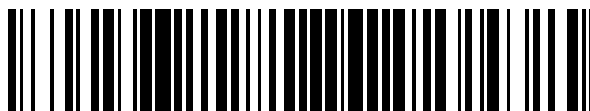


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 714 535**

51 Int. Cl.:

C07K 14/00 (2006.01)

C12N 15/86 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **10.10.2014** **PCT/US2014/060163**

87 Fecha y número de publicación internacional: **16.04.2015** **WO15054653**

96 Fecha de presentación y número de la solicitud europea: **10.10.2014** **E 14789476 (0)**

97 Fecha y número de publicación de la concesión europea: **05.12.2018** **EP 3060575**

54 Título: **Métodos para predecir secuencias de virus ancestrales y usos de los mismos**

30 Prioridad:

11.10.2013 US 201361889827 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
28.05.2019

73 Titular/es:

**MASSACHUSETTS EYE & EAR INFIRMARY
(50.0%)**

243 Charles Street

Boston, MA 02114, US y

SCHEPENS EYE RESEARCH INSTITUTE (50.0%)

72 Inventor/es:

**VANDENBERGHE, LUK H. y
ZINN, ERIC**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 714 535 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

DESCRIPCIÓN

Métodos para predecir secuencias de virus ancestrales y usos de los mismos

5 CAMPO TÉCNICO

Generalmente esta divulgación se refiere a virus.

10 ANTECEDENTES

Eludir y evitar una respuesta inmunitaria neutralizante o tóxica contra un vector de terapia génica es un desafío importante con todos los tipos de vectores de transferencia génica. Hasta la fecha, la transferencia génica se logra de manera más eficiente utilizando vectores basados en virus que circulan en seres humanos y animales, por ejemplo, adenovirus y virus adenoasociados (AAV). Sin embargo, si los sujetos han sido infectados de forma natural con un virus, un tratamiento posterior con un vector basado en ese virus conlleva un aumento de los riesgos de seguridad y una menor eficiencia de la transferencia génica debido a las respuestas inmunitarias celulares y humorales. Los antígenos de cápside son los principales responsables de la inmunidad innata y/o adaptativa hacia las partículas víricas, sin embargo, los polipéptidos codificados por genes víricos también pueden ser inmunógenos.

20 Calcedo et al., 2009 divulga la seroprevalencia mundial hacia varios tipos de AAV diferentes.

RESUMEN

La invención se define por las reivindicaciones.

25 Esta divulgación describe métodos para predecir y sintetizar secuencias víricas ancestrales o porciones de las mismas, y también describe partículas víricas que contienen tales secuencias víricas ancestrales. Los métodos descritos en el presente documento se aplicaron al virus adenoasociados (AAV); por lo tanto, esta divulgación describe secuencias de AAV ancestrales predichas y partículas de virus de AAV que contienen tales secuencias de AAV ancestrales. Esta divulgación también describe la reducida seroprevalencia presentada por partículas víricas que contienen secuencias ancestrales en relación con partículas víricas que contienen secuencias contemporáneas. En un aspecto, esta divulgación incluye polipéptidos de cápside de virus adenoasociado (AAV), por ejemplo, polipéptidos de cápside de AAV sintéticos y/o artificiales, que tienen una secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15 y 17. En algunas implementaciones, los polipéptidos de cápside de AAV o partículas víricas que comprenden los polipéptidos de cápside de AAV presentan una menor seroprevalencia que polipéptido de cápside de AAV2 o una partícula vírica que comprende un polipéptido de cápside de AAV2, y los polipéptidos de cápside de AAV o partículas víricas que comprenden los polipéptidos de cápside de AAV presentan una seroprevalencia igual o inferior a la de un polipéptido de cápside de AAV8 o una partícula vírica que comprende un polipéptido de cápside de AAV8. En algunas formas de realización, los polipéptidos de cápside de AAV o partículas víricas que comprenden los polipéptidos de cápside de AAV se neutralizan en una menor medida por suero humano que un polipéptido de cápside de AAV2 o una partícula vírica que comprende un polipéptido de cápside de AAV2, y los polipéptidos de cápside de AAV o partículas víricas que comprenden los polipéptidos de cápside de AAV se neutralizan en una menor medida o similar por suero humano que un polipéptido de cápside de AAV8 o una partícula vírica que comprende un polipéptido de cápside de AAV8. En algunas formas de realización, los polipéptidos de cápside de AAV están purificados. Los polipéptidos de cápside de AAV proporcionados en el presente documento pueden codificarse por una secuencia de ácidos nucleicos seleccionada del grupo que consiste en las SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, y 18.

50 En un aspecto, la divulgación proporciona moléculas de ácido nucleico, por ejemplo, moléculas de ácido nucleico sintéticas y/o artificiales, que codifican un polipéptido de cápside de virus adenoasociado (AAV) que tiene una secuencia de ácido nucleico seleccionada del grupo que consiste en las SEQ ID NOs: 2, 4, 6, 8, 10, 12, 14, 16, y 18. También se proporcionan vectores que incluyen tal ácido nucleico, y una célula huésped que incluye tal vector.

55 En otro aspecto, la divulgación proporciona partículas víricas purificadas que incluyen un polipéptido de cápside de AAV descrito en el presente documento. En algunas formas de realización, las partículas víricas incluyen un transgén.

En otros aspectos, la divulgación proporciona polipéptidos de cápside de virus adenoasociado (AAV), por ejemplo, polipéptidos de cápside de AAV sintéticos y/o artificiales, que tienen al menos el 95% (por ejemplo, 97, 98, 99 o el

100%) de identidad de secuencia con una secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NOs: 19, 20, 21, 22, 23, 24, 25 y 26. En algunas formas de realización, los polipéptidos de cápside de AAV o partículas víricas que comprenden el polipéptido de cápside de AAV presentan una menor seroprevalencia que un polipéptido de cápside de AAV2 o una partícula vírica que comprende un polipéptido de cápside de AAV2, y el polipéptido de cápside de AAV o una partícula vírica que comprende el polipéptido de cápside de AAV presentan una seroprevalencia aproximadamente igual o inferior a la de un polipéptido de cápside de AAV8 o una partícula vírica que comprende un polipéptido de cápside de AAV8. En algunas formas de realización, los polipéptidos de cápside de AAV o partículas víricas que comprenden el polipéptido de cápside de AAV se neutralizan en una menor medida por suero humano que un polipéptido de cápside de AAV2 o una partícula vírica que comprende un polipéptido de cápside de AAV2, y el polipéptido de cápside de AAV o una partícula vírica que comprende el polipéptido de cápside de AAV se neutraliza en una menor medida o similar por suero humano que un polipéptido de cápside de AAV8 o una partícula vírica que comprende un polipéptido de cápside de AAV8. En algunas formas de realización, los polipéptidos de cápside de AAV están purificados.

En otro aspecto, los polipéptidos de cápside de AAV descritos en el presente documento pueden codificarse por secuencias de ácidos nucleicos como se describe en el presente documento. En una implementación, la divulgación proporciona moléculas de ácido nucleico que codifican un polipéptido de cápside de virus adenoasociado (AAV), en las que las moléculas de ácido nucleico tienen al menos el 95% (por ejemplo, 97, 98, 99 o el 100%) de identidad de secuencia con una secuencia de ácido nucleico como se muestra en el presente documento. La divulgación también proporciona vectores que incluyen dichas moléculas de ácido nucleico, como son las células huésped que incluyen tal vector.

En un aspecto, la divulgación proporciona partículas víricas que incluyen al menos uno de los polipéptidos de cápside de AAV descritos en el presente documento. En algunas formas de realización, las partículas víricas incluyen un transgén.

En ciertos aspectos, la divulgación proporciona una partícula vírica como se describe en el presente documento para su uso en un método de administración a un sujeto que necesita transferencia génica o vacunación. En algunas formas de realización, las partículas víricas presentan menos seroprevalencia que una partícula vírica AAV2. En algunas formas de realización, las partículas víricas presentan aproximadamente la misma o menos seroprevalencia que una partícula vírica AAV8. En algunas formas de realización, las partículas víricas se neutralizan en una menor medida por suero humano que una partícula vírica de AAV2, y las partículas víricas de AAV se neutralizan en una menor medida o similar por suero humano que una partícula de virus de AAV8.

En un aspecto, la divulgación proporciona un antígeno diana unido operativamente a un polipéptido de cápside de AAV como se describe en el presente documento para su uso en un método de administración a un sujeto que necesita vacunación. En algunas formas de realización, los polipéptidos de cápside de AAV presentan menos seroprevalencia que un polipéptido de cápside de AAV2. En algunas formas de realización, el polipéptido de cápside de AAV presenta aproximadamente la misma o menos seroprevalencia que un polipéptido de cápside de AAV8. En algunas formas de realización, los polipéptidos de cápside de AAV se neutralizan en una menor medida por suero humano que un polipéptido de cápside de AAV2, y el polipéptido de cápside de AAV se neutraliza en una menor medida o similar en suero humano que un polipéptido de cápside de AAV8.

En otro aspecto, la divulgación proporciona métodos *in silicio* para predecir una secuencia de un virus ancestral o una porción del mismo. Dichos métodos típicamente incluyen proporcionar secuencias de nucleótidos o aminoácidos de una pluralidad de virus contemporáneos o porciones de los mismos; alinear las secuencias utilizando un algoritmo de alineamiento múltiple de secuencias (MSA); modelar la evolución para obtener una filogenia ancestral predicha de la pluralidad de virus contemporáneos o porciones de los mismos; estimar, en un nodo filogenético de la filogenia ancestral predicha, la probabilidad evolutiva de un nucleótido particular o un residuo de aminoácido en cada posición de la secuencia; y predecir, basándose en la probabilidad estimada en cada posición, una secuencia de un virus ancestral o una porción del mismo.

En algunas formas de realización de la presente divulgación, una o más, o todas, las etapas se realizan usando un procesador informático. En algunas formas de realización de la presente divulgación, el algoritmo MSA usa información filogenética para predecir si un hueco en el alineamiento es resultado de una eliminación o una inserción. En algunas formas de realización de la presente divulgación, el algoritmo MSA es un kit de alineamiento probabilístico (PRANK). En algunas formas de realización de la presente divulgación, el modelo utilizado para modelar la evolución se selecciona utilizando el criterio de información de Akaike (AIC). En algunas formas de realización de la presente divulgación, la filogenia ancestral predicha se obtiene utilizando un modelo JTT con un

modelo de distribución gamma ("G") y un cálculo de frecuencia de π ("F"). En algunas formas de realización de la presente divulgación, la modelación de la etapa de evolución se realiza utilizando un modelo JTT+G+F. En algunas formas de realización de la presente divulgación, los métodos incluyen la síntesis, basada en la secuencia predicha, del virus ancestral o una porción del mismo. En algunas formas de realización de la presente divulgación, los métodos incluyen ensamblar el virus ancestral o una porción del mismo en una partícula de virus ancestral.

En algunas formas de realización, los métodos también incluyen cribar la partícula de virus ancestral para al menos uno de los siguientes: (a) replicación; (b) propiedades de transferencia génica; (c) unión al receptor; o (d) seroprevalencia. En algunas formas de realización, las partículas de virus ancestrales presentan menos seroprevalencia que una partícula vírica ensamblada de al menos una de la pluralidad de virus contemporáneos o porciones de los mismos. En algunas formas de realización, la partícula vírica ancestral se neutraliza en menor medida por suero humano que una partícula vírica ensamblada a partir de al menos uno de la pluralidad de virus contemporáneos o porciones de los mismos. En algunas formas de realización, la pluralidad de virus contemporáneos o porciones de los mismos, pertenecen a una familia seleccionada del grupo que consiste en adenovirus (AV), virus de inmunodeficiencia humana (VIH), retrovirus, lentivirus, virus del herpes simple (VHS), virus vaccinia, virus de la viruela, virus de influenza, virus sincitial de las vías respiratorias, virus parainfluenza y virus espumoso.

Por lo tanto, la presente divulgación proporciona virus ancestrales o porciones de los mismos que muestran una susceptibilidad reducida a la inmunidad preexistente en las poblaciones humanas actuales en comparación con los virus contemporáneos o porciones de los mismos. Generalmente, la susceptibilidad reducida a la inmunidad preexistente mostrada por los virus ancestrales o porciones de los mismos en las poblaciones humanas actuales se refleja como una susceptibilidad reducida a anticuerpos neutralizantes.

A menos que se defina de otro modo, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que entiende comúnmente un experto en la técnica a la que pertenecen los métodos y composiciones de la materia. Aunque se pueden usar métodos y materiales similares o equivalentes a los descritos en el presente documento en la práctica o en el ensayo de los métodos y composiciones de materia, a continuación, se describen los métodos y materiales adecuados. Además, los materiales, métodos y ejemplos son solo ilustrativos y no pretenden ser limitantes.

DESCRIPCIÓN DE LOS DIBUJOS

El archivo de patente o solicitud contiene al menos un dibujo realizado a color. La Oficina proporcionará copias de esta patente o publicación de solicitud de patente con dibujos a color previa solicitud y pago de la tarifa necesaria.

La Figura 1 es un esquema que muestra las relaciones entre las infecciones víricas ancestrales/contemporáneas y la respuesta inmunitaria del huésped ancestral/contemporánea.

Las Figuras 2A a 2D son una serie de esquemas que muestran un ejemplo de un procedimiento de reconstrucción ancestral. Los datos mostrados están extraídos de un conjunto de datos completo y representan los residuos 564-584 (numeración VP1 de AAV2).

Las Figuras 3A a 3C juntas ilustran un árbol filogenético de secuencias contemporáneas de AAV generadas usando los métodos descritos en el presente documento.

Las Figuras 4A y 4B juntas ilustran un alineamiento de polipéptidos de VP1 de AAV ancestrales.

Las Figuras 5A y 5B juntas ilustran un alineamiento de polipéptidos de VP1 de AAV ancestrales funcionales y polipéptidos de VP1 de AAV contemporáneos.

La Figura 6 es un gel electroforético que demuestra que las secuencias de VP1 de AAV ancestrales se transcriben y se someten a corte y empalme alternativamente de una manera similar a la de las secuencias de VP1 de AAV contemporáneas.

La Figura 7 es un gráfico que muestra la actividad de la luciferasa en células HEK293 transducidas con vectores de AAV ancestrales.

La Figura 9 es un gráfico que muestra la comparación de secuencias (% por encima de la diagonal, N.º de diferencias de aa a continuación) entre la biblioteca Anc80 y Anc80L65.

Las Figuras 10A-D son imágenes de resultados experimentales que demuestran que Anc80L65 es capaz de ensamblar y producir partículas de alto título. El panel A muestra que Anc80L65 es capaz de producir rendimientos vectoriales equivalentes a AAV2; el panel B es una imagen TEM de partículas víricas que incluyen Anc80L65; el panel C muestra que las partículas víricas que incluyen Anc80L65 son capaces de producir proteínas VP1, 2 y 3 de cap de AAV basadas en gel SDS-PAGE en condiciones de

desnaturalización; y el panel D muestra una transferencia Western de Anc80L65 utilizando el anticuerpo de cápside AAV, B1.

Las Figuras 11A-C son imágenes de resultados experimentales que demuestran que Anc80L65 es capaz de infectar células *in vitro* en células HEK293 usando GFP como lectura (Panel A) o luciferasa (Panel B) en comparación con los controles de AAV2 y/o AAV8 y también es eficaz en el direccionamiento al hígado después de una inyección IV de AAV que codifica un transgén LacZ nuclear (fila superior, Panel C: hígado), después de la inyección IM directa de un AAV que codifica una GFP (fila central, Panel C: músculo), y después de la inyección subretiniana con un AAV que codifica GFP (fila inferior, Panel C: retina).

Las Figuras 13A y 13B son matrices de identidad de secuencia que producen el uso de MAFFT que muestran las secuencias de aminoácidos de las proteínas VP1 de vectores ancestrales alineadas con las de los AAV existentes representativos (Figura 13A), y las secuencias de aminoácidos de las proteínas VP3 de vectores ancestrales alineadas con las de AAV existentes representativos (Figura 13B).

La Figura 14 es un gráfico que demuestra que los vectores de AAV se produjeron por triplicado en pequeña escala (placas de 6 pocillos). Los virus en bruto se evaluaron mediante qPCR para determinar la producción absoluta de cada vector.

La Figura 15 es una tabla que muestra los títulos de cada vector, promediados y comparados, con los de AAV8.

La Figura 16 son fotografías que muestran los resultados de los experimentos en los que se añadieron 1,9E3 GC/célula de cada vector a las células HEK293 (excepto para Anc126, en cuyo caso se lograron MOI de 2,5E2 - 3,1E2 GC/célula). Sesenta horas más tarde, la infectividad se evaluó utilizando microscopía de fluorescencia.

La Figura 17 es un gráfico que muestra los resultados de experimentos en los que las mismas células de la Figura 16 se lisaron y se analizaron para determinar la expresión de luciferasa. Como en la Figura 16, Anc126 no se controló de título con los otros vectores, sino que varió desde una MOI de 2,5E2 - 3,1E2 GC/celda.

La Figura 18 es una tabla que muestra la luminiscencia de las células transducidas por cada vector, que se promediaron y compararon con las de AAV8.

La Figura 19 es un gráfico que proporciona un resumen de experimentos *in vitro* para determinar la producción relativa y la infectividad de los vectores de AAV ancestrales descritos en el presente documento.

DESCRIPCIÓN DETALLADA

La transferencia génica, ya sea con fines experimentales o terapéuticos, se basa en un vector o sistema vectorial para transferir información genética a las células diana. El vector o sistema de vectores se considera el principal determinante de la eficiencia, especificidad, respuesta del huésped, farmacología y longevidad de la reacción de transferencia génica. Actualmente, la forma más eficiente y eficaz de lograr la transferencia génica es a través del uso de vectores o sistemas de vectores basados en virus que se han hecho defectuosos en la replicación.

Sin embargo, los estudios de seroprevalencia indican que se han expuesto previamente proporciones significativas de poblaciones humanas en todo el mundo (por ejemplo, por infección natural) a un gran número de los virus utilizados actualmente en la transferencia génica y, por lo tanto, albergan inmunidad preexistente. Se sabe que los anticuerpos neutralizantes hacia el vector viral en estos individuos expuestos previamente limitan, a veces significativamente, la extensión de la transferencia génica o incluso redirigen el virus lejos de la diana. Véanse, por ejemplo, Calcedo et al. (2009, J. Infect. Dis., 199:381-90) y Boutin et al. (2010, Human Gene Ther., 21:704-12). Por lo tanto, la presente divulgación se basa en el reconocimiento de que los virus ancestrales o porciones de los mismos presentan una susceptibilidad reducida a la inmunidad preexistente (por ejemplo, una susceptibilidad reducida a anticuerpos neutralizantes) en las poblaciones humanas actuales en comparación con los virus contemporáneos o porciones de los mismos.

La Figura 1 es un esquema que muestra las relaciones entre las infecciones víricas ancestrales y contemporáneas y la respuesta inmunitaria del huésped ancestral y contemporánea. La Figura 1 muestra cómo los AAV ancestrales pueden ser refractarios a la inmunidad preexistente contemporánea. Se presume que un virus contemporáneo existente (Vc) ha evolucionado a partir de una especie ancestral (Vanc), principalmente bajo presiones evolutivas de la inmunidad del huésped a través de mecanismos de escape inmunológico. Cada una de estas especies, Vanc y Vc, tienen la capacidad de inducir inmunidad adaptativa, incluyendo la inmunidad de linfocitos B y T (Ianc y Ic, respectivamente). Se formuló la hipótesis, y se confirmó en el presente documento, de que la inmunidad inducida por virus contemporáneos no necesariamente reacciona de forma cruzada con una especie vírica ancestral, que puede ser sustancialmente diferente en términos de la composición de epítipo que el virus existente.

Esta divulgación proporciona métodos para predecir la secuencia de un virus ancestral o una porción del mismo. Una o más de las secuencias de virus ancestrales predichas utilizando los métodos descritos en el presente documento se pueden generar y ensamblar en una partícula vírica. Como se demuestra en el presente documento, las partículas víricas ensambladas a partir de secuencias víricas ancestrales predichas pueden presentar menos, a veces significativamente menos, seroprevalencia que las partículas víricas contemporáneas actuales. Por lo tanto, las secuencias de virus ancestrales divulgadas en el presente documento son adecuadas para su uso en vectores o sistemas de vectores para la transferencia génica.

Métodos para predecir y sintetizar una secuencia vírica ancestral

Para predecir una secuencia vírica ancestral, las secuencias de nucleótidos o aminoácidos primero se compilan a partir de una pluralidad de virus contemporáneos o porciones de los mismos. Si bien los métodos descritos en el presente documento se ilustraron utilizando secuencias de cápside de virus adenoasociado (AAV), los mismos métodos se pueden aplicar a otras secuencias de AAV (por ejemplo, el genoma completo, secuencias rep, secuencias ITR) o cualquier otro virus o porción del mismo. Los virus distintos de AAV incluyen, sin limitación, adenovirus (AV), virus de inmunodeficiencia humana (VIH), retrovirus, lentivirus, virus del herpes simple (HSV), sarampión, virus vaccinia, virus de la viruela, virus de la influenza, virus sincitial de las vías respiratorias, virus parainfluenza, virus espumoso, o cualquier otro virus para el cual la inmunidad preexistente se considere un problema.

Se pueden usar secuencias de tan solo dos virus contemporáneos o porciones de los mismos, sin embargo, se entiende que es deseable un mayor número de secuencias de virus contemporáneos o porciones de los mismos para incluir la mayor parte del paisaje de la diversidad de secuencia diaria moderna posible, pero también porque un número mayor de secuencias puede aumentar las capacidades predictivas de los algoritmos descritos y utilizados. Por ejemplo, se pueden usar secuencias de 10 o más virus contemporáneos o porciones de los mismos, secuencias de 50 o más virus contemporáneos o porciones de los mismos, o secuencias de 100 o más virus contemporáneos o porciones de los mismos.

Dichas secuencias se pueden obtener, por ejemplo, de cualquier número de bases de datos públicas incluyendo, sin limitación, GenBank, UniProt, EMBL, Colaboración internacional de bases de datos de secuencias de nucleótidos (INSDC), o el Archivo europeo de nucleótidos. Adicionalmente, o como alternativa, dichas secuencias pueden obtenerse de una base de datos que sea específica de un organismo en particular (por ejemplo, la base de datos de VIH). Las secuencias contemporáneas pueden corresponder a todo el genoma, o solo se puede usar una porción del genoma, tal como, sin limitación, secuencias que codifican uno o más componentes de la cápside vírica, la proteína de replicación, o las secuencias ITR.

A continuación, las secuencias contemporáneas se alinean utilizando un algoritmo de alineamiento múltiple de secuencias (MSA). La Figura 2(a) es un esquema que muestra un alineamiento múltiple de secuencias. Los algoritmos MSA se conocen bien en la técnica y generalmente están diseñados para aplicarse a conjuntos de datos de diferentes tamaños y diferentes entradas (por ejemplo, ácido nucleico o proteína), y para alinear las secuencias de una manera particular (por ejemplo, programación dinámica, progresiva, heurística) y aplicar diferentes esquemas de puntuación en el alineamiento (por ejemplo, basado en matriz o basado en consistencia, por ejemplo, mínima entropía, suma de pares, matriz de similitud, puntuaciones de huecos). Los algoritmos MSA ya conocidos incluyen, por ejemplo, ClustalW (Thompson et al., 1994, Nuc. Acids Res., 22:4673-90), Kalign (Lassmann et al., 2006, Nuc. Acids Res., 34:W596-99), MAFFT (Katoh et al., 2005, Nuc. Acids Res., 33:511-8), MUSCLE (Edgar, 2004, BMC Bioinform., 5:113), y T-Coffee (Notredame et al., 2000, J. Mol. Biol., 302:205-17).

Como se describe en el presente documento, una de las características principales al seleccionar un algoritmo MSA para su uso en los métodos descritos en el presente documento es la manera en que el algoritmo trata un hueco en el alineamiento. A los huecos en un alineamiento de secuencias se les puede asignar un valor de penalización que es dependiente o independiente del tamaño del hueco. En los presentes métodos, se prefiere que el algoritmo MSA utilizado en los métodos descritos en el presente documento aplique información filogenética para predecir si un hueco en el alineamiento es el resultado de una eliminación o una inserción en oposición a un tratamiento sesgado, no filogenético de huecos debido, por ejemplo, a inserciones y/o eliminaciones. Un método adecuado para tratar los huecos en los alineamientos y el análisis evolutivo se describe en Loytynoja y Goldman, 2008, Science, 320:1632-5, y los algoritmos disponibles comercialmente que aplican huecos en los alineamientos de una manera que es adecuada para su uso en los métodos descritos en el presente documento son un kit de alineamiento probabilístico (PRANK; Goldman Group Software; Loytynoja y Goldman, 2005, PNAS USA, 102:10557-62), y variaciones del algoritmo PRANK.

Luego se aplica un modelo evolutivo al alineamiento resultante para obtener una filogenia ancestral predicha (véase la Figura 2(b)). Existe una serie de modelos evolutivos disponibles en la técnica, cada uno de los cuales aplica matrices de tasas de reemplazo ligeramente diferentes para los aminoácidos. Sin limitación, los algoritmos para aplicar modelos de evolución incluyen los modelos de Dayhoff (por ejemplo, PAM120, PAM160, PAM250; Dayhoff et al., 1978, In Atlas of Protein Sequence and Structure (ed. Dayhoff), págs. 345-52, National Biomedical Research Foundation, Washington D.C.), el modelo JTT (Jones et al., 1992, Comp. Appl. Biosci., 8:275-82), el modelo WAG (Whelan y Goldman, 2001, Mol. Biol. Evol., 18:691-9), y los modelos Blosom (por ejemplo, Blosom45, Blosom62, Blosom80; Henikoff y Henikoff, 1992, PNAS USA, 89:10915-9).

Además, las restricciones que la estructura y la función imponen en un modelo evolutivo pueden modelarse, por ejemplo, considerando que algunas posiciones son invariantes ("+I"; Reeves, 1992, J. Mol. Evol., 35:17-31), que algunas posiciones experimentan diferentes tasas de cambio ("+G"; Yang, 1993, Mol. Biol. Evol., 10:1396-1401), y/o que las frecuencias de equilibrio de nucleótidos o aminoácidos son las mismas que las del alineamiento ("+F"; Cao et al., 1994, J. Mol. Evol., 39:519-27).

La idoneidad de uno o más modelos de evolución puede evaluarse utilizando el criterio de información de Aikake (AIC; Akaike, 1973, In Second International Symposium on Information Theory, Petrov y Csaki, eds., págs. 267-81, Budapest, Akademiai Kiado), el criterio de información bayesiano (BIC; Schwarz, 1978, Ann. Statist. 6:461-4), o variaciones o combinaciones de los mismos. Además, se puede usar AIC, BIC, o variaciones o combinaciones de los mismos para evaluar la importancia relativa de incluir uno o más parámetros (por ejemplo, las restricciones analizadas anteriormente) en el modelo evolutivo.

Como se explica en la sección de Ejemplos a continuación, ProTest3 (Darriba et al., 2011, Bioinformatics, 27(8):1164-5) se puede usar para determinar, en función del AIC más bajo, que un algoritmo JTT+G+F fue el modelo más adecuado para la evolución de AAV. También se entenderá por un experto en la técnica que un algoritmo JTT+G+F también se puede usar para predecir secuencias víricas ancestrales distintas de los polipéptidos de cápside de AAV, sin embargo, un experto en la técnica también entenderá que, dependiendo del conjunto de datos y la puntuación de aptitud, puede ser más adecuado un modelo diferente de evolución.

Una vez que se ha seleccionado un modelo de evolución y se ha determinado su aptitud, se puede construir un árbol filogenético de las secuencias víricas o porciones de las mismas. La construcción de árboles filogenéticos es conocida en la técnica y típicamente emplea métodos de máxima probabilidad tales como los implementados por PhyML (Guindon y Gascuel, 2003, Systematic Biology, 52:696-704), MOLPHY (Adachi y Hasegawa, 1996, ed. Tokyo Institute of Statistical Mathematics), BioNJ (Gascuel, 1997, Mol. Biol. Evol., 14:685-95), o PHYLIP (Felsenstein, 1973, Systematic Biology, 22:240-9). Un experto en la materia entenderá que es deseable un equilibrio entre la complejidad computacional y la bondad del ajuste en un modelo de sustituciones de aminoácidos.

Si se desea, el árbol filogenético puede evaluarse para determinar su importancia. Se dispone de varios métodos estadísticos y se utilizan de forma rutinaria para evaluar la importancia de un modelo incluyendo, sin limitación, muestreo con reemplazo, jackknife, validación cruzada, pruebas de permutación o combinaciones o variaciones de los mismos. La importancia también puede evaluarse utilizando, por ejemplo, una prueba de razón de verosimilitud aproximada (aLRT; Anisimova y Gascuel, 2006, Systematic Biology, 55:539-52)).

En cualquier nodo filogenético de la filogenia (por ejemplo, un nodo filogenético interior), la secuencia se puede reconstruir estimando la probabilidad evolutiva de un residuo particular de nucleótido o de aminoácido en cada posición de la secuencia (Figura 2(c)). Un nodo filogenético se refiere a un punto de ramificación evolutivo intermedio dentro de la filogenia ancestral predicha. Como se usa en el presente documento, la "probabilidad evolutiva" se refiere a la probabilidad de la presencia de un nucleótido o aminoácido particular en una posición particular basada en un modelo evolutivo en oposición a un modelo que no tiene en cuenta, por ejemplo, un cambio evolutivo en el uso del codón. Los modelos ejemplares que tienen en cuenta la probabilidad evolutiva de un residuo de nucleótido de aminoácido particular en una posición particular pueden estimarse utilizando, por ejemplo, cualquier número de métodos de máxima verosimilitud que incluyen, sin limitación, el análisis filogenético por máxima verosimilitud (PAML; Yang, 1997, Comp. Applic. BioSci., 13:555-6) o análisis filogenético usando parsimonia (PAUP; Sinauer Assoc., Inc., Sunderland, MA).

En base a la probabilidad evolutiva estimada de un residuo de nucleótido o aminoácido particular en cada posición, la secuencia predicha de un virus ancestral o una porción del mismo, puede ensamblarse para formar una secuencia completa o parcial sintética de ácido nucleico o polipéptido. Si se desea, puede calcularse la verosimilitud de que

cualquier residuo se encuentre en un estado determinado en un nodo determinado a lo largo del nodo, y se puede identificar cualquier posición a lo largo de la secuencia que tenga una verosimilitud posterior calculada por debajo de un umbral particular (Figura 2(d)). De esta manera, se puede generar una secuencia de andamiaje ancestral, que puede incluir variaciones en aquellas posiciones que tienen una probabilidad por debajo del umbral particular.

Si la secuencia ancestral que se predice usando los métodos en el presente documento es una secuencia de ácido nucleico, la secuencia puede optimizarse entonces por codón de modo que se pueda traducir de manera eficiente en una secuencia de aminoácidos. Se conocen en la técnica las tablas de uso de codones para diferentes organismos. Opcionalmente, sin embargo, se puede diseñar una tabla de uso de codones basada en una o más secuencias contemporáneas que tiene homología (por ejemplo, al menos un 90% de identidad de secuencia) con la secuencia de andamiaje ancestral, y una secuencia ancestral como se describe en el presente documento puede optimizarse por codones para el uso de codones mamíferos (por ejemplo, un ser humano).

Cualquiera o todas las etapas descritas en el presente documento para predecir una secuencia vírica ancestral se pueden realizar o simular en un ordenador (por ejemplo, *in silicio*) usando un procesador o un microprocesador.

Secuencias de andamiaje de virus adenoasociados (AAV) ancestrales

Los métodos descritos en el presente documento se aplicaron a virus adenoasociados (AAV) utilizando secuencias de cápside contemporáneas (descritas en detalle en los Ejemplos a continuación). El AAV es ampliamente considerado como un vector de transferencia génica terapéutico y un vehículo de vacuna genético, pero presenta una alta seroprevalencia en poblaciones humanas. Usando los métodos descritos en el presente documento, un árbol filogenético se ensambló utilizando secuencias de AAV contemporáneas (véanse las Figuras 3A-3C) y se obtuvieron secuencias de andamiaje ancestrales predichas en el nodo filogenético designado (Tabla 1). Como se usa en el presente documento, una secuencia de andamiaje ancestral se refiere a una secuencia que se construye usando los métodos descritos en el presente documento (por ejemplo, usando probabilidades evolutivas y modelos evolutivos) y no se sabe que hayan existido en la naturaleza. Como se usa en el presente documento, las secuencias de andamiaje ancestrales son diferentes de las secuencias consenso, que se construyen típicamente usando la frecuencia de los nucleótidos o residuos de aminoácidos en una posición particular.

Tabla 1.

Nodo	Polipéptido (SEQ ID NO)	Ácido nucleico (SEQ ID NO)
Anc80	1	2
Anc81	3	4
Anc82	5	6
Anc83	7	8
Anc84	9	10
Anc94	11	12
Anc113	13	14
Anc126	15	16
Anc127	17	18

La secuencia de andamiaje del polipéptido Anc80 se muestra en la SEQ ID NO:1, que está codificada por la secuencia de andamiaje del ácido nucleico Anc80 que se muestra en la SEQ ID NO:2. La secuencia de andamiaje de Anc80 contiene 11 posiciones en las que cualquiera de los dos residuos era probable. Por lo tanto, la secuencia de andamiaje de Anc80 representa 2048 (2^{11}) secuencias diferentes.

Para demostrar la eficacia de los métodos descritos en el presente documento para predecir la secuencia ancestral de un virus o porción del mismo, se generó una biblioteca de las 2048 secuencias ancestrales predichas en el nodo Anc80 de AAV y, como se describe en el presente documento, se demostró que forman partículas víricas viables que presentan menos seroprevalencia, en algunos casos, significativamente menos seroprevalencia, que las partículas víricas ensambladas con polipéptidos de cápside contemporáneos.

Métodos para hacer partículas víricas ancestrales

Después de que se haya obtenido la secuencia ancestral predicha de un virus o porción del mismo, la molécula y/o el polipéptido o polipéptidos de ácido nucleico real pueden generarse, por ejemplo, sintetizarse. Los métodos para generar una molécula o polipéptido de ácido nucleico artificial basado en una secuencia obtenida, por ejemplo, *in*

silicio, se conocen en la técnica e incluyen, por ejemplo, síntesis química o clonación recombinante. Se conocen en la técnica métodos adicionales para generar moléculas o polipéptidos de ácido nucleico y se describen con más detalle a continuación.

Una vez que se ha producido un polipéptido ancestral, o una vez que se ha generado y expresado una molécula de ácido nucleico ancestral para producir un polipéptido ancestral, el polipéptido ancestral puede ensamblarse en una partícula de virus ancestral usando, por ejemplo, una célula huésped empaquetadora. Los componentes de una partícula vírica (por ejemplo, secuencias rep, secuencias cap, secuencias de repetición terminales invertidas (ITR)) pueden introducirse, de forma transitoria o estable, en una célula huésped empaquetadora utilizando uno o más vectores como se describe en el presente documento. Uno o más de los componentes de una partícula vírica se pueden basar en una secuencia ancestral predicha como se describe en el presente documento, mientras que los componentes restantes se pueden basar en secuencias contemporáneas. En algunos casos, toda la partícula vírica puede basarse en secuencias ancestrales predichas.

Dichas partículas víricas ancestrales pueden purificarse usando métodos de rutina. Como se usa en el presente documento, las partículas víricas "purificadas" se refieren a partículas víricas que se eliminan de los componentes en la mezcla en la que se fabricaron, tales como, pero sin limitación, componentes víricos (por ejemplo, secuencias rep, secuencias cap), células huésped empaquetadoras, y partículas víricas ensambladas parcialmente o de forma incompleta.

Una vez ensambladas, las partículas víricas ancestrales pueden cribarse, por ejemplo, para determinar la capacidad de replicación; las propiedades de la transferencia génica; la capacidad de unión al receptor; y/o la seroprevalencia en una población (por ejemplo, una población humana). Determinar si una partícula vírica puede replicarse es rutinario en la técnica y típicamente incluye infectar una célula huésped con una cantidad de partículas víricas y determinar si las partículas víricas aumentan en número con el tiempo. Determinar si una partícula vírica es capaz de realizar la transferencia génica también es rutinario en la técnica e incluye típicamente infectar células huésped con partículas víricas que contienen un transgén (por ejemplo, un transgén detectable tal como un gen indicador, analizado con más detalle más adelante a continuación). Tras la infección y la eliminación del virus, las células huésped pueden evaluarse para determinar la presencia o ausencia del transgén. La determinación de si una partícula de virus se une a su receptor es habitual en la técnica, y dichos métodos pueden realizarse *in vitro* o *in vivo*.

La determinación de la seroprevalencia de una partícula vírica se realiza de forma rutinaria en la técnica y típicamente incluye el uso de un inmunoensayo para determinar la prevalencia de uno o más anticuerpos en muestras (por ejemplo, muestras de sangre) de una población particular de individuos. La seroprevalencia se entiende en la técnica que se refiere a la proporción de sujetos en una población que es seropositiva (es decir, ha sido expuesta a un patógeno o inmunógeno en particular), y se calcula como el número de sujetos en una población que producen un anticuerpo contra un patógeno o inmunógeno particular dividido por el número total de individuos en la población examinada. Los inmunoensayos se conocen bien en la técnica e incluyen, sin limitación, una inmunotransferencia, transferencia Western, inmunoensayos enzimáticos (EIA), inmunoensayo enzimático ligado a enzimas (ELISA), o radioinmunoensayo (RIA). Como se indica en el presente documento, las partículas víricas ancestrales presentan menos seroprevalencia que las partículas víricas contemporáneas (es decir, partículas víricas ensambladas que usan secuencias víricas contemporáneas o porciones de las mismas). Simplemente a modo de ejemplo, véanse Xu et al. (2007, Am. J. Obstet. Gynecol., 196:43.e1-6); Paul et al. (1994, J. Infect. Dis., 169:801-6); Sauerbrei et al. (2011, Eurosurv., 16(44):3); y Sakhría et al. (2013, PLoS Negl. Trop. Dis., 7:e2429), cada uno de los cuales determinó la seroprevalencia para un anticuerpo en particular en una población dada.

Como se describe en el presente documento, las partículas víricas ancestrales se neutralizan por el sistema inmunitario de una persona, por ejemplo, del paciente, en menor medida que las partículas víricas contemporáneas. Se dispone de varios métodos para determinar la extensión de los anticuerpos neutralizantes en una muestra de suero. Por ejemplo, un ensayo de anticuerpos neutralizantes mide el título en el que una muestra experimental contiene una concentración de anticuerpos que neutraliza la infección en un 50% o más en comparación con una muestra de control sin anticuerpo. Véanse, también, Fisher et al. (1997, Nature Med., 3:306-12) y Manning et al. (1998, Human Gene Ther., 9:477-85).

Con respecto a los polipéptidos de cápside de AAV ancestrales ilustrados en el presente documento, la seroprevalencia y/o el grado de neutralización pueden compararse, por ejemplo, con un polipéptido de cápside de AAV8 o partícula vírica que incluye un polipéptido de cápside de AAV8, o un polipéptido de cápside de AAV2 o partícula vírica que incluye un polipéptido de cápside de AAV2. Generalmente, en la técnica se entiende que los

polipéptidos de cápside de AAV8 o partículas víricas presentan una seroprevalencia, y una neutralización resultante, en la población humana que se considera baja, mientras que el polipéptido de cápside de AAV2 o partículas víricas presentan una seroprevalencia y una neutralización resultante, en la población humana que se considera alta. Obviamente, la seroprevalencia particular dependerá de la población examinada, así como los métodos inmunológicos utilizados, pero hay informes de que AAV8 presenta una seroprevalencia de aproximadamente el 22% hasta aproximadamente el 38%, mientras que AAV2 presenta una seroprevalencia de aproximadamente el 43,5% hasta aproximadamente el 72%. Véanse, por ejemplo, Boutin et al., 2010, "Prevalence of serum IgG and neutralizing factors against AAV types 1, 2, 5, 6, 8 and 9 in the healthy population: implications for gene therapy using AAV vectors", Hum. Gene Ther., 21:704-12. Véase, también, Calcedo et al., 2009, J. Infect. Dis., 199:381-90.

Ácido nucleico ancestral predicho y secuencias de polipéptidos ancestrales de virus adenoasociado (AAV)

Se secuenciaron varios clones diferentes de la genoteca que codificaba polipéptidos de cápside ancestrales predichos del nodo Anc80, y las secuencias de aminoácidos de polipéptidos de cápside ancestrales predichos de AAV representativos se muestran en la SEQ ID NO: 19 (Anc80L27); SEQ ID NO: 20 (Anc80L59); SEQ ID NO: 21 (Anc80L60); SEQ ID NO: 22 (Anc80L62); SEQ ID NO: 23 (Anc80L65); SEQ ID NO: 24 (Anc80L33); SEQ ID NO: 25 (Anc80L36); y SEQ ID NO:26 (Anc80L44). Los expertos en la técnica apreciarán que la secuencia de ácido nucleico que codifica cada secuencia de aminoácidos puede determinarse fácilmente.

Además de los polipéptidos de cápside ancestral predichos que tienen las secuencias mostradas en las SEQ ID NOs: 19, 20, 21, 22, 23, 24, 25 o 26, se proporcionan polipéptidos que tienen al menos un 95% de identidad de secuencia (por ejemplo, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99% o un 100% de identidad de secuencia) con los polipéptidos de cápside ancestral predichos que tienen las secuencias que se muestran en las SEQ ID NOs: 19, 20, 21, 22, 23, 24, 25, o 26. De forma similar, se proporcionan moléculas de ácido nucleico que tienen al menos un 95% de identidad de secuencia (por ejemplo, al menos un 96%, al menos un 97%, al menos un 98%, al menos un 99% o un 100% de identidad de secuencia) con las moléculas de ácido nucleico que codifican los polipéptidos de cápside ancestrales (es decir, que tienen al menos un 95% de identidad de secuencia).

En el cálculo del porcentaje de identidad de secuencia, se alinean dos secuencias y se determina el número de coincidencias idénticas de nucleótidos o residuos de aminoácidos entre las dos secuencias. El número de coincidencias idénticas se divide por la longitud de la región alineada (es decir, el número de nucleótidos o residuos de aminoácidos alineados) y se multiplica por 100 para llegar a un porcentaje de valor de identidad de secuencia. Se apreciará que la longitud de la región alineada puede ser una porción de una o ambas secuencias hasta el tamaño de longitud completa de la secuencia más corta. También se apreciará que una única secuencia puede alinearse con más de una secuencia diferente y, por lo tanto, puede tener diferentes valores de porcentaje de identidad de secuencia en cada región alineada.

El alineamiento de dos o más secuencias para determinar el porcentaje de identidad de secuencia puede realizarse utilizando el algoritmo descrito por Altschul et al. (1997, Nucleic Acids Res., 25:3389 3402) como se incorpora en los programas BLAST (herramienta de búsqueda de alineamiento local básico), disponibles en ncbi.nlm.nih.gov en Internet. Se pueden realizar búsquedas de BLAST para determinar el porcentaje de identidad de secuencia entre una secuencia (ácido nucleico o aminoácido) y cualquier otra secuencia o porción de la misma alineada usando el algoritmo de Altschul et al. BLASTN es el programa usado para alinear y comparar la identidad entre secuencias de ácidos nucleicos, mientras que BLASTP es el programa usado para alinear y comparar la identidad entre secuencias de aminoácidos. Cuando se utilizan programas BLAST para calcular el porcentaje de identidad entre una secuencia y otra secuencia, generalmente se usan los parámetros predeterminados de los programas respectivos.

Se muestran alineamientos representativos en las Figuras 4A y 4B y en las Figuras 5A y 5B. Las Figuras 4A y 4B muestran un alineamiento de polipéptidos de cápside de VP1 de AAV ancestrales, designados como Anc80L65 (SEQ ID NO: 23), Anc80L27 (SEQ ID NO: 19), Anc80L33 (SEQ ID NO: 24), Anc80L36 (SEQ ID NO: 25), Anc80L44 (SEQ ID NO: 26), Anc80L59 (SEQ ID NO: 20), Anc80L60 (SEQ ID NO: 21), y Anc80L62 (SEQ ID NO: 22). El alineamiento mostrado en las Figuras 4A y 4B confirma la variación predicha en cada uno de los 11 sitios, y una única mutación no sinónima en la posición 609E de Anc80L60 (SEQ ID NO: 21), que puede ser un artefacto de clonación. Las Figuras 5A y 5B muestran un alineamiento entre los polipéptidos de cápside de VP1 de AAV ancestrales (Anc80L65 (SEQ ID NO: 23), Anc80L27 (SEQ ID NO: 19), Anc80L33 (SEQ ID NO: 24), Anc80L36 (SEQ ID NO: 25), Anc80L60 (SEQ ID NO: 21), Anc80L62 (SEQ ID NO: 22), Anc80L44 (SEQ ID NO: 26), y Anc80L59 (SEQ ID NO: 20)) y los polipéptidos de cápside de VP1 de AAV contemporáneos (AAV8 (SEQ ID NO: 27), AAV9 (SEQ ID NO: 28), AAV6 (SEQ ID NO: 29), AAV1 (SEQ ID NO: 30), AAV2 (SEQ ID NO: 31), AAV3 (SEQ ID NO: 32), AAV3B (SEQ ID NO: 33), y AAV7 (SEQ ID NO: 34)). El alineamiento en las Figuras 5A y 5B muestra que las secuencias de

AAV ancestrales tienen entre aproximadamente el 85% y el 91% de identidad de secuencia con las secuencias de AAV contemporáneas.

También se proporcionan vectores que contienen moléculas de ácido nucleico que codifican polipéptidos. Los vectores, incluidos los vectores de expresión, están disponibles comercialmente o pueden producirse mediante tecnología recombinante. Un vector que contiene una molécula de ácido nucleico puede tener uno o más elementos para la expresión unidos operativamente a tal molécula de ácido nucleico, y además puede incluir secuencias tales como las que codifican un marcador seleccionable (por ejemplo, un gen de resistencia a antibióticos), y/o aquellas que se pueden usar en la purificación de un polipéptido (por ejemplo, la etiqueta 6xHis). Los elementos para la expresión incluyen secuencias de ácido nucleico que dirigen y regulan la expresión de secuencias codificantes de ácido nucleico. Un ejemplo de un elemento de expresión es una secuencia promotora. Los elementos de expresión también pueden incluir uno o más de intrones, secuencias potenciadoras, elementos de respuesta o elementos inducibles que modulan la expresión de una molécula de ácido nucleico. Los elementos de expresión pueden ser de origen bacteriano, levadura, insecto, mamífero o vírico y los vectores pueden contener una combinación de elementos de expresión de diferentes orígenes. Como se usa en el presente documento, unido operativamente significa que los elementos para la expresión se colocan en un vector con respecto a una secuencia codificante de tal manera que dirijan o regulen la expresión de la secuencia codificante.

Una molécula de ácido nucleico, por ejemplo, una molécula de ácido nucleico en un vector (por ejemplo, un vector de expresión, un vector vírico) puede introducirse en una célula huésped. El término "célula huésped" se refiere no solo a la célula o células en particular en las que se ha introducido la molécula de ácido nucleico, sino también a la progenie o posible progenie de tal célula. Los expertos en la técnica conocen muchas células huésped adecuadas; las células huésped pueden ser células procariotas (por ejemplo, *E. coli*) o células eucariotas (por ejemplo, células de levadura, células de insecto, células vegetales, células de mamífero). Las células huésped representativas pueden incluir, sin limitación, A549, WEHI, 3T3, 10T1/2, BHK, MDCK, COS 1, COS 7, BSC 1, BSC 40, BMT 10, VERO, WI38, HeLa, células 293, Saos, C2C12, células L, HT1080, HepG2 y fibroblastos primarios, hepatocitos y células de mioblastos derivadas de mamíferos, incluyendo seres humanos, monos, ratones, ratas, conejos y hámsteres. Los métodos para introducir moléculas de ácido nucleico en células huésped se conocen bien en la técnica e incluyen, sin limitación, precipitación con fosfato de calcio, electroporación, choque térmico, lipofección, microinyección y transferencia de ácido nucleico mediada por virus (por ejemplo, transducción).

Con respecto a los polipéptidos, "purificado" se refiere a un polipéptido (es decir, un péptido o un polipéptido) que se ha separado o purificado a partir de componentes celulares que lo acompañan de forma natural. Típicamente, el polipéptido se considera "purificado" cuando está al menos el 70% (por ejemplo, al menos el 75%, 80%, 85%, 90%, 95% o el 99%) en peso seco, libre de los polipéptidos y moléculas de origen natural con las que se asocia de forma natural. Dado que un polipéptido que se sintetiza químicamente está, por naturaleza, separado de los componentes que lo acompañan de forma natural, un polipéptido sintético se considera "purificado", pero puede eliminarse adicionalmente de los componentes utilizados para sintetizar el polipéptido (por ejemplo, residuos de aminoácidos). Con respecto a las moléculas de ácido nucleico, "aislado" se refiere a una molécula de ácido nucleico que está separada de otras moléculas de ácido nucleico que usualmente están asociadas con ésta en el genoma. Además, una molécula de ácido nucleico aislada puede incluir una molécula de ácido nucleico modificada por ingeniería genética, tal como una molécula de ácido nucleico recombinante o sintética.

Los polipéptidos se pueden obtener (por ejemplo, purificar) a partir de fuentes naturales (por ejemplo, una muestra biológica) mediante métodos conocidos tales como intercambio iónico DEAE, filtración en gel, y/o cromatografía de hidroxiapatita. También se puede obtener un polipéptido purificado, por ejemplo, expresando una molécula de ácido nucleico en un vector de expresión o mediante síntesis química. El grado de pureza de un polipéptido se puede medir utilizando cualquier método apropiado, por ejemplo, cromatografía en columna, electroforesis en gel de poliacrilamida, o análisis por HPLC. De manera similar, las moléculas de ácido nucleico se pueden obtener (por ejemplo, aislar) utilizando métodos de rutina tales como, sin limitación, tecnología de ácido nucleico recombinante (por ejemplo, digestión con enzimas de restricción y ligadura) o la reacción en cadena de la polimerasa (PCR; véase, por ejemplo, PCR Primer: A Laboratory Manual, Dieffenbach & Dveksler, Eds., Cold Spring Harbor Laboratory Press, 1995). Además, las moléculas de ácido nucleico aisladas se pueden sintetizar químicamente.

Métodos de uso de virus ancestrales o porciones de los mismos

Un virus ancestral o una porción del mismo como se describe en el presente documento, particularmente aquellos que presentan una seroprevalencia reducida en relación con los virus contemporáneos o porciones de los mismos, se puede usar en varias aplicaciones de investigación y/o terapéuticas. Por ejemplo, un virus ancestral o una porción

del mismo, como se describe en el presente documento, se puede usar en medicina humana o animal para terapia génica (por ejemplo, en un vector o sistema de vectores para transferencia génica) o para vacunación (por ejemplo, para la presentación de antígenos). Más específicamente, un virus ancestral o una porción del mismo, como se describe en el presente documento, se puede usar para la adición de genes, el aumento de genes, el suministro genético de un polipéptido terapéutico, la vacunación genética, el silenciamiento de genes, la edición del genoma, la terapia génica, el suministro de ARNi, el suministro de ADNc, el suministro de ARNm, el suministro de ARNm, esponjamiento de ARNm, la inmunización genética, la terapia génica optogenética, la transgénesis, la vacunación con ADN, o la inmunización con ADN.

Una célula huésped puede transducirse o infectarse con un virus ancestral o una porción del mismo *in vitro* (por ejemplo, creciendo en cultivo) o *in vivo* (por ejemplo, en un sujeto). Las células huésped que se pueden transducir o infectar con un virus ancestral o una porción del mismo *in vitro* se describen en el presente documento; las células huésped que pueden transducirse o infectarse con un virus ancestral o una porción del mismo *in vivo* incluyen, sin limitación, cerebro, hígado, músculo, pulmón, ojo (por ejemplo, retina, epitelio pigmentario de la retina), riñón, corazón, gónadas (por ejemplo, testículos, útero, ovarios), piel, fosas nasales, sistema digestivo, páncreas, células de los islotes, neuronas, linfocitos, oído (por ejemplo, oído interno), folículos pilosos, y/o glándulas (por ejemplo, tiroides).

Un virus ancestral o una porción del mismo como se describe en el presente documento puede modificarse para incluir un transgén (en *cis* o *trans* con otras secuencias víricas). Un transgén puede ser, por ejemplo, un gen indicador (por ejemplo, beta-lactamasa, beta-galactosidasa (LacZ), fosfatasa alcalina, timidina cinasa, polipéptido verde fluorescente (GFP), cloranfenicol acetiltransferasa (CAT), o luciferasa, o polipéptidos de fusión que incluyen un dominio de etiqueta de antígeno tal como hemaglutinina o Myc) o un gen terapéutico (por ejemplo, genes que codifican hormonas o receptores de los mismos, factores de crecimiento o receptores de los mismos, factores de diferenciación o receptores de los mismos, reguladores del sistema inmunitario (por ejemplo, citocinas e interleucinas) o receptores de los mismos, enzimas, ARN (por ejemplo, ARN inhibitorios o ARN catalíticos), o antígenos diana (por ejemplo, antígenos oncogénos, antígenos autoinmunitarios).

El transgén particular dependerá, al menos en parte, de la enfermedad o deficiencia particular que se está tratando. Simplemente a modo de ejemplo, la transferencia génica o la terapia génica pueden aplicarse al tratamiento de la hemofilia, retinitis pigmentosa, fibrosis quística, amaurosis congénita de Leber, trastornos de almacenamiento lisosomal, errores innatos del metabolismo (por ejemplo, errores innatos del metabolismo de los aminoácidos, incluyendo fenilcetonuria, errores innatos del metabolismo de los ácidos orgánicos, incluyendo academia propiónica, errores innatos del metabolismo de los ácidos grasos, incluida la deficiencia de acil-CoA deshidrogenasa de cadena media (MCAD), cáncer, acromatopsia, distrofias de conos y bastones, degeneraciones maculares (por ejemplo, degeneración macular relacionada con la edad), deficiencia de lipopolipéptido lipasa, hipercolesterolemia familiar, atrofia muscular espinal, distrofia muscular de Duchenne, enfermedad de Alzheimer, enfermedad de Parkinson, obesidad, trastorno inflamatorio del intestino, diabetes, insuficiencia cardíaca congestiva, hipercolesterolemia, pérdida de la audición, enfermedad coronaria, amiloidosis familiar, síndrome de Marfan, insomnio familiar fatal, enfermedad de Creutzfeldt-Jakob, enfermedad de células falciformes, enfermedad de Huntington, degeneración del lóbulo fronto-temporal, síndrome de Usher, intolerancia a la lactosa, trastornos de almacenamiento de lípidos (por ejemplo, enfermedad de Niemann-Pick, tipo C), enfermedad de Batten, coroideremia, enfermedad de almacenamiento de glucógeno tipo II (enfermedad de Pompe), ataxia telangiectasia (síndrome de Louis-Bar), hipotiroidismo congénito, inmunodeficiencia combinada severa (SCID), y/o esclerosis lateral amiotrófica (ALS).

Un transgén también puede ser, por ejemplo, un inmunógeno que es útil para inmunizar a un sujeto (por ejemplo, un ser humano, un animal (por ejemplo, un animal de compañía, un animal de granja, un animal en peligro de extinción). Por ejemplo, los inmunógenos pueden obtenerse de un organismo (por ejemplo, un organismo patógeno) o una porción inmunógena o un componente del mismo (por ejemplo, un polipéptido de toxina o un subproducto del mismo). A modo de ejemplo, los organismos patógenos de los cuales se pueden obtener polipéptidos inmunógenos incluyen virus (por ejemplo, picornavirus, enterovirus, ortomixovirus, reovirus, retrovirus), procariotas (por ejemplo, neumococos, estafilococos, listeria, pseudomonas) y eucariotas (por ejemplo, amebiasis, malaria, leishmaniasis, nematodos). Se entenderá que los métodos descritos en el presente documento y las composiciones producidas por dichos métodos no se limitan por ningún transgén en particular.

Un virus ancestral o una porción del mismo, usualmente suspendido en un vehículo fisiológicamente compatible, puede administrarse a un sujeto (por ejemplo, un mamífero humano o no humano). Los vehículos adecuados incluyen una solución salina, que se puede formular con una diversidad de soluciones tamponantes (por ejemplo, solución salina tamponada con fosfato), lactosa, sacarosa, fosfato de calcio, gelatina, dextrano, agar, pectina y agua.

El virus ancestral o porción del mismo, se administra en cantidades suficientes para transducir o infectar las células y para proporcionar niveles suficientes de transferencia y expresión génica para proporcionar un beneficio terapéutico sin demasiados efectos adversos. Las vías de administración convencionales y farmacéuticamente aceptables incluyen, pero sin limitación, la administración directa a un órgano tal como, por ejemplo, el hígado o el pulmón, por vía oral, intranasal, intratraqueal, por inhalación, intravenosa, intramuscular, intraocular, subcutánea, intradérmica, transmucosa, o por otras vías de administración. Las rutas de administración se pueden combinar, si se desea.

La dosis del virus ancestral o porción del mismo administrada a un sujeto dependerá principalmente de factores tales como la afección que se está tratando, y la edad, el peso y la salud del sujeto. Por ejemplo, una dosis terapéuticamente eficaz de un virus ancestral o una porción del mismo a administrar a un sujeto humano generalmente está en el intervalo de aproximadamente 0,1 ml a aproximadamente 10 ml de una solución que contiene concentraciones de aproximadamente 1×10^1 a 1×10^{12} copias genómicas (GC) de virus ancestrales (por ejemplo, aproximadamente 1×10^3 a 1×10^9 GC). La transducción y/o la expresión de un transgén se pueden controlar en diversos puntos de tiempo después de la administración por ADN, ARN o ensayos de proteínas. En algunos casos, los niveles de expresión del transgén se pueden controlar para determinar la frecuencia y/o la cantidad de dosis. Los regímenes de dosificación similares a los descritos con fines terapéuticos también pueden utilizarse para la inmunización.

Los métodos descritos en el presente documento también se pueden usar para modelar la evolución directa, para modificar o eliminar uno o más dominios inmunógenos de un virus o porción del mismo.

De acuerdo con la presente invención, pueden emplearse técnicas convencionales de biología molecular, microbiología, bioquímica y ADN recombinante dentro de los conocimientos de la técnica. Dichas técnicas se explican completamente en la bibliografía. La invención se describirá adicionalmente en los siguientes ejemplos, que no limitan el alcance de los métodos y composiciones de materia que se describen en las reivindicaciones.

EJEMPLOS

Ejemplo 1 - Predicción computacional de secuencias ancestrales

Se obtuvo un conjunto de 75 secuencias de aminoácidos diferentes de cápsides de AAV a partir de varias bases de datos públicas, incluyendo GenBank, y las secuencias se alinearon utilizando el algoritmo PRANK-MSA, versión 121002, con la opción "-F".

Se usó ProtTest3 (véase, por ejemplo, Darriba et al., 2011, Bioinformatics, 27(8):1164-5; disponible en darwin.uvigo.es/software/prottest3 en Internet) para evaluar diferentes modelos de la evolución polipeptídica (por ejemplo, los incluidos en ProTest3, concretamente, JTT, LG, WAG, VT, CpRev, RtRev, Dayhoff, DCMut, FLU, Blosum62, VT, HIVb, MtArt, MtMam) en condiciones diferentes (por ejemplo, las incluidas en ProTest3, concretamente, "+I", "+F", "+ G", y combinaciones de los mismos). El modelo JTT (Jones et al., 1992, Comp. Appl. Biosci., 8:275-82) con +G y +F (Yang, 1993, Mol. Biol. Evol., 10:1396-1401; y Cao et al., 1994, J. Mol. Evol., 39:519-27) se seleccionó basándose en su puntuación del criterio de información de Aikake (AIC; Hirotugu, 1974, IEEE Transactions on Automatic Control, 19:716-23) de acuerdo como se implementó en ProTest3.

Se construyó una filogenia de la evolución de AAV utilizando PhyML (Guindon y Gascuel, 2003, Systematic Biology, 52:696-704)). Véase la Figura 3. El árbol se generó utilizando el modelo de sustitución JTT + F con 4 categorías de sustitución discretas y un parámetro de forma Gamma estimado. Los árboles resultantes se mejoraron a través del intercambio de vecinos más cercanos (NNI) y la poda y reinjerto de subárbol (SPR), y se evaluó su importancia a través de muestreo por repetición y la prueba de razón de similitud aproximada (aLRT; Anisimova y Gascuel, 2006, Systematic Biology, 55:539-52)) usando la variante "de tipo SH".

El árbol filogénico construido anteriormente se usó entonces para estimar los estados ancestrales de la cápside de AAV en cada nodo interior a la filogenia. Las secuencias de cápside ancestral se reconstruyeron utilizando los principios de máxima verosimilitud mediante el software Phylogenetic Analysis by Maximum Likelihood (PAML) (Yang, 1997, Comp. Applic. BioSci., 13:555-6; disponible en abacus.gene.ucl.ac.uk/software/paml.html en Internet) envuelto en Lazarus (Sourceforge en sf.net). Más específicamente, la reconstrucción de Lazarus/PAML se configuró para generar una reconstrucción de aminoácidos usando el modelo de sustitución JTT+F utilizando 4 categorías distribuidas por gamma. Se usó AAV5 como un grupo externo. Finalmente, se añadió la opción "I" para colocar los indeles (es decir, se codificó de forma binaria y se colocó a través de parsimonia máxima utilizando el algoritmo de Fitch) después de que se realizó la reconstrucción de PAML.

Debido a que la reconstrucción se realizó en un modo de máxima verosimilitud, se puede calcular la verosimilitud de que cualquier residuo se encuentre en una posición dada en un nodo dado. Para ello, se escribió un guión adicional para identificar todas las posiciones a lo largo de la secuencia con una probabilidad posterior calculada por debajo de un cierto umbral. Se seleccionó un umbral de 0,3, lo que significa que cualquier aminoácido con una probabilidad posterior calculada de más de 0,3 se incluyó en la síntesis de la genoteca. Estos residuos se seleccionaron como variantes de interés en la genoteca.

Para finalizar la secuencia, se tuvo que codificar una utilidad adicional para seleccionar los codones. Se escribió un guión para derivar codones similares a los de otra secuencia de AAV (AVVRh10, que tiene aproximadamente un 92% de identidad de secuencia con la secuencia de andamiaje de Anc80) y se aplica un algoritmo novedoso para sustituir codones en los que existían emparejamientos incorrectos de secuencia basados en una matriz de sustitución de codones. El nuevo algoritmo se muestra a continuación:

Dado: secuencia de aminoácidos, Pt, con la secuencia de nucleótidos correspondiente, Nt, donde Nt codifica Pt; y secuencia de proteínas, Pi, donde Pi presenta una fuerte homología con Pt.
Alinear Pi con Pt usando Needleman-Wunsch usando la tabla Blosom62 para la puntuación. Generar una nueva secuencia de nucleótidos, Ni, pasando por el alineamiento de proteínas, utilizando el codón correspondiente de Nt,
donde el aminoácido en Pt coincide exactamente con el de Pi,
el codón de "mejor puntuación" de la matriz Codon-PAM (Schneider et al., 2005, BMC Bioinform., 6: 134) donde hay una sustitución,
un hueco donde existe un hueco en Pi alineada contra un aminoácido en Pt, y
el nucleótido que ocurre con mayor frecuencia en el Nt (que codifica un aminoácido dado) donde existe un aminoácido en Pi alineada contra un hueco en Pt.

Además, se realizaron dos cambios de un solo nucleótido para suprimir la transcripción de la proteína activadora del ensamblaje (AAP), que se codifica fuera del marco dentro del gen de la cápside del AAV en el AAV natural. Dado que la codificación de la AAP (contemporánea o ancestral) no fue parte de esta reconstrucción, la expresión de la AAP se eliminó haciendo una mutación sinónima en la secuencia cap, y la secuencia de AAP se proporcionó en trans durante la producción vírica.

Ejemplo 2 - Expresión de secuencias de VP1 de AAV ancestrales

Se realizaron experimentos para determinar si las secuencias de cápside de AAV ancestrales predichas se pueden usar para hacer vectores víricos.

Se clonaron varias de las secuencias de cápside de AAV ancestrales predichas. La genoteca de cápsides ancestrales se transfirió a un plásmido de expresión rep-cap para permitir la formación de partículas víricas en la transfección transitoria. Para mantener los niveles de expresión apropiados y el corte y empalme de VP1, VP2 y VP3, se clonaron los genes de cap de la genoteca cortando *HindIII*, ubicado en 5' de cap en la secuencia codificante rep, y *SpeI*, que se modificó entre el codón de parada cap y la señal de poliadenilación. En consecuencia, para clonar las cápsides ancestrales en una construcción "REP/CAP" más convencional, el plásmido de paso se digirió con *HindIII* y *SpeI*, se purificó en gel y se ligó en un plásmido rep/cap digerido de manera similar.

Los polipéptidos expresados se resolvieron en un gel de SDS al 10%. Como se muestra en la Figura 6, los polipéptidos de la cápside se expresaron y sometieron a corte y empalme adecuadamente en VP1, VP2 y VP3 a partir de varias secuencias de AAV ancestrales (Anc80L44, Anc80L27 y Anc80L65), así como de una secuencia de AAV contemporánea, AAV2/8.

Ejemplo 3 - Valoración vírica

El AAV se produjo en células HEK293 a través de la cotransfección transitoria de plásmidos que codifican todos los elementos necesarios para el ensamblaje de partículas víricas. Brevemente, las células HEK293 se hicieron crecer al 90% de confluencia y se transfectaron con (a) el plásmido del genoma vírico que codifica el transgén de luciferasa (expresado por el promotor de CMV) flanqueado por ITR de AAV2, (b) el plásmido de empaquetamiento de AAV que codifica rep de AAV2 y las proteínas de la cápside sintetizadas divulgadas en el presente documento, (c) cápside que expresa AAP de AAV2, y (d) genes auxiliares adenovirales necesarios para el empaquetado y el ensamblaje de AAV. Las células se incubaron a 37°C durante 2 días y las células y los medios se recolectaron y recogieron.

La suspensión de medio celular se lisó mediante 3 ciclos consecutivos de congelación y descongelación. A continuación, el lisado se eliminó por centrifugación y se trató con una enzima en condiciones para realizar una digestión exhaustiva del ADN, aquí Benzonase™, para digerir cualquier ADN presente fuera de la partícula vírica. La preparación de AAV se diluyó hasta estar dentro del intervalo de medición lineal de una plantilla de ADN de control, en este caso, un plásmido linealizado con el cebador TaqMan™ idéntico y la secuencia de unión de la sonda en comparación con el genoma del vector. Se realizó la PCR TaqMan™ con cebadores e hibridación de la sonda al genoma del vector vírico de elección. El título se calculó basándose en la medición de TaqMan™ en copias genómicas (GC) por mililitro (ml) como se muestra en la Tabla 2 a continuación.

Tabla 2

Títulos (GC/ml)	Escala pequeña N.º 1	Escala pequeña N.º 2
AAV2/2	$1,12 \times 10^9$	$1,99 \times 10^9$
AAV2/8	$4,17 \times 10^{10}$	$5,91 \times 10^{10}$
Anc80L27	$8,01 \times 10^8$	$1,74 \times 10^9$
Anc80L44	$1,52 \times 10^9$	$1,43 \times 10^9$
Anc80L65	$1,42 \times 10^9$	$2,05 \times 10^9$
Sin control de cápside	$5,23 \times 10^5$	$7,25 \times 10^5$

Los resultados de la producción de vectores a pequeña escala en partículas de cápside de AAV reconstruidas de forma ancestral demostraron rendimientos similares a AAV2, pero reducidos en relación con AAV8, ambos de los cuales son preparaciones de vectores basadas en AAV contemporáneos.

Ejemplo 4 - Transducción vírica *in vitro*

Se realizaron transducciones víricas *in vitro* para evaluar la capacidad de los virus que contienen las secuencias de AAV ancestrales predichas para infectar células.

Después de la producción vectorial de alto rendimiento utilizando la biblioteca de secuencias de Anc80, las células HEK293 se transdujeron con cada vector vírico. Además de una secuencia Anc80, cada vector vírico contenía un transgén de luciferasa. La luciferasa se midió mediante la cuantificación de la bioluminiscencia en un lector de placas de 96 pocillos después de la adición del sustrato de luciferina a las células transducidas o al lisado celular. Después de la cuantificación, se produjo un mapa de calor de la expresión de luciferasa en cuatro placas de 96 pocillos concatenadas (excluyendo una columna de controles en cada placa). Debido a la gran cantidad de inserciones, eliminaciones y transiciones asociadas con el proceso de producción de vectores de alto rendimiento, muchos de los vectores no eran funcionales. Para los fines del presente documento, solo se evaluaron adicionalmente los virus que eran funcionales en este ensayo (es decir, capaces de transducir células HEK293 y expresar el transgén).

Se transdujeron células HEK293, a igual multiplicidad de infección (MOI) de 1×10^4 copias genómicas (GC) por célula, con dos vectores de AAV contemporáneos (AAV2/2 y AAV2/8) y tres vectores de AAV ancestrales predichos (Anc80L27, Anc80L44, y Anc80L65). Cada vector contenía un transgén que codificaba luciferasa o un transgén que codificaba eGFP. Se tomaron imágenes de las células 60 horas después utilizando el canal de GFP de un microscopio óptico AMG EvosF1. La Figura 7 muestra la expresión de luciferasa después de la transducción *in vitro*. Cada uno de los virus AAV ancestrales demostró una transducción eficiente de células HEK293.

Ejemplo 5 - Transducción retiniana *in vivo*

Se realizaron transducciones retinianas para determinar si los vectores AAV ancestrales son capaces de dirigirse a las células retinianas murinas *in vivo*.

Los ojos murinos se transdujeron con 2×10^8 copias genómicas (GC) de tres AAV ancestrales diferentes (Anc80L27, Anc80L44 y Anc80L65) y un AAV contemporáneo (AAV2/8), todos los cuales incluyen un transgén que codifica eGFP. Para las transducciones, cada vector AAV se administró quirúrgicamente debajo de la retina al generar un espacio entre el fotorreceptor y la capa del epitelio pigmentario de la retina a través de la administración de un bolo vectorial con un dispositivo de inyección. El bolo vectorial se dejó en el espacio subretiniano y el desprendimiento subretiniano se resolvió con el tiempo. La expresión de GFP se controló de forma no invasiva mediante fotografía de

fondo de la retina del animal después de la dilatación de la pupila con Tropicamide™. Todas las retinas presentadas demostraron diversos grados de direccionamiento exitoso de AAV ancestrales a la retina.

La histología retiniana también se realizó y se visualizó bajo microscopía fluorescente para identificar el tipo o tipos de células transducidos. La histología se realizó en una retina murina transducida con el vector de AAV ancestral Anc80L65 como se describe anteriormente. La expresión de eGFP mediada por Anc80L65 fue evidente en la capa nuclear externa (ONL), los segmentos internos (IS) y el epitelio pigmentario de la retina (RPE), lo que indica que el vector Anc80L65 ancestral se dirige a fotorreceptores murinos y células del epitelio pigmentario retiniano.

Ejemplo 6 - Ensayo de anticuerpos neutralizantes

Se realizaron ensayos de anticuerpos neutralizantes para evaluar si un virus AAV ancestral es más resistente a la neutralización de anticuerpos que un virus AAV contemporáneo. Los ensayos de anticuerpos neutralizantes miden la concentración de anticuerpos (o el título en el que una muestra experimental contiene una concentración de anticuerpos) que neutraliza una infección en un 50% o más en comparación con un control en ausencia del anticuerpo.

Las muestras de suero o la solución madre de IVIG (200 mg/ml) se diluyeron en serie dos veces, y las muestras sin diluir y diluidas se incubaron conjuntamente con un virus AAV ancestral, Anc80L65, y un virus AAV contemporáneo, AAV2/8, a una MOI de 10^4 durante aproximadamente 30 minutos a 37°C. Cada virus incluía un transgén de luciferasa. El vector mezclado y una muestra de anticuerpo se transdujeron entonces en células HEK293. Para estos experimentos, la muestra de anticuerpos utilizada fue inmunoglobulina intravenosa (IGIV), IgG agrupadas extraídas del plasma de más de mil donantes de sangre (vendidos comercialmente, por ejemplo, como Gammagard™ (Baxter Healthcare; Deerfield, IL) o Gamunex™ (Grifols; Los Angeles, CA)). 48 horas después del inicio de la transducción, las células se analizaron por bioluminiscencia para detectar luciferasa. El título de anticuerpos neutralizantes se determinó identificando la dilución de la muestra para la cual se alcanzó un 50% o más de neutralización (transducción de muestra/transducción del virus de control en ausencia de muestra).

Ejemplo 7 - Caracterización de Anc80

En base a los métodos descritos en el presente documento, se obtuvo la secuencia Anc80 más probable (de acuerdo como se determinó a través de la probabilidad posterior) y se designó Anc80L1 (la SEQ ID NO:35 muestra la secuencia de ácido nucleico de la cápside de Anc80L1 y la SEQ ID NO:36 muestra la secuencia de aminoácidos del polipéptido de VP1 de Anc80L1). La genoteca probabilística de Anc80 también se sintetizó utilizando las secuencias descritas en el presente documento por una empresa comercial y se subclonó en vectores de expresión.

La genoteca de Anc80 se evaluó clonalmente para determinar el rendimiento vectorial y la infectividad en ensayos combinados. Fuera de este cribado, se caracterizaron adicionalmente Anc80L65 (SEQ ID NO:23), así como varias otras variantes.

La genoteca de Anc80 y Anc80L65 se compararon en términos de diferencia de secuencia (Figura 9; % por encima de la diagonal, N.º de diferencias de aminoácidos a continuación). Usando NCBI-BLAST, la secuencia más cercana disponible públicamente para Anc80L65 es rh10 (GenBank N.º de acceso AAO88201.1).

La Figura 10 muestra que Anc80L65 produjo rendimientos vectoriales equivalentes a AAV2 (Panel A), generó partículas víricas bajo electroscopia de transmisión (TEM) (Panel B), y produjo bioquímicamente la cap de AAV y las proteínas VP1, 2 y 3 basadas en SDS page en condiciones de desnaturalización (Panel C) y transferencia Western utilizando el anticuerpo de cápside de AAV, B1 (Panel D). Estos experimentos se describen con más detalle en los párrafos siguientes.

Brevemente, los vectores AAV2/8, AAV2/2, AAV2/Anc80L27, AAV2/Anc80L44, y AAV2/Anc80L65 se produjeron a pequeña escala que contenían una construcción de indicador compuesta por eGFP y luciferasa de luciérnaga bajo un promotor de CMV a pequeña escala. Los títulos de estas preparaciones de virus a pequeña escala se obtuvieron después a través de qPCR. Basándose en estos experimentos, se encontró que los vectores Anc80L27, Anc80L44 y Anc80L65 producen niveles víricos comparables a los de AAV2 (Figura 10A).

Para confirmar que las proteínas de la cápside Anc80L65 se ensamblaron en partículas pseudovíricas intactas del tamaño y conformación apropiados, se obtuvieron micrografías utilizando microscopía electrónica de transmisión (TEM). Una preparación purificada a gran escala de Anc80-L065 se cargó en rejillas de cobre recubiertas con

polivinilo formal (Formvar®) y luego se tiñó con acetato de uranilo. Las micrografías revelaron partículas hexagonales intactas con diámetros entre 20 y 25 nm (Figura 10B).

Para determinar si los genes de cápside ancestrales sintéticos se procesaron adecuadamente (es decir, se sometieron a corte y empalme y se expresaron), se cargaron preparaciones purificadas a gran escala de vectores AAV2/8, AAV2/2 y AAV2/Anc80L65 en un gel SDS-PAGE (1E10 GC/pocillo) en condiciones de desnaturalización. Las bandas que representaban las proteínas de cápside vírica VP1, VP2 y VP3 estaban claramente presentes para cada preparación de vector (Figura 10C). La transferencia de Western con el anticuerpo de cápside de AAV B1 confirmó además que estas bandas representaban las proteínas predichas (Figura 10D).

Además, la Figura 11 muestra que Anc80L65 infectó tejidos y células de mamífero *in vitro* en células HEK293 en MOI 10E4 GC/célula usando GFP como lectura (Panel A) o luciferasa (Panel B) frente a controles de AAV2 y/o AAV8. Anc80L65 también fue eficaz en el direccionamiento al hígado después de una inyección IV del AAV indicado que codifica un transgén LacZ nuclear (fila superior, Panel C), después de la inyección intramuscular (IM) directa del AAV indicado que codifica GFP (fila central, Panel C), y después de la inyección subretiniana con el AAV indicado que codifica GFP (fila inferior, Panel C). Estos experimentos se describen con más detalle en los párrafos siguientes.

Para obtener una medida relativa de la infectividad de los viriones ancestrales, se produjeron preparaciones en bruto de AAV2/2, AAV2/8, AAV2/Anc80L65, AAV2/Anc80L44, AAV2/Anc80L27, AAV2/Anc80L121, AAV2/Anc80L122, AAV2/Anc80L123, AAV2/Anc80L124, y AAV2/Anc80L125 que contienen una construcción indicadora bi-cistrónica que incluye una eGFP y secuencias de luciferasa de luciérnaga bajo el control de un promotor de CMV. Las placas de 96 pocillos confluentes con células HEK293 se sometieron luego a transducción con cada vector a una MOI de 1E4 GC/célula (títulos obtenidos a través de qPCR como anteriormente). 48 horas después, la microscopía de fluorescencia confirmó la presencia de GFP en células transducidas (Figura 11A). Después, las células se analizaron para determinar la presencia de luciferasa (Figura 11B), que determinó que la expresión de luciferasa en células transducidas con vectores derivados de Anc80 estaba en medio de células transducidas con AAV8 (nivel inferior de transducción) y AAV2 (nivel superior de transducción).

Para evaluar la eficacia relativa de la transferencia génica en un contexto *in vivo*, se obtuvieron preparaciones purificadas de alto título de AAV2/2, AAV2/8 y AAV2/Anc80L65. Se inyectaron 3,9E10 GC de cada vector, encapsidando un transgén que codifica LacZ nuclear bajo el control de un promotor TBG, en ratones C57BL/6 (3 ratones por condición) a través de una inyección IP después de la anestesia general. 28 días después de la inyección, se sacrificaron los ratones y se recogieron los tejidos. Los hígados se seccionaron mediante técnicas histológicas estándar y se tiñeron para determinar la beta-galactosidasa. Luego, se tomaron imágenes de las secciones bajo un microscopio y se muestran imágenes representativas en la Figura 11C, fila superior.

Después, se obtuvieron vectores de los mismos serotipos que contenían un transgén bicistrónico que codificaba eGFP y hA1AT bajo el control de un promotor pCASI. Para evaluar la capacidad de Anc80L65 para transducir el músculo esquelético murino, se inyectó 1E10 GC de cada vector en el músculo esquelético de ratones C57BL/6 (5 ratones por condición) después de la anestesia general. 28 días después de la inyección, los ratones se sacrificaron, los tejidos se sometieron a criosección y se evaluó la presencia de eGFP utilizando microscopía confocal fluorescente (azul es DAPI, verde es eGFP). Las imágenes representativas se muestran en la Figura 11C, fila central. Estos experimentos demostraron que los vectores Anc80L65 eran capaces de transducir el músculo esquelético murino mediante inyección intramuscular.

Se obtuvieron vectores de los mismos serotipos, esta vez encapsidando construcciones que codificaban solo un transgén de eGFP bajo el control de un promotor de CMV. Las partículas de 2E9 se inyectaron por vía subretiniana en ratones C57BL/6 después de la anestesia general. 28 días después de la inyección, se sacrificaron los ratones y se recogieron los ojos, se sometieron a criosección, y se evaluó la presencia de eGFP utilizando microscopía confocal fluorescente (azul es DAPI, verde es eGFP). Se muestran imágenes representativas en la Figura 11C, fila inferior. Estos experimentos demuestran que los vectores Anc80L65 pueden transducir la retina murina a un nivel que es comparable a los vectores AAV8.

Brevemente, se obtuvieron preparaciones purificadas y de alto título de vectores víricos AAV2/8, AAV2/2, AAV2/rh32.33 y AAV2/Anc80L65 que encapsidaban un transgén bicistrónico que incluye eGFP y luciferasa de luciérnaga bajo el control de un promotor de CMV. Estos vectores se incubaron luego con diluciones seriadas de dos veces de IVIG (10 mg, 5 mg, 2,5 mg, etc.) o se incubaron sin IVIG (1E9 GC por condición). Después de la incubación, se usaron vectores para transducir células HEK293 a una MOI de 1E4 por pocillo (una dilución por pocillo).

Ejemplo 8 - Generación de cápsides de AAV ancestrales adicionales

Las secuencias de cápside de AAV ancestrales más probables (de acuerdo con lo determinado por la probabilidad posterior) se sintetizaron luego a través de un laboratorio comercial (Gen9) y se proporcionaron como ADNds lineal. Estas secuencias de aminoácidos se compararon entonces con las de los AAV existentes para determinar el grado en que difieren (Figura 13). Cada proteína VP1 ancestral difiere de las de los AAV existentes representativos seleccionados entre un 3,6% y un 9,3% (Figura 13A), mientras que las proteínas VP3 ancestrales difieren entre un 4,2 y un 9,4% (Figura 13B). Estas cápsides se subclonaron en plásmidos de producción de AAV (pAAVector2/Vacio) a través de la digestión con enzimas de restricción (HindIII y SpeI) y la ligadura de T4. Estos clones se confirmaron a través de digestión de restricción y secuenciación de Sanger, y luego se produjeron preparaciones a escala media de ADN plasmídico.

Cada uno de estos plásmidos se usó luego para producir vectores de AAV que contenían un gen indicador que codificaba tanto eGFP como luciferasa de luciérnaga. Estos vectores se produjeron por triplicado en pequeña escala como se describe anteriormente. Las preparaciones en bruto del virus se titularon luego a través de qPCR y se encontró que producían entre un 2,71% y un 183,1% de partículas víricas en relación con AAV8 (Figuras 14 y 15). Estos títulos se utilizaron entonces para configurar un experimento controlado por títulos para evaluar la infectividad relativa. Anc126 no estaba controlado por título debido a su producción significativamente reducida y, en consecuencia, los datos sobre la infectividad de Anc126 no se pueden comparar con precisión con la infectividad de los otros virus en el experimento. Los otros vectores se usaron para transducir células HEK293 a una multiplicidad de infección (MOI) de 1,9E3 GC/célula.

60 horas después de la transducción, las células se evaluaron para determinar la expresión de GFP a través de microscopía de fluorescencia. Se detectaron células positivas para eGFP en cada una de las condiciones, excepto para el control negativo (Figura 16). Esto indica que cada una de las secuencias ancestrales predichas, sintetizadas y clonadas es capaz de producir partículas víricas infecciosas y viables. Para tener una idea de los niveles relativos de infectividad, los ensayos de luciferasa también se realizaron en las mismas células. Los resultados indican que cada uno de los vectores ancestrales es capaz de transducir células HEK293 entre un 28,3% y un 850,8% en relación con AAV8 (Figuras 17 y 18). Se observa que Anc126 se excluyó del análisis de la transducción relativa ya que no se controló por título.

En resumen, se sintetizaron ocho nuevos genes de cápside de AAV ancestrales y se usaron en la producción de vectores víricos funcionales junto con AAV8, AAV2 y los vectores Anc80L65 descritos previamente. La producción y la infectividad se evaluaron *in vitro* y se muestra un resumen de estos hallazgos en la Figura 19.

Ejemplo 9 - Inmunoprofilaxis vectorizada

En la inmunoprofilaxis vectorizada, los vehículos de terapia génica (tal como el AAV) se utilizan para administrar transgenes que codifican anticuerpos ampliamente neutralizantes contra agentes infecciosos. Véanse, por ejemplo, Balazs et al. (2013, Nat. Biotechnol., 31:647-52); Limberis et al. (2013, Sci. Transl. Med., 5:187ra72); Balazs et al. (2012, Nature, 481:81-4); y Deal et al. (2014, PNAS USA, 111:12528-32). Una ventaja de este tratamiento es que el huésped produce los anticuerpos en sus propias células, lo que significa que una única administración tiene el potencial de conferir una protección de por vida contra los agentes etiológicos.

Ejemplo 10 - Vehículos de administración farmacológica

LUCENTIS® (ranibizumab) y AVASTIN® (bevacizumab) son ambos agentes anti-angiogénesis basados en los mismos anticuerpos monoclonales de ratón humanizados contra el factor de crecimiento endotelial vascular A (VEGF-A). Aunque el bevacizumab es un anticuerpo completo y el ranibizumab es un fragmento (Fab), ambos actúan para tratar la degeneración macular relacionada con la edad húmeda a través del mismo mecanismo: antagonizando el VEGF. Véanse, por ejemplo, Mao et al. (2011, Hum. Gene Ther., 22:1525-35); Xie et al. (2014, Gynecol. Oncol., doi: 10.1016/j.ygyno.2014.07.105); y Watanabe et al. (2010, Gene Ther., 17:1042-51). Debido a que ambas de estas moléculas son proteínas, pueden codificarse por el ADN y producirse en células transducidas con vectores que contienen un transgén, y son lo suficientemente pequeñas como para ser empaquetadas en vectores de AAV.

OTRAS FORMAS DE REALIZACIÓN

Debe entenderse que, si bien los métodos y las composiciones de materia se han descrito en el presente documento junto con varios aspectos diferentes, la descripción anterior de los diversos aspectos pretende ilustrar y no limitar el alcance de los métodos y composiciones de la materia. Otros aspectos, ventajas y modificaciones están dentro del alcance de las siguientes reivindicaciones.

Se divulgan métodos y composiciones que se pueden usar para, se pueden usar junto con, se pueden usar en la preparación para, o son productos de los métodos y composiciones que se divulgan. Estos y otros materiales se divulgan en el presente documento, y se entiende que se describen combinaciones, subconjuntos, interacciones, grupos, etc. de estos métodos y composiciones. Es decir, aunque la referencia específica a cada una de las diversas combinaciones y permutaciones individuales y colectivas de estas composiciones y métodos no se puede divulgar explícitamente, cada una se contempla y se describe específicamente en el presente documento. Por ejemplo, si se describe y se analiza una composición particular de materia o un método particular y se analizan varias composiciones o métodos, todas y cada una de las combinaciones y permutaciones de las composiciones y los métodos se contemplan específicamente a menos que se indique específicamente de otro modo. Asimismo, también se contempla específicamente y se divulga cualquier subconjunto o combinación de estas.

APÉNDICE A**SEQ ID NO:1: Polipéptido Anc80**

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFNGLDKGEPVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
GKKRPVEQSPQEPDSSSGIGKKGQQPAX1KRLNFGQTGDSESVDPQPPLGEPAAAPSGVGSNTMX2AGGG
APMADNNEGADGVGNASGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISSQSGX3STNDNTYFGYS
TPWGYFDFNRFHCHFSPRDWQRLINNNWGFPRKX4LNFKLENIQVKEVTNDGTTTIANNLSTVQVFTD
SEYQLPYVLGSAHQGCLPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFX5FSYT
FEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRTQTTSCTAGNRX6LQFSQAGPSSMANQAKNWLPGPCY
RQQRVSKTX7NQNNNSNFAWTGATKYHLNGRDSLVPNPGPAMATHKDDEDKFFPMSGVLI FGKQGAGNSNV
DLDNVMITX8EEEIKTTNPVATEX9YGT VATNLQ SX10NTAPATGTVNSQGALPGMVWQX11RDVYLQGP
IWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTVPANPPTTFSPAKFASFITQYSTGQVSVEIEWEL
QKENS KRWNPEIQYTSN YNKSTNVDFAVDTNGVYSEPRPIGTRYLTRNL

X1 = K/R; X2 = A/S; X3 = A/G; X4 = R/K; X5 = E/Q; X6 = T/E; X7 = A/T; X8 = S/N; X9 = Q/E; X10 = S/A; X11 = N/D

SEQ ID NO:2: ADN Anc80

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTTCGCGAGTGGTGGG
 ACTTGAACCTGGAGCCCCGAAACCCAAAGCCAACAGCAAAAGCAGGACGACGGCCGGGGTCTGGTGGT
 TCCTGGCTACAAGTACCTCGGACCCCTTCAACGGACTCGACAAGGGGGAGCCCGTCAACGCGGCGGACGCA
 GCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATA
 ACCACGCCGACGCCGAGTTTCAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGC
 AGTCTTCCAGGCCAAGAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT
 GGAAAGAAGAGACCGGTAGAGCAATCACCCAGGAACCAGACTCCTCTTCGGGCATCGGCAAGAAAGGCC
 AGCAGCCCCGCGXXX1AAGAGACTCAACTTTGGGCAGACAGGCGACTCAGAGTCAGTGCCCCGACCCCTCAAC
 CACTCGGAGAACCCCCCGAGCCCCCTCTGGTGTGGGATCTAATAACAATGXXX2GCAGGCGGTGGCGCTC
 CAATGGCAGACAATAACGAAGGCGCCGACGAGTGGGTAAACGCCTCAGGAAATTGGCATTGCGATTCCAC
 ATGGCTGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTCCCCACCTACAACAACACCTC
 TACAAGCAATCTCCAGCCAATCGGGAXXX3AGCACCAACGACAACACCTACTTCGGCTACAGCACCCCC
 TGGGGGTATTTTGACTTTTACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAACA
 ACAACTGGGGATTCCGGCCCAAGXXX4CTCAACTTCAAGCTCTTCAACATCCAGGTCAAGGAGGTACGCA
 CGAATGATGGCACCACGACCATCGCCAATAACCTTACCAGCAGGTTTCAAGTCTTTACGGACTCGGAATA
 CCAGCTCCCGTACGTCTCGGCTCTGCGCACAGGGCTGCTGCCTCCGTTCGGGCGGACGTCTTCATG
 ATTCTCAGTACGGGTACCTGACTCTGAACAATGGCAGTCAGGCGGTGGGCGGTCTCTCTTCTACTGCC
 TGGAGTACTTTCCTTCTCAAATGCTGAGAACGGGCAACAACCTTTXXX5TTCAGCTACACGTTTGAGGACG
 TGCCTTTTTCACAGCAGCTACGCGCACAGCCAAAGCCTGGACCGGTGATGAACCCCTCATCGACCAGTA
 CCTGTACTACCTGTCTCGGACTCAGACCAGAGTGGTACCGCAGGAAATCGGXXX6TTGCAATTTTCTCA
 GGCCGGGCTTAGTAGCATGGCGAATCAGGCCAAAACTGGCTACCCGGGCCCTGCTACCGGCAGCAACGC
 GTCTCCAAGACAXXX7AATCAAAATAACAACAGCAACTTGCCTGGACCGGTGCCACCAAGTATCATCTG
 AATGGCAGAGACTCTCTGGTAAATCCCGGTCCCGCTATGGCAACCCACAAGGACGACGAAGACAAATTTT
 TTCCGATGAGCGGAGTCTTAATATTTGGGAAACAGGGAGCTGGAAATAGCAACGTGGACCTTGACAACGT
 TATGATAACCCXXX8GAGGAAGAAATTAACCACCAACCCAGTGGCCACAGAAAXXX9TACGGCACGGTGG
 CCACTAACCTGCAATCGXXX10AACACCGCTCCTGCTACAGGGACCGTCAACAGTCAAGGAGCCTTACCT
 GGCATGGTCTGGCAGXXX11CGGGACGTGTACCTGCAGGGTCTATCTGGGCCAAGATTCTCACACGGA
 CGGACACTTTCATCCCTCGCCGCTGATGGGAGGCTTTGGACTGAAACACCCGCTCCTCAGATCCTGATT
 AAGAATACACCTGTTCCCGCAATCCTCCAACCTACCTTCACTCCAGCTAAGTTTGCCTCGTTCATCACGC
 AGTACAGACCCGGACAGGTGAGCGTGGAAATTGAATGGGAGCTGCAGAAAGAAAACAGCAAACGCTGGAA
 CCCAGAGATTCAATACACTTCCAACCTACAACAATCTACAAATGTGGACTTGTCTGTTGACACAAATGGC
 GTTTATTCTGAGCTCGCCCATCGGCACCCGTTACCTACCCGTAATCTG
 XXX1 = AAG/AAA; XXX2 = GCA/AGC; XXX3 = GCA/GGC; XXX4 = AGA/AAG; XXX5 = GAG/CAG; XXX6 =
 ACG/GAG; XXX7 = GCG/ACC; XXX8 = AGT/AAC; XXX9 = CAG/GAG; XXX10 = TCA/GCC; XXX11 =
 AAC/GAC

5 **SEQ ID NO:3: Polipéptido Anc81**
 MAADGYLPDWLEDNLSEGIREWWDLPKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFNGLDKGEPVNAADA
 AALEHDKAYDQQLKAGDNPYLRYNHADAQEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
 GKKRPVEQSPQEPDSSX1GIGKKQQPAX2KRLNFGQTGDSESVDPQPLGEPAPSGVGSNTMAAGGG
 APMADNNEGADGVGNASGNWHCDSTWLGDRVITSTRTWALPTYNNHLYKQISX3X4QSGGSTNDNTYFG
 YSTPWGYFDFNRFHCHFSRPRDQRLINNNWGFPRKX5LNFKLFNIQVKEVTNDGTTTIANNLSTVQVF
 TDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFSPQMLRTGNNEFX6FS
 YTFEDVPFHSSYAHSQSLDRLMNLIDQYLYLSRTQTTGGTAGNX7X8LQFSQAGPSSMANQAKNWLPG
 PCYRQQRVSKTTNQNNSNFAWTGATKYHLNGRDSLVPNGVAMATHKDDDRFFPSSGVLI FGKQAGNX
 9NVDX10X11NVMITX12EEEIKTTNPVATEEYX13VATNLQX14NTAPQTGTVNSQGALPGMVWQNR
 DVYLQGPWIWAKIPHDTGNFHPSPMLGGFGLKHPPQILIKNTVPANPPTTFX15PAKFASFITQYSTGQ
 VSVEIEWELQKENS KRWNPEIQYTSNYNKSTNVDFAVDTEGVYSEPRPIGTRYLTRNL
 X1=T/S; X2=K/R; X3=N/S; X4=S/H; X5=R/K; X6=E/Q; X7=R/Q; X8=T/E; X9=D/S; X10=L/Y; X11=D/S;
 X12=S/N; X13=V/I; X14=A/S; X15=S/T
SEQ ID NO:4: ADN Anc81

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGCGAGTGGTGGG
 ACTTGAAACCTGGAGCCCCGAAACCCAAAGCCAACAGCAAAAGCAGGACGACGGCCGGGGTCTGGTGTCT
 TCCTGGCTACAAGTACCTCGGACCCCTTCAACGGACTCGACAAGGGGGAGCCCGTCAACGGCGGGACGCA
 GCGGCCCTCGAGCACGACAAGGCTTACGACCAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATA
 ACCACGCCGACGCCGAGTTTCAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGC
 AGTCTTCCAGGCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT
 GGAAAGAAGAGACCGGTAGAGCAATCACCCAGGAACAGACTCCTCTXXX1GGCATCGGCAAGAAAGGC
 CAGCAGCCCGCGXXX2AAGAGACTCAACTTTGGGCAGACTGGCGACTCAGAGTCAGTGCCCGACCCCTCAA
 CCACTCGGAGAACCCCCGACGCCCTCTGGTGTGGGATCTAATACAATGGCTGCAGGCGGTGGCGCTC
 CAATGGCAGACAATAACGAAGCGCCGACGGAGTGGGTAATGCCTCAGGAAATTGGCATTGCGATTCCAC
 ATGGCTGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTCCCCACCTACAACAACCACCTC
 TACAAGCAAATCTCCXXX3XXX4CAATCGGGAGGAAGACCAACGACAACACCTACTTTCGGCTACAGCAC
 CCCCTGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATC
 AACAACAACCTGGGGATTCCGGCCCCAAGXXX5CTCAACTTCAAGCTCTTCAACATCCAGGTCAAGGAGGT
 ACGACGAATGATGGCACCACGACCATCGCCAATAACCTTACCAGCAGGTTTACGGTCTTTACGGACTCGG
 AATACCAGCTCCCGTACGTCTCGGCTCTGCGCACCAAGGCTGCCTGCCTCCGTTCGCCGCGGACGTCTT
 CATGATTCTCTCAGTACGGGTACCTGACTCTGAACAATGGCAGTCAGGCGGTGGGCCGTTCCTCTTCTAC
 TGCCTGGAGTACTTTCTTCTCAAATGCTGAGAAGGGCAACAACCTTTXXX6TTCAGTACACGTTTGAG
 GACGTGCCTTTTTCACAGCAGCTACGCGCACAGCCAAAGCCTGGACCGGCTGATGAACCCCTCATCGACC
 AGTACCTGTACTACCTGTCTCGGACTCAGACCACGGGAGGTACCGCAGGAAATXXX7XXX8TTGCAATTT
 TCTCAGGCCGGGCTAGTAGCATGGCGAATCAGGCCAAAACCTGGCTACCCGGGCCCTGTACCGGCAGC
 AACGCGTCTCCAAGACAACGAATCAAAATAACAACAGCAACTTTGCCTGGACCGGTGCCACCAAGTATCA
 TCTGAATGGCAGAGACTCTCTGGTAAATCCCGGTGTGCTATGGCAACCCACAAGGACGACGAAGACCGA
 TTTTTTCCGTCCAGCGGAGTCTTAATATTTGGGAAACAGGGAGCTGGAAATXXX9AACGTGGACXXX10X
 XX11AACGTTATGATAACGXXX12GAGGAAGAAATTAACCACCAACCCAGTGGCCACAGAAGAGTACG
 GCXXX13GTGGCCACTAACCTGCAATCGXXX14AACACCGCTCCTCAAACAGGGACCGTCAACAGTCAAG
 GAGCCTTACCTGGCATGGTCTGGCAGAACCGGACGTGTACCTGCAGGGTCTTATCTGGGCCAAGATTCC
 TCACACGGACGGAACCTTTCATCCCTCGCCGCTGATGGGAGGCTTTGGACTGAAACACCCGCTCCTCAG
 ATCCTGATTAAGAATACACCTGTTCCCGCAATCCTCCAACCTACCTCXXX15CCAGCTAAGTTTGCCTC
 GTTCATCACGCAGTACAGCACCGGACAGGTACGCGTGGAAATTGAATGGGAGCTGCAGAAAGAAAACAGC

AAACGCTGGAACCCAGAGATTCAATACACTTCCAACCTACAACAAATCTACAATGTGGACTTTGCTGTTG
 ACACAGAAGGCGTTTATTCTGAGCCTCGCCCATCGGCACCCGTACCTCACCCGTAATCTG

XXX1 = ACG/AGC; XXX2 = AAA/AAG; XXX3 = AAC/AGT; XXX4 = AGC/CAC; XXX5 = AGA/AAG; XXX6 =
 GAG/CAG; XXX7 = CGG/CAG; XXX8 = ACG/GAG; XXX9 = GAC/AGC; XXX10 = CTT/TAC; XXX11 =
 GAC/AGC; XXX12 = AGT/AAC; XXX13 = GTG/ATC; XXX14 = GCA/AGC; XXX15 = AGT/ACC

SEQ ID NO:5: Polipéptido Anc82

MAADGYLPDWLEDNLSEGIREWWDLPKGPAPKANQQKQDDGRGLVLPGYKYLPGFNGLDKGEFVNAADA
 AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKRVLEPLGLVEEGAKTAP
 GKRPVEQSPQREPDSSX1GIGKKGQQPAX2KRLNFGQTGDSSESVPDPQLGEPPAAPSGVGSNTMAAGG
 GAPMADNNEGADGVGNSSGNWHCDSTWLGDVITTTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGY
 STPWGYFDFNRFHCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVKEVTTNEGTKTIANNLTSTVQVFTD
 SEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTF
 EDVFFHSSYAHQSILDRLMNPLIDQYLYLSRTQTGGTAGTQTLQFSQAGPSSMANQAKNWLPGPCYRQ
 QRVSTTTNQNNSNFAWTGATKYHLNGRDSLVPNGVAMATHKDDERFFPSSGVLI FGKQAGNDNVDS
 NVMITX3EEEIKTTNPVATEEYGVVATNLQSANAPQTGTVNSQGALPGMVWQNRDYYLQGPWAKIPHT
 DGNFHPSPLMGFGFLKHPPPQILIKNTPVPADPPPTTFNQAKLNSFITQYSTGQVSVEIEWELQKENSRRW
 NPEIQYTSNYYKSTNVDFAVNTEGVYSEPRPIGTRYLTRNL

X1=T/S; X2=K/R; X3=S/N

SEQ ID NO:6: ADN Anc82

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTTCGCGAGTGGTGGG
ACCTGAAACCTGGAGCCCCGAAACCCAAAGCCAACCAGCAAAAGCAGGACGACGCGCCGGGTCTGGTGCT
TCCTGGCTACAAGTACCTCGGACCCTTCAACGGACTCGACAAGGGGGAGCCCGTCAACGCGGCGGACGCA
GCGGCCCTCGAGCAGACAAGGCCTACGACCAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATA
ATCACGCCGACGCCGAGTTTCAGGAGCGTCTGCAAGAAGATAACGTCTTTTGGGGGCAACCTCGGGCGAGC
AGTCTTCCAGGCCAAGAAGCGGGTCTCGAACCTCTCGGTCTGGT'TGAGGAAGGCGCTAAGACGGCTCCT
GGAAAGAAGAGACCGGTAGAGCAGTCACCACAGCGTGAGCCCGACTCCTCCXXX1GGCATCGGCAAGAAA
GGCCAGCAGCCCCGCCXXX2AAGAGACTCAATTTTCGGTCAGACTGGCGACTCAGAGTCAGTCCCCGACCT
CAACCTCTCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGATCTAATAACAATGGCTGCAGGCGGTGGCG
CACCAATGGCAGACAATAACGAAGGTGCCGACGGAGTGGTAATTCCTCGGAAATTGGCATTGCGATT
CACATGGCTGGGCGACAGAGTCATCACCAACAGCACCCGAACCTGGGCCCTGCCACCTACAACAACCA
CTCTACAAGCAAATCTCCAACGGGACCTCGGGAGGACAGCAACGACAACACCTACTTTGGGTACAGCA
CCCCCTGGGGGTATTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCAT
CAACAACAACCTGGGGATTCCGGCCCCAAGAGACTCAACTTCAAGCTCTTCAACATCCAGGTCAAAGAGGT
ACGACGAATGAAGGCACCAAGACCATCGCCAATAACCTCACCAGCACCCTCCAGGTGTTTACGGACTCGG
AATACCAGCTGCCGTACGTCTCGGCTCTGCCCACCAGGGCTGCCTGCCTCCGTTCCCGGCGGACGTCTT
CATGATTCTCAGTACGGCTACCTGACTCTCAACAACGGTAGTCAGGCCGTGGGACGTTCTCTCTTCTAC
TGCCTGGAGTACTTCCCTCTCAGATGCTGAGAACGGGCAACAACCTTCAATTACGCTACACTTTCGAGG
ACGTGCCTTTCCACAGCAGCTACGCGCACAGCCAGAGTTTGGACAGGCTGATGAATCCTCTCATCGACCA
GTACCTGTACTACCTGTCAAGAACCCAGACTACGGGAGGACAGCGGGAAACCCAGACGTTGCAGTTTTCT
CAGGCCGGGCTAGCAGCATGGCGAATCAGGCCAAAACTGGCTGCCTGGACCCTGCTACAGACAGCAGC
GCGTCTCCACGACAACGAATCAAAACAACAACAGCAACTTGCCTGGACTGGTGGCACAAGTATCATCT
GAACGGCAGAGACTCTCTGGTGAATCCGGGCGTCCCATGGCAACCCACAAGGACGACGAGGACCGCTT
TTCCATCCAGCGGCGTCTCATATTTGGCAAGCAGGGAGCTGGAAATGACAACGTGGACTATAGCAACG
TGATGATAACXXX3GAGGAAGAAATCAAGACCACCAACCCGTGGCCACAGAAGAGTATGGCGTGGTGG
CTACTAACCTACAGTCGGCAACACCGCTCCTCAAACGGGGACCGTCAACAGCCAGGGAGCCTTACCTGG
CATGGTCTGGCAGAACCGGACGTGTACCTGCAGGGTCTATTTGGGCCAAGATTCTCTCACACAGATGGC
AACTTTACCCGCTCTCTTTAATGGGCGCTTTGGACTTAAACATCCGCTCCTCAGATCCTCATCAAAA

ACACTCCTGTTTCTCGGGATCCTCCAACAACGTTCAACCAGGCCAAGCTGAATCTTTTCATCACGCAGTA
CAGACCCGGACAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAGAACAGCAAGCGCTGGAACCCA
GAGATTAGTATACTTCCAATACTACAAATCTACAAATGTGGACTTTGCTGTTAATACTGAGGGTGT
ACTCTGAGCCTCGCCCCATTGGCACTCGTTACCTACCCGTAATCTG

XXX1 = ACG/AGC; XXX2 = AAA/AGA; XXX-3 = AGC/AAC

5

SEQ ID NO:7: Polipéptido Anc83

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
GKKRPVEQSPQREPDSX1GIGKKGQQPAX2KRLNFGQTGDSESVDPQPLGEPFAAPSGVGSNTMAAGG
GAPMADNNEGADGVGSSSGNWHCDSTWLGDRIITSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGY
STPWGYFDNRFHCHFSRDPWQRLINNNWGRPKRLX3FKLFNIQVKEVTQNEGKTIANNLTSTIQVFT
DSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTNNGSQAVGRSSFYCLEYFPSQMLRTGNNFX4FSY
TFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRTQTGGTAGTQTLQFSQAGPSX5MANQAKNWLPGPC
YRQQRVSTTTSONNNSNFAWTGATKYHLNGRDSLVPVAMATHKDDX7RFFPSSGX7LIFGKQGAGKD
NVDYSNVMILTSEEEIKTTNPVATEEYGVVADNLQQNTAPQX8GTVNSQALPGMVWQNRDVLQGPPIWA
KIPHTDGNFHPSPMLMGFGLKHPPQILIKNTPVPADPPTTFNQAKLNSFITQYSTGQVSVEIEWELQKE
NSKRWNPEIQYTSNYKSTNVDFAVNTEGVYSEPRPIGTRYLTRNL

X1=T/S; X2=R/K; X3=N/S; X4=Q/E; X5=N/T/S; X6=D/E; X7=I/V; X8=I/V

SEQ ID NO:8: ADN Anc83

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGCGAGTGGTGGG
 ACCTGAAACCTGGAGCCCCGAAACCCAAAGCCAACCAGCAAAAGCAGGACGACGGCCGGGGTCTGGTGCT
 TCCTGGCTACAAGTACCTCGGACCCTTCAACGGACTCGACAAGGGGGAGCCCGTCAACGCGGCGGACGCA
 GCGGCCCTCGAGCACGACAAGGCCCTACGACCAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATA
 ATCAGCCGACGCGGAGTTTTCAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGC
 AGTCTTCCAGGCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT
 GGAAAGAAGAGACCGGTAGAGCAGTCACCACAGCGTGAGCCCGACTCCTCCXX1GGCATCGGCAAGAAA
 GGCCAGCAGCCCGCCXX2AAGAGACTCAATTCGGTCAGACTGGCGACTCAGAGTCAGTCCCCGACCCCT
 CAACCTCTCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGATCTAATAACAATGGCTGCAGGCGGTGGCG
 CACCAATGGCAGACAATAACGAAGGTGCCGACGGAGTGGGTAGTTCCTCGGGAAATTGGCATTGCGATTTC
 CACATGGCTGGGCGACAGAGTCATCACCACCAGCACC CGAACCTGGGCCCTGCCACCTACAACAACCAC
 CTCTACAAGCAAATCTCCAACGGGACCTCGGGAGGACGACCAACGACAACACCTACTTTGGCTACAGCA
 CCCCCGGGGGTATTTTGAATTTTAAAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCAT
 CAACAACAACCTGGGGATTCCGGGCCAAGAGACTCXX3TTCAAGCTCTTCAACATCCAGGTCAAAGAGGT
 CACGAGAATGAAGGCACCAAGACCATCGCCAATAACCTCACCAGCACCATCCAGGTGTTTACGGACTCG
 GAATACCAGCTGCCGTACGTCTCGGCTCTGCCACCAGGGCTGCCTGCCTCCGTTCGCCGGCGGACGTCT
 TCATGATTCTCAGTACGGCTACCTGACTCTCAACAACGGTAGTCAGGCCGTGGGACGTTCTCTCTTCTA
 CTGCCTGGAGTACTTCCCCTCTCAGATGCTGAGAACGGGCAACAACCTTTXX4TTCAGCTACACTTTCGA
 GGACGTGCCTTTCCACAGCAGCTACGCGCACAGCCAGAGTTTGGACAGGCTGATGAATCCTCTCATCGAC
 CAGTACCTGTACTACCTGTCAAGAACCAGACTACGGGAGGCACAGCGGGAACCCAGACGTTGCAGTTTT
 CTCAGGCCGGGCTAGCXX5ATGGCGAATCAGGCCAAAACCTGGCTGCCTGGACCTGCTACAGACAGC
 AGCGCGTCTCCACGACAACGTGCGCAAAACAACAACAGCAACTTGCCTGGACTGGTGCCACCAAGTATCA
 TCTGAACGGCAGAGACTCTCTGGTGAATCCGGGCGTCGCCATGGCAACCCACAAGGACGACGAGXX6CG
 CTTCTTCCCATCCAGCGGCXX7CTCATATTTGGCAAGCAGGGAGCTGGAAAAGACAACGTGGACTATAG
 CAACGTGATGCTAACAGCAGGAGGAAGAAATCAAGACCACCAACCCCGTGGCCACAGAAGAGTATGGCGTG
 GTGGCTGATAACCTACAGCAGCAAAACACCGCTCCTCAAXX8GGGACCGTCAACAGCCAGGGAGCCTTA
 CCTGGCATGGTCTGGCAGAACCAGGACGTGTACCTGCAGGGTCTATTGGGGCAAGATTCTCTACACAG
 ATGGCAACTTTACCCGTCTCCTTTAATGGGCGGCTTTGGACTTAAACATCCGCTCCTCAGATCCTCAT
 CAAAAACACTCCTGTTCTGCGGATCCTCCAACAACGTTCACCAGGCCAAGCTGAATTCTTTCATCAGC

CAGTACAGCACCGGACAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAGAACAGCAAGCGCTGGA
 ACCCAGAGATTCACTATACTTCCAATACTACAAATCTACAAATGTGGACTTTGCTGTTAATACTGAGGG
 TGTCTACTCTGAGCCTCGCCCCATTGGCACTCGTTACCTCACCCTGAATCTG
 XXX1 = ACG/AGC; XXX2 = AGA/AAG; XXX3 = AAC/AGC; XXX4 = CAA/GAA; XXX5 = AAC/ACC/AGC;
 XXX6 = GAC/GAG; XXX7 = ATC/GTC; XXX8 = ATA/GTA

SEQ ID NO:9: Polipéptido Anc84

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFNGLDKGEPVNAADA
 AALEHDKAYDQQLKAGDNPYLRYNHADAEFQERLQEDTSGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
 GKRRPVEPSPQRS PDSSTGIGKKQQPAX1KRLNFGQTGDSESVDPQPIGEPPAAPSGVSGSMTAAGGG
 APMADNNEGADGVGSSSGNWHCDSTWLGDRVITSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYS
 TPWGYFDNFRFHCHFSRDPWQRLINNNWGFPRKRLX2FKLFNIQVKEVTQNEGKTIANNLTSTIQVFTD
 SEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTNLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFEFSYTF
 EDVFPFHSSYAHQS LDRMLNPLIDQYLYLSRTQSTGGTAGTQQLLFSQAGPSNMSAQAKNWLPGPCYRQ
 QRVSTTSLQNNNSNFAWTGATKYHLNGRDSLVPVAMATHKDDX3RFFPSSGX4LMFGKQAGKDNVD
 YSNVMTLSEEEIKTTNPVATEQYGVVADNLQQNTAPIVGAVNSQALPGMVWQNRDVYLQGPWAKIPH
 TDGNFHPSPMLMGFGFLKHPPQILIKNTVPADPPTTFNQAKLNSFITQYSTGQVSVEIEWELQKENS KR
 WNPEIQYTSNYYKSTNVDFAVNTEGVYSEPRPIGTRYLTRNL

X1=R/K; X2=N/S; X3=D/E; X4=I/V

SEQ ID NO:10: ADN Anc84

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTCGCGAGTGGTGGG
 ACCTGAAACCTGGAGCCCCGAAACCCAAAGCCAACCAGCAAAAGCAGGACGACGGCCGGGGTCTGGTGCT
 TCCTGGCTACAAGTACCTCGGACCCCTCAACGGACTCGACAAGGGGGAGCCCGTCAACGCGGCGGACGCA
 GCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAAAGCGGGTGAACAATCCGTACCTGCGGTATA
 ATCAGCCCGACGCCGAGTTTCAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGC
 AGTCTTCCAGGCCAAGAAGCGGGTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT
 GGAAAGAAGAGACCGGTAGAGCCGTACCACAGCGTCCCCCGACTCCTCCACGGGCATCGGCAAGAAAG
 GCCAGCAGCCCCGCCXXX1AAGAGACTCAATTTTCGGTCAGACTGGCGACTCAGAGTCAGTCCCCGACCCCTC
 AACCTATCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGATCTGGTACAATGGCTGCAGGCGGTGGCGC
 ACCAATGGCAGACAATAACGAAGGTGCCGACGGAGTGGGTAGTTCTCGGGAAATTGGCATTGCGATTCC
 ACATGGCTGGGCGACAGAGTCATCACCACAGCACCCGAACCTGGGCCCTGCCACCTACAACAACACACC
 TCTACAAGCAAATCTCCAACGGGACCTCGGGAGGACGACCAACGACAACACCTACTTTGGCTACAGCAC
 CCCCTGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATC
 AACAACTTGGGGATTCCGGCCCAAGAGACTCXXX2TTCAAGCTCTTCAACATCCAGGTCAAAGAGGTG
 ACGCAGAATGAAGGCACCAAGACCATCGCCAATAACCTCACCAGCACCATCCAGGTGTTTACGGACTCGG
 AATACAGCTGCGGTACGTCTCGGCTCTGCCACAGGGCTGCCTGCCTCCGTTCGGCGGACGTCTT
 CATGATTCTCTCAGTACGGCTACCTGACTCTCAACAACGGTAGTCAAGCCGTGGGACGTTCTCTCTAC
 TGCCTGGAGTACTTCCCCTCTCAGATGCTGAGAACGGGCAACAACCTTTGAGTTTCACTTTCGAGG
 ACGTGCCCTTTCCACAGCAGCTACGCGCACAGCCAGAGTTTGGACAGGCTGATGAATCCTCTCATCGACCA
 GTACCTGTACTACCTGTCAAGAACCAGTCTACGGGAGGCACAGCGGGAACCCAGCAGTTGCTGTTTCT
 CAGGCCGGGCTAGCAACATGTCGGCTCAGGCCAAAACTGGCTGCCTGGACCCCTGCTACAGACAGCAGC
 GCGTCTCCACGACACTGTGCAAAACAACAACAGCAACTTTGCCTGGACTGGTGCCACCAAGTATCATCT
 GAACGGCAGAGACTCTCTGGTGAATCCGGGCGTCCCATGGCAACCCACAAGGACGACGAGXXX3CGCTT
 CTTCCATCCAGCGGCXXX4CTCATGTTTGGCAAGCAGGGAGCTGGAAAAGACAACGTGGACTATAGCAA
 CGTGATGCTAACCGAGGAGGAAGAAATCAAGACCACCAACCCCGTGGCCACAGAACAGTATGGCGTGGTG
 GCTGATAACCTACAGCAGCAAAACACCGCTCCTATTGTGGGGGCCGTC AACAGCCAGGGAGCCTTACCTG
 GCATGGTCTGGCAGAACCGGGACGTGTACCTGCAGGGTCTTATTTGGGCCAAGATTCTCACACAGATGG
 CAACCTTTCACCCGTCTCCTTTAATGGGCGGCTTTGGACTTAAACATCCGCCTCCTCAGATCCTCATCAA
 AACACTCCTGTTCTGCGGATCCTCCAACAACGTTCAACCAGGCCAAGCTGAATTCTTTCATCACGAGT

ACAGCACCGGACAAGTCAGCGTGGAGATCGAGTGGGAGCTGCAGAAGGAGAACAGCAAGCGCTGGAACCC
 AGAGATTCACTATACTTCCAATACTACAAATCTACAAATGTGGACTTTGCTGTTAATACTGAGGGTGT
 TACTCTGAGCCTCGCCCCATTGGCACTCGTTACCTCACCCGTAATCTG
 XXX-1 = AGA/AAA; XXX2 = AAC/AGC; XXX3 = GAC/GAG; XXX4 = ATC/GTC

5

SEQ ID NO:11: Polipéptido Anc94

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEFVNAADA
 AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFAQKKRVLEPLGLVEEGAKTAP
 GKRRPVEFSPQSRPDSSTGIGKKGQQPAKKRLNFGQTGDSSEVPDPQPIGEPPAGPSGLSGTMAAGGGA
 PMADNNEGADGVGSSSGNWHCDSTWLGDVITTTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYST
 PWGYFDNRFHCHFSPRDWQRLINNNWGFPRKRLNFKLFNIQVKEVTQNEGKTIANNLSTSIQVFTDSE
 YQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFESYTFED
 VPFHSSYAHSQLDRLMNPLIDQYLYLSRTQSTGGTAGTQQLLFSQAGPX1NMSAQAKNWLPGPCYRQQ
 RVSTTSLQNNNSNFAWTGATKYHLNGRDSLVPNGVAMATHKDDEERFFPSSGVLMFGKQAGKDNVDYSS
 VMLTSEEEIKTTNPVATEQYGVVADNLQQNTAPIVGAVNSQALPGMVWQNRDVYLQGPWAKIPHTDG
 NFHPSPLMGGFGLKHPPQILIKNTPVPADEPTTFSQAKLASFITQYSTGQVSVEIEWELQKENS KRWN
 EIQYTSNYYKSTNVDFAVNTEGTYSEPRPIGTRYLTRNL
 X1=S/N

SEQ ID NO:12: ADN Anc94

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTTCGCGAGTGGTGGG
 ACTTGAAACCTGGAGCCCCGAAACCCAAAGCCAACCAGCAAAAGCAGGACGACGGCCGGGTCTGGTGCCT
 TCCTGGCTACAAGTACCTCGGACCCCTTCAACGGACTCGACAAGGGGGAGCCCGTCAACGCGGCGGACGCA
 GCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATA
 ACCACGCGGACGCGGAGTTTCAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGC
 AGTCTTCCAGGCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT
 GGAAAGAAGAGACCGGTAGAGCCATCACCCAGCGTTCTCCAGACTCCTCTACGGGCATCGGCAAGAAAG
 GCCAGCAGCCCGCAAAAAGAGACTCAACTTTGGGCAGACTGGCGACTCAGAGTCAGTGCCCGACCCCTCA
 ACCAATCGGAGAACCCCCCGAGGCCCTCTGGTCTGGGATCTGGTACAATGGCTGCAGGCGGTGGCGCT
 CCAATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGTAGTTTCTCAGGAAATTGGCATTGCGATTCCA
 CATGGCTGGGCGACAGAGTCATCACCACCAGCACCCGAACCTGGGCCCTCCCCACCTACAACAACCACCT
 CTACAAGCAAATCTCCAACGGGACTTCGGGAGGAAGCACCAACGACAACACCTACTTCGGCTACAGCACC
 CCCTGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCA
 ACAACAACCTGGGGATTCCGGCCCAAGAGACTCAACTTCAAGCTCTTCAACATCCAGGTCAAGGAGGTCAC
 GCAGAATGAAGGCACCAAGACCATCGCCAATAACCTTACCAGCAGATTAGGTCTTTACGGACTCGGAA
 TACCAGTCCCGTACGTCTCGGCTCTGCGCACACGAGGCTGCCTGCCTCCGTCCCGGCGGACGTCTTCA
 TGATTCTCAGTACGGGTACCTGACTCTGAACAATGGCAGTCAAGGCCGTGGGCCGTTCCCTCTTCTACTG
 CCTGGAGTACTTTCTTCTCAAATGCTGAGAACGGGCAACAACCTTTGAGTTCAGCTACACGTTTGAGGAC
 GTGCCTTTTTCACAGCAGCTACGCGCACAGCCAAAGCCTGGACCGGTGATGAACCCCTCATCGACCAGT
 ACCTGTACTACCTGTCTCGGACTCAGTCCACGGGAGGTACCGCAGGAACCTCAGCAGTTGCTATTTTCTCA
 GGCCGGGGCTXXXAACATGTCGGCTCAGGCCAAAAAAGTGGCTACCCGGGCCCTGCTACCCGGCAGCAACGC
 GTCTCCACGACACTGTGCGAAAAATAACAACAGCAACTTTGCTGACCGGTGCCACCAAGTATCATCTGA
 ATGGCAGAGACTCTCTGGTAAATCCCGGTGTCGCTATGGCAACCCACAAGGACGACGAAGAGCGATTTTT
 TCCGTCCAGCGGAGTCTTAATGTTTGGGAAACAGGGAGCTGGAAAAGACAACGTGGACTATAGCAGCGTT
 ATGCTAACCAAGTGAAGGAAGAAATTAACCAACCAACCCAGTGGCCACAGAACAGTACGGCGTGGTGGCCG
 ATAACCTGCAACAGCAAAACACCGCTCCTATTGTAGGGGCCGTCAACAGTCAAGGAGCCCTACCTGGCAT
 GGTCTGGCAGAACCGGGACGTGTACCTGCGGGTCCTATCTGGGCCAAGATTCTCTACACGACGGAAC
 TTTTATCCCTCGCGCTGATGGGAGGCTTTGGACTGAAACACCCGCCTCCTCAGATCCTGATTAAGAATA
 CACCTGTTCGCGGATCCTCAACTACCTTCAGTCAAGCTAAGCTGGCGTCGTTTCATCAGCAGTACAG
 CACCGACAGGTCAAGCTGGAATTTGAATGGGAGCTGCAGAAAGAAAACAGCAACGCTGGAACCCAGAG

ATTCAATACACTTCCAATACTACTACAAATCTACAAATGTGGACTTTGCTGTTAACACAGAAGGCATTATT
 CTGAGCCTCGCCCCATCGGCACCCGTTACCTCACCCGTAATCTG

XXX1 = AGT/AAT

5

SEQ ID NO:13: Polipéptido Anc113

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFNGLDKGEPVNAADA
 AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
 GKRRPVEX1SPQRSPTSSTGIGKKGQQPAX2KRLNFGQTGDSESVDPQPLGEPAPPSGVSGTMAAGG
 GAPMADNNEGADGVGNASGNWHDSTWLGDRTVITTTSTRTWALPTYNNHLYKQISSQSAGSTNDNTYFGYS
 TPWGYFDENRFHCHFSPRDWQRLINNNWGFPRPKKLX3FKLFNIQVKEVTNDGVTTIANNLTSTVQVFS
 SEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLTLNNGSQSVGRSSFYCLEYFSPQMLRTGNNFEFSYTF
 EDVPFHSSYAHQSGLDRMLNPLIDQYLYLARTQSTTGGTGNRELQFX4QAGPSTMAEQAKNWLPGPCY
 RQQRVSKTLDQNNNSNFAWTGATKYHLNGRNSLVNPGVAMATHKDDERFFPSSGVLI FGKTGAANKTTL
 ENVLMTX5EEEIKTTNPVATEEYGX6VSSNLQX7NTAPQTQTVNSQGALPGMVWQNRDQVYLQGPWIWAKI
 PHTDGNFHPSPLMGGFGLKHPPPQILIKNTVPANPPEVFTPAKFASFITQYSTGQVSVEIEWELQKENS
 KRWNPEIQYTSNYDKSTNVDFAVDSEGVYSEPRPIGTRYLTRNL

X-1=P/Q; X-2=K/R; X-3=R/N; X4=Y/H; X-5=N/S; X-6=V/I; X-7=A/S

SEQ ID NO:14: ADN Anc113

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACCTCTCTGAGGGCATTTCGCGAGTGGTGGG
 ACCTGAAACCTGGAGCCCCGAAACCCAAAGCCAACCAGCAAAAGCAGGACGACGGCCGGGTCTGGTGCT
 TCCTGGCTACAAGTACCTCGGACCCTTCAACGGACTCGACAAGGGGAGCCCGTCAACGCGGCGGACGCA
 GCGGCCCTCGAGACGACAAGGCCTACGACCAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATA
 ACCACGCCGACGCCGAGTTTCAGGAGCGTCTGCAAGAAGATACGTCATTTGGGGGCAACCTCGGCGGAGC
 AGTCTTCCAGGCCAAGAAGCGGGTTCTCGAACCTCTCGGTCTGGTTGAGGAAGGCGCTAAGACGGCTCCT
 GGAAAGAAGAGACCGGTAGAGXXX1TCACCTCAGCGTTCCCCCGACTCCTCCACGGGCATCGGCAAGAAA
 GGCCAGCAGCCCCGCCXXX2AAGAGACTCAATTTGGTTCAGACTGGCGACTCAGAGTCAGTCCCCGACCCT
 CAACCTCTCGGAGAACCTCCAGCAGCGCCCTCTGGTGTGGGATCTGGTACAATGGCTGCAGGCGGTGGCG
 CACCAATGGCAGACAATAACGAAGGTGCCGACGGAGTGGGTAATGCCTCAGGAAATTGGCATTGCGATTTC
 CACATGGCTGGGCGACAGAGTCATTACCACCAGCACCAGAACCTGGGCCCTGCCACCTACAACAACCAC
 CTCTACAAGCAAATCTCCAGTCAAAGTGCAGGTAGTACCAACGACAACACCTACTTCGGCTACAGCACCC
 CCTGGGGGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATCAA
 CAACAACCTGGGGATTCCGGCCCAAGAAGCTGXXX3TTCAAGCTCTTCAACATCCAGGTCAAGGAGGTCAC
 GACGAATGACGGCGTTACGACCATCGCTAATAACCTTACCAGCACGGTTTCAAGTATTCTCGGACTCGGAA
 TACCAGCTGCCGTACGTCCTCGGCTCTGCGCACAGGGCTGCCTGCCTCCGTTCGCGGCGGACGCTTTCA
 TGATTCTCAGTACGGCTACCTGACTCTCAACAATGGCAGTCAGTCTGTGGGACGTTCTCTCTTCTACTG
 CCTGGAGTACTTCCCCTCTCAGATGCTGAGAACGGGCAACAACCTTTGAGTTTCACTACACCTTCGAGGAC
 GTGCCTTTCCACAGCAGCTACGCACACAGCCAGAGCCTGGACCGGCTGATGAATCCCCTCATCGACCAGT
 ACTTGTACTACCTGGCCAGAACACAGAGTACCACAGGAGGCACAGCTGGCAATCGGGAACCTGCAGTTTXX
 X4CAGGCCGGGCTTCAACTATGGCCGAACAAGCCAAGAATTGGTTACCTGGACCTTGCTACCGGCAACA
 AAGAGTCTCCAAAACGCTGGATCAAAACAACAACAGCAACTTTGCTTGGACTGGTGCCACCAAAATATCAC
 CTGAACGGCAGAACTCGTTGGTTAATCCCGGCGTCGCCATGGCAACTCACAAGGACGACGAGGACCGCT
 TTTTCCCATCCAGCGGAGTCTGATTTTGGAAAACTGGAGCAGCTAACAAAACCTACATTGGAAAAATGT
 GTTAATGACAXXX5GAAGAAGAAATTAAACTACTAATCCTGTAGCCACGGAAGAATACGGGXXX6GTCA
 GCAGCAACTTACAATCGXXX7AATACTGCACCCCAGACACAACTGTCAACAGCCAGGGAGCCTTACCTG
 GCATGGTCTGGCAGAACCGGGACGTGTACCTGCAGGGTCCCCTCTGGGCCAAGATTCTCACACGGATGG
 CAACTTTACCCGCTCTCTTTGATGGGCGGCTTTGGACTTAAACATCCGCCTCCTCAGATCCTGATCAAG
 AACACTCCCGTTCGCGTAATCCTCCGAGGTGTTTACTCCTGCCAAGTTTGTCTCGTTTCATCACACAGT
 ACAGCACCGGACAAGTCAGCGTGGAAATCGAGTGGGAGCTGCAGAAGGAAAACAGCAAGCGCTGGAACCC

GGAGATTCAGTACACCTCCAACATATGATAAGTCGACTAATGTGGACTTTGCCGTTGACAGCGAGGGTGT
 TACTCTGAGCCTCGCCCTATTGGCACTCGTTACCTCACCCGTAATCTG

XXX1 = CCG/CAG; XXX2 = AAA/AGA; XXX3 = CGG/AAC; XXX4 = TAC/CAC; XXX5 = AAT/AGT; XXX6 =
 GTA/ATA; XXX7 = GCT/TCT

SEQ ID NO:15: Polipéptido Anc126

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEFVNAADA
 AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
 GKRPVEQSPQEPDSSS GIGKX1GQQPAX2KRLNFGQTGDSESVDPQPLGEPPAAPSGVGSNTMASGGG
 APMADNNEGADGVGNX3SGNWHCDSTWLGDRVITTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYST
 PWGYFDNRFHCHFSPRDWQRLINNNWGFPRKX4LNFKLFNIQVKEVTNDGTTTIANNLTSTVQVFTDS
 EYQLPYVLGSAHQGCLPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFX5FSYTF
 EDVPPFHSSYAHQS LDRMLNPLIDQYLYYLX6RTQTTSGTAQNRELX7FSQAGPSSMX8NQAKNWLPGPC
 YRQQRVSKTANDNNNSNFAWTGATKYHLNGRDSLVPNGPAMASHKDDKDFPMSGVLIFGKQGAGASNV
 DLDNVMITDEEEIKTTNPVATEQYGTVATNLQSSNTAPATGTVNSQGALPGMVWQDRDVYLQGPWAKIP
 HTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQKENS
 RWNPEIQYTSNYSX9NVDFTVDTNGVYSEPRPIGTRYLTRNL

X1=S/T; X2=K/R; X3=A/S; X4=R/K; X5=T/Q; X6=S/N; X7=Q/L; X8=A/S; X9=A/T

SEQ ID NO:16: ADN Anc126

ATGGCTGCCGATGGTTATCTTCCAGATTGGCTCGAGGACAACTCTCTGAGGGCATTTCGCGAGTGGTGGG
 ACTTGAAACCTGGAGCCCCGAAACCCAAAGCCAACAGCAAAAGCAGGACGACGGCCGGGTCTGGTGC
 TCCTGGCTACAAGTACCTCGGACCCCTCAACGGACTCGACAAGGGGGAGCCCGTCAACGCGGCGGATGCA
 GCGGCCCTCGAGCACGACAAGGCCTACGACCAGCAGCTCAAAGCGGGTGACAATCCGTACCTGCGGTATA
 ACCACGCGGACGCCGAGTTTCAGGAGCGTCTGCAAGAAGATACGTCTTTTGGGGGCAACCTCGGGCGAGC
 AGTCTTCCAGGCCAAGAAGAGGGTTCTCGAACCTCTTGGTCTGGTTGAGGAAGGTGCTAAGACGGCTCCT
 GGAAAGAAACGTCCGGTAGAGCAGTCGCCACAAGAGCCAGACTCCTCCTCGGGCATTGGCAAGXXX1GGC
 CAGCAGCCCGCTXXX2AAGAGACTCAATTTTGGTCAGACTGGCGACTCAGAGTCAGTCCCCGACCCACAA
 CCTCTCGGAGAACCTCCAGCAGCCCCCTCTGGTGTGGGATCTAATACAATGGCTTCAGGCGGTGGCGCAC
 CAATGGCAGACAATAACGAAGGCGCCGACGGAGTGGGTAATXXX3TCAGGAAATTGGCATTTGCGATTCCA
 CATGGCTGGGCGACAGAGTCATCACCACCAGCACCCGAACATGGGCCTTGCCACCTATAACAACACCT
 CTACAAGCAAATCTCCAGTCAATCAGGGGCCAGCAACGACAACCACTACTTCGGCTACAGCACCCCTGG
 GGGTATTTTGATTTCAACAGATTCCACTGCCATTTCTCACCACGTGACTGGCAGCGACTCATCAACAACA
 ATTGGGGATTCCGGCCCAAGXXX4CTCAACTTCAAGCTCTTCAACATCCAAGTCAAGGAGGTACGACGA
 ATGATGGCACCACGACCATCGCTAATAACCTTACCAGCAGGTTCAAGTCTTCACGGACTCGGAGTACCA
 GTTGGCGTACGTCTCGGCTCTGCGCACAGGGCTGCCTCCCTCCGTTCCCGGCGGACGTGTTTCATGATT
 CCGCAGTACGGCTACCTAACGCTCAACAATGGCAGCCAGGCAAGTGGGACGGTCATCCTTTTACTGCCTGG
 AATATTTCCCATCGCAGATGCTGAGAACGGGCAATAACTTTXXX5TTCAGCTACACCTTCGAGGACGTGC
 CTTTCCACAGCAGCTACGCGCACAGCCAGAGCCTGGACCGGTGATGAATCCTCTCATCGACAGTACCT
 GTATTACCTGXXX6AGAAGTACAGTACGTCCGGAAGTGGCCAAAACAGGGAGTTGXXX7TTTAGCCAGG
 CGGGTCCATCTAGCATGXXX8AATCAGGCCAAAACCTGGTACCTGGACCTGTTACCGGCGAGCAGCGCG
 TTTCTAAAACAGCAAATGACAACAACACAGCAACTTTGCCTGGACTGGTGCTACAAAATATCACCTTAA
 TGGGCGTGATTCTTTAGTCAACCCTGGCCCTGCTATGGCCTCACACAAAGACGACGAAGACAAGTTCTTT
 CCCATGAGCGGTGTCTTGATTTTGGAAAGCAGGGCGCGGAGCTTCAAACGTTGATTTGGACAATGTCA
 TGATCAGACGAAGAGGAAATCAAACCACTAACCCCGTGGCCACCGAACAATATGGGACTGTGGCAAC
 CAATCTCCAGAGCAGCAACACAGCCCCTGCGACCGGAAGTGTGAATTTCTCAGGGAGCCTTACCTGGAATG
 GTGTGGCAAGACAGAGACGTATACCTGCAGGGTCTTATTTGGGCCAAAATCCTCACACGGATGGACACT
 TTCACCCGTCTCCTCTCATGGGCGGCTTTGGACTTAAGCACCCGCTCCTCAGATCCTCATCAAAAACAC
 GCCTGTTCTGCGAATCCTCCGACAACGTTTTCGCCTGCAAAGTTTGCTTCATTATCACCCAGTATTC
 ACAGGACAAGTGAGCGTGAGATTGAATGGGAGCTGCAGAAAGAAAACAGCAAACGCTGGAATCCCGAAA

TACAGTATACATCTAACTATAATAAACTCTXXX9AACGTTGATTTCACTGTGGACACCAATGGAGTTTATA
 GTGAGCCTCGCCCCATTGGCACCCGTTACCTACCCGTAACCTG

XXX1 = TCA/ACA; XXX2 = AAA/AGA; XXX3 = GCC/TCC; XXX4 = AGA/AAA; XXX5 = ACC/CAG; XXX6 =
 AGC/AAC; XXX7 = CAG/CTG; XXX8 = GCT/TCT; XXX9 = GCC/ACC

SEQ ID NO:17: Polipéptido Anc127

MAADGYLPDWLEDNLSEGIREWDLKPGAPQPKANQQHQDDX1RGLVLPGYKYLGPFNGLDKGEPVNEAD
 AAALHDKAYDQQLKAGDNPYLKYNHADADEFQERLQEDTSFGGNLGRAVFAQAKRVLEPLGLVEEAAKTA
 PGKKRPVEQSPQEPDSSSGIGKSGQPAX2KRLNFGQTGDSESVDPQPLGEPAPPSGVGSNTMASGGG
 APMADNNEGADGVGNSSGNWHCDSTWLGDRIITSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTP
 WGYFDNRFHCHFSRWDQRLINNNWGFPRKX3LNFKLFIQVKEVTQNDGTTTIANNLSTVQVFTDSE
 YQLPYVLGSAHQGLPFPFADVFMIPOYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNF4FSYTFE
 DVPFHSSYAHQSGLDRLMNPIDQYLYYLX5RTQTSTGTTQQSRLX6FSQAGPSSMX7QQAX8NWLPGPC
 YRQQRVSKTANDNNNSNFAWTX9ATKYHLNGRDSLVPNPGPAMASHKDDEEKFFPMHG10LIFGKQGTGA
 SNVDLNDVMITDEEEIRTNPVATEQYGTVATNLQSSNTAPATGTVNSQGALPGMVWQDRDVYLQGPPIWA
 KIPHTDGHFHPSPLMGGFGLKHPPPILIKNTPVANPPTTFSPAKFASFITQYSTGQVSVEIEWELQKE
 NSKRWNPEIQYTSNYNKSXVNVDFTVDTNGVYSEPRPIGTRYLTRNL

X1=G/S; X2=R/K; X3=K/R; X4=T/Q; X5=S/R; X6=Q/L; X7=A/S; X8=K/R; X9=G/A; X10=V/N

SEQ ID NO:18: ADN Anc127

ATGGCTGCTGACGGTTATCTTCCAGATTGGCTCGAGGACAACCTTTCTGAAGGCATTCGTGAGTGGTGGG
ATCTGAAACCTGGAGCCCCTCAACCCAAAGCGAACCACCAACACCAGGACGACXXX1CGGGGTCTTGTGC
TTCCGGGTACAAATACCTCGGACCCTTTAACGGACTCGACAAAGGAGAGCCGGTCAACGAGGCGGACGC
GGCAGCCCTCGAACACGACAAAGCTTACGACCAGCAGCTCAAGGCCGGTGACAACCCGTACCTCAAGTAC
AACCACGCCGACGCCGAGTTTCAGGAGCGTCTTCAAGAAGATACGTCTTTTGGGGCAACCTTGGCAGAG
CAGTCTTCCAGGCCAAAAAGAGGGTCTTGGCTCTTGGTCTGGTTGAGGAAGCAGCTAAAACGGCTCC
TGGAAAGAAGAGGCCTGTAGAACAGTCTCCTCAGGAACCGGACTCATCATCTGGTATTGGCAAATCGGGC
CAACAGCCTGCCXXX2AAAAGACTAAATTTCCGTGAGACTGGAGACTCAGAGTCAGTCCCAGACCCCTCAA
CCTCTCGGAGAACCACAGCAGCCCCCTCAGGTGTGGGATCTAATACAATGGCTTCAGGCGGTGGCGCAC
CAATGGCAGACAATAACGAGGGTGCCGATGGAGTGGGTAATTCCTCAGGAAATTTGGCATTGGCATTCCAC
ATGGCTGGGCGACAGAGTCATCACCACCAGCACCAGAACCTGGGCCCTGCCACTTACAACAACCATCTC
TACAAGCAAATCTCCAGCCAATCAGGAGCTTCAAACGACAACCACTACTTTGGCTACAGCACCCCTTGGG
GGTATTTTGACTTTAACAGATTCCACTGCCACTTCTCACCACGTGACTGGCAGCGACTCATTAAACACAA
CTGGGGATTCCGGCCCAAGXXX3CTCAACTTCAAGCTCTTCAACATCCAAGTTAAAGAGGTACGCGAGAA
CGATGGCAGCAGCACTATTGCCAATAACCTTACCAGCAGGTTCAAGTGTTCAGGACTCGGAGTATCAG
CTCCCGTACGTGCTCGGGTCGGCGCACCAAGGCTGTCTCCCGCGTTCAGCGCGAGCTCTTCATGATCC
CTCAGTATGGATACCTCACCTGAACAACGGAAGTCAAGCGGTGGGACGCTCATCTTTTACTGCCTGGA
GTACTTCCCTTCGAGATGCTAAGGACTGGAAATAACTTCXXX4TTCAGCTATACCTTCGAGGATGTACC
TTTTTACAGCAGCTACGCTCACAGCCAGAGTTTGGATCGCTTGATGAATCCTCTTATTGATCAGTATCTG
TACTACCTGXXX5AGAACGCAAACAACCTCTGGAACAACCCAACAATCACGGCTGXXX6TTTAGCCAGGC
TGGCCCTTCGTCTATGXXX7CAGCAGGCCXXX8AATTGGCTACCTGGGCCCTGCTACCGGCAACAGAGAG
TTTCAAAGACTGCTAACGACAACAACAACAGTAACCTTTGCTTGGACAXXX9GCCACCAAATATCATCTCA
ATGGCCGCGACTCGCTGGTGAATCCAGGACCAGCTATGGCCAGTCACAAGGACGATGAAGAAAAATTTT
CCCTATGCACGGCXXX10CTAATATTTGGCAAACAAGGGACAGGGGCAAGTAACGTAGATTTAGATAATG
TAATGATTACGGATGAAGAAGAGATTCTGACCACCAATCCTGTGGCAACAGAGCAGTATGGAAGTGTGGC
AACTAAGTTCAGAGCTCAAATACAGCTCCCGCGACTGGAAGTGTCAATAGTCAGGGGGCCTTACCTGGC
ATGGTGTGGCAAGATCGTGACGTGTACCTTCAAGGACCTATCTGGGCAAAGATTCCTCACACGGATGGAC
ACTTTCATCCTTCTCCTCTGATGGGAGGCTTTGGACTGAAACATCCGCCTCCTCAAATCTTGATCAAAAA

TACTCCGGTACCGGCAAATCCTCCGACGACTTTCAGCCCGGCCAAGTTTGCTTCATTTATCACTCAGTAC
TCCACTGGACAGGTACGCTGGAATTTGAGTGGGAGCTACAGAAAGAAAACAGCAAACGTTGGAATCCAG
AGATTCAGTACACTTCCAATAACAAGTCTGTTAATGTGGACTTTACTGTAGACACTAATGGTGTGTTA
TAGTGAACCTCGCCCTATTGGAACCCGCTATCTCACACGAACTTG
XXX-1 = GGT/AGT; XXX2 = AGA/AAA; XXX3 = AAA/AGA; XXX4 = ACA/CAG; XXX5 = -AGC/AGA; XXX6 =
CAA/CTC; XXX7 = GCT/TCT; XXX8 = AAA/AGA; XXX9 = GGG/GCG; XXX10 = GTT/GAC

SEQ ID NO:19: Polipéptido L0027

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADAQERLQEDTSFGGNLGRAVFAQAKRVLEPLGLVEEGAKTAP
GKKRPVEQSPQEPDSSSGIGKKGQPPARKRLNFGQTGDSSEVPDPQPLGEPPAAPSGVGSNTMAAGGGAP
MADNNEGADGVGNASGNWHCDSTWLGDRIITSTRTWALPTYNNHLYKQISSQSGGSTNDNTYFGYSTPW
GYFDNFRFHCHFSRDPWQRLINNNWGFPRKRLNFKLFNIQVKEVTNDGTTTIANNLSTVQVFTDSEYQ
LPYVLGSAHQCLPPFPADVFMIPQYGYLTNNNGSQAVGRSSFYCLEYFPSQMLRTGNNEFFSYTFEDVP
FHSSYAHSQSLDRMLNPLIDQYLYLSRTQTTSGTAGNRLQFSQAGPSSMANQAKNWLPGPCYRQQRVS
KTANQNNNSNFAWTGATKYHLNGRDSLVPNPGPAMATHKDDKFFPMSGVLIFGKQGAGNSNVDLDNMI
TNEEEIKTTNPVATEQYGTVATNLQSANTAPATGTVNSQALPGMVWQDRDVLQGPWAKIPIHTDGHFH
PSPLMGGFGLKHPPPIILIKNTVPANPPTTFSPAKFASFITQYSTGQVSVIEWELQKENSKRWNPEIQ
YTSNYNKSTNVDFAVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO:20: Polipéptido L0059

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADAQERLQEDTSFGGNLGRAVFAQAKRVLEPLGLVEEGAKTAP
GKKRPVEQSPQEPDSSSGIGKKGQPPAKKRLNFGQTGDSSEVPDPQPLGEPPAAPSGVGSNTMASGGAP
MADNNEGADGVGNASGNWHCDSTWLGDRIITSTRTWALPTYNNHLYKQISSQSGASTNDNTYFGYSTPW
GYFDNFRFHCHFSRDPWQRLINNNWGFPRKRLNFKLFNIQVKEVTNDGTTTIANNLSTVQVFTDSEYQ
LPYVLGSAHQCLPPFPADVFMIPQYGYLTNNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVP
FHSSYAHSQSLDRMLNPLIDQYLYLSRTQTTSGTAGNRELQFSQAGPSSMANQAKNWLPGPCYRQQRVS
KTANQNNNSNFAWTGATKYHLNGRDSLVPNPGPAMATHKDDKFFPMSGVLIFGKQGAGNSNVDLDNMI
TNEEEIKTTNPVATEEYGTVATNLQSANTAPATGTVNSQALPGMVWQNRDVLQGPWAKIPIHTDGHFH
PSPLMGGFGLKHPPPIILIKNTVPANPPTTFSPAKFASFITQYSTGQVSVIEWELQKENSKRWNPEIQ
YTSNYNKSTNVDFAVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO:21: Polipéptido L0060

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEPVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
GKKRPVEQSPQEPDSSSGIGKKGQQPARKRLNFGQTGDSESVDPDPQLGEPAPAPSGVGSNTMAAGGGAP
MADNNEGADGVGNASGNWHCDSTWLGDRIITTSTRTWALPTYNNHLYKQISSQSGGSTNDNTYFGYSTPW
GYFDNRFHCHFSPRDWQRLINNNWGFPRKLNFKLFNIQVKEVTNDGTTTIANNLTSTVQVFTDSEYQ
LPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFSPQMLRTGNNFEFSYTFEDVP
FHSSYAHSQSLDRLMNPLIDQYLYLSRTQTTSGTAGNRELQFSQAGPSSMANQAKNWLPGPCYRQQRVS
KTTNQNNNSNFAWTGATKYHLNGRDSLVPNGPAMATHKDDDEKFFPMSGVLI FGKQAGNSNVDLDNVM
TSEEEIKTTNPVATEEYGT VATNLQSSNTAPATGT VNSQ GALPGMVWQERDVYLQGP I WAKI PHTDGHFH
PSPLMGGFGLKHPPPQILIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQ
YTSNKNSTNVDFAVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO:22: Polipéptido L0062

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEPVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP

5

GKKRPVEQSPQEPDSSSGIGKKGQQPARKRLNFGQTGDSESVDPDPQLGEPAPAPSGVGSNTMASGGGAP
MADNNEGADGVGNASGNWHCDSTWLGDRIITTSTRTWALPTYNNHLYKQISSQSGGSTNDNTYFGYSTPW
GYFDNRFHCHFSPRDWQRLINNNWGFPRKLNFKLFNIQVKEVTNDGTTTIANNLTSTVQVFTDSEYQ
LPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFSPQMLRTGNNFEFSYTFEDVP
FHSSYAHSQSLDRLMNPLIDQYLYLSRTQTTSGTAGNRELQFSQAGPSSMANQAKNWLPGPCYRQQRVS
KTTNQNNNSNFAWTGATKYHLNGRDSLVPNGPAMATHKDDDEKFFPMSGVLI FGKQAGNSNVDLDNVM
TSEEEIKTTNPVATEEYGT VATNLQSSNTAPATGT VNSQ GALPGMVWQDRDVYLQGP I WAKI PHTDGHFH
PSPLMGGFGLKHPPPQILIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQ
YTSNKNSTNVDFAVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO:23: Polipéptido L0065

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEPVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
GKKRPVEQSPQEPDSSSGIGKKGQQPARKRLNFGQTGDSESVDPDPQLGEPAPAPSGVGSNTMAAGGGAP
MADNNEGADGVGNASGNWHCDSTWLGDRIITTSTRTWALPTYNNHLYKQISSQSGGSTNDNTYFGYSTPW
GYFDNRFHCHFSPRDWQRLINNNWGFPRKLNFKLFNIQVKEVTNDGTTTIANNLTSTVQVFTDSEYQ
LPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFSPQMLRTGNNFQFSYTFEDVP
FHSSYAHSQSLDRLMNPLIDQYLYLSRTQTTSGTAGNRTLQFSQAGPSSMANQAKNWLPGPCYRQQRVS
KTTNQNNNSNFAWTGATKYHLNGRDSLVPNGPAMATHKDDDEKFFPMSGVLI FGKQAGNSNVDLDNVM
TNEEEIKTTNPVATEEYGT VATNLQSSNTAPATGT VNSQ GALPGMVWQDRDVYLQGP I WAKI PHTDGHFH
PSPLMGGFGLKHPPPQILIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQ
YTSNKNSTNVDFAVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO:24: Polipéptido L0033

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEPVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
GKKRPVEQSPQEPDSSSGIGKKGQQPAKKRLNFGQTGDSESVDPDPQLGEPAPAPSGVGSNTMAAGGGAP
MADNNEGADGVGNASGNWHCDSTWLGDRIITTSTRTWALPTYNNHLYKQISSQSGGSTNDNTYFGYSTPW
GYFDNRFHCHFSPRDWQRLINNNWGFPRKLNFKLFNIQVKEVTNDGTTTIANNLTSTVQVFTDSEYQ
LPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFSPQMLRTGNNFEFSYTFEDVP
FHSSYAHSQSLDRLMNPLIDQYLYLSRTQTTSGTAGNRTLQFSQAGPSSMANQAKNWLPGPCYRQQRVS
KTANQNNNSNFAWTGATKYHLNGRDSLVPNGPAMATHKDDDEKFFPMSGVLI FGKQAGNSNVDLDNVM
TSEEEIKTTNPVATEEYGT VATNLQSSNTAPATGT VNSQ GALPGMVWQNRDVYLQGP I WAKI PHTDGHFH
PSPLMGGFGLKHPPPQILIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQ
YTSNKNSTNVDFAVDTNGVYSEPRPIGTRYLTRNL

SEQ ID NO:25: Polipéptido L0036

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEPVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP
GKKRPVEQSPQEPDSSSGIGKKGQQPAKKRLNFGQTGDSESVDPDPQLGEPAPAPSGVGSNTMASGGGAP
MADNNEGADGVGNASGNWHCDSTWLGDRIITTSTRTWALPTYNNHLYKQISSQSGGSTNDNTYFGYSTPW
GYFDNRFHCHFSPRDWQRLINNNWGFPRKLNFKLFNIQVKEVTNDGTTTIANNLTSTVQVFTDSEYQ
LPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFSPQMLRTGNNFEFSYTFEDVP
FHSSYAHSQSLDRLMNPLIDQYLYLSRTQTTSGTAGNRTLQFSQAGPSSMANQAKNWLPGPCYRQQRVS
KTANQNNNSNFAWTGATKYHLNGRDSLVPNGPAMATHKDDDEKFFPMSGVLI FGKQAGNSNVDLDNVM
TSEEEIKTTNPVATEEYGT VATNLQSSNTAPATGT VNSQ GALPGMVWQNRDVYLQGP I WAKI PHTDGHFH
PSPLMGGFGLKHPPPQILIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQ
YTSNKNSTNVDFAVDTNGVYSEPRPIGTRYLTRNL

10

SEQ ID NO:26: Polipéptido L0044

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEPVNAADA
AALEHDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAP

GKKRPVEQSPQEPDSSSGIGKKGQQPAKKRLNFGQTGDSESVDPQPPLGEPPAAPSGVGSNTMASGGGAP
MADNNEGADGVGNASGNWHCDSTWLGDRVITTSTRTWALPTYNNHLYKQISSQSGGSTNDNTYFGYSTPW
GYFDFNRHCHFSRPRDWQRLINNNWGFPRPKLNFKLFNIQVKEVTNDGTTTIANNLTSTVQVFTDSEYQ
LPYVLGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVP
FHSSYAHSQSLDRMLNPLIDQYLYLSRTQTTSGTAGNRELQFSQAGPSSMANQAKNWLPGPCYRQQRV
KTTNQNNSNFAWTGATKYHLNGRDSLVPNGPAMATHKDDKFFPMSGVLI FGKQAGNSNVDLDNVM
TNEEEIKTTNPVATEQYGT VATNLQSANTAPATGTVNSQ GALPGMVWQDRDVYLQGP I WAKI PHTDGHFH
PSPLMGGFGLKHPPPQILIKNTVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQ
YTSNYNKSTNVDFAVDTNGVYSEPRPIGTRYLTRNL

5

SEQ ID NO:27: Polipéptido VP1 de AAV8

SEQ ID NO:28: Polipéptido VP1 de AAV9

SEQ ID NO:29: Polipéptido VP1 de AAV6

SEQ ID NO:30: Polipéptido VP1 de AAV1

SEQ ID NO:31: Polipéptido VP1 de AAV2

10

SEQ ID NO:32: Polipéptido VP1 de AAV3

SEQ ID NO:33: Polipéptido VP1 de AAV3B

SEQ ID NO:34: Polipéptido VP1 de AAV7

REIVINDICACIONES

1. Un polipéptido de cápside de virus adenoasociado (AAV) que tiene una secuencia de aminoácidos que consiste en la SEQ ID NO: 1.

2. El polipéptido de cápside de AAV de la reivindicación 1, en el que el polipéptido de cápside de AAV o una partícula vírica que comprende el polipéptido de cápside de AAV presenta una menor seroprevalencia que un polipéptido de cápside de AAV2 o una partícula vírica que comprende un polipéptido de cápside de AAV2, y en el que el polipéptido de cápside de AAV o una partícula vírica que comprende el polipéptido de cápside de AAV presenta una seroprevalencia aproximadamente igual o inferior a la de un polipéptido de cápside de AAV8 o una partícula vírica que comprende un polipéptido de cápside de AAV8.

3. El polipéptido de cápside de AAV de la reivindicación 1, en el que el polipéptido de cápside de AAV o una partícula vírica que comprende el polipéptido de cápside de AAV se neutralizan en menor medida por suero humano que un polipéptido de cápside de AAV2 o una partícula vírica que comprende un polipéptido de cápside de AAV2, y en el que el polipéptido de cápside de AAV o una partícula vírica que comprende el polipéptido de cápside de AAV se neutraliza en una menor medida o similar por suero humano, que un polipéptido de cápside de AAV8 o una partícula vírica que comprende un polipéptido de cápside de AAV8.

4. El polipéptido de cápside de AAV de una cualquiera de las reivindicaciones 1 a 3, en el que el polipéptido de cápside de AAV está purificado.

5. El polipéptido de cápside de AAV de la reivindicación 1 codificado por una secuencia de ácido nucleico que consiste en la SEQ ID NO: 2.

6. Una molécula de ácido nucleico que codifica un polipéptido de cápside de virus adenoasociado (AAV) que tiene una secuencia de ácido nucleico que consiste en la SEQ ID NO: 2.

7. Un vector que comprende la molécula de ácido nucleico de la reivindicación 6.

8. Una célula huésped que comprende el vector de la reivindicación 7.

9. Una partícula de virus purificada que comprende el polipéptido de cápside de AAV de una cualquiera de las reivindicaciones 1 a 5.

10. La partícula de virus purificada de la reivindicación 9, que comprende además un transgén.

11. Una partícula vírica de la reivindicación 10 para su uso en un método de transferencia génica y/o vacunación con un transgén, comprendiendo el método administrar la partícula vírica a un sujeto que necesita transferencia génica o vacunación, en la que la partícula vírica presenta menos seroprevalencia que una partícula de virus AAV2.

12. La partícula vírica para el uso de la reivindicación 11, en la que la partícula vírica presenta aproximadamente la misma o menos seroprevalencia que una partícula de virus AAV8, o en la que la partícula vírica se neutraliza en menor medida por suero humano que una partícula de virus AAV2, y en la que la partícula del virus AAV se neutraliza en menor medida o similar por suero humano que una partícula de virus AAV8.

13. Un antígeno diana unido operativamente al polipéptido de cápside de AAV de la reivindicación 1 para su uso en un método de vacunación de un sujeto, comprendiendo el método administrar el antígeno diana unido operativamente al polipéptido de cápside de AAV de la reivindicación 1 a un sujeto que necesita vacunación, en el que el polipéptido de cápside de AAV presenta menos seroprevalencia que un polipéptido de cápside de AAV2.

14. El antígeno diana unido operativamente al polipéptido de cápside de AAV de la reivindicación 1 para el uso de la reivindicación 13, en el que el polipéptido de cápside de AAV presenta aproximadamente la misma o menos seroprevalencia que un polipéptido de cápside de AAV8.

15. El antígeno diana para el uso de la reivindicación 13, en el que el polipéptido de cápside de AAV se neutraliza en menor medida por suero humano que un polipéptido de cápside de AAV2, y en el que el polipéptido de

cápside de AAV se neutraliza en una menor medida o similar por suero humano que un polipéptido de cápside AAV8.

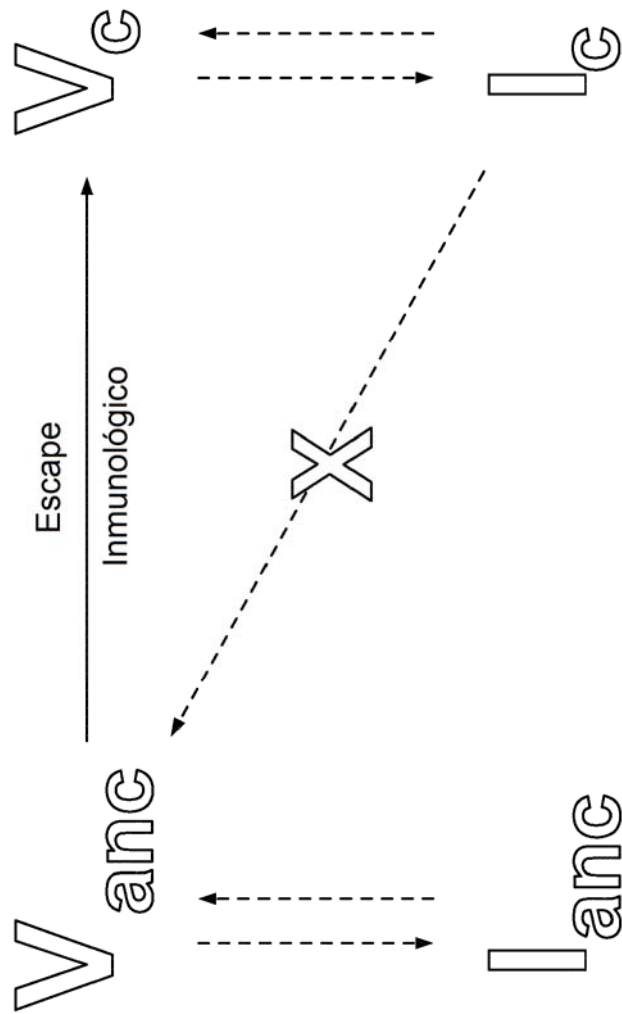
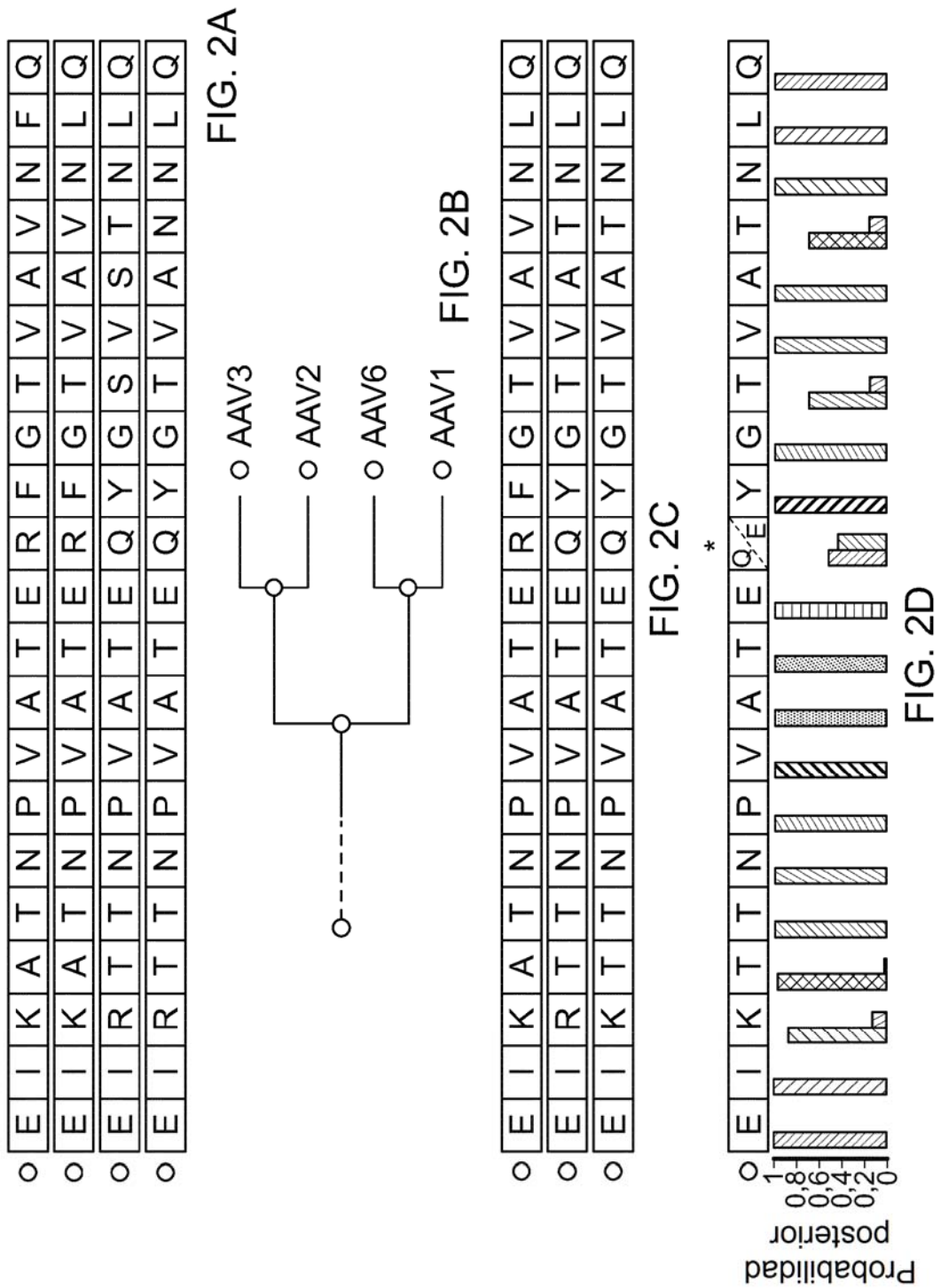


FIG. 1



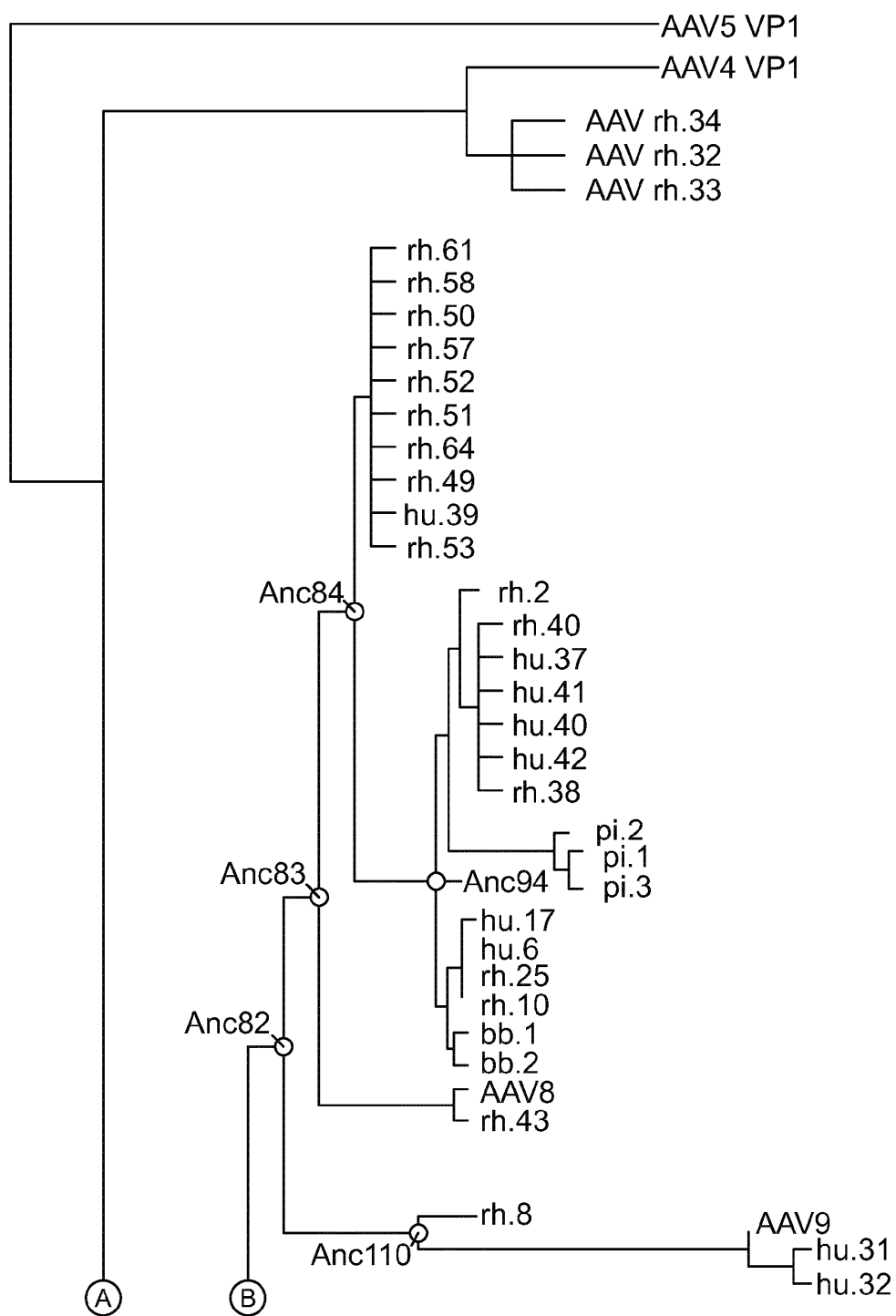


FIG. 3

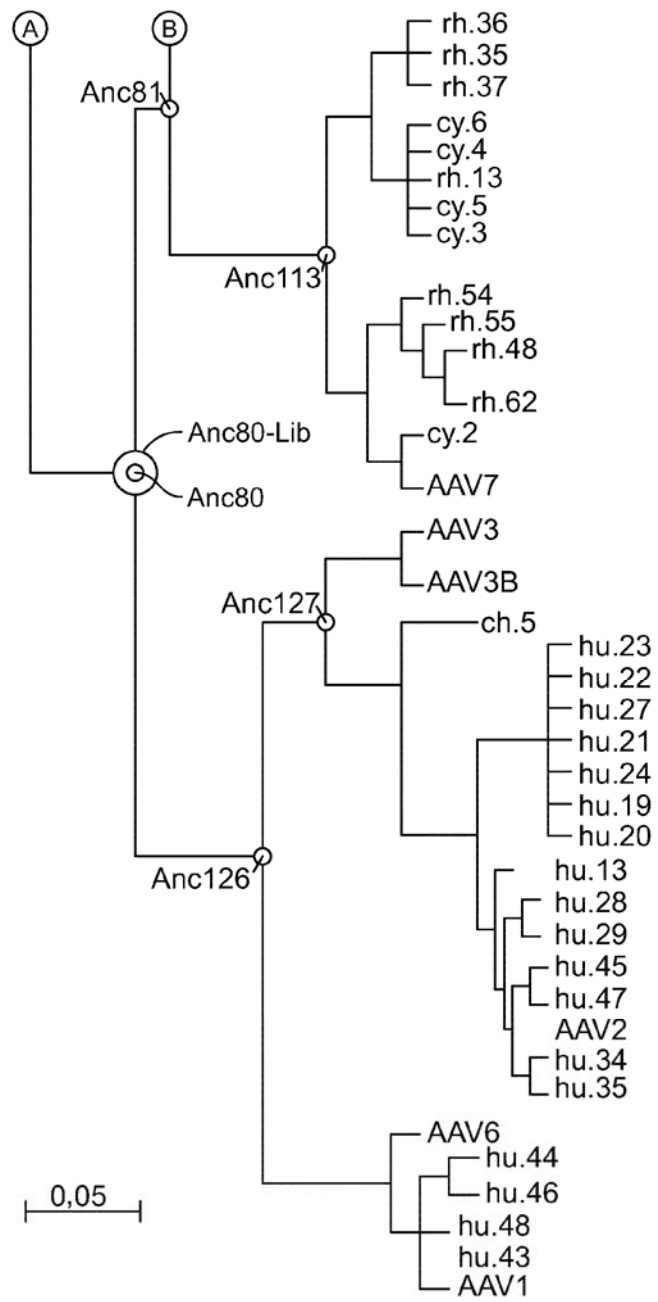


FIG. 3 (Continuación)

L0065	10	20	30	40	50	60	70	80	90	100
L0027	MAADGYLPDWLE	DNLS	EGIR	EWDLKPG	AKPKANQ	KQDD	GRGLVLP	GYKYL	GPFGN	LDKGF
L0033	MAADGYLPDWLE	DNLS	EGIR	EWDLKPG	AKPKANQ	KQDD	GRGLVLP	GYKYL	GPFGN	LDKGF
L0036	MAADGYLPDWLE	DNLS	EGIR	EWDLKPG	AKPKANQ	KQDD	GRGLVLP	GYKYL	GPFGN	LDKGF
L0044	MAADGYLPDWLE	DNLS	EGIR	EWDLKPG	AKPKANQ	KQDD	GRGLVLP	GYKYL	GPFGN	LDKGF
L0059	MAADGYLPDWLE	DNLS	EGIR	EWDLKPG	AKPKANQ	KQDD	GRGLVLP	GYKYL	GPFGN	LDKGF
L0060	MAADGYLPDWLE	DNLS	EGIR	EWDLKPG	AKPKANQ	KQDD	GRGLVLP	GYKYL	GPFGN	LDKGF
L0062	MAADGYLPDWLE	DNLS	EGIR	EWDLKPG	AKPKANQ	KQDD	GRGLVLP	GYKYL	GPFGN	LDKGF
L0065	110	120	130	140	150	160	170	180	190	200
L0027	QERLQ	EDTS	FGNL	GRAV	FQAK	KRVLEP	LGVE	EAGTAP	GKKRP	VEQSP
L0033	QERLQ	EDTS	FGNL	GRAV	FQAK	KRVLEP	LGVE	EAGTAP	GKKRP	VEQSP
L0036	QERLQ	EDTS	FGNL	GRAV	FQAK	KRVLEP	LGVE	EAGTAP	GKKRP	VEQSP
L0044	QERLQ	EDTS	FGNL	GRAV	FQAK	KRVLEP	LGVE	EAGTAP	GKKRP	VEQSP
L0059	QERLQ	EDTS	FGNL	GRAV	FQAK	KRVLEP	LGVE	EAGTAP	GKKRP	VEQSP
L0060	QERLQ	EDTS	FGNL	GRAV	FQAK	KRVLEP	LGVE	EAGTAP	GKKRP	VEQSP
L0062	QERLQ	EDTS	FGNL	GRAV	FQAK	KRVLEP	LGVE	EAGTAP	GKKRP	VEQSP
L0065	210	220	230	240	250	260	270	280	290	300
L0027	NTMA	AGG	GAP	WAD	NNE	GAD	GV	GN	W	H
L0033	NTMA	AGG	GAP	WAD	NNE	GAD	GV	GN	W	H
L0036	NTMA	AGG	GAP	WAD	NNE	GAD	GV	GN	W	H
L0044	NTMA	AGG	GAP	WAD	NNE	GAD	GV	GN	W	H
L0059	NTMA	AGG	GAP	WAD	NNE	GAD	GV	GN	W	H
L0060	NTMA	AGG	GAP	WAD	NNE	GAD	GV	GN	W	H
L0062	NTMA	AGG	GAP	WAD	NNE	GAD	GV	GN	W	H
L0065	310	320	330	340	350	360	370	380	390	400
L0027	INN	WG	FR	PK	LN	FK	LN	FK	LN	FK
L0033	INN	WG	FR	PK	LN	FK	LN	FK	LN	FK
L0036	INN	WG	FR	PK	LN	FK	LN	FK	LN	FK
L0044	INN	WG	FR	PK	LN	FK	LN	FK	LN	FK
L0059	INN	WG	FR	PK	LN	FK	LN	FK	LN	FK
L0060	INN	WG	FR	PK	LN	FK	LN	FK	LN	FK
L0062	INN	WG	FR	PK	LN	FK	LN	FK	LN	FK

FIG. 4

I0065	410	420	430	440	450	460	470	480	490	500
I0027										
I0033	SQMLRTGNNFQFSYTFEDVPPFHSSYAHSCSLDRLMPLIDQYLYLSRTQTTSGTAGNRTLQFSQAGPSSMANCAKNWLPGFCYRQGVSKTTNQNNNSN									
I0036	E.....									
I0044	E.....									
I0059	E.....									
I0060	E.....									
I0062	E.....									
I0065	510	520	530	540	550	560	570	580	590	600
I0027										
I0033	FAWTGATKYHLNGRDSLVNPGPAMATHKDEDEKFFPNSGVLIIFGQCGAGNSVDLDNVMITNEEEIKTNPVATEEYGTVATNLQSANATAPATGTVNSQG									
I0036										
I0044										
I0059										
I0060										
I0062										
I0065	610	620	630	640	650	660	670	680	690	700
I0027										
I0033	ALPGMVWQDRDVIYLGPIWAKIPIHTDGHFHPSPLMGGFGLKHPFPQILIKNTPVPANPPTTFSPAXFASFITQYSTGQVSVIEIEWELQKENSKRWNPEIQ									
I0036										
I0044	N.....									
I0059	N.....									
I0060	E.....									
I0062										
I0065	710	720	730							
I0027										
I0033	YTSNYSKSTNVDFAVDTNGVYSEPRPIGTRYLTRNL*									
I0036										
I0044										
I0059										
I0060										
I0062										

FIG. 4 (Cont.)

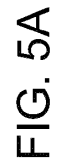


FIG. 5A (Cont.)

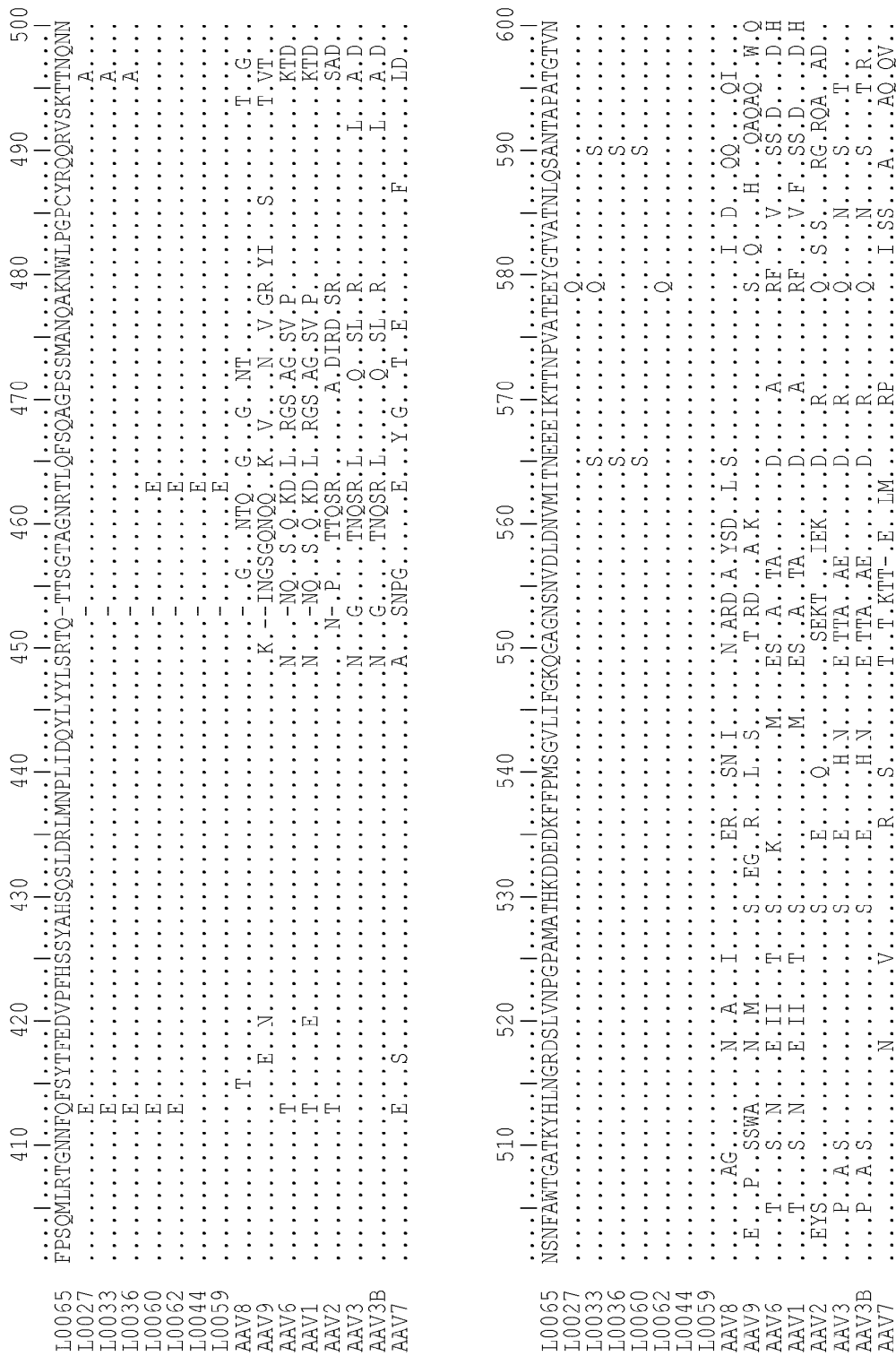


FIG. 5B

	610	620	630	640	650	660	670	680	690	700	
L0065	SQALPCMWQDRDVI	QGP	WAKI	PHTD	GHFHP	SPIMGG	FGIK	HP	PPQ	LIKNT	PPV
L0027											
L0033											
L0036											
L0060											
L0062											
L0044											
L0059											
AAV8											
AAV9											
AAV6											
AAV1											
AAV2											
AAV3											
AAV3B											
AAV7											

	710	720	730
L0065	EQYTSN	YNKSTNV	DFAVDTNGVYSEPR
L0027			
L0033			
L0036			
L0060			
L0062			
L0044			
L0059			
AAV8			
AAV9			
AAV6			
AAV1			
AAV2			
AAV3			

FIG. 5B (Cont.)

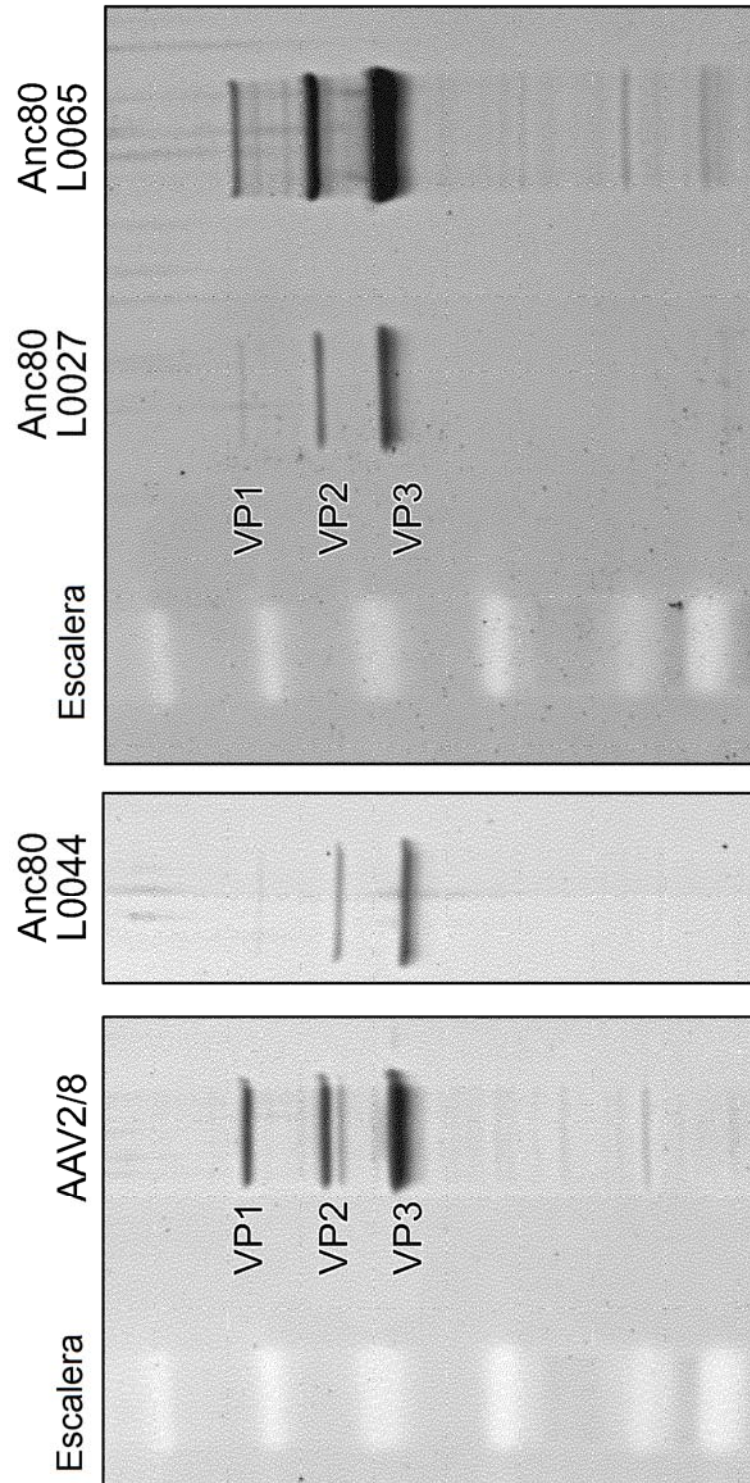


FIG. 6

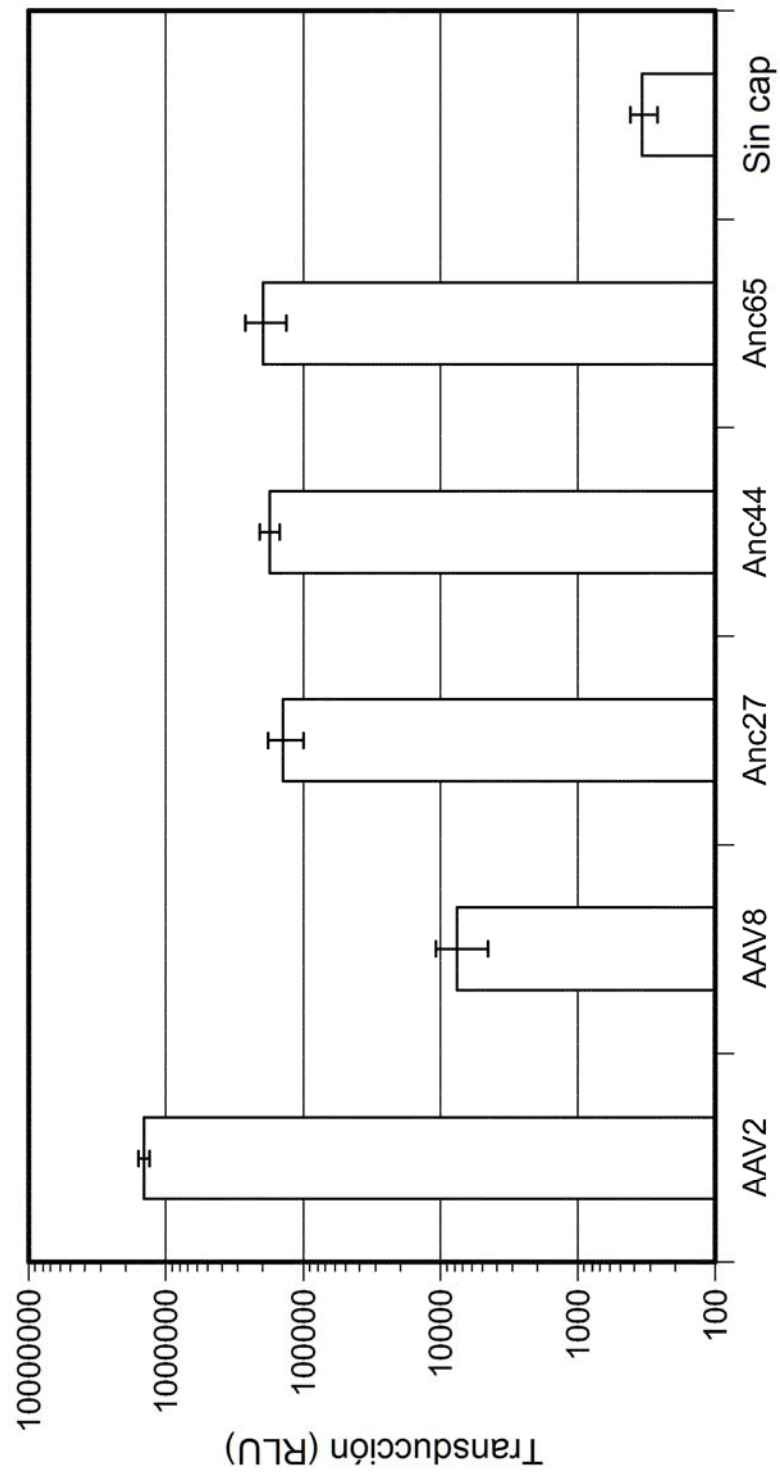


FIG. 7

VP1 seq.Δ	AAV5	AAV2	AAV8	rh10	Anc80L65	Anc80Lib
AAV5	ID	43,0%	42,6%	43,0%	40,5%	39,9-40,5%
AAV2	320	ID	17,1%	15,9%	12,2%	12,0-12,8%
AAV8	317	127	ID	6,5%	9,1%	8,7-10,1%
rh10	320	118	48	ID	8,6%	7,8-8,9%
L0065	301	91	68	64		0-1,5%
Anc80Lib	297-302	89-95	65-75	58-66	0-11	ID

FIG. 9

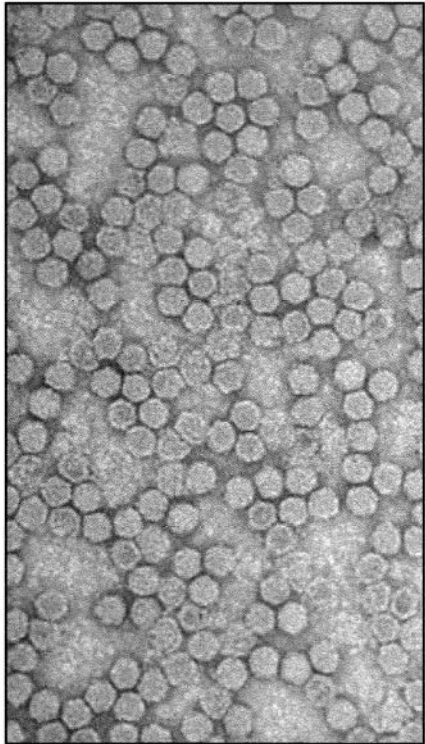
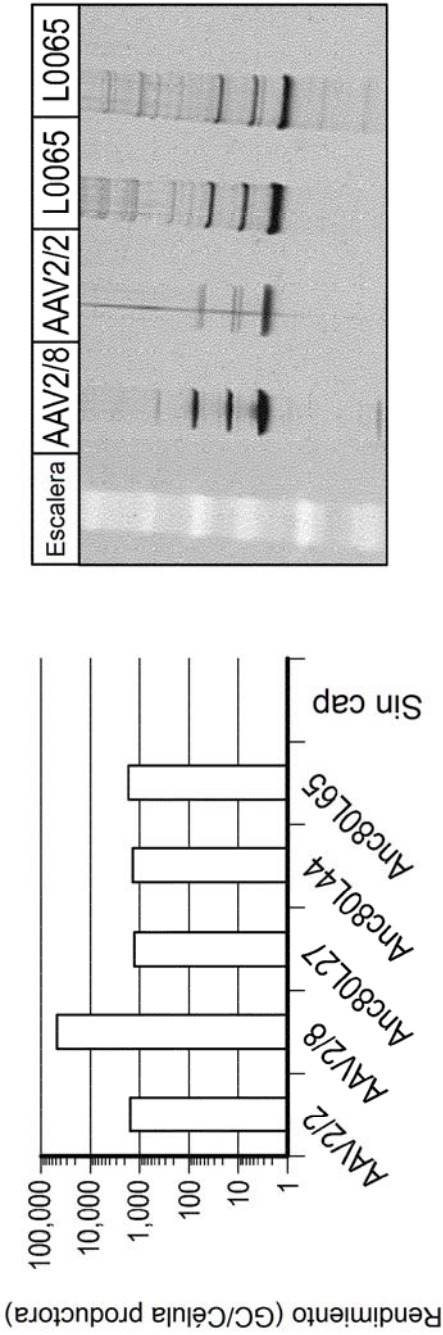
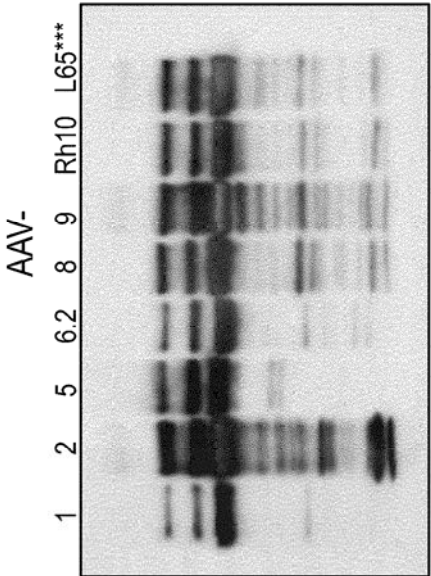


FIG. 10B



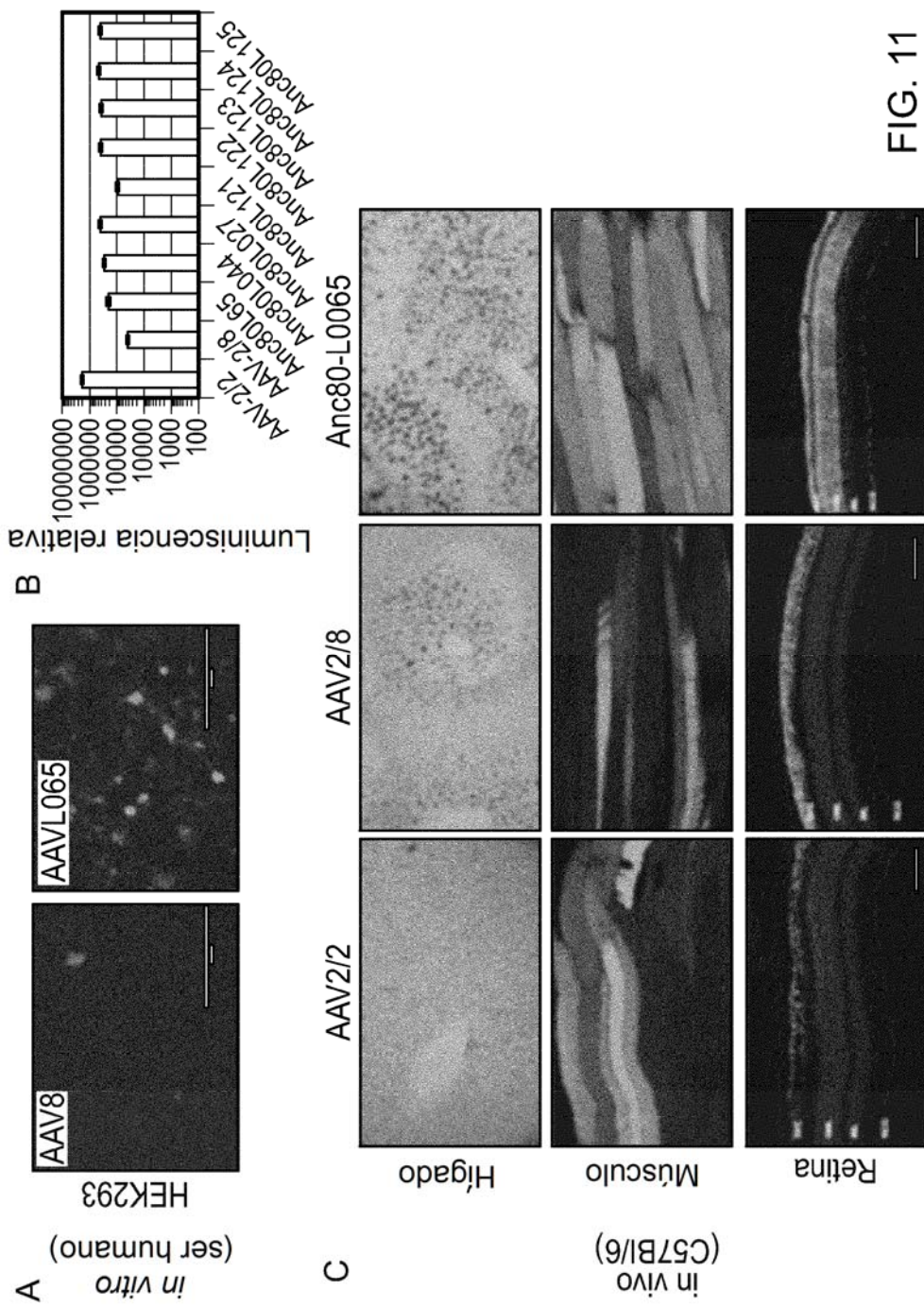


FIG. 11

	Anc81	Anc127	Anc83	Anc110	Anc113	AAV3	AAV1	AAV2	AAV9						
L0065	Anc126	Anc82	Anc84	AAV8	AAV7										
L0065	ID	97,8%	97,2%	95,5%	95,5%	94,0%	92,1%	94,0%	90,7%	93,4%	90,7%	90,5%	89,5%	87,9%	86,1%
Anc81	ID	95,9%	93,6%	97,4%	95,6%	93,9%	95,1%	92,1%	94,7%	91,4%	88,6%	88,3%	86,7%	86,5%	
Anc126	ID	97,4%	93,4%	92,0%	90,7%	89,7%	92,5%	89,1%	91,8%	88,8%	91,9%	91,5%	89,2%	85,8%	
Anc127	ID	91,7%	90,7%	89,7%	90,7%	88,3%	89,8%	87,3%	94,4%	89,4%	90,7%	86,0%			
Anc82	ID	98,2%	96,2%	97,4%	94,7%	92,8%	89,7%	86,8%	86,4%	85,2%	87,8%				
Anc83	ID	97,8%	95,6%	96,0%	91,7%	89,3%	86,1%	85,3%	84,8%	87,1%					
Anc84	ID	93,7%	94,5%	91,3%	89,1%	85,5%	85,0%	84,1%	86,3%						
Anc110	ID	92,5%	90,5%	88,2%	86,3%	86,5%	85,0%	84,1%	86,3%						
AAV8	ID	89,7%	87,8%	85,3%	84,0%	82,9%	85,2%								
Anc113	ID	96,4%	85,9%	86,7%	83,8%	83,3%									
AAV7	ID	84,2%	85,0%	82,3%	81,8%										
AAV3	ID	86,4%	87,3%	83,7%											
AAV1	ID	83,2%	82,4%												
AAV2	ID	81,9%													
AAV9	ID														

FIG. 13A

	Anc81	Anc127	Anc83	Anc84	Anc110	AAV8	AAV3	Anc113	AAV7	AAV2	AAV1	AAV9			
L0065	Anc126														
L0065	ID	97,3%	96,6%	94,9%	94,3%	92,1%	90,2%	92,1%	88,5%	90,6%	92,1%	88,7%	88,7%	86,8%	84,1%
Anc81	ID	94,7%	92,7%	96,6%	94,3%	92,7%	93,6%	90,4%	88,4%	88,4%	93,4%	89,7%	87,4%	85,2%	84,4%
Anc126	ID	97,5%	91,5%	89,7%	88,7%	88,7%	89,9%	86,7%	92,6%	92,6%	89,9%	86,5%	90,8%	89,5%	83,3%
Anc127	ID	90,2%	88,7%	88,0%	88,5%	86,3%	94,9%	88,2%	85,2%	92,1%	87,6%	83,1%			
Anc82	ID	97,7%	95,7%	97,0%	93,8%	86,1%	90,6%	87,1%	85,6%	82,8%	86,3%				
Anc83	ID	97,7%	94,7%	95,5%	93,4%	88,4%	88,4%	85,8%	84,4%	81,6%	84,6%				
Anc84	ID	92,8%	93,4%	88,4%	88,4%	85,0%	88,0%	85,6%	85,7%	82,4%	87,8%				
Anc110	ID	91,4%	85,0%	86,7%	84,5%	80,1%	83,5%								
AAV8	ID	84,5%	86,7%	84,5%	88,9%	85,7%	80,9%								
AAV3	ID	85,4%	82,8%	88,9%	85,7%	80,9%									
Anc113	ID	95,8%	84,1%	83,5%	80,5%										
AAV7	ID	82,4%	81,8%	78,9%											
AAV2	ID	83,3%	81,4%												
AAV1	ID	79,6%													
AAV9	ID														

FIG. 13B

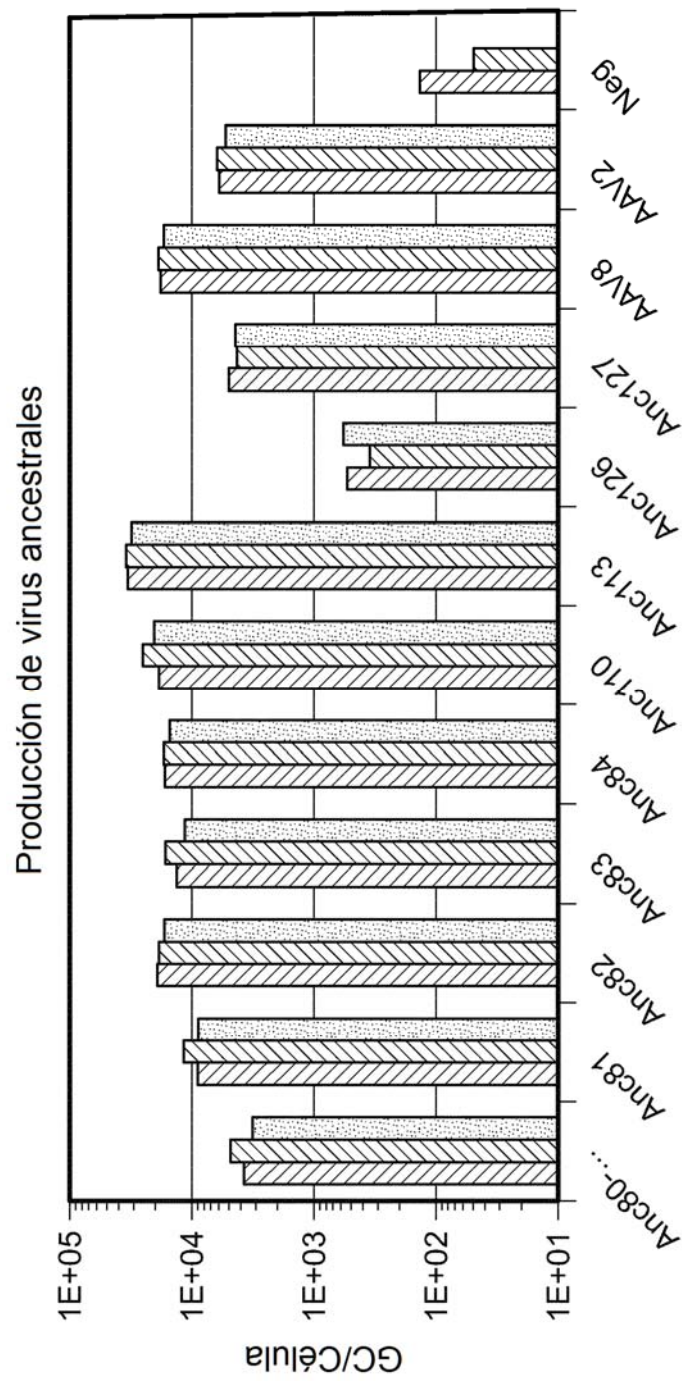


FIG. 14

Producción vírica con respecto a AAV8 (%)	
Anc80-L0065	21,75%
Anc81	54,14%
Anc82	100,99%
Anc83	76,86%
Anc84	91,20%
Anc110	118,97%
Anc113	183,10%
Anc126	2,71%
Anc127	25,34%
AAV2	32,56%
Neg	0,27%

FIG. 15

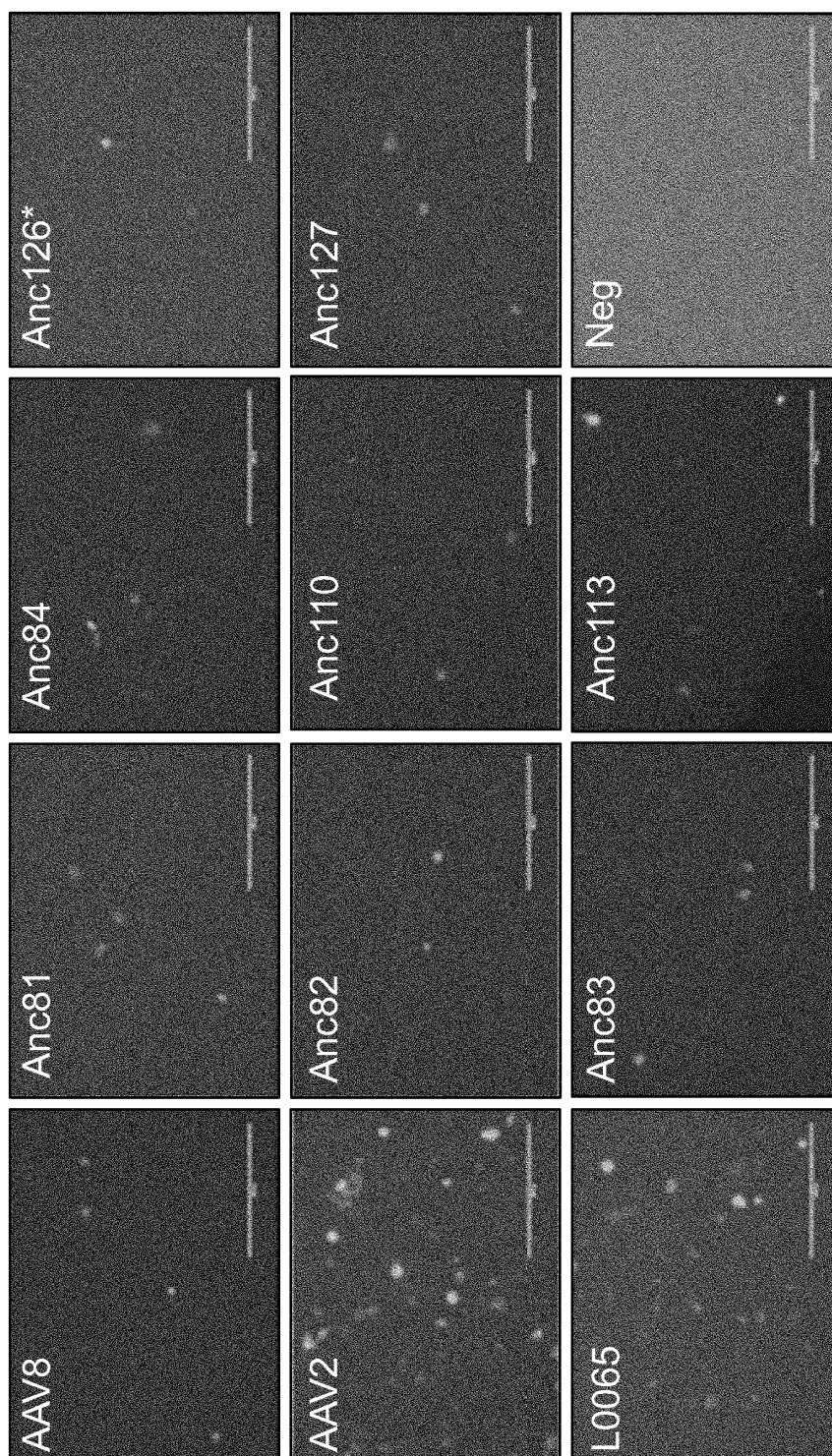


FIG. 16

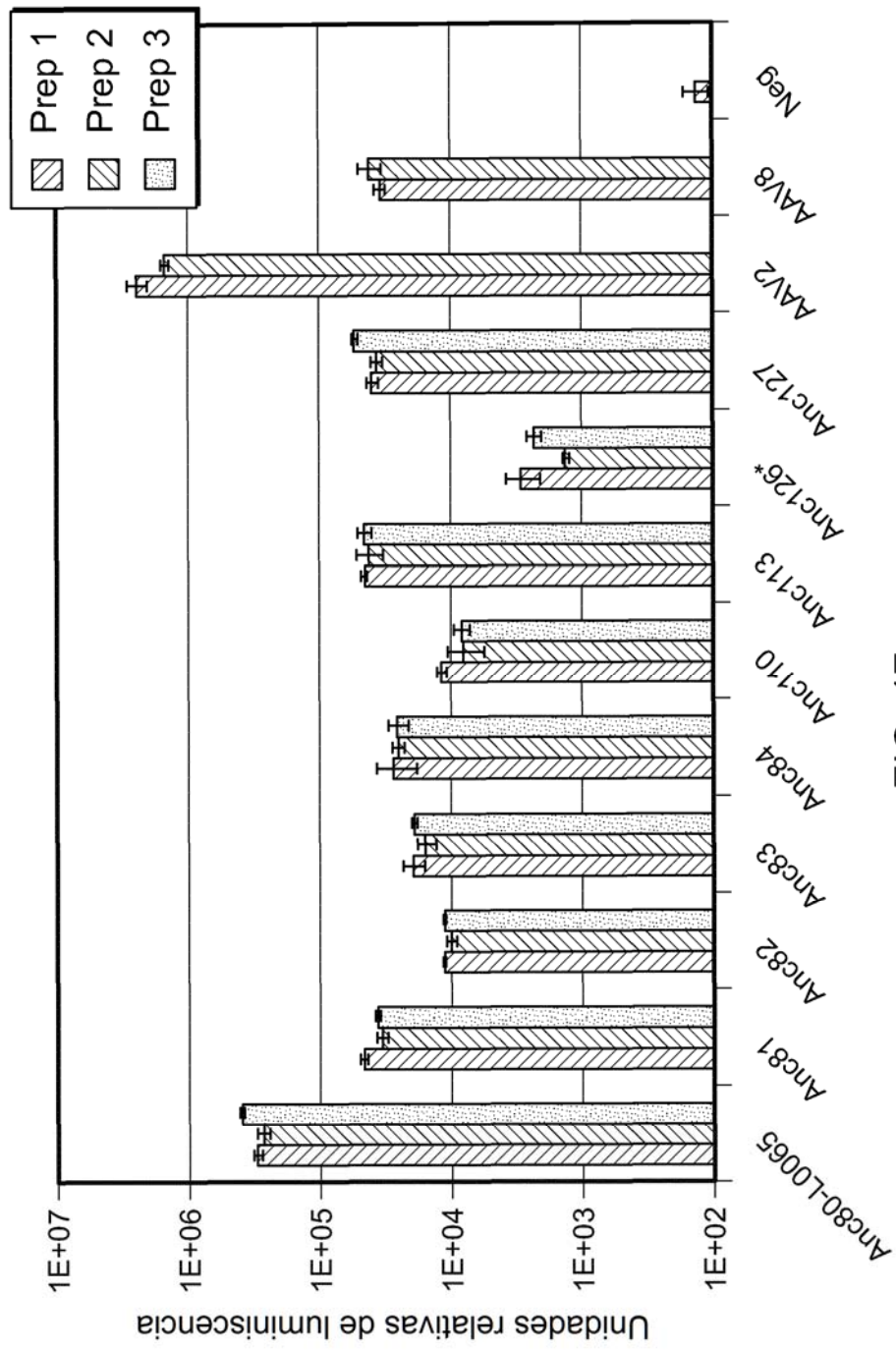


FIG. 17

Producción vírica con respecto a AAV8 (%)	
Anc80-L0065	850,83%
Anc81	102,26%
Anc82	28,31%
Anc83	47,36%
Anc84	68,78%
Anc110	24,56%
Anc113	117,12%
Anc127	113,78%
AAV2	5225,61%
Neg	0,35%

FIG. 18

AAV8 AAV2		Anc80- L0065	Anc81	Anc82	Anc83	Anc84	Anc110	Anc113	Anc126	Anc127
Infectividad	=	+++	+++	=	-	-	-	-	=	ND
Producción	=	-	--	-	=	=	=	++	--	-

Clave
+++ : > 225%
++ : 175% - 224,5%
+ : 125 - 174,9%
= : 75% - 124,9%
- : 25-74,9%
-- : 0-24,9%
ND : No determinado

FIG. 19