA no empalmado que codifica las proteínas Rep78 y Rep52 es suficiente para la producción del vector de AAV en células mamíferas (véase US20160032254).

Los términos “virus adeno-asociado”, “AAV” se utilizan de forma indistinta en el presente documento.

Los términos “vector con base en virus adeno-asociado”, “vector con base en AAV”, “rAAV” se utilizan de forma indistinta en el presente documento.

En algunas realizaciones de la invención, el plásmido ayudante para producir vectores con base en virus adeno-asociados comprende un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp, 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, y comprende la secuencia de SEQ ID NO: 2 o una secuencia homóloga a la secuencia de SEQ ID NO: 2 con identidad de al menos 62%, respectivamente. Como se describiera previamente, la secuencia homóloga a la secuencia de SEQ ID NO: 2 con un grado de identidad de al menos 62% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 2, o una variante optimizada de codón de la secuencia de SEQ ID NO: 2.

En algunas realizaciones de la invención, el plásmido ayudante para producir vectores con base en virus adeno-asociados comprende un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende una supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, y comprende la secuencia de SEQ ID NO: 3 o una secuencia homóloga a la secuencia de SEQ ID NO: 3 con identidad de al menos 61%, respectivamente. Como se describiera previamente, la secuencia homóloga a la secuencia de SEQ ID NO: 3 con un grado de identidad de al menos 61% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 3, o una variante optimizada de codón de la secuencia de SEQ ID NO: 3.

En algunas realizaciones de la invención, el plásmido ayudante para producir vectores con base en virus adeno-asociados comprende un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende una supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, y comprende la secuencia de SEQ ID NO: 4 o una secuencia homóloga a la secuencia de SEQ ID NO: 4 con identidad de al menos 64%, respectivamente. Como se describiera previamente, la secuencia homóloga a la secuencia de SEQ ID NO: 4 con identidad de al menos 64% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 4, o una variante optimizada de codón de la secuencia de SEQ ID NO: 4.

En algunas realizaciones de la invención, el plásmido ayudante para producir vectores con base en virus adeno-asociados comprende un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende una supresión de una región de secuencia correspondiente a la región de 2400-2634 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, y comprende la secuencia de SEQ ID NO: 5 o una secuencia homóloga a la secuencia de SEQ ID NO: 5 con identidad de al menos 64%, respectivamente. Como se describiera previamente, la secuencia homóloga a la secuencia de SEQ ID NO: 5 con identidad de al menos 64% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 5, o una variante optimizada de codón de la secuencia de SEQ ID NO: 5.

En algunas realizaciones de la invención, el plásmido ayudante para producir vectores con base en virus adeno-asociados comprende un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende una supresión de una región de secuencia correspondiente a la región de 2839-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, y comprende la secuencia de SEQ ID NO: 6 o una secuencia homóloga a la secuencia de SEQ ID NO: 6 con identidad de al menos 61%, respectivamente. Como se describiera previamente, la secuencia homóloga a la secuencia de SEQ ID NO: 6 con identidad de al menos 61% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 6, o una variante optimizada de codón de la secuencia de SEQ ID NO: 6.

En algunas realizaciones de la invención, el plásmido ayudante para producir vectores con base en virus adeno-asociados comprende un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp y 2400-2634 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, y comprende la secuencia de SEQ ID NO: 7 o una secuencia homóloga a la secuencia de SEQ ID NO: 7 con identidad de al menos 63%, respectivamente. Como se describiera previamente, la secuencia homóloga a la secuencia de SEQ ID NO: 7 con un grado de identidad de al menos 63% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 7, o una variante optimizada de codón de la secuencia de SEQ ID NO: 7.

En algunas realizaciones de la invención, el plásmido ayudante para producir vectores con base en virus adeno-asociados comprende un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp y 2839-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, y comprende la secuencia de SEQ ID NO: 8 o una secuencia homóloga a la secuencia de SEQ ID NO: 8 con identidad de al menos 62%, respectivamente. Como se describiera previamente, la secuencia homóloga a la secuencia de SEQ ID NO: 8 con identidad de al menos 62% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 8, o una variante optimizada de codón de la secuencia de SEQ ID NO: 8.

En algunas realizaciones de la invención, el plásmido ayudante para producir vectores con base en virus adeno-asociados comprende un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende supresiones de regiones de secuencia correspondientes a las regiones de 2400-2634 bp y 2839-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, y comprende la secuencia de SEQ ID NO: 9 o una secuencia homóloga a la secuencia de SEQ ID NO: 9 con identidad de al menos 61%, respectivamente. Como se describiera previamente, la secuencia homóloga a la secuencia de SEQ ID NO: 9 con identidad de al menos 61% es una secuencia del gen E2a de un serotipo de adenovirus que no sea el serotipo con SEQ ID NO: 9, o una variante optimizada de codón de la secuencia de SEQ ID NO: 9.

En un aspecto, la presente invención se relaciona con un plásmido ayudante para producir vectores con base en virus adeno-asociados que comprenden un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende la secuencia de SEQ ID NO: 2 o una secuencia homóloga a la secuencia de SEQ ID NO: 2 con identidad de al menos 62%.

En un aspecto, la presente invención se relaciona con un plásmido ayudante para producir vectores con base en virus adeno-asociados que comprenden un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende la secuencia de SEQ ID NO: 3 o una secuencia homóloga a la secuencia de SEQ ID NO: 3 con identidad de al menos 61%.

En un aspecto, la presente invención se relaciona con un plásmido ayudante para producir vectores con base en virus adeno-asociados que comprenden un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende la secuencia de SEQ ID NO: 4 o una secuencia homóloga a la secuencia de SEQ ID NO: 4 con identidad de al menos 64%.

En un aspecto, la presente invención se relaciona con un plásmido ayudante para producir vectores con base en virus adeno-asociados que comprenden un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende la secuencia de SEQ ID NO: 5 o una secuencia homóloga a la secuencia de SEQ ID NO: 5 con identidad de al menos 64%.

En un aspecto, la presente invención se relaciona con un plásmido ayudante para producir vectores con base en virus adeno-asociados que comprenden un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende la secuencia de SEQ ID NO: 6 o una secuencia homóloga a la secuencia de SEQ ID NO: 6 con identidad de al menos 61%.

En un aspecto, la presente invención se relaciona con un plásmido ayudante para producir vectores con base en virus adeno-asociados que comprenden un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende la secuencia de SEQ ID NO: 7 o una secuencia homóloga a la secuencia de SEQ ID NO: 7 con identidad de al menos 63%.

En un aspecto, la presente invención se relaciona con un plásmido ayudante para producir vectores con base en virus adeno-asociados que comprenden un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende la secuencia de SEQ ID NO: 8 o una secuencia homóloga a la secuencia de SEQ ID NO: 8 con identidad de al menos 62%.

En un aspecto, la presente invención se relaciona con un plásmido ayudante para producir vectores con base en virus adeno-asociados que comprenden un ácido nucleico que a su vez comprende una secuencia modificada del gen E2a de adenovirus, que a su vez comprende la secuencia de SEQ ID NO: 9 o una secuencia homóloga a la secuencia de SEQ ID NO: 9 con identidad de al menos 61%.

En algunas realizaciones de la invención, el plásmido ayudante para producir vectores con base en virus adeno-relacionados comprende un ácido nucleico que comprende una secuencia de gen E2a de adenovirus (cualquiera de las variantes anteriores), ácido nucleico que comprende una secuencia del gen ARN VA de adenovirus, un ácido nucleico que comprende una secuencia del gen E4 de adenovirus.

En algunas realizaciones de la invención, la secuencia del gen ARN VA de adenovirus comprende la SEQ ID NO: 28.

En algunas realizaciones de la invención, la secuencia del gen E4 de adenovirus es una secuencia que ocurre en la naturaleza de longitud completa del gen E4, secuencia de E4 orf6/7, o secuencia de E4 orf6.

En algunas realizaciones de la invención, la secuencia que ocurre en la naturaleza de longitud completa del gen E4 de adenovirus comprende la SEQ ID NO: 29.

En algunas realizaciones de la invención, la secuencia E4 orf6/7 de adenovirus es una secuencia que codifica SEQ ID NO: 30.

En algunas realizaciones de la invención, la secuencia E4 orf6 de adenovirus es una secuencia que codifica SEQ ID NO: 31.

Uso

En un aspecto, la presente invención se relaciona con el uso de un ácido nucleico que comprende cualquiera de las variantes de la secuencia modificada del gen E2a de adenovirus descriptas en el presente invención, o de un plásmido ayudante que comprende un ácido nucleico que comprende cualquiera de las variantes de la secuencia modificada del gen E2a de adenovirus descrito en el presente documento para producir vectores en función de virus adeno-asociados.

En un aspecto, la presente invención se relaciona con el uso de un ácido nucleico que comprende cualquiera de las variantes de la secuencia modificada del gen E2a de adenovirus descriptas en el presente documento para producir vectores con base en virus adeno-asociados.

En un aspecto, la presente invención se relaciona con el uso de un plásmido ayudante que comprende un ácido nucleico que comprende cualquiera de las variantes de la secuencia

modificada del gen E2a de adenovirus descritas en el presente documento para producir vectores con base en virus adeno-asociados.

En algunas variantes del uso, el plásmido ayudante que comprende un ácido nucleico que comprende cualquiera de las variantes de la secuencia modificada del gen E2a de adenovirus descritas en el presente documento se utiliza para producir vectores con base en virus adeno-asociados mediante la transfección de una célula anfitriona con su plásmido ayudante, y también con plásmidos que comprenden secuencias del gen de interés y genes que codifican proteínas de AAV.

Los términos “producción” y “obtención” se utilizan de forma indistinta en el presente documento.

En algunas realizaciones de la presente invención, pueden integrarse variantes de la secuencia modificada del gen E2a de adenovirus de acuerdo con la invención al genoma de una célula anfitriona para producir una línea celular de empaquetamiento o una línea celular productora para vectores virales. Estos enfoques para producir vectores virales son reconocidos por las personas versadas en la técnica y se han descrito en la literatura (Imre Kovesdi et al., Adenoviral producer cells, *Viruses*. 2010 Aug;2(8):1681-1703).

En un aspecto, la presente invención se relaciona con el uso de variantes de la secuencia modificada del gen E2a de adenovirus de acuerdo con la invención para producir vectores virales recombinantes con base en adenovirus mediante la integración de una parte del genoma de adenovirus, que comprende cualquiera de las variantes de la secuencia modificada del gen E2a de adenovirus descritas en el presente documento en el genoma de una célula anfitriona, y también mediante la integración, en el genoma de un adenovirus recombinante incapaz de replicarse, del gen de interés.

En algunos aspectos, la presente invención se relaciona con el uso de variantes de la secuencia modificada del gen E2a de adenovirus de acuerdo con la invención para producir vectores virales recombinantes con base en adenovirus mediante la integración, en el genoma de la célula anfitriona, de una parte del genoma de adenovirus, que no comprende la secuencia del gen E2a, y también mediante la integración, en el genoma del adenovirus recombinante incapaz de replicarse, del gen de interés y ácido nucleico que comprende cualquiera de las variantes de la secuencia modificada del gen E2a de adenovirus de acuerdo con la invención. Estos enfoques son reconocidos por las personas versadas en la técnica y se han divulgado en la literatura (M I Gorziglia et al., Generation of an adenovirus vector lacking E1, e2a, E3, and all of E4 except open reading frame 3, *J Virol*. 1999 Jul;73(7):6048-55; H Zhou et al., A new vector system with inducible E2a cell line for production of higher titer and safer adenoviral vectors, *Virology*. 2000 Sep 30;275(2):348-57).

En un aspecto, la presente invención se relaciona con el uso de variantes de la secuencia modificada del gen E2a de adenovirus de acuerdo con la invención para producir vectores virales recombinantes con base en virus adeno-asociados mediante la integración, en el genoma de la célula anfitriona, de una parte del genoma de adenovirus que comprende un ácido nucleico que comprende cualquiera de las variantes de la secuencia modificada del gen E2a de adenovirus de

acuerdo con la invención, y/o una parte del genoma de virus adeno-asociado y/o de un casete de expresión con el gen de interés e ITR. Estos enfoques son reconocidos por las personas versadas en la técnica y se han descrito en la literatura (Satoki Nakamura et al., Development of packaging cell lines for generation of adeno-associated virus vectors by lentiviral gene transfer of trans-complementary components, 2004 Oct;73(4): 285-94, véase Resumen; Nagarathinam Selvaraj et al., Detailed Protocol for the Novel and Scalable Viral Vector Upstream Process for AAV Gene Therapy Manufacturing, Hum Gene Ther. August 2021; 32(15-16): 850–861).

Breve descripción de los dibujos

La **Figura 1** es un gráfico que muestra la productividad de la generación de serotipos de rAAV 2, 5, 6, y 9 en células HEK293 luego de la transfección de todas las secuencias necesarias para la generación de rAAV, inclusive las secuencias que codifican el gen E2a y elementos regulatorios correspondientes a los siguientes ácidos nucleicos: a un ácido nucleico que comprende la secuencia natural del gen E2a de adenovirus (SEQ ID NO: 1), un ácido nucleico que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp, 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO: 1 (SEQ ID NO: 2), un ácido nucleico que comprende una supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1 (SEQ ID NO: 3).

Productividad de generación de rAAV luego de la transfección de tres plásmidos:

1 – al usar el plásmido de pAyudante que comprende las secuencias que ocurren en la naturaleza de los genes E4, ARN VA y E2a, en la que la secuencia que ocurre en la naturaleza del gen E2a de adenovirus comprende la SEQ ID NO: 1, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep y Cap del serotipo 2; el plásmido que codifica un transgén con base en un anticuerpo a VEGF/C5,

2 – al usar el plásmido de pAyudante que comprende secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y una secuencia modificada del gen E2a, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp, 2400-2634 bp, 2839-3492 bp. n. en la secuencia de SEQ ID NO: 1 e incluye la secuencia de SEQ ID NO: 2, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep y Cap del serotipo 2; el plásmido que codifica un transgén con base en un anticuerpo a VEGF/C5,

3 – al usar el plásmido de pAyudante que comprende secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y una secuencia modificada del gen E2a, que comprende una supresión de una región de secuencia correspondiente a la región de 1696-3492 bp. n. en la secuencia de SEQ ID NO: 1 e incluye la secuencia de SEQ ID NO: 3, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep y Cap del serotipo 2; el plásmido que codifica un transgén con base en un anticuerpo a VEGF/C5,

4 – al usar el plásmido de pAyudante que comprende las secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y E2a, en la que la secuencia que ocurre en la naturaleza del gen E2a de adenovirus comprende la SEQ ID NO: 1, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep de serotipo 2 y Cap de serotipo 5; el plásmido que codifica un transgén con base en una proteína factor IX de coagulación,

5 – al usar el plásmido de pAyudante que comprende las secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y una secuencia modificada del gen E2a, que comprende supresiones de las regiones de secuencia correspondientes a las regiones de 1696-2321 bp, 2400-2634 bp, 2839-3492 bp. n. en la secuencia de SEQ ID NO: 1 e incluye la secuencia de SEQ ID NO: 2, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep de serotipo

2 y Cap de serotipo 5; el plásmido que codifica un transgén con base en una proteína factor IX de coagulación,

6 – al usar el plásmido de pAyudante que comprende las secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y una secuencia modificada del gen E2a, que comprende una supresión de una región de secuencia correspondiente a las regiones de 1696-3492 bp. n. en la secuencia de SEQ ID NO: 1 e incluye la secuencia de SEQ ID NO: 3, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep de serotipo 2 y Cap de serotipo 5; el plásmido que codifica un transgén con base en una proteína factor IX de coagulación,

7 – al usar el plásmido de pAyudante que comprende las secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y E2a, en la que la secuencia que ocurre en la naturaleza del gen E2a de adenovirus comprende la SEQ ID NO: 1, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep de serotipo 2 y Cap de serotipo 6; el plásmido que codifica un transgén con base en una proteína factor VIII de coagulación;

8 – al usar el plásmido de pAyudante que comprende las secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y una secuencia modificada del gen E2a, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp, 2400-2634 bp, 2839-3492 bp. n. en la secuencia de SEQ ID NO: 1 e incluye la secuencia de SEQ ID NO: 2, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep de serotipo 2 y Cap de serotipo 6; el plásmido que codifica un transgén con base en una proteína factor VIII de coagulación,

9 – al usar el plásmido de pAyudante que comprende una secuencia que ocurre en la naturaleza de los genes E4, ARN VA, y una secuencia modificada del gen E2a, que comprende una supresión de una región de secuencia correspondiente a las regiones de 1696-3492 bp. n. en la secuencia de SEQ ID NO: 1 e incluye la secuencia de SEQ ID NO: 3, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep de serotipo 2 y Cap de serotipo 6; el plásmido que codifica un transgén con base en una proteína factor VIII de coagulación,

10 – al usar el plásmido de pAyudante que comprende secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y E2a, en la que la secuencia que ocurre en la naturaleza del gen E2a de adenovirus comprende SEQ ID NO: 1, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep de serotipo 2 y Cap de serotipo 9; el plásmido que codifica un transgén con base en una proteína de supervivencia de motoneuronas SMN1,

11 – al usar el plásmido de pAyudante que comprende secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y una secuencia modificada del gen E2a, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp, 2400-2634 bp, 2839-3492 bp. n. en la secuencia de SEQ ID NO: 1 e incluye la secuencia de SEQ ID NO: 2, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep de serotipo 2 y Cap de serotipo 9; el plásmido que codifica un transgén con base en una proteína de supervivencia de motoneuronas SMN1,

12 – al usar el plásmido de pAyudante que comprende secuencias que ocurren en la naturaleza de los genes E4, ARN VA, y una secuencia modificada del gen E2a, que comprende una supresión de una región de secuencia correspondiente a las regiones de 1696-3492 bp. n. en la secuencia de SEQ ID NO: 1 e incluye la secuencia de SEQ ID NO: 3, con elementos regulatorios; el plásmido de pRepCap que codifica los genes Rep de serotipo 2 y Cap de serotipo 9; el plásmido que codifica un transgén con base en una proteína de supervivencia de motoneuronas SMN1,

La **Figura 2** es una vista esquemática de una región del genoma del serotipo de adenovirus 2 que comprende la secuencia del gen E2a con elementos regulatorios (A; corresponde a N° de Ref. NCBI AC_000007.1 de 22393 bp a 27214 bp) y una región del genoma del serotipo de adenovirus 54 que comprende la secuencia del gen E2a con elementos regulatorios (B; corresponde a la Secuencia de Referencia NCBI: NC_012959.1 de 21166 bp a 25369 bp) que indican supresiones de las regiones de secuencia descritas en el presente documento y su correspondencia entre sí. Las regiones de supresión de secuencia se enumeran en relación a SEQ ID NO: 1 (para A) y SEQ ID NO: 19 (para B).

Ejemplos (implementación de la invención)

Los siguientes ejemplos se proveen para un mejor entendimiento de la invención. Estos ejemplos solamente sirven a fines ilustrativos y no deberán interpretarse como limitantes del alcance de la invención de ninguna manera.

A pesar de que se ha descrito la invención previa en cierto detalle a modo ilustrativo y ejemplificativo a fines de ser claros en su entendimiento, las personas versadas en la técnica común comprenderán rápidamente que, a la luz de las enseñanzas de la presente invención, se podrán realizar ciertos cambios y modificaciones a ella sin salirse del espíritu o alcanza de las realizaciones adjuntas.

Materiales y procedimientos generales

Técnicas de ADN recombinante

Se utilizaron procedimientos estándar para manipular el ADN tal como se describe en Sambrook, J. et al, Molecular cloning: A laboratory manual; Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 2012. Los reactivos biológicos moleculares se utilizaron de acuerdo con los protocolos del fabricante. En pocas palabras, se produjo ADN de plásmido para una mayor manipulación en células de E. coli cultivadas en presión antibiótica selectiva para que los plásmidos no se perdieran en la población celular. Se aisló el ADN de plásmido de las células con el uso de kits comerciales, se midió la concentración, y se utilizó para clonar mediante tratamiento de endonucleasa de restricción o amplificación PCR. Los fragmentos de ADN se ligaron entre sí con el uso de ligasas y se transformaron en células bacterianas para la selección de clones y mayor producción. Todas las construcciones genéticas resultantes fueron confirmadas por patrones de restricción y secuenciación Sanger completa.

Determinación de secuencia de ADN

Se determinaron las secuencias de ADN mediante la secuenciación Sanger. Las secuencias de ADN y proteína se analizaron y se procesaron los datos de secuencia en SnapGene Viewer 4.2 o posterior para la creación de secuencia, mapeado, análisis, anotación, e ilustración.

Acción de cultivar los cultivos celulares

Los experimentos utilizaron líneas celulares de HEK293 (Clon 293 de Riñón Embrionario Humano). Las células suspendidas de HEK293 utilizadas para producir AAV se cultivaron en condiciones estándar a 37°C y 5% CO₂ en un medio de cultivo completo sin FBS y antibiótico. Se evaluó la viabilidad celular con el uso de colorante Azul de Tripano y cámaras de conteo celular descartables con el uso de un contador automático Countess II.

Generación de partículas virales de vectores recombinantes de AAV

Para producir partículas virales de AAV recombinantes que comprenden un transgén con base en varias proteínas, se utilizaron células productoras de HEK293 que se transfectaron con 3 plásmidos:

- Un plásmido que comprende un casete de expresión de AAV para la expresión del transgén;
- Un plásmido que comprende genes Cap de serotipo 2, 5, 6, o 9 y genes Rep de serotipo AAV2. Cada gen, con el uso de marcos de lectura alternativos, codifica varios productos de proteína;
- Un plásmido que comprende genes Ad2 de adenovirus necesarios para el armado y empaquetamiento de cápsides de AAV: E2a, E4 y ARN VA.

El título de las partículas virales se determinó mediante PCR cuantitativo con cebadores y una muestra que fue específica para la región del genoma viral recombinante y se expresó como el número de copia de genomas virales por 1 ml.

Análisis de datos estadísticos

Los resultados indican un valor promedio $\pm$ desviación estándar (SD), análisis unidireccional de varianza (ANOVA) seguido de varias comparaciones por pares de Dunnett se utilizó para comparar los resultados del experimento, y se determinó que fueron estadísticamente significativos.

Ejemplo 1. Evaluación de la productividad de generación de vectores virales recombinantes con base en AAV al utilizar secuencia de gen E2a de adenovirus modificada y plásmido ayudante con base en ellos

A la secuencia que ocurre en la naturaleza del gen E2a de adenovirus con elementos regulatorios que comprenden SEQ ID NO: 1 o una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63% se introdujeron las supresiones

i. 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o

ii. 1696-3492 bp

Se produjeron variantes de secuencias modificadas del gen E2a de adenovirus con elementos regulatorios, que se analizaron posteriormente en cuanto al nivel de productividad al generar vectores con base en los serotipos de virus adeno-asociados rAAV2, rAAV5, rAAV6 y rAAV9.

Para mostrar un nivel mayor de productividad al generar rAAV, se produjeron plásmidos que codifican construcciones genéticas a las que se apunta, en particular que comprenden

- a. una secuencia de gen E2a con elementos regulatorios que incluye SEQ ID NO: 1 o una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%,
- b. una secuencia modificada del gen E2a con elementos regulatorios, que comprende supresiones de regiones de secuencia correspondientes a las regiones de 1696-2321 bp, 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, e incluye la secuencia de SEQ ID NO: 2 o una secuencia homóloga a la secuencia de SEQ ID NO: 2 con identidad de al menos 62%, y
- c. una secuencia del gen E2a con elementos regulatorios, que comprende una supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1 o en una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, e incluye la secuencia de SEQ ID NO: 3 o una secuencia homóloga a la secuencia de SEQ ID NO: 3 con identidad de al menos 61%.

También se produjeron los plásmidos restantes necesarios para generar vectores virales

i. con base en el serotipo 2 de virus adeno-asociado que lleva una secuencia de nucleótido que codifica un anticuerpo a VEGF/C5,

ii. con base en el serotipo 5 de virus adeno-asociado que lleva una secuencia de nucleótido que codifica la proteína factor IX(FIX) de coagulación,

iii. con base en el serotipo 6 de virus adeno-asociado que lleva una secuencia de nucleótido que codifica la proteína factor VIII(FVIII) de coagulación, y

iv. con base en el serotipo 9 de virus adeno-asociado que lleva una secuencia de nucleótido que codifica la proteína de supervivencia de motoneuronas (SMN1).

Se ha demostrado que las secuencias modificadas desarrolladas que codifican el gen E2a y sus elementos regulatorios llevaron a una mayor productividad en la generación de rAAV si se compara con una secuencia inicial análoga que codifica el gen E2a y elementos regulatorios con la SEQ ID NO: 1 (Figura 1):

- al generar serotipo 2 de rAAV con un transgén con base en un anticuerpo a VEGF/C5, por 2,41 veces en el caso del uso de la secuencia del gen E2a de adenovirus con supresiones de las regiones de secuencia correspondientes a las regiones 1696-2321 bp, 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO: 1, que incluye SEQ

- ID NO: 2, y por 3,02 veces en el caso del uso de la secuencia del gen E2a de adenovirus con supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1, que incluye SEQ ID NO: 3;
- al generar serotipo 5 de rAAV con un transgén con base en una proteína factor IX de coagulación, por 1,45 veces en el caso del uso de la secuencia del gen E2a de adenovirus con supresiones de las regiones de secuencia correspondientes a las regiones 1696-2321 bp, 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO: 1, que incluye SEQ ID NO: 2, y por 1,69 veces en el caso del uso de la secuencia del gen E2a de adenovirus con supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1, que incluye SEQ ID NO: 3;
 - al generar serotipo 6 de rAAV con un transgén con base en una proteína factor VII de coagulación, por 1,81 veces en el caso del uso de la secuencia del gen E2a de adenovirus con supresiones de las regiones de secuencia correspondientes a las regiones 1696-2321 bp, 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO: 1, que incluye SEQ ID NO: 2, y por 1,83 veces en el caso del uso de la secuencia del gen E2a de adenovirus con supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1, que incluye SEQ ID NO: 3;
 - al generar serotipo 9 de rAAV con un transgén con base en una proteína de supervivencia de motoneuronas (SMN1), por 1,8 veces en el caso del uso de la secuencia del gen E2a de adenovirus con supresiones de las regiones de secuencia correspondientes a las regiones 1696-2321 bp, 2400-2634 bp, 2839-3492 bp en la secuencia de SEQ ID NO: 1, que incluye SEQ ID NO: 2, y por 1,97 veces en el caso del uso de la secuencia del gen E2a de adenovirus con supresión de una región de secuencia correspondiente a la región de 1696-3492 bp en la secuencia de SEQ ID NO: 1, que incluye SEQ ID NO: 3;

Se obtuvieron datos similares luego de comparar la variante optimizada de codón del gen E2a de adenovirus con SEQ ID NO: 10 con sus variantes modificadas, que comprenden las supresiones de i. 1696-2321 bp, 2400-2634 bp, 2839-3492 bp (SEQ ID NO: 11) o ii. 1696-3492 bp (SEQ ID NO: 12), y luego de comparar la secuencia del gen E2a de adenovirus de un serotipo que no sea SEQ ID NO: 1 con SEQ ID NO: 19 con sus variantes modificadas, que comprenden las supresiones de i. 1534-2036 bp, 2115-2340 bp, 2545-3198 bp (SEQ ID NO: 20) o ii. 1534-3198 bp (SEQ ID NO: 21).

De esta manera, los ácidos nucleicos producidos que codifican el gen E2a con sus elementos regulatorios y comprenden supresiones de regiones de secuencias correspondientes a las regiones de

- i. 1696-2321 bp, 2400-2634 bp, 2839-3492 bp; o
- ii. 1696-3492 bp

en la secuencia de SEQ ID NO: 1 o en la secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%, se provee una mayor productividad de vectores virales recombinantes con base en AAV de varios serotipos en comparación con la secuencia natural del gen E2a de adenovirus que comprende SEQ ID NO: 1 o con una secuencia homóloga a la secuencia de SEQ ID NO: 1 con un grado de identidad de al menos 63%.

