

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 898 271**

51 Int. Cl.:

**C07K 14/16**

(2006.01)

12

## TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **15.12.2016** **E 19186284 (6)**

97 Fecha y número de publicación de la concesión europea: **25.08.2021** **EP 3584252**

54 Título: **Antígenos del virus de la inmunodeficiencia humana, vectores, composiciones y métodos de uso de los mismos**

30 Prioridad:

**15.12.2015 EP 15200138**

**17.10.2016 EP 16194124**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**04.03.2022**

73 Titular/es:

**JANSSEN VACCINES & PREVENTION B.V.**  
**(100.0%)**

**Archimedesweg 4**  
**2333 CN Leiden, NL**

72 Inventor/es:

**LANGEDIJK, JOHANNES, PETRUS, MARIA;**  
**CALLENDRET, BENOIT, C.,S.;**  
**VAN MANEN, DANIELLE;**  
**KRARUP, ANDERS;**  
**STITZ, JÖRN;**  
**WEGMANN, FRANK y**  
**VELLINGA, JORT**

74 Agente/Representante:

**LEHMANN NOVO, María Isabel**

ES 2 898 271 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

Antígenos del virus de la inmunodeficiencia humana, vectores, composiciones y métodos de uso de los mismos

5 Antecedentes de la invención

El virus de la inmunodeficiencia humana (VIH) afecta a millones de personas en todo el mundo y la prevención del VIH mediante una vacuna eficaz sigue siendo una prioridad fundamental, incluso en la época del uso generalizado del tratamiento antirretroviral. El VIH-1 es la cepa más común y patógena del virus, donde más del 90 % de los

10 casos de VIH/SIDA corresponden a infecciones con el grupo M del VIH-1. El grupo M se subdivide a su vez en clados o subtipos. Idealmente, una vacuna eficaz sería capaz de provocar tanto respuestas celulares potentes como anticuerpos ampliamente neutralizantes capaces de neutralizar cepas del VIH-1 de diferentes clados.

15 La gran variabilidad genética del VIH-1 convierte el desarrollo de la vacuna contra el VIH-1 en un desafío sin precedentes. Con el fin de mejorar la cobertura de los posibles epítomos de los linfocitos T, y mejorar las respuestas celulares, se han descrito y desarrollado antígenos Gag, Pol y Env del VIH-1 "mosaico", derivados de las proteínas Antígeno específico de grupo (Gag), Polimerasa (Pol) y Envoltura (Env) del VIH, en un intento de proporcionar la máxima cobertura de posibles epítomos de linfocitos T (p. ej., Barouch et al, Nat Med 2010, 16: 319-323). Los

20 antígenos mosaico tienen una longitud y estructura del dominio similares a los antígenos del VIH-1 de tipo silvestre, de origen natural.

Por ejemplo, los antígenos del VIH mosaico descritos y utilizados en las vacunas incluyen los descritos en Barouch et al, *supra*, y en el documento WO 2010/059732 tales como:

25 (a) antígenos mosaico Gag incluyendo:

(a)(i) una primera secuencia de Gag mosaico ("mos1Gag") que tiene la secuencia de aminoácidos como la expuesta en el presente documento en la SEQ NO: 1, y

30 (a)(ii) una segunda secuencia de Gag mosaico ("mos2Gag") que tiene la secuencia de aminoácidos como la expuesta en el presente documento en la SEQ NO: 2;

(b) antígenos mosaico Pol incluyendo:

35 (b)(i) una primera secuencia de Pol mosaico ("mos1Pol") que tiene la secuencia de aminoácidos como la expuesta en el presente documento en la SEQ ID NO: 3, y

(b)(ii) una segunda secuencia de Pol mosaico ("mos2Pol") que tiene la secuencia de aminoácidos como la expuesta en el presente documento en la SEQ ID NO: 4; y

40 (c) antígenos mosaico Env incluyendo:

(c)(i) una primera secuencia de Env mosaico ("mos1Env") que tiene la secuencia de aminoácidos como la expuesta en el presente documento en la SEQ ID NO: 5, y

45 (c)(ii) una segunda secuencia de Env mosaico ("mos2Env") que tiene la secuencia de aminoácidos como la expuesta en el presente documento en la SEQ ID NO: 6.

Las secuencias que codifican estos antígenos se han clonado en vectores, por ejemplo, tales como en vectores adenovíricos, por ejemplo, adenovirus recombinante de serotipo 26 (rAd26), y estos vectores recombinantes se utilizaron previamente como vacunas para generar respuestas inmunitarias contra los antígenos (véase p. ej., Barouch et al, *supra*; y el documento WO 2010/059732. Por ejemplo, las secuencias de los antígenos mosaico mos1Gag y mos1Pol se combinan normalmente en una proteína de fusión de Gag y Pol ("mos1GagPol"), cuya

50 secuencia codificante se clona en un primer vector Ad26 ("rAd26.mos1GagPol"); y las secuencias de los antígenos mos2Gag y mos2Pol se combinan en otra proteína de fusión de Gag y Pol ("mos2GagPol"), cuya secuencia codificante se clona en un segundo vector Ad26 ("rAd26.mos2GagPol"). Las construcciones que codifican mos1Env y mos2Env se clonan normalmente en vectores Ad26 independientes ("rAd26.mos1Env" y "rAd26.mos2Env", respectivamente).

Un conjunto de tales antígenos mosaico como los descritos anteriormente proporciona una buena cobertura global de los aislados de VIH-1 del grupo M, donde los vectores rAd26 que codifican las secuencias de los antígenos 1 mosaico (p. ej., rAd26.mos1GagPol y rAd26.mos1Env) favorecen el clado B y los subtipos del VIH-1 CRF01 y los

60 vectores rAd26 que codifican las secuencias de los antígenos 2 mosaico (p. ej., rAd26.mos2GagPol y rAd26.mos2Env) favorecen las cepas del clado C. Los antígenos Gag, Pol y Env del VIH-1 mosaico expresados en los vectores rAd26 se pueden utilizar para mejorar tanto la amplitud como la profundidad de las respuestas por linfocitos T específicos del antígeno en monos rhesus, sin comprometer el alcance de las respuestas celular y humoral cuando se comparan con los antígenos del VIH-1 de secuencia consenso o natural (Barouch et al, *supra*; y el documento WO 2010/059732.

Sin embargo, tras esfuerzos adicionales de desarrollo de los componentes de las vacunas descritas anteriormente, se observó que rAd26.mos2Env mostraba una expresión en la superficie celular y una respuesta inmunitaria no óptimas en primates no humanos, pero además mostró una estabilidad genética no descrita hasta la fecha, inesperada e impredecible no óptima durante el proceso de fabricación en comparación con otros vectores rAd26, tales como rAd26.mos1Env. Así pues, las vacunas que contienen rAd26.mos2Env pueden dar como resultado respuestas inmunitarias no óptimas contra los subtipos del VIH-1 del Clado C, ya que el antígeno mosaico mos2Env favorece a las cepas del VIH-1 del clado C. En consecuencia, se necesita una alternativa al antígeno mos2Env en las vacunas contra el VIH que se pueda utilizar para inducir respuestas inmunitarias mejoradas contra el clado C del VIH-1.

#### Breve resumen de la invención

La invención se refiere a proteínas de la envoltura del virus de la inmunodeficiencia humana (VIH) sintéticas novedosas que tienen una expresión en la superficie celular y una estabilidad genética mejorada en comparación con el antígeno mos2Env descrito previamente. La invención también se refiere a composiciones y métodos para utilizar tales proteínas de la envoltura del VIH sintéticas novedosas y/o sus secuencias codificantes para inducir mayores respuestas inmunitarias contra el VIH-1, en particular contra el clado C del VIH-1, preferiblemente cuando se utilizan combinadas con otros antígenos del VIH.

En un aspecto general, la invención se refiere a un ácido nucleico que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8, o la SEQ ID NO: 8 que tiene una o más mutaciones seleccionadas del grupo que consiste en (i) I529P (es decir, una sustitución de Ile por Pro en la posición 529 de la SEQ ID NO: 8), (ii) K480E (es decir, una sustitución de Lys por Glu en la posición 480 de la SEQ ID NO: 8) y (iii) una combinación de EK479-480RRRR (es decir, un reemplazo de Glu-Lys en las posiciones 479-480 de la SEQ ID NO: 8 por cuatro restos consecutivos de Arg), I529P, A471C (es decir, una sustitución de Ala por Cys en la posición 471 de la SEQ ID NO: 8) y T575C (es decir, una sustitución de Thr por Cys en la posición 575 de la SEQ ID NO: 8). En una realización, la proteína de la envoltura del VIH sintética comprende además una secuencia señal, por ejemplo, una secuencia señal que tiene la secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NO: 9-12. En una realización, la secuencia señal tiene la secuencia de aminoácidos de la SEQ ID NO: 9.

En ciertas realizaciones, la proteína de la envoltura del VIH sintética comprende además un dominio transmembrana, preferiblemente un dominio transmembrana que tiene la secuencia de aminoácidos de la SEQ ID NO: 13.

En ciertas realizaciones, la proteína de la envoltura de VIH sintética comprende además un fragmento de un dominio citoplasmático, preferiblemente un fragmento de un dominio citoplasmático que comprende la secuencia de aminoácidos de la SEQ ID NO: 14 o los aminoácidos 1-4 de esta (es decir, NRVR). En las realizaciones en las que la proteína de la envoltura del VIH sintética comprende además un dominio transmembrana y un fragmento de un dominio citoplasmático, se prefiere que la proteína comprenda también la secuencia de aminoácidos de la SEQ ID NO: 37, que está fusionada con el extremo carboxilo (extremo C) de la SEQ ID NO: 8 y el extremo amino (extremo N) de la región transmembrana.

En otra realización, la proteína de la envoltura del VIH sintética comprende un dominio de trimerización, por ejemplo, un dominio de trimerización que comprende la secuencia de aminoácidos de la SEQ ID NO: 15 (GCN4) o la SEQ ID NO: 16 (dominio foldon). En una realización preferida, el dominio de trimerización comprende la secuencia de aminoácidos de la SEQ ID NO: 15. Tales realizaciones con los dominios de trimerización son útiles para las proteínas de la envoltura del VIH sintéticas solubles (es decir, no unidas a la membrana) basadas en las secuencias del ectodominio proporcionadas en el presente documento, tales como la que comprende la secuencia de aminoácidos de la SEQ ID NO: 8, en la que el dominio de trimerización está localizado en el extremo C de la proteína de la envoltura del VIH sintética.

En otras realizaciones más, la proteína de la envoltura del VIH sintética comprende la SEQ ID NO: 8 con las siguientes mutaciones: EK479-480RRRR, I529P, A471C y T575C. La introducción de 6 restos de arginina consecutivos (las posiciones 478 y 481 de la secuencia natural de la SEQ ID NO: 8 ya son restos de Arg) da como resultado un sitio de escisión de furina más optimizado, de manera que se obtiene un ectodominio con un procesado (es decir, escindido) mejorado. Las tres mutaciones de I529P, A471C y T575C son conocidas como mutaciones SOSIP, en las que las dos últimas mutaciones dan como resultado la introducción de un posible puente disulfuro entre los restos de cisteína recién creados. En general, estas mutaciones pueden dar como resultado una proteína de la envoltura del VIH sintética soluble, trimerizada, sin que sea necesario un dominio de trimerización.

En una realización preferida, la invención se refiere a un ácido nucleico que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 17, la SEQ ID NO: 18 o los aa 1-686 de la SEQ ID NO: 19. Lo más preferiblemente, la proteína de la envoltura del VIH sintética codificada por el ácido nucleico comprende o está constituida por la secuencia de aminoácidos de la SEQ ID NO: 18.

En otro aspecto general, la invención se refiere a un vector que comprende un ácido nucleico que codifica una

proteína de la envoltura del VIH sintética de acuerdo con una realización de la invención. En una realización, el vector es un vector vírico. En una realización preferida, el vector vírico es un vector adenovírico. En una realización preferida, el vector adenovírico es un vector de adenovirus 26.

Otro aspecto general de la invención se refiere a una composición, preferiblemente una composición de vacuna, que comprende una cantidad inmunogénicamente eficaz de un vector de acuerdo con una realización de la invención y un vehículo, donde el ácido nucleico que codifica la proteína de la envoltura del VIH sintética está ligado operativamente a una secuencia promotora. En una realización, la composición comprende un vector de adenovirus, preferiblemente un vector de adenovirus 26, que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 18.

En otro aspecto general, la invención se refiere a una combinación de vacuna para inducir una respuesta inmunitaria contra un virus de la inmunodeficiencia humana (VIH) en un sujeto que lo necesite. La combinación de vacuna comprende una primera composición que comprende una cantidad inmunogénicamente eficaz de un vector, preferiblemente un vector de adenovirus, más preferiblemente un vector de adenovirus 26, que codifica una proteína de la envoltura del VIH sintética que tiene la secuencia de aminoácidos de la SEQ ID NO: 18, una segunda composición que comprende una cantidad inmunogénicamente eficaz de un segundo vector, preferiblemente un segundo vector de adenovirus, más preferiblemente un segundo vector de adenovirus 26, que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5 y, opcionalmente, al menos una composición adicional que comprende una cantidad inmunogénicamente eficaz de al menos uno seleccionado del grupo que consiste en un vector que codifica un polipéptido antigénico que tiene la secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NO: 1-4, 28 y 29, y un polipéptido que comprende una cantidad inmunogénicamente eficaz de un polipéptido antigénico del VIH aislado, incluyendo, aunque no de forma limitativa, un polipéptido que tiene los restos 30-708 de la secuencia de aminoácidos de la SEQ ID NO: 7 o un polipéptido que tiene los restos 30-724 de la SEQ ID NO: 36, en donde la primera composición, la segunda composición y la composición adicional opcional están presentes en la misma composición o en una o más composiciones diferentes.

Otro aspecto general más de la invención se refiere a métodos para inducir una respuesta inmunitaria contra un virus de la inmunodeficiencia humana (VIH) en un sujeto que lo necesite, que comprende administrar al sujeto una composición o combinación de vacuna de acuerdo con una realización de la invención. La invención también se refiere a métodos para inducir una respuesta inmunitaria contra un VIH que comprende la sensibilización y refuerzo de la respuesta inmunitaria utilizando una composición o una combinación de vacuna de acuerdo con una realización de la invención.

Un aspecto adicional más de la invención se refiere a una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8, o la SEQ ID NO: 8 que tiene una o más mutaciones seleccionadas del grupo que consiste en (i) I529P, (ii) K480E, (iii) una combinación de EK479-480RRRR, I529P, A471C y T575C. En una realización, la proteína de la envoltura del VIH sintética comprende la SEQ ID NO: 8 con las mutaciones EK479-480RRRR, I529P, A471C y T575C. En otra realización, la proteína de la envoltura del VIH sintética comprende los restos 30-704 o 30-711 de la secuencia de aminoácidos de la SEQ ID NO: 18. En otra realización más, la proteína de la envoltura del VIH sintética comprende los restos 30-686 de la secuencia de aminoácidos de la SEQ ID NO: 19.

Otro aspecto de la invención se refiere a una célula, preferiblemente una célula aislada, que comprende un vector de acuerdo con una realización de la invención.

Breve descripción de varias vistas de los dibujos

El resumen anterior, así como la siguiente descripción detallada de la invención, se entenderán mejor al leerlos junto con los dibujos adjuntos. Se debe sobreentender que la invención no se limita a las realizaciones precisas que se muestran en los dibujos.

En los dibujos:

Las FIGS. 1A-1C son representaciones esquemáticas de la estructura de las proteínas de la envoltura del VIH; La FIG. 1A muestra una proteína de la envoltura del VIH completa; La FIG. 1B muestra la estructura de una proteína de la envoltura del VIH de una sola cadena soluble de acuerdo con una realización de la invención en la cual el dominio transmembrana (TM) está reemplazado por un dominio de trimerización GCN4 y el sitio de escisión de furina está mutado (sC4); La FIG. 1C muestra la estructura de una proteína de la envoltura del VIH unida a la membrana de acuerdo con una realización de la invención que comprende un dominio transmembrana y un fragmento de un dominio citoplasmático (C4D7); La FIG. 2 muestra los niveles de expresión de la proteína de la envoltura sC1 soluble, que se basa en la secuencia del antígeno mosaico mos2Env con un dominio de trimerización C-terminal adicional y una proteína de la envoltura del VIH sintética soluble (sC4) de acuerdo con una realización de la invención; la expresión se midió mediante transferencia Western cuantitativa utilizando un anticuerpo policlonal contra gp120; los plásmidos que codifican sC1 o sC4 se expresaron de manera transitoria dos veces y cada transfección se cuantificó dos veces

mediante densitometría; la proteína sC1 mostró unos niveles de expresión muy bajos en comparación con la proteína de la envoltura del VIH sintética sC4, que mostró unos niveles de expresión relativamente elevados;

Las FIGS. 3A y 3B muestran la unión de las proteínas de la envoltura del VIH sintéticas con el anticuerpo monoclonal 17b (mAb17b) en presencia (gris claro) y ausencia (gris oscuro) de CD4 soluble según se determina mediante un ensayo ELISA; La FIG. 3A muestra la unión de sC1; La FIG. 3B muestra la unión de sC4;

La FIG. 4 es una imagen de una transferencia Western a partir de una electroforesis en gel de poliacrilamida en condiciones nativas de la proteína sC1 y la proteína de la envoltura del VIH sintética sC4;

La FIG. 5 muestra los niveles de expresión en la superficie celular relativos de las proteínas de la envoltura del VIH sintéticas C1, C1D7, C4 y C4D7 unidas a la membrana mediante un análisis FACS de las células que expresan estas proteínas utilizando un anticuerpo policlonal anti-gp120 (GP120) y mediante la unión a los anticuerpos ampliamente neutralizantes PG9 (PG9) y PG16 (PG16) que presentan una dependencia de la estructura cuaternaria y que se unen preferiblemente al trímero de Env plegado correctamente;

La FIG. 6 es una representación gráfica de la estabilidad de los vectores de adenovirus que contienen secuencias que codifican proteínas de la envoltura del VIH sintéticas de la invención incluida la C4 completa (FLC4), C4D7 y sC4. después de múltiples pases víricos; los vectores de adenovirus 26 recombinante se generaron en células PER.C6; después de 3 pases iniciales para la transfección y la purificación de placas, se seleccionaron 5 placas y se aumentó de escala durante 10 pases en un formato T25, lo que dio como resultado un número de pases víricos total (npv) de 13; se muestra la estabilidad después de 3, 5, 10 y 13 npv según se determina mediante la reacción en cadena de la polimerasa (PCR) del casete del transgén E1; por ejemplo, 3/5 significa que de 5 placas examinadas 3 placas fueron estables y 5/5 significa que de 5 placas examinadas 5 placas fueron estables; y

Las FIGS. 7A y 7B muestran los títulos de neutralización de los virus contra partículas víricas pseudotipadas de la envoltura del VIH-1 (PVE) en un ensayo de neutralización con células TZM-bl en conejos; los valores de la  $CI_{50}$  con una transformación  $\log_{10}$  de los grupos a los que se administraron dosis elevadas del vector adenovírico se midieron frente a PVE VSV-G (control negativo) y MW965.26 (Nivel 1A clado C) en las semanas 1, 8, 14 y 20; cada punto representa el valor de la  $CI_{50}$  con una transformación  $\log_{10}$  de un conejo individual, indicándose la media del grupo con una línea horizontal; HD: dilución más elevada examinada (línea continua superior); LD: dilución más baja examinada (línea continua inferior); LOB: límite de fondo, valor del percentil 95 de las muestras negativas recopiladas (línea discontinua); los valores  $\log_{10}$  de la  $CI_{50}$  que superaron el umbral LD o HD se colocaron en la línea correspondiente; se realizó una comparación no paramétrica unidireccional con control utilizando el método de Dunn para una clasificación conjunta para cada punto temporal; las diferencias estadísticamente significativas se indican en los gráficos: \* =  $P < 0,05$ , \*\* =  $P < 0,01$ , y \*\*\* =  $P < 0,001$ ; La FIG. 7A muestra los resultados con VSV-G (control negativo); y la FIG. 7B muestra los resultados con MW965.26 (Nivel 1A clado C).

#### Descripción detallada de la invención

El análisis de documentos, actos, materiales, dispositivos, artículos o similares, el cual se ha incluido en la presente memoria descriptiva, tiene el fin de proporcionar contexto para la invención. Tal análisis no es una admisión de que todos o cualquiera de estos temas formen parte de la técnica anterior respecto a cualesquiera invenciones descritas o reivindicadas.

A menos que se defina lo contrario, todos los términos científicos y técnicos utilizados en el presente documento tienen los mismos significados que los que normalmente conoce un experto habitual en la materia a la que pertenece esta invención. Por otro lado, ciertos términos utilizados en el presente documento tienen los significados que se exponen en la memoria descriptiva. Cabe destacar que tal como se usa en el presente documento y en las reivindicaciones adjuntas, las formas en singular "un", "una", y "el/la" incluyen la referencia en plural a menos que el contexto estipule claramente otra cosa.

Como se usa en el presente documento, "sujeto" significa cualquier animal, preferiblemente un mamífero, lo más preferiblemente un ser humano, al cual se administrará o se ha administrado un vector, composición o combinación de vacuna de acuerdo con las realizaciones de la invención. El término "mamífero" tal y como se usa en el presente documento, abarca cualquier mamífero. Los ejemplos de mamíferos incluyen, aunque no de forma limitativa, vacas, caballos, ovejas, cerdos, gatos, perros, ratones, ratas, conejos, cobayas, monos, seres humanos, etc., más preferiblemente, un ser humano.

La invención se refiere en general a proteínas de la envoltura del VIH sintéticas, ácidos nucleicos y vectores que codifican las proteínas de la envoltura del VIH sintéticas y métodos para inducir una respuesta inmunitaria contra el VIH con vectores que codifican las proteínas de la envoltura del VIH sintéticas o las proteínas de la envoltura del VIH sintéticas, solos o combinados con uno o más vectores adicionales que codifican uno o más polipéptidos antigénicos del VIH adicionales y/o combinados con uno o más polipéptidos antigénicos del VIH aislados adicionales.

El virus de la inmunodeficiencia humana (VIH) es un miembro del género Lentivirinae, que forma parte de la familia Retroviridae. Dos especies del VIH infectan a los seres humanos: VIH-1 y VIH-2. El VIH-1 es la cepa más común del virus VIH y se sabe que es más patógena que el VIH-2. Como se usa en el presente documento, la expresión "virus de la inmunodeficiencia humana" y el término "VIH" se refieren, aunque no de forma limitativa, al VIH-1 y VIH-2.

el VIH se clasifica en múltiples clados con un alto grado de divergencia genética. Como se usa en el presente documento, la expresión "clado del VIH" o "subtipo del VIH" se refiere a virus de la inmunodeficiencia humana relacionados clasificados de acuerdo con su grado de similitud genética. En la actualidad existen tres grupos de aislados del VIH-1: M, N y O. El grupo M (cepas principales) consiste en al menos diez clados, de la A a la J. El grupo O (otras cepas) puede consistir en un número similar de clados. El grupo N es un aislado del VIH-1 nuevo que no se ha clasificado ni en el grupo M ni en el O.

Como se usa en el presente documento, las expresiones "polipéptido antigénico del VIH", "proteína antigénica del VIH", e "inmunógeno del VIH" se refieren a un polipéptido capaz de inducir una respuesta inmunitaria, por ejemplo, una respuesta de tipo celular y/o humoral, contra el VIH en un sujeto. El polipéptido antigénico puede ser una proteína del VIH, un fragmento o un epítipo de esta, o una combinación de múltiples proteínas del VIH o porciones de estas, que pueden inducir una respuesta inmunitaria o producir inmunidad, por ejemplo, inmunidad protectora, contra el VIH en un sujeto.

Preferiblemente, un polipéptido antigénico es capaz de provocar en un hospedador una respuesta inmunitaria protectora, por ejemplo, inducir una respuesta inmunitaria contra una infección o enfermedad vírica, y/o producir una inmunidad (es decir, vacunados) en un sujeto contra una infección o enfermedad vírica, que protege al sujeto contra la infección o enfermedad vírica. Por ejemplo, el polipéptido antigénico puede comprender una proteína o fragmentos de esta del virus de la inmunodeficiencia de simios (VIS) o de un VIH, tal como la proteína de la envoltura gp160 del VIS o del VIH, las proteínas de la matriz/cápside del VIS o del VIH y los productos de los genes *gag*, *pol* y *env*.

Un polipéptido antigénico del VIH puede ser cualquier antígeno o fragmento de este del VIH-1 o VIH-2. Ejemplos de antígenos del VIH incluyen, pero sin limitación, los productos de los genes *gag*, *pol*, y *env*, que codifican proteínas estructurales y enzimas esenciales. Los productos de los genes *Gag*, *pol* y *env* se sintetizan como poliproteínas, que son procesadas adicionalmente para obtener múltiples productos proteicos diferentes. El producto proteico primario del gen *gag* es la proteína estructural vírica, la poliproteína *gag*, que se procesa adicionalmente en los productos proteicos MA, CA, SP1, NC, SP2 y P6. El gen *pol* codifica enzimas víricas (Pol, polimerasa) y el producto proteico primario se procesa adicionalmente en los productos proteicos RT, RNasa H, IN y PR. El gen *env* codifica proteínas estructurales, específicamente glucoproteínas de la envoltura del virión. El producto proteico primario del gen *env* es gp160, que se procesa adicionalmente en gp120 y gp41. Otros ejemplos de antígenos del VIH incluyen las proteínas reguladoras de genes Tat y Rev; las proteínas complementarias Nef, Vpr, Vif y Vpu; las proteínas de la cápside, las proteínas de la nucleocápside y la proteína vírica p24.

En ciertas realizaciones, el polipéptido antigénico del VIH comprende un antígeno Gag, Env o Pol del VIH, o cualquier porción antigénica o epítipo o combinación de estos, preferiblemente un antígeno Gag, Env o Pol del VIH-1 o cualquier porción antigénica o epítipo o combinación de estos.

Los polipéptidos antigénicos del VIH también pueden ser antígenos del VIH mosaico. Como se usa en el presente documento, "antígeno mosaico" se refiere a una proteína recombinante ensamblada a partir de fragmentos de secuencias naturales. Los antígenos mosaico se parecen a los antígenos naturales, pero están optimizados para maximizar la cobertura de posibles epítipos de los linfocitos T que se observan en las secuencias naturales, lo que mejora la amplitud y cobertura de la respuesta inmunitaria. Los antígenos del VIH mosaico para su uso con la invención son preferiblemente los antígenos Gag, Pol y/o Env mosaico y, más preferiblemente, un antígeno Gag, Pol y/o Env del VIH-1 mosaico. Como se usa en el presente documento, "un antígeno Gag, Pol y/o Env mosaico" se refiere específicamente a un antígeno mosaico que comprende múltiples epítipos obtenidos a partir de una o más de las secuencias de las poliproteínas Gag, Pol y/o Env del VIH.

En una realización, un antígeno del VIH mosaico para su uso con la invención es un antígeno Gag del VIH mosaico con epítipos obtenidos a partir de las secuencias de los productos del gen *gag* (se proporcionan ejemplos en las SEQ ID NO: 1, 2); un antígeno Pol del VIH mosaico con epítipos obtenidos a partir de las secuencias de los productos del gen *pol* (se proporcionan ejemplos en las SEQ ID NO: 3, 4); o un antígeno Env del VIH mosaico con epítipos obtenidos a partir de las secuencias de los productos del gen *env* (se proporcionan ejemplos en las SEQ ID NO: 5, 6; los antígenos novedosos de la invención, p. ej. en las SEQ ID NO: 8, 17, 18, 19, también se pueden considerar antígenos Env del VIH mosaico). En ciertas realizaciones, un antígeno del VIH mosaico para su uso con la invención puede comprender una combinación de epítipos obtenidos a partir de secuencias de los productos de los genes *gag*, *pol*, y/o *env*. Los ejemplos ilustrativos y no limitantes incluyen los antígenos Env-Pol mosaico con epítipos obtenidos a partir de las secuencias de los productos de los genes *env* y *pol*; antígenos Gag-Pol mosaico con epítipos obtenidos a partir de las secuencias de los productos de los genes *gag* y *pol*; antígenos Gag-Env mosaico con epítipos obtenidos a partir de las secuencias de los productos de los genes *gag* y *env*. Las secuencias de los productos de los genes *gag*, *pol*, y *env* se pueden obtener a partir de uno o más clados.

Ejemplos de antígenos Gag, Pol y/o Env del VIH mosaico que se pueden usar en la invención incluyen los descritos en, por ejemplo, US20120076812; Barouch et al., Nat Med 2010, 16:319-323; y Barouch et al., Cell 155:1-9, 2013, todas las cuales se incorporan en el presente documento por referencia en su totalidad. Preferiblemente, los antígenos Gag, Pol y/o Env del VIH mosaico para su uso con la presente invención incluyen, aunque no de forma

limitativa, mos1Gag (SEQ ID NO: 1), mos2Gag (SEQ ID NO: 2), mos1Pol (SEQ ID NO: 3), mos2Pol (SEQ ID NO: 4), mos1Env (SEQ ID NO: 5), mos2Env (SEQ ID NO: 6), mos1GagPol (SEQ ID NO: 28), mos2GagPol (SEQ ID NO: 29), y combinaciones de estos.

Como se usa en el presente documento, cada una de las expresiones "proteína de la envoltura del VIH", "proteína env", y "Env" se refiere a una proteína que se expresa en la envoltura de un virión de un VIH y que permite a un VIH dirigirse y unirse a la membrana plasmática de las células infectadas por VIH, o un fragmento o un derivado de esta que puede inducir una respuesta inmunitaria o producir inmunidad contra el VIH en un sujeto que lo necesite. El gen *env* del VIH codifica la proteína precursora gp160, que es escindida proteolíticamente en dos glucoproteínas de la envoltura maduras, gp120 y gp41. La reacción de escisión está mediada por una proteasa de la célula hospedadora, la furina, en una secuencia sumamente conservada en los precursores glucoproteicos de la envoltura retroviral. Más específicamente, gp160 trimeriza a (gp160)<sub>3</sub> y a continuación, experimenta la escisión en dos gp120 y gp41 asociadas de manera no covalente. La entrada del virus está mediada posteriormente por un trímero de los heterodímeros gp120/gp41. Gp120 es el fragmento de unión del receptor y se une al receptor CD4 en una célula diana que tiene un receptor de este tipo, tal como, por ejemplo, un linfocito T cooperador. Gp41, que está unido a gp120 de manera no covalente, es el fragmento de fusión y proporciona la segunda etapa mediante la cual el VIH entra en la célula. Gp41 se encuentra en un principio enterrado en la envoltura vírica, pero cuando gp120 se une al receptor CD4, gp120 cambia su conformación lo que provoca que gp41 quede expuesto, de modo que puede facilitar la fusión con la célula hospedadora. Gp140 es el ectodominio no escindido de gp160 trimérico, es decir, (gp160)<sub>3</sub>, que se ha utilizado como un sustituto para el estado nativo de la protuberancia vírica escindida.

Según realizaciones de la invención, una "proteína de la envoltura del VIH" puede ser una proteína gp160, gp140, gp120, gp41, combinaciones, fusiones, truncamientos o derivados de estos. Por ejemplo, una "proteína de la envoltura del VIH" puede incluir una proteína gp120 asociada de manera no covalente con una proteína gp41. También puede incluir una proteína gp140 trimérica estabilizada que puede tener o haberse modificado para incluir un dominio de trimerización que estabilice trímeros de gp140. Los ejemplos de dominios de trimerización incluyen, aunque no de forma limitativa, el dominio de trimerización "foldon" de la fibritina T4; el dominio de trimerización de la superhélice obtenido a partir de GCN4; y la subunidad catalítica de la aspartato-transcarbamoilasa de *E. coli* como una etiqueta trimérica. Una "proteína de la envoltura del VIH" también puede ser una proteína de la envoltura del VIH truncada incluyendo, aunque no de forma limitativa, proteínas de la envoltura que comprenden un truncamiento C-terminal en el ectodominio (es decir, el dominio que se extiende en el espacio extracelular), un truncamiento en gp41, tal como un truncamiento en el dominio transmembrana de gp41 o un truncamiento en el dominio citoplasmático de gp41. Una "proteína de la envoltura del VIH" puede además ser un derivado de una proteína de la envoltura del VIH de origen natural que tiene mutaciones en la secuencia, por ejemplo, en los sitios de escisión de furina, y/o las denominadas mutaciones SOSIP.

Preferiblemente, una "proteína de la envoltura del VIH" es una "proteína de la envoltura del VIH sintética". Como se usa en el presente documento, la expresión "proteína de la envoltura del VIH sintética" se refiere a una proteína de la envoltura del VIH que no tiene un origen natural que está optimizada para inducir una respuesta inmunitaria o producir inmunidad contra una o más cepas del VIH de origen natural en un sujeto que lo necesite. Las proteínas Env del VIH mosaico son ejemplos de proteínas Env del VIH sintéticas y la invención proporciona antígenos Env del VIH sintéticos novedosos, por ejemplo, los comprendidos en las SEQ ID NO: 8, 17, 18, o 19.

#### Proteínas de la envoltura del VIH sintéticas y secuencias que codifican las mismas

Las realizaciones de la invención se refieren a proteínas de la envoltura del VIH sintéticas novedosas y moléculas de ácido nucleico que las codifican.

En una realización, la invención se refiere a una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8, o la SEQ ID NO: 8 que tiene una o más mutaciones seleccionadas del grupo que consiste en (i) I529P, (ii) K480E, y (iii) una combinación de EK479-480RRRR, I529P, A471C y T575C. La SEQ ID NO: 8 comprende una gp120 madura sintética y una gp41 truncada sintética sin la región transmembrana, ni el dominio citoplasmático. La SEQ ID NO: 8 es una secuencia que no tiene un origen natural que comprende una quimera de secuencias del antígeno mosaico mos2Env (SEQ ID NO: 6) y otras secuencias de proteínas de la envoltura del VIH. La secuencia del antígeno Env sintético novedoso que comprende la SEQ ID NO: 8 está optimizada para proporcionar una cobertura amplia y una respuesta de linfocitos T potenciada contra el clado C del VIH (en comparación con el antígeno mos2Env (SEQ ID NO: 6)). En ciertas realizaciones, se pueden añadir más aminoácidos a la SEQ ID NO: 8 o una de sus variantes definidas en el presente documento.

En ciertas realizaciones, la proteína de la envoltura del VIH sintética comprende además una secuencia señal. La proteína de la envoltura del VIH sintética se sintetiza con una secuencia señal que es escindida de la cadena polipeptídica naciente durante su transporte al lumen del retículo endoplasmático (RE). En principio, se podría utilizar cualquier secuencia señal conocida. Preferiblemente, se utiliza una secuencia señal Env del VIH o una variante de esta. En la técnica se han utilizado diferentes secuencias señal para proteínas Env del VIH (véase p. ej. el documento WO 2014/107744). En ciertas realizaciones, las secuencias señal comprenden la SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11 o SEQ ID NO: 12. En una realización preferida, la secuencia señal comprende la SEQ ID NO: 9.

En ciertas realizaciones, la proteína de la envoltura del VIH sintética comprende además un dominio transmembrana. El dominio transmembrana ancla la proteína de la envoltura del VIH sintética a la membrana del RE y contribuye al ensamblaje de la membrana y a la función de la envoltura del VIH. Preferiblemente, el dominio transmembrana comprende la SEQ ID NO: 13.

En otra realización, la proteína de la envoltura del VIH sintética comprende una gp41 que tiene un dominio citoplasmático truncado. La gp41 tiene un dominio citoplasmático inusualmente largo en su extremo carboxilo, normalmente de aproximadamente 150 aminoácidos (Edwards et al., J. Virology, 2002, 76:2683-2691). Se ha descrito que el truncamiento del dominio citoplasmático induce la exposición de regiones conservadas en el ectodominio de la proteína Env del VIH-1 (*Id.*). El dominio citoplasmático truncado en una envoltura del VIH sintética de la invención puede variar desde uno hasta aproximadamente 140 aminoácidos, tal como 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, o 140 aminoácidos de un dominio citoplasmático completo. En ciertas realizaciones, el dominio citoplasmático truncado se obtiene a partir de los aminoácidos 704-862 de la SEQ ID NO: 17 (es decir, a partir del dominio citoplasmático de la molécula C4 de la invención), mediante un truncamiento después de un aminoácido dado hasta el extremo C. En una realización preferida, la proteína de la envoltura del VIH sintética comprende un dominio citoplasmático truncado que tiene de 1 a 10 restos de aminoácidos, más preferiblemente de 4 a 8 restos de aminoácidos y, lo más preferiblemente, 7 restos de aminoácidos de un dominio citoplasmático de gp41 del VIH. El dominio citoplasmático, o el fragmento de este, de una proteína de la envoltura del VIH sintética tiene una ubicación C- terminal respecto al dominio extracelular (ectodominio) y cuando la proteína de la envoltura del VIH sintética también comprende un dominio transmembrana, el dominio citoplasmático, o el fragmento de este, tiene una ubicación C-terminal respecto al dominio transmembrana. Véase, por ejemplo, las FIGS. 1A y 1C. En una realización particular, la proteína de la envoltura del VIH sintética comprende una gp41 con un dominio citoplasmático truncado que tiene la secuencia de aminoácidos de la SEQ ID NO: 14 o un fragmento de esta, tal como los restos 1-4 de esta (es decir, NRVR). Se han descrito y se podrían utilizar otros dominios citoplasmáticos truncados (p. ej. Schiernle et al., PNAS 1997; Abrahamyan et al., J Virol 2005).

En las realizaciones en las que la proteína de la envoltura del VIH sintética comprende además un dominio transmembrana y un fragmento de un dominio citoplasmático, se prefiere que la proteína comprenda también la secuencia de aminoácidos de la SEQ ID NO: 37, que contiene los restos 655-682 de la SEQ ID NO: 18, en donde la secuencia de aminoácidos de la SEQ ID NO: 37 está fusionada al extremo C de la SEQ ID NO: 8 y al extremo N del dominio transmembrana.

En una realización particularmente preferida de la invención, la proteína de la envoltura del VIH sintética comprende además un dominio transmembrana, tal como el que tiene la secuencia de aminoácidos de la SEQ ID NO: 13, y un dominio citoplasmático truncado o un fragmento del dominio citoplasmático, tal como el que tiene la secuencia de aminoácidos de la SEQ ID NO: 14 o los restos 1-4 de la SEQ ID NO: 14 (es decir, NRVR). Lo más preferiblemente, la proteína de la envoltura del VIH sintética comprende o consiste en la secuencia de aminoácidos de la SEQ ID NO: 18, con o sin la secuencia señal (es decir, los restos de aminoácidos 1-29 de la SEQ ID NO: 18).

En otra realización, la proteína de la envoltura del VIH sintética comprende un dominio de trimerización que sustituye a una región transmembrana Env. El dominio de trimerización aumenta la estabilidad de una estructura trimérica Env. Preferiblemente, la proteína de la envoltura del VIH sintética comprende un polipéptido gp140 que se modifica para incluir un dominio de trimerización que estabiliza trímeros de gp140. Los ejemplos de dominios de trimerización incluyen, aunque no de forma limitativa, el dominio de trimerización "foldon" de la fibrina T4, tal como el que comprende la secuencia de aminoácidos de la SEQ ID: 16; el dominio de trimerización de la superhélice obtenido a partir de GCN4, tal como el que comprende la secuencia de aminoácidos de la SEQ ID: 15; la subunidad catalítica de la aspartato transcarbamoylase de *E. coli* como una etiqueta trimérica; o motivos de trimerización basados en la matrilina. Si está presente, el dominio de trimerización tiene normalmente una ubicación C-terminal respecto al dominio extracelular (véase la FIG. 1B). En ciertas realizaciones preferidas donde la proteína de la envoltura del VIH sintética comprende un dominio de trimerización, la proteína de la envoltura del VIH sintética comprende la secuencia de aminoácidos de la SEQ ID NO: 19, con o sin la secuencia señal (es decir, los restos de aminoácidos 1-29 de la SEQ ID NO: 19). Estas realizaciones con dominios de trimerización son principalmente útiles para las variantes de ectodominios solubles de la proteína de la envoltura del VIH sintética. En ciertas realizaciones de tales variantes solubles de la invención, es posible mutar el sitio de escisión de furina (p. ej., mutación de Lys en Glu en la posición 480 de la SEQ ID NO: 8) para inactivar este sitio de escisión, de manera que la proteína tenga una sola cadena; esto combina bien con un dominio de trimerización, especialmente con el dominio de trimerización GCN4 de la SEQ ID NO: 19.

Las versiones alternativas de tales variantes de ectodominios solubles de la proteína de la envoltura del VIH sintética sin utilizar dominios de trimerización también son realizaciones de la invención y se pueden preparar a partir de la SEQ ID NO: 8 combinando mutaciones que optimicen el sitio de escisión de furina (reemplazando el dipéptido Gly-Lys en las posiciones 479-480 por cuatro restos de Arg) así como también las denominadas mutaciones SOSIP (mutación I>P en la posición 529, e introducción de un puente disulfuro entre las posiciones 471 y 575 reemplazando



las respectivas Ala y Thr en esas posiciones en la SEQ ID NO: 8 con un resto de Cys cada una). Esto da lugar a una proteína que tiene la secuencia de aminoácidos de la SEQ ID NO: 8 con la siguiente combinación de mutaciones: EK479-480RRRR, I529P, A471C y T575C.

- 5 Una modificación posible para aumentar aún más el contenido trimérico de una proteína de la envoltura del VIH sintética de la invención (que comprende la SEQ ID NO: 8) es la modificación de Ile en Pro en la posición 529. Esta puede ser eficaz tanto para las variantes solubles como las unidas a la membrana.

#### Vectores

- 10 Otro aspecto general de la invención se refiere a vectores que comprenden ácido nucleico que codifica una proteína de la envoltura del VIH sintética. Según realizaciones de la invención, los vectores pueden comprender cualquiera de las proteínas de la envoltura del VIH sintéticas descritas en el presente documento. En una realización preferida de la invención, el vector comprende un ácido nucleico que codifica una proteína de la envoltura del VIH sintética que  
15 comprende la secuencia de aminoácidos de la SEQ ID NO: 8, SEQ ID NO: 17, SEQ ID NO: 18, o SEQ ID NO: 19, y más preferiblemente SEQ ID NO: 18.

- Según realizaciones de la invención, el ácido nucleico que codifica la proteína de la envoltura del VIH sintética está ligado operativamente a un promotor, lo que significa que el ácido nucleico está bajo el control de un promotor. El  
20 promotor puede ser un promotor homólogo (es decir, obtenido a partir de la misma fuente genética que el vector) o un promotor heterólogo (es decir, obtenido a partir de un vector o una fuente genética diferentes). Ejemplos de promotores adecuados incluyen el promotor del citomegalovirus (CMV) y el promotor del virus del sarcoma de Rous (RSV). Preferiblemente, el promotor está ubicado en dirección 5' respecto al ácido nucleico dentro de un casete de expresión. Un ejemplo de secuencia promotora de CMV que puede estar ligada operativamente al ácido nucleico  
25 que codifica la proteína de la envoltura del VIH sintética se muestra en la SEQ ID NO: 24.

- Según realizaciones de la invención, un vector puede ser un vector de expresión. Los vectores de expresión incluyen, aunque no de forma limitativa, vectores para la expresión de proteínas recombinantes y un vector para la administración de ácido nucleico a un sujeto para la expresión en un tejido del sujeto, tal como un vector vírico. Los  
30 ejemplos de vectores víricos adecuados para su uso con la invención incluyen, pero no se limitan a vectores adenovíricos, vectores de virus adenoasociados, vectores del virus de la viruela, vectores MVA, vectores de virus entéricos, vectores del virus de la encefalitis equina de Venezuela, vectores del virus del bosque Semliki, vectores del virus del mosaico del tabaco, vectores lentivíricos, etc. El vector también puede ser un vector no vírico. Los ejemplos de vectores no víricos incluyen, pero sin limitación, plásmidos, cromosomas bacterianos artificiales,  
35 cromosomas de levadura artificiales, bacteriófagos, etc.

- En determinadas realizaciones de la invención, el vector es un vector adenovírico. Un adenovirus de acuerdo con la invención pertenece a la familia de los Adenoviridae y preferiblemente es uno que pertenece al género Mastadenovirus. Puede ser un adenovirus humano, pero también un adenovirus que infecte a otras especies, que  
40 incluyen, pero sin limitación, un adenovirus bovino (p. ej., adenovirus bovino 3, BAdV3), un adenovirus canino (p. ej., CAdV2), un adenovirus porcino (p. ej., PAdV3 o 5) o un adenovirus de simio (que incluye un adenovirus de monos y un adenovirus de simios, tal como un adenovirus de chimpancé o un adenovirus de gorila). Preferiblemente, el adenovirus es un adenovirus humano (HadV o AdHu) o un adenovirus de simio tal como un adenovirus de chimpancé o gorila (ChAd, AdCh o SAdV). En la invención, se entiende que es un adenovirus humano si se hace referencia a Ad sin indicar la especie, p. ej., la notación breve "Ad26" significa lo mismo que HadV26, que es el  
45 adenovirus humano de serotipo 26. Como también se usan en el presente documento, la notación "rAd" significa adenovirus recombinante, por ejemplo, "rAd26" se refiere a un adenovirus 26 humano recombinante.

- Los estudios más avanzados se han realizado utilizando adenovirus humanos y, de acuerdo con ciertos aspectos de la invención, se prefieren adenovirus humanos. En determinadas realizaciones preferidas, un adenovirus recombinante de acuerdo con la invención está basado en un adenovirus humano. En realizaciones preferidas, el  
50 adenovirus recombinante está basado en un adenovirus humano del serotipo 5, 11, 26, 34, 35, 48, 49, 50, 52, etc. De acuerdo con una realización particularmente preferida de la invención, un adenovirus es un adenovirus humano del serotipo 26. Una ventaja de este serotipo es una baja seroprevalencia y/o títulos bajos de anticuerpos neutralizantes  
55 preexistentes en la población humana y la experiencia con su uso en sujetos humanos en ensayos clínicos.

- Los adenovirus de simios por lo general también tienen una baja seroprevalencia y/o títulos bajos de anticuerpos neutralizantes preexistentes en la población humana y se ha publicado una cantidad significativa de investigaciones que utilizan vectores de adenovirus de chimpancé (p. ej., los documentos US6083716; WO 2005/071093; WO  
60 2010/086189; WO 2010085984; Farina et al, 2001, J Virol 75: 11603-13 [13]; Cohen et al, 2002, J Gen Virol 83: 151-55 [69]; Kobinger et al, 2006, Virology 346: 394-401 [70]; Tatsis et al., 2007, Molecular Therapy 15: 608-17 [71]; véase también la revisión de Bangari y Mittal, 2006, Vaccine 24: 849-62 [72]; y la revisión de Lasaro y Ertl, 2009, Mol Ther 17: 1333-39 [73]). Por lo tanto, en otras realizaciones, el adenovirus recombinante de acuerdo con la invención está basado en un adenovirus de simio, por ejemplo, un adenovirus de chimpancé. En ciertas realizaciones, el  
65 adenovirus recombinante se basa en el adenovirus de simios de tipo 1, 7, 8, 21, 22, 23, 24, 25, 26, 27.1, 28.1, 29, 30, 31.1, 32, 33, 34, 35.1, 36, 37.2, 39, 40.1, 41.1, 42.1, 43, 44, 45, 46, 48, 49, 50 o SA7P.

Preferiblemente, el vector de adenovirus es un vector vírico recombinante deficiente en la replicación, tales como rAd26, rAd35, rAd48, rAd5HVR48, etc.

En una realización preferida de la invención, los vectores adenovíricos comprenden proteínas de la cápside procedentes de serotipos raros que incluyen Ad26. En la realización típica, el vector es un virus rAd26. Una "proteína de la cápside de adenovirus" se refiere a una proteína de la cápside de un adenovirus (p. ej., los vectores Ad26, Ad35, rAd48, rAd5HVR48) que están implicados en la determinación del serotipo y/o tropismo de un adenovirus particular. Las proteínas de la cápside normalmente incluyen las proteínas de la fibra, pentona y/o hexona. Como se usa en el presente documento una "proteína de la cápside" para un adenovirus particular, tal como una "proteína de la cápside de Ad26" puede ser, por ejemplo, una proteína de la cápside quimérica que incluye al menos una parte de una proteína de la cápside de Ad26. En ciertas realizaciones, la proteína de la cápside es una proteína de la cápside completa de Ad26. En ciertas realizaciones, la hexona, la pentona y la fibra son de Ad26.

Un experto habitual en la materia comprenderá que se pueden combinar elementos obtenidos a partir de múltiples serotipos en un único vector de adenovirus recombinante. Así pues, se puede producir un adenovirus quimérico que combina propiedades deseables de diferentes serotipos. Así pues, en algunas realizaciones, un adenovirus quimérico de la invención podría combinar la ausencia de inmunidad preexistente de un primer serotipo con características tales como la estabilidad térmica, ensamblaje, anclaje, rendimiento de producción, infección redirigida o mejorada, estabilidad del ADN en la célula diana y similares.

En ciertas realizaciones, el vector de adenovirus recombinante útil en la invención se obtiene principal o totalmente de Ad26 (es decir, el vector es rAd26). En algunas realizaciones, el adenovirus es deficiente en la replicación, p. ej., debido a que contiene una delección en la región E1 del genoma. Para los adenovirus obtenidos a partir de adenovirus que no pertenecen al grupo C, tal como Ad26 o Ad35, es habitual intercambiar la secuencia codificante E4-orf6 del adenovirus por la E4-orf6 de un adenovirus del subgrupo C humano, tal como Ad5. Esto permite la propagación de tales adenovirus en líneas celulares complementarias muy conocidas que expresan los genes E1 de Ad5, tales como, por ejemplo, las células 293, células PER.C6 y similares (véase, p. ej., Havenga, et al., 2006, J Gen Virol 87: 2135-43; WO 03/104467). Sin embargo, tales adenovirus no serán capaces de replicarse en células no complementarias que no expresen los genes E1 de Ad5.

La preparación de vectores adenovíricos recombinantes es muy conocida en la técnica. La preparación de los vectores de rAd26 se describe, por ejemplo, en el documento WO 2007/104792 y en Abbink et al., (2007) Virol 81(9): 4654-63. Ejemplos de secuencias del genoma de Ad26 se encuentran en GenBank Acceso EF 153474 y en la SEQ ID NO: 1 del documento WO 2007/104792. Los ejemplos de vectores útiles para la invención incluyen, por ejemplo, los descritos en el documento WO2012/082918, cuya divulgación se incorpora al presente documento por referencia en su totalidad.

Normalmente, un vector útil en la invención se produce utilizando un ácido nucleico que comprende el genoma adenovírico recombinante completo (p. ej., un plásmido, cósmido o vector de baculovirus). Así pues, la invención también proporciona moléculas de ácido nucleico aislado que codifican los vectores adenovíricos de la invención. Las moléculas de ácido nucleico de la invención pueden estar en forma de ARN o en forma de ADN obtenido por clonación o producido sintéticamente. El ADN puede ser bicatenario o monocatenario.

Los vectores de adenovirus útiles en la invención son normalmente deficientes en la replicación. En estas realizaciones, el virus se vuelve deficiente en la replicación mediante la delección o inactivación de regiones vitales para la replicación del virus, tales como la región E1. Las regiones pueden eliminarse o inactivarse sustancialmente, por ejemplo, insertando un gen de interés, tal como un gen que codifica una proteína de la envoltura del VIH sintética (normalmente ligado a un promotor) o un gen que codifica un polipéptido antigénico del VIH (normalmente ligado a un promotor) dentro de la región. En algunas realizaciones, los vectores de la invención pueden contener delecciones en otras regiones, tales como las regiones E2, E3 o E4, o inserciones de genes heterólogos ligados a un promotor dentro de una o más de estas regiones. Para los adenovirus mutados en E2 y/o E4, se utilizan generalmente líneas celulares complementarias E2 y/o E4 para generar adenovirus recombinantes. Las mutaciones en la región E3 del adenovirus no necesitan estar complementadas por la línea celular, ya que no se requiere E3 para la replicación.

Normalmente se utiliza una línea celular empaquetadora para producir cantidades suficientes de vectores de adenovirus para su uso en la invención. Una célula empaquetadora es una célula que comprende aquellos genes que han sido eliminados o inactivados en un vector deficiente en la replicación, permitiendo de esta manera que el virus se replique en la célula. Las líneas celulares empaquetadoras para los adenovirus con una delección en la región E1 incluyen, por ejemplo, PER.C6, 911.293, y E1 A549.

De acuerdo con las realizaciones de la invención, y como se ha señalado anteriormente, cualquiera de las proteínas de la envoltura del VIH sintéticas descritas en el presente documento se pueden expresar en los vectores de la invención. Dada la degeneración del código genético, el experto es completamente consciente de que se pueden diseñar varias secuencias de ácido nucleico que codifiquen la misma proteína, de acuerdo con métodos totalmente rutinarios en la técnica. El ácido nucleico que codifica la proteína de la envoltura del VIH sintética puede

opcionalmente tener los codones optimizados para garantizar una expresión adecuada en la célula hospedadora tratada (p. ej., humana). La optimización de codones es una tecnología ampliamente aplicada en la técnica. Algunos ejemplos no limitantes de secuencias que codifican una proteína de la envoltura del VIH sintética de la invención se proporcionan en las SEQ ID NO: 25, 26 y 27. Normalmente, el ácido nucleico que codifica la proteína de la envoltura del VIH sintética se clona en la región E1 y/o E3 del genoma adenovírico.

En una realización preferida de la invención, el vector es un vector de adenovirus y, más preferiblemente, un vector de rAd26, o lo más preferiblemente un vector de rAd26 con al menos una delección en la región E1 del genoma adenovírico, p. ej., como el descrito en Abbink, J Virol, 2007. 81(9): p. 4654-63, que se incorpora en el presente documento por referencia.

La invención también proporciona células, preferiblemente células aisladas, que comprenden cualquiera de los vectores descritos en el presente documento. Las células se pueden utilizar para la producción de proteínas recombinantes o para la producción de partículas víricas.

Por lo tanto, las realizaciones de la invención también se refieren a un método para producir un polipéptido antigénico del VIH sintético. El método comprende transfectar una célula hospedadora con un vector de expresión que comprende ácido nucleico que codifica el polipéptido antigénico del VIH sintético ligado operativamente a un promotor, cultivar la célula transfectada en condiciones adecuadas para la expresión del polipéptido antigénico del VIH sintético y aislar el polipéptido antigénico del VIH sintético de la célula. El polipéptido antigénico del VIH sintético se puede aislar o recoger a partir de la célula mediante cualquier método conocido en la técnica incluida la cromatografía por afinidad, etc. Las técnicas utilizadas para la expresión de proteínas recombinantes serán evidentes para el experto en la materia en vista de la presente divulgación.

La invención también incluye un método para producir un vector que codifica un polipéptido antigénico del VIH sintético de la invención, comprendiendo el método cultivar una célula que comprende el vector, propagar y multiplicar el vector durante dicho cultivo y aislar el vector que codifica el polipéptido antigénico del VIH sintético de la invención a partir del cultivo de células, p. ej., de las células, del medio de cultivo o de ambos. El vector se puede purificar adicionalmente de acuerdo con métodos conocidos en la técnica.

En ciertas realizaciones, la invención proporciona un vector de acuerdo con una realización de la invención que comprende un ácido nucleico que codifica un polipéptido antigénico del VIH sintético y, en ciertas realizaciones ilustrativas, el ácido nucleico tiene una secuencia de nucleótidos seleccionada del grupo que consiste en la SEQ ID NO: 25, 26 y 27.

#### Composiciones

En otro aspecto general, la invención se refiere a una composición que comprende un vector que comprende un ácido nucleico que codifica una proteína de la envoltura del VIH sintética y un vehículo. Según realizaciones de la invención, cualquiera de los vectores descritos en el presente documento se pueden incluir en la composición. Preferiblemente, el vector es un vector vírico, más preferiblemente un vector de adenovirus y, aún más preferiblemente, un vector de adenovirus 26. En una realización preferida, una composición comprende un vector de adenovirus, preferiblemente un vector de adenovirus 26 que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8, la SEQ ID NO: 18, o la SEQ ID NO: 19, y más preferiblemente la secuencia de aminoácidos de la SEQ ID NO: 18.

En un aspecto, la invención proporciona una combinación de vacuna que comprende uno o más vectores juntos que comprenden secuencias de ácido nucleico que codifican (i) una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8 (p. ej., la SEQ ID NO: 18 o 19) y (ii) una segunda proteína de la envoltura del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5. Cada uno de los vectores puede estar en una composición independiente o se pueden combinar en una única composición. Se pretende que ambos ácidos nucleicos en el vector o vectores se administren a un sujeto, lo que dará como resultado una respuesta inmunitaria frente al VIH que es más amplia que la respuesta inmunitaria que se obtendría después de la administración de cualquiera de los vectores solos. Ambas secuencias de ácido nucleico pueden estar presentes en un único vector.

Según realizaciones de la invención, una composición comprende una cantidad inmunogénicamente eficaz de un vector, tal como un vector vírico. Como se usa en el presente documento, "cantidad inmunogénicamente eficaz" o "cantidad inmunológicamente eficaz" significa una cantidad de una composición suficiente para inducir una respuesta inmunitaria o efecto inmunitario deseados en un sujeto que lo necesite. En una realización, una cantidad inmunogénicamente eficaz significa una cantidad suficiente para inducir una respuesta inmunitaria en un sujeto que lo necesite. En otra realización, una cantidad inmunogénicamente eficaz significa una cantidad suficiente para producir inmunidad en un sujeto que lo necesite, por ejemplo, proporcionar un efecto protector contra una enfermedad tal como una infección vírica. Una cantidad inmunogénicamente eficaz puede variar dependiendo de varios factores, tales como el estado físico del sujeto, edad, peso, salud, etc.; la aplicación concreta, si se va a inducir una respuesta inmunitaria o proporcionar inmunidad protectora; el vector recombinante específico

administrado; el inmunógeno o polipéptido antigénico codificado por el vector recombinante administrado; el polipéptido antigénico específico administrado; y la enfermedad particular, por ejemplo, infección vírica, para la cual se desea la inmunidad. El experto en la materia puede determinar fácilmente una cantidad inmunogénicamente eficaz en vista de la presente divulgación.

Como guía general, una cantidad inmunogénicamente eficaz cuando se utiliza haciendo referencia a un vector vírico recombinante tal como un vector adenovírico puede estar comprendida entre aproximadamente  $10^8$  partículas víricas y aproximadamente  $10^{12}$  partículas víricas, por ejemplo  $10^8$ ,  $10^9$ ,  $10^{10}$ ,  $10^{11}$ , o  $10^{12}$  partículas víricas. Se puede administrar una cantidad inmunogénicamente eficaz en una única composición o en múltiples composiciones, tales como 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 (p. ej., comprimidos, cápsulas o composiciones inyectables), en donde la administración de las múltiples cápsulas o inyecciones proporciona globalmente al sujeto la cantidad inmunogénicamente eficaz. En general, cuando se utiliza haciendo referencia a un polipéptido, tal como un polipéptido antigénico aislado, una cantidad inmunogénicamente eficaz puede variar desde, p. ej., aproximadamente 0,3 a aproximadamente 3000 microgramos ( $\mu\text{g}$ ), p. ej., 1-1000  $\mu\text{g}$ , p. ej., 10-500  $\mu\text{g}$ , p. ej., aproximadamente 10, 50, 100, 150, 200, 250, 300, 350, 400, 450 o 500  $\mu\text{g}$ . Como un ejemplo no limitante, es posible combinar la administración del vector que codifica el antígeno Env del VIH sintético de la invención (que tiene la SEQ ID NO: 8) con la administración de un polipéptido Env, p. ej., 250  $\mu\text{g}$  de la proteína trimérica Env del Clado C del VIH que tiene los aminoácidos 30-708 de la SEQ ID NO: 7. También es posible administrar una cantidad inmunogénicamente eficaz a un sujeto y, posteriormente, administrar otra dosis de una cantidad inmunogénicamente eficaz al mismo sujeto, en un régimen denominado de sensibilización-refuerzo. Este concepto general de un régimen de sensibilización-refuerzo es muy conocido por el experto en el campo de las vacunas. Opcionalmente, se pueden añadir administraciones de refuerzo adicionales al régimen, según sea necesario.

Las composiciones de la invención comprenden además un vehículo. Un vehículo puede incluir uno o más excipientes farmacéuticamente aceptables tales como aglutinantes, disgregantes, agentes de hinchamiento, agentes de suspensión, agentes emulsionantes, agentes humectantes, lubricantes, saborizantes, edulcorantes, conservantes, colorantes, solubilizantes y agentes de recubrimiento. La naturaleza precisa del vehículo u otro material puede depender de la vía de administración, por ejemplo, intramuscular, subcutánea, oral, intravenosa, cutánea, intramucosal (p. ej., intestino), intranasal o intraperitoneal. Para las preparaciones inyectables líquidas, por ejemplo, suspensiones y soluciones, los vehículos y aditivos adecuados incluyen agua, glicoles, aceites, alcoholes, conservantes, agentes colorantes y similares. Para las preparaciones sólidas orales, por ejemplo, polvos, cápsulas, pastillas, cápsulas de gelatina y comprimidos, los vehículos y aditivos adecuados incluyen almidones, azúcares, diluyentes, agentes de granulación, lubricantes, aglutinantes, agentes disgregantes y similares. Para las mezclas de aerosoles/inhaladores nasales, la solución/suspensión acuosa puede comprender agua, glicoles, aceites, emolientes, estabilizantes, agentes humectantes, conservantes, aromatizantes, sabores y similares como vehículos y aditivos adecuados.

Las composiciones de la invención se pueden formular en cualquier materia adecuada para su administración a un sujeto para facilitar la administración y mejorar la eficacia, incluyendo, aunque no de forma limitativa, la administración oral (enteral) e inyecciones parenterales. Las inyecciones parenterales incluyen la inyección o perfusión intravenosa, inyección intraarterial, inyección subcutánea, inyección intramuscular e inyección intraarticular. Las composiciones de la invención también se pueden formular para otras vías de administración incluidas la administración transmucosal, ocular, rectal, implantes de acción prolongada, administración sublingual, debajo de la lengua, desde la mucosa oral para evitar la circulación portal, por inhalación o intranasal.

De acuerdo con determinadas realizaciones de la invención, una composición comprende una cantidad inmunogénicamente eficaz de un vector de adenovirus purificado o parcialmente purificado, tal como un vector de adenovirus 26, que comprende un ácido nucleico que codifica una proteína de la envoltura del VIH sintética de la invención. Dichas composiciones se pueden formular como una vacuna (también denominada como una "composición inmunogénica") de acuerdo con métodos muy conocidos en la técnica.

Las composiciones de la invención pueden comprender además opcionalmente un adyuvante para potenciar las respuestas inmunitarias. El término "adyuvante" y la expresión "estimulante inmunitario" se utilizan indistintamente en el presente documento y se definen como una o más sustancias que provocan la estimulación del sistema inmunitario. En este contexto, se utiliza un ayudante para potenciar una respuesta inmunitaria frente a los vectores que codifican proteínas de la envoltura del VIH sintéticas de la invención y/o polipéptidos antigénicos del VIH utilizados en combinación con vectores que codifican proteínas de la envoltura del VIH sintéticas de la invención.

Los adyuvantes adecuados para su uso con la invención deberían ser aquellos que son potencialmente seguros, bien tolerados y eficaces en las personas, tales como por ejemplo QS-21, Detox-PC, MPL-SE, MoGM-CSF, TiterMax-G, CRL-1005, GERBU, TERamide, PSC97B, Adjuver, PG-026, GSK-I, GcMAF, B-aletina, MPC-026, Adjuvax, CpG ODN, Betafectina, sales de aluminio (p. ej., AdjuPhos), AdjuPLEX, y MF59. Las proporciones óptimas de cada componente en la formulación se pueden determinar mediante técnicas muy conocidas por los expertos en la materia en vista de la presente divulgación.

En una realización preferida, el adyuvante es una sal de aluminio, tal como AdjuPhos.

Los expertos en la materia están muy familiarizados con la preparación y el uso de composiciones inmunogénicas. Las composiciones farmacéuticas líquidas incluyen, en general, un vehículo líquido tal como agua, vaselina, aceites animales o vegetales, aceite mineral o aceite sintético. También se puede incluir solución salina fisiológica, dextrosa u otra solución de sacáridos o glicoles tales como etilenglicol, propilenglicol o polietilenglicol.

Por ejemplo, el vector de adenovirus recombinante se puede conservar en el tampón que también se utiliza en el Adenovirus World Standard (Hoganson et al., 2002, Bioprocessing J 1: 43-8): Tris 20 mM pH 8, NaCl 25 mM, glicerol al 2,5 %. Otro tampón de formulación de adenovirus útil adecuado para la administración a seres humanos es Tris 20 mM, MgCl<sub>2</sub> 2 mM, NaCl 25 mM, sacarosa 10 % p/v, polisorbato-80 0,02 % p/v. Otro tampón de formulación que es adecuado para el adenovirus recombinante comprende tampón citrato 10-25 mM pH 5,9-6,2, 4-6 % (p/p) de hidroxipropil-beta-ciclodextrina (HBCD), NaCl 70-100 mM, 0,018-0,035 % (p/p) de polisorbato-80, y opcionalmente 0,3-0,45 % (p/p) de etanol. Obviamente, se pueden utilizar muchos otros tampones y existen varios ejemplos de formulaciones adecuadas para el almacenamiento y para la administración farmacéutica de vectores purificados.

Según realizaciones de la invención, se puede utilizar una composición de la invención junto con uno o más vectores adicionales que codifican uno o más polipéptidos antigénicos del VIH adicionales y/o uno o más polipéptidos antigénicos del VIH aislados. Los vectores adicionales y/o polipéptidos antigénicos del VIH pueden estar presentes en la misma composición que comprende una proteína Env del VIH sintética de la invención. También pueden estar presentes en una o más composiciones diferentes que se pueden utilizar junto con una composición que comprende una proteína Env del VIH sintética de la invención en una combinación de vacuna. Preferiblemente, el vector o los vectores adicionales son vectores víricos, tales como vectores de adenovirus y, lo más preferiblemente, son vectores de adenovirus 26. El vector o los vectores adicionales pueden codificar cualquier polipéptido antigénico del VIH que identificarán los expertos en la materia en vista de la presente divulgación.

En una realización, una composición o una combinación de vacuna comprende además un segundo vector de adenovirus, preferiblemente un vector de adenovirus 26, que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5. Una ventaja de tales realizaciones es una mayor amplitud de la respuesta inmunitaria (que abarca las cepas de los Clados B y C).

En otra realización, una composición o una combinación de vacuna de la invención comprende además un vector de adenovirus, preferiblemente un vector de adenovirus 26, que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 28 (mos1GagPol).

En otra realización, una composición o una combinación de vacuna de la invención comprende además un vector de adenovirus, preferiblemente un vector de adenovirus 26, que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 29 (mos2GagPol).

En una realización particular, una composición o una combinación de vacuna de la invención comprende además un segundo vector de adenovirus, preferiblemente un vector de adenovirus 26, que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5 y uno o más vectores de adenovirus adicionales, preferiblemente vectores de adenovirus 26, que codifican uno o más polipéptidos antigénicos del VIH que comprenden la secuencia de aminoácidos seleccionada del grupo que consiste en la SEQ ID NO: 28 o la SEQ ID NO: 29. Por ejemplo, una composición o combinación de vacuna de acuerdo con una realización de la invención puede comprender cuatro vectores de adenovirus, preferiblemente vectores de adenovirus 26, donde un primer vector codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8 (p. ej., la SEQ ID NO: 18); un segundo vector que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5; un tercer vector que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 28; y un cuarto vector que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 29.

En algunas realizaciones, la composición o combinación de vacuna comprende además uno o más polipéptidos antigénicos del VIH aislados. Cualquier polipéptido antigénico del VIH que identifiquen los expertos en la materia en vista de la presente divulgación puede estar incluido además en una composición o combinación de vacuna de la invención, incluyendo, pero sin limitación, una proteína de la envoltura del VIH (p. ej., gp160, gp140, gp120 o gp41), preferiblemente una proteína gp140 trimérica estabilizada, tal como una proteína gp140 del clado C o del clado A estabilizada. En una realización preferida, el polipéptido antigénico del VIH aislado es una proteína gp140 trimérica del clado C del VIH estabilizada, tal como la que comprende los restos 30-708 de la secuencia de aminoácidos de la SEQ ID NO: 7 (los restos 1-29 de la SEQ ID NO: 7 están en la secuencia señal). Un polipéptido Env del VIH alternativo o adicional que se podría utilizar solo o además de la proteína gp140 del clado C, es una proteína trimérica Env mosaico, por ejemplo, que tiene la secuencia de aminoácidos tal como se divulga en los aminoácidos 30-724 de la SEQ ID NO: 36 (correspondiente a la SEQ ID NO: 2 del documento WO 2014/107744, los restos 1-29 de la SEQ ID NO: 36 están en la secuencia señal).

De acuerdo con una realización particular de la invención, una proteína antigénica del VIH puede ser una proteína de la envoltura del VIH sintética de la invención. Así pues, se puede utilizar una proteína de la envoltura sintética de la

invención en forma aislada y/o purificada para inducir una respuesta inmunitaria o proporcionar una inmunidad protectora, etc. contra el VIH en un sujeto que lo necesite. Cualquiera de las proteínas de la envoltura sintética descritas en el presente documento que comprenden la secuencia de aminoácidos de la SEQ ID NO: 8 se puede utilizar como una proteína antigénica del VIH en forma aislada y/o purificada. En una realización preferida, cuando se utiliza en forma aislada como una proteína antigénica del VIH, la proteína de la envoltura sintética comprende los restos 30-711 de la secuencia de aminoácidos de la SEQ ID NO: 18 o los restos 30-686 de la secuencia de aminoácidos de la SEQ ID NO: 19 y, más preferiblemente, los restos 30-704 de la secuencia de aminoácidos de la SEQ ID NO: 18. El polipéptido antigénico del VIH aislado también puede comprender la SEQ ID NO: 8 con las siguientes mutaciones: EK479-480RRRR, I529P, A471C y T575C.

Las realizaciones de la invención también se refieren a composiciones o combinaciones de vacuna que comprenden una proteína de la envoltura del VIH sintética aislada que comprende la secuencia de aminoácidos de la SEQ ID NO: 8. Se puede utilizar cualquiera de las proteínas de la envoltura del VIH sintéticas descritas en el presente documento. En realizaciones en particular de la invención, la proteína de la envoltura del VIH sintética aislada comprende los restos 30-704 o 30-711 de la secuencia de aminoácidos de la SEQ ID NO: 18, los restos 30-686 de la secuencia de aminoácidos de la SEQ ID NO: 19 o la secuencia de aminoácidos de la SEQ ID NO: 8 con las siguientes mutaciones: EK479-480RRRR, I529P, A471C y T575C. Tales composiciones o combinaciones de vacuna pueden comprender además uno o más vectores de expresión, por ejemplo, vectores adenovíricos tales como vectores de adenovirus 26, que codifican uno o más polipéptidos antigénicos del VIH adicionales, tales como las proteínas de la envoltura del VIH sintéticas de la invención, u otras proteínas antigénicas del VIH tales como las expuestas en las SEQ ID NO: 4, 5, 7, 28 o 29 o fragmentos de estas.

La invención también se refiere a un método para producir una composición o combinación de vacuna de la invención. Según realizaciones de la invención, un método para producir una composición o una combinación comprende combinar un vector que comprende ácido nucleico que codifica la proteína de la envoltura del VIH sintética de la invención con un vehículo y opcionalmente uno o más vectores adicionales que codifican uno o más polipéptidos antigénicos del VIH adicionales y/o uno o más polipéptidos antigénicos del VIH aislados. El experto en la materia está familiarizado con las técnicas habituales utilizadas para preparar tales composiciones.

#### Vacuna y combinaciones de vacuna

Según realizaciones de la invención, una composición puede ser una vacuna. Como se usa en el presente documento, el término "vacuna" se refiere a una composición que comprende un vector de expresión, preferiblemente, un vector vírico, que codifica una proteína de la envoltura del VIH sintética de la invención que puede proporcionar inmunidad protectora o una respuesta inmunitaria protectora a un sujeto, o vacunar a un sujeto. Según realizaciones de la invención, tras la administración de la composición a un sujeto, el vector de expresión expresa la proteína de la envoltura del VIH sintética codificada y la proteína de la envoltura del VIH sintética expresada se presenta al sistema inmunitario del sujeto, induciéndose de esta manera la respuesta requerida para producir la inmunidad o para inducir una respuesta inmunitaria.

Así pues, en otro aspecto general, la invención proporciona una vacuna para inducir una respuesta inmunitaria contra un virus de la inmunodeficiencia humana (VIH) en un sujeto. Según realizaciones de la invención, la vacuna comprende una composición que comprende una cantidad inmunogénicamente eficaz de un vector de expresión que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8 y preferiblemente la secuencia de aminoácidos de la SEQ ID NO: 18. Preferiblemente, el vector de expresión es un vector vírico, más preferiblemente un vector de adenovirus y, lo más preferiblemente, un vector de adenovirus 26.

Según realizaciones de la invención, "inducir una respuesta inmunitaria", cuando se utiliza haciendo referencia a los métodos y composiciones descritos en el presente documento, abarca proporcionar inmunidad protectora y/o vacunar a un sujeto contra una infección, tal como una infección por VIH, con fines profilácticos, así como también provocar una respuesta o efecto inmunitarios deseados contra una infección en un sujeto que los necesite, tal como una infección por VIH, con fines terapéuticos. Preferiblemente, los métodos de la invención tienen fines profilácticos, tales como proporcionar inmunidad protectora. La respuesta inmunitaria puede ser una respuesta inmunitaria celular y/o una respuesta inmunitaria humoral.

Como se usa en el presente documento, "inmunidad protectora" o "respuesta inmunitaria protectora" significa que el sujeto vacunado es capaz de controlar una infección con el agente patógeno contra el que se realizó la vacunación. Habitualmente, el sujeto que ha desarrollado una "respuesta inmunitaria protectora" desarrolla únicamente síntomas clínicos leves o moderados o ningún síntoma en absoluto. Habitualmente, un sujeto que tiene una "respuesta inmunitaria protectora" o "inmunidad protectora" contra un cierto agente no fallecerá como resultado de la infección con dicho agente.

Según realizaciones de la invención, las composiciones de vacuna pueden comprender además uno o más vectores adicionales, por ejemplo, vectores víricos, tales como vectores de adenovirus, preferiblemente vectores de adenovirus 26, que codifican uno o más polipéptidos antigénicos del VIH adicionales y/o uno o más polipéptidos

antigénicos del VIH aislados. La proteína de la envoltura del VIH sintética, los vectores adicionales y/o uno o más polipéptidos antigénicos del VIH aislados se pueden formular en la misma composición o en una o más composiciones diferentes en la vacuna.

La invención también se refiere a combinaciones de vacuna para sensibilizar y reforzar una respuesta inmunitaria frente a uno o más clados del VIH en un sujeto que lo necesite utilizando uno o más vectores combinados con un polipéptido antigénico aislado. Así pues, en otro aspecto general, la invención proporciona una combinación de vacuna para inducir una respuesta inmunitaria contra un VIH en un sujeto. Según realizaciones de la invención, la combinación de vacuna comprende:

- (i) una primera composición que comprende una cantidad inmunogénicamente eficaz de un vector de expresión que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8 o la SEQ ID NO: 8 que tiene una o más mutaciones seleccionadas del grupo que consiste en (a) I529P, (b) K480E, y (c) una combinación de EK479-480RRRR, I529P, A471C y T575C, y un vehículo; y
- (ii) una segunda composición que comprende una cantidad inmunogénicamente eficaz de un polipéptido antigénico del VIH aislado y un vehículo,

en donde una de la primera y segunda composiciones se destina a la inmunización de sensibilización y la otra composición se destina a la inmunización de refuerzo.

Según realizaciones de la invención, la combinación de vacuna comprende además opcionalmente una cantidad inmunogénicamente eficaz de uno o más vectores de expresión adicionales que codifican uno o más polipéptidos antigénicos del VIH adicionales. El vector o los vectores de expresión adicionales pueden estar incluidos en la primera composición o en la segunda composición, o el vector o los vectores de expresión adicionales pueden estar incluidos en una o más composiciones adicionales que se administrarán junto con la primera y/o la segunda composición.

Como se usa en el presente documento, los términos "cosuministro", "coadministración" o "administrado junto con" se refieren a la administración simultánea de dos o más componentes, tales como un vector de expresión vírico y un polipéptido antigénico aislado o múltiples vectores de expresión víricos. La "administración simultánea" puede ser la administración de los dos o más componentes al menos en el mismo día. Cuando dos componentes se "administran junto con", se pueden administrar en composiciones independientes de manera secuencial en un periodo de tiempo corto, tal como 24, 20, 16, 12, 8 o 4 horas o en 1 hora o menos o se pueden administrar en una única composición a la vez.

En las realizaciones particulares de una combinación de vacuna de la invención, la primera composición comprende un vector de adenovirus, preferiblemente un vector de adenovirus 26, que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 18; y el polipéptido antigénico del VIH aislado comprende los restos 30-708 de la secuencia de aminoácidos de la SEQ ID NO: 7 o los restos 30-724 de la SEQ ID NO: 36. En una realización particular, la primera composición comprende además un vector de adenovirus, preferiblemente un vector de adenovirus 26, que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5. En otra realización particular, la primera composición comprende además uno o más vectores de adenovirus adicionales, preferiblemente vectores de adenovirus 26, que codifican uno o más polipéptidos antigénicos del VIH adicionales que comprenden las secuencias de aminoácidos seleccionadas del grupo que consiste en las SEQ ID NO: 28 y 29.

Otro aspecto general de la invención se refiere a un kit que comprende una combinación de vacuna de acuerdo con una realización de la invención.

Otras realizaciones de la proteína de la envoltura del VIH sintética, vectores de expresión, vectores de expresión adicionales, polipéptidos antigénicos del VIH codificados por los vectores de expresión y polipéptido antigénico del VIH aislado, etc. que se pueden utilizar en las combinaciones de vacuna de la invención se han analizado detalladamente anteriormente y en los ejemplos ilustrativos a continuación.

#### Método para inducir inmunidad protectora contra la infección por VIH

La invención también se refiere a un método para inducir una respuesta inmunitaria contra uno o más clados del VIH en un sujeto que lo necesite. Los métodos descritos en el presente documento incluyen métodos de sensibilización y refuerzo de una respuesta inmunitaria utilizando uno o más vectores de expresión en combinación con uno o más polipéptidos antigénicos aislados.

En un aspecto general, un método para inducir una respuesta inmunitaria contra un virus de la inmunodeficiencia humana (VIH) en un sujeto comprende administrar al sujeto una composición que comprende una cantidad inmunogénicamente eficaz de un vector de expresión que comprende un ácido nucleico que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8. Se pueden utilizar cualquiera de las composiciones descritas en el presente documento en un método para inducir una respuesta

inmunitaria contra el VIH en un sujeto. Preferiblemente, la composición comprende un vector de adenovirus, preferiblemente un vector de adenovirus 26, comprendiendo el vector un ácido nucleico que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 18. La composición puede comprender además uno o más vectores adicionales que codifican uno o más polipéptidos antigénicos del VIH adicionales y/o uno o más polipéptidos antigénicos del VIH aislados adicionales.

En otro aspecto general, un método para inducir una respuesta inmunitaria contra un virus de la inmunodeficiencia humana (VIH) en un sujeto comprende:

- (i) administrar al sujeto una primera composición que comprende una cantidad inmunogénicamente eficaz de un vector de expresión que codifica una proteína de la envoltura del VIH mosaico que tiene la secuencia de aminoácidos de la SEQ ID NO: 8 o la SEQ ID NO: 8 que tiene una o más mutaciones seleccionadas del grupo que consiste en (a) I529P, (b) K480E, y (c) una combinación de EK479-480RRRR, I529P, A471C y T575C, y un vehículo;
- (ii) administrar al sujeto una segunda composición que comprende una cantidad inmunogénicamente eficaz de un polipéptido antigénico del VIH aislado y un vehículo; y
- (iii) opcionalmente, administrar al sujeto una cantidad inmunogénicamente eficaz de uno o más vectores de expresión adicionales que codifican uno o más polipéptidos antigénicos del VIH adicionales,

en donde las etapas (i) y (ii) se llevan a cabo en cualquier orden, siendo una de las etapas para la inmunización de sensibilización y la otra etapa para la inmunización de refuerzo. Según realizaciones de la invención, la cantidad eficaz, opcional del vector o de los vectores de expresión adicionales se administra junto con la primera composición o la segunda composición. En una realización preferida, la cantidad eficaz opcional del vector o los vectores de expresión adicionales se administra junto con la primera composición.

En una realización particular de un método para inducir una respuesta inmunitaria, la primera composición comprende un vector de adenovirus, preferiblemente un vector de adenovirus 26, que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 8 y un segundo vector de adenovirus, preferiblemente un vector de adenovirus 26, que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5; la segunda composición comprende un polipéptido antigénico del VIH aislado que tiene los restos 30-708 de la secuencia de aminoácidos de la SEQ ID NO: 7 o los restos 30-724 de la SEQ ID NO: 36; y el vector o los vectores de expresión adicionales son vectores de adenovirus, preferiblemente vectores de adenovirus 26, que codifican uno o más polipéptidos antigénicos del VIH adicionales que comprenden las secuencias de aminoácidos seleccionadas del grupo que consiste en las SEQ ID NO: 28 y 29; en donde la primera composición se administra al sujeto, junto con el vector o los vectores de expresión adicionales, una o más veces para la inmunización de sensibilización y la segunda composición se administra al sujeto una o más veces para la inmunización de refuerzo.

La administración de las composiciones inmunogénicas que comprenden los vectores de expresión y/o los polipéptidos antigénicos es normalmente intramuscular, intradérmica o subcutánea. Sin embargo, también se pueden contemplar otros modos de administración tales como la administración intravenosa, rectal, cutánea, oral, nasal, etc. La administración intramuscular de las composiciones inmunogénicas se puede conseguir utilizando una aguja para inyectar una suspensión de los vectores de expresión, p. ej., vectores de adenovirus, y/o polipéptidos antigénicos. Una alternativa es el uso de un dispositivo de inyección sin aguja para administrar la composición (utilizando, por ejemplo, Biojector™) o un polvo liofilizado que contiene la vacuna.

Para la inyección intramuscular, intravenosa, cutánea o subcutánea, o la inyección en el sitio de la dolencia, el vector estará en forma de una solución acuosa parenteralmente aceptable sin pirógenos y con un pH, isotonicidad y estabilidad adecuados. De manera análoga, el polipéptido antigénico aislado estará en forma de una solución parenteralmente aceptable con un pH, isotonicidad y estabilidad adecuados. Los expertos en la materia estarán capacitados para preparar soluciones adecuadas utilizando, por ejemplo, vehículos isotónicos tales como inyección de cloruro de sodio, inyección de Ringer, inyección de Ringer lactato. Pueden incluirse conservantes, estabilizantes, tampones, antioxidantes y/u otros aditivos, según sea necesario. También se puede emplear una formulación de liberación lenta.

Normalmente, la administración de las composiciones de vacuna de acuerdo con las realizaciones de la invención tendrá un objetivo profiláctico para generar una respuesta inmunitaria contra un antígeno del VIH antes de la infección o el desarrollo de los síntomas. En otras realizaciones, los vectores de expresión, por ejemplo, los vectores de adenovirus, y/o polipéptidos antigénicos del VIH se pueden administrar para una profilaxis posterior a la exposición.

Las composiciones inmunogénicas que contienen los vectores de expresión, por ejemplo, vectores de adenovirus, y/o polipéptidos antigénicos se administran a un sujeto, provocando una respuesta inmunitaria en el sujeto. Una cantidad de una composición suficiente para inducir una respuesta inmunitaria detectable se define como una "dosis inmunogénicamente eficaz" o "cantidad inmunogénicamente eficaz". En una realización típica de la invención, la respuesta inmunitaria es una respuesta inmunitaria protectora.



La cantidad real administrada, y la tasa y la pauta de administración, dependerán de la naturaleza y la gravedad de lo que se está tratando. La prescripción del tratamiento, por ejemplo, decisiones sobre la dosificación, etc., es responsabilidad de los médicos generalistas y otros médicos, o en un contexto veterinario del veterinario, y normalmente tiene en cuenta el trastorno que se va a tratar, el estado del paciente individual, el sitio de suministro, el método de administración y otros factores conocidos por los facultativos. Se pueden encontrar ejemplos de las técnicas y protocolos mencionados anteriormente en Remington's Pharmaceutical Sciences, 16ª edición, Osol, A. ed., 1980.

Después de la producción de vectores de adenovirus y la formulación opcional de tales partículas en composiciones, los vectores se pueden administrar a un individuo, especialmente a un ser humano u otro primate. El suministro a un mamífero no humano no necesita tener un fin terapéutico, sino que se puede utilizar en un contexto experimental, por ejemplo, en la investigación de los mecanismos de las respuestas inmunitarias a la proteína de la envoltura del VIH sintética expresada por los vectores de adenovirus de la invención.

En algunas realizaciones de los métodos divulgados, se utilizan uno o más vectores de adenovirus que codifican uno o más polipéptidos antigénicos del VIH para sensibilizar y provocar la respuesta inmunitaria. Se pueden utilizar uno o más polipéptidos antigénicos del VIH aislados junto con el vector o los vectores de adenovirus para la inmunización de sensibilización. La inmunización de sensibilización se puede administrar solo una vez, pero también se puede administrar opcionalmente múltiples veces, por ejemplo, la administración de sensibilización inicial en tiempo 0, seguida por otra administración de sensibilización aproximadamente 4-14 semanas, p. ej., 4, 5, 6, 7, 8, 9, 10, 11, 12, 13 o 14 semanas, o en cualquier momento intermedio, después de la administración de sensibilización inicial. Se pueden utilizar uno o más polipéptidos antigénicos del VIH aislados opcionalmente junto con uno o más adenovirus u otros vectores adicionales que codifican uno o más polipéptidos antigénicos del VIH adicionales para reforzar la respuesta inmunitaria. También se puede administrar una inmunización de refuerzo una o múltiples veces, por ejemplo, la primera aproximadamente 18-36, p. ej., 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35 o 36 semanas, o en cualquier momento intermedio, después de la administración de sensibilización inicial, seguida por otra administración de refuerzo aproximadamente 36-52, p. ej., 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59 o 60 semanas, o en cualquier momento intermedio, después de la administración de sensibilización inicial. Se hace un seguimiento de la respuesta inmunitaria inducida por la inmunización.

Las realizaciones de los métodos divulgados también contemplan pautas más cortas de sensibilización-refuerzo, lo que significa que la inmunización de refuerzo final se administra aproximadamente 22-26 semanas después de la administración de sensibilización inicial. La inmunización de sensibilización se puede administrar en la semana 0. La inmunización de refuerzo se puede administrar múltiples veces, por ejemplo, en primer lugar aproximadamente 7-9 semanas u 11-13 semanas o aproximadamente 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, o 14 semanas, o cualquier momento intermedio, después de la administración de sensibilización inicial, seguida por otra administración de refuerzo aproximadamente 22-26 semanas o aproximadamente 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, o 28 semanas, o cualquier momento intermedio, después de la administración de sensibilización inicial. En ciertas realizaciones, se administran uno o más polipéptidos antigénicos del VIH aislados junto con el vector o los vectores de adenovirus para la inmunización de sensibilización y/o refuerzo.

Los expertos en la materia comprenderán fácilmente que la pauta para las administraciones de sensibilización y refuerzo se puede ajustar en función de las respuestas inmunitarias medidas después de las administraciones. Por ejemplo, las composiciones de refuerzo se administran generalmente semanas o meses después de la administración de la composición de sensibilización, por ejemplo, aproximadamente 2-3 semanas o 4 semanas u 8 semanas o 16 semanas o 20 semanas o 24 semanas o 28 semanas o 30 semanas o 32 semanas o de uno a dos años después de la administración de la composición de sensibilización.

Según realizaciones de la invención, se puede administrar un adyuvante junto con el polipéptido antigénico del VIH aislado como parte de la inmunización de sensibilización y/o de refuerzo. Se puede utilizar cualquier adyuvante en vista de la presente divulgación, y en ciertas realizaciones el adyuvante es una sal de aluminio, tal como AdjuPhos.

En una realización preferida de la invención, los vectores de adenovirus utilizados en los métodos divulgados en el presente documento incluyen un vector de rAd26. Preferiblemente, un vector de rAd26 que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 18 o la SEQ ID NO: 19), lo más preferiblemente la SEQ ID NO: 18, se utiliza para sensibilizar la respuesta inmunitaria, solo o en combinación con uno o más vectores de rAd26 adicionales que codifican uno o más polipéptidos antigénicos del VIH adicionales, tales como mos1Env que tiene la secuencia de aminoácidos de la SEQ ID NO: 5, y un polipéptido antigénico del VIH aislado, tal como el que comprende los restos 30-708 de la secuencia de aminoácidos de la SEQ ID NO: 7 o los restos 30-724 de la SEQ ID NO: 36, se utiliza para reforzar la respuesta inmunitaria o viceversa.

En una realización a modo de ejemplo, un vector de rAd26 que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 18 se utiliza para sensibilizar la respuesta inmunitaria en combinación con un vector de rAd26 que codifica un polipéptido antigénico del VIH que tiene la

secuencia de aminoácidos de la SEQ ID NO: 5. Uno o más vectores de rAd26 adicionales que codifican uno o más polipéptidos antigénicos del VIH adicionales que tienen las secuencias de aminoácidos seleccionadas del grupo que consiste en las SEQ ID NO: 1-4, 28 y 29 también se pueden administrar junto con los otros vectores de rAd26 para sensibilizar la respuesta inmunitaria. La administración de sensibilización en ciertas realizaciones se administra dos veces antes de que se administre cualquier inmunización de refuerzo. Un polipéptido antigénico del VIH aislado, tal como el que comprende los restos 30-708 de la secuencia de aminoácidos de la SEQ ID NO: 7 (preferiblemente), o el que comprende los restos 30-724 de la secuencia de aminoácidos de la SEQ ID NO: 36, o una combinación de al menos dos de tales polipéptidos antigénicos del VIH aislados, se administra a continuación para reforzar la respuesta inmunitaria y preferiblemente, se administra más de una vez. Preferiblemente, se administra además un ayudante con el polipéptido antigénico del VIH aislado en la inmunización de refuerzo.

En una realización particular, la respuesta inmunitaria se sensibiliza mediante la administración de cuatro antígenos del VIH codificados en los vectores adenovíricos, preferiblemente vectores de rAd26, siendo los cuatro antígenos codificados: (i) una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 18, (ii) el polipéptido que tiene la secuencia de aminoácidos de la SEQ ID NO: 5, (iii) el polipéptido que tiene la secuencia de aminoácidos de la SEQ ID NO: 28 y (iv) el polipéptido que tiene la secuencia de aminoácidos de la SEQ ID NO: 29. Cada uno de estos cuatro antígenos puede estar codificado en un vector adenovírico independiente, preferiblemente un vector de rAd26, administrado con una dosis total de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9 o  $10 \times 10^{10}$  partículas víricas (pv), p. ej., aproximadamente  $5 \times 10^{10}$  pv (para todos los vectores juntos). Los vectores pueden estar premezclados, p. ej., en una proporción de 1:1:1:1. La administración de sensibilización se puede repetir después de la administración de sensibilización inicial, p. ej., 8, 9, 10, 11, 12, 13, 14, 15 o 16 semanas después de la administración de sensibilización inicial. En esta realización, se refuerza una respuesta inmunitaria administrando la misma vacuna con el vector adenovírico utilizada para la administración de sensibilización junto con una proteína Env gp140 del VIH aislada, p. ej., proteína gp140 del clado C (que comprende los restos 30-708 de la secuencia de aminoácidos de la SEQ ID NO: 7), o la proteína gp140 mosaico (que comprende los restos 30-724 de la secuencia de aminoácidos de la SEQ ID NO: 36), o la proteína gp140 del clado C y la proteína gp140 mosaico, con una dosis total de aproximadamente 50-300 µg de proteína, 50, 100, 150, 200, 250 o 300 microgramos, o cualquier cantidad intermedia, de la proteína gp140 del clado C, o, p. ej., 50, 100, 150, 200, 250 o 300 microgramos, o cualquier cantidad intermedia, de la proteína gp140 mosaico, o, p. ej., 50, 100, 150, 200, 250 o 300 microgramos, o cualquier cantidad intermedia, de una combinación de la proteína gp140 del clado C y la proteína gp140 mosaico (p. ej., en una proporción 1:1, ya sea mezcladas o administradas por separado). Preferiblemente, la proteína gp140 se administra junto con un adyuvante, p. ej., fosfato de aluminio. La administración del adenovirus más la proteína gp140 para reforzar la respuesta inmunitaria se puede realizar aproximadamente 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29 o 30 semanas, o en cualquier momento intermedio, después de la administración de sensibilización inicial. La administración de refuerzo se puede repetir, p. ej., aproximadamente 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53 o 54 semanas, o cualquier momento intermedio, después de la administración de sensibilización inicial. Todas las administraciones de acuerdo con esta realización se realizan preferiblemente mediante la vía intramuscular.

## Ejemplos

### Ejemplo 1: Diseño de las secuencias de antígenos de la envoltura del VIH

Se diseñaron varias secuencias de antígenos de la envoltura del VIH que tenían similitud de secuencia con el antígeno del VIH mosaico mos2Env (SEQ ID NO: 6; descrito también previamente en el documento WO 2010/059732). Las secuencias unidas a la membrana, recién diseñadas, se basaron en (una combinación de) secuencias de tipo silvestre totalmente naturales de proteínas de la envoltura del VIH o una quimera de la secuencia de mos2Env y secuencias de proteínas de la envoltura del VIH de tipo silvestre. Además de las secuencias de proteínas de la envoltura completas (véase la FIG. 1A), también se diseñaron secuencias que tenían un truncamiento C-terminal del dominio citoplasmático (véase, por ejemplo, la FIG. 1C). Véase también p. ej., Schiernle et al., PNAS 1997; Abrahamyan et al., J Virol 2005; Edwards et al., J. Virology, 2002, 76:2683-2691. También se prepararon variantes solubles mediante un truncamiento C-terminal antes de la región transmembrana (TM), que fue reemplazada por un dominio de trimerización, tal como un dominio de trimerización GCN4 (véase, por ejemplo, la FIG. 1B). Estas variantes solubles se convirtieron además en una variante monocatenaria mediante la mutación del sitio de escisión de furina, inhibiendo así el procesamiento del dominio extracelular de la proteína de la envoltura para obtener las subunidades gp120 y gp41.

De todas las construcciones generadas y probadas, las construcciones basadas en C4 tenían las propiedades más óptimas, por ejemplo, una buena procesabilidad, plegado, inmunogenicidad, etc. y estas se seleccionaron para estudios adicionales. También se generaron una variante soluble de la construcción C4 que tenía un dominio de trimerización GCN4 en lugar del dominio transmembrana (sC4, FIG. 1B) y una variante que comprendía un fragmento de 7 aminoácidos del dominio citoplasmático (C4D7, FIG. 1C) y se examinaron en estudios adicionales. Las secuencias de aminoácidos de C4, sC4, y C4D7 se muestran en las SEQ ID NO: 17, 19, y 18, respectivamente. Las secuencias que codifican estas se muestran en las SEQ ID NO: 25, 27, y 26, respectivamente. La construcción C1 tiene una secuencia del dominio extracelular basada en la secuencia de mos2Env (SEQ ID NO: 6). También se generaron una variante soluble de la construcción C1 que tiene un dominio de trimerización GCN4 en lugar del dominio transmembrana (sC1) y una variante que comprende un fragmento de 7 aminoácidos del dominio

citoplasmático (C1D7), similar a sC4 y C4D7 tal como se muestra en las FIGS. 1B y 1C, respectivamente. La construcción C1 y sus variantes se utilizaron en estudios adicionales a efectos de comparación, ya que están basadas esencialmente en la secuencia de mos2Env de la técnica anterior. Las secuencias de aminoácidos de C1, sC1, y C1D7 se muestran en las SEQ ID NO: 31, 30, y 32, respectivamente. Las secuencias de ácidos nucleicos que codifican estas se muestran en las SEQ ID NO: 34, 33, y 35, respectivamente. Otras construcciones que se examinaron resultaron ser menos óptimas que las basadas en la construcción C4 y no se utilizaron para un desarrollo adicional.

#### Ejemplo 2: Expresión y plegamiento de las proteínas de la envoltura del VIH sintéticas

Se midieron el nivel de expresión, plegamiento y expresión en la superficie celular de las proteínas de la envoltura del VIH sintéticas.

#### Niveles de expresión

Las células HEK293F se transfectaron de manera transitoria con un plásmido que codifica las proteínas de la envoltura del VIH sintéticas solubles sC1 y sC4 tal como se describen en el Ejemplo 1. Se midieron los niveles de expresión de la proteína soluble en el sobrenadante utilizando una transferencia Western cuantitativa (QWB). Los resultados se muestran en la FIG. 2. Los bajos niveles de expresión para sC1 (que corresponde esencialmente a mos2Env con un dominio transmembrana añadido) coinciden con los recientes conocimientos respecto a mos2Env de los autores. Tal como demuestran los resultados, la variante sC4 de la invención mostró unos niveles de expresión significativamente más elevados que la variante sC1 (control).

#### Plegamiento de las proteínas

Se examinó el plegamiento de las proteínas midiendo la unión de las proteínas de la envoltura del VIH sintéticas solubles a un anticuerpo (MAb 17b) que se sabe que se une al sitio de unión del correceptor de la proteína de la envoltura del VIH, que está expuesto únicamente después de la unión de CD4, mediante un ensayo inmunoabsorbente ligado a enzima (ELISA). En particular, se examinó la unión de sC4 purificada para determinar la unión a MAb 17b con una unión anterior de sC4 a CD4 y sin una unión anterior de sC4 a CD4. Se utilizó sC1 purificada como control. La unión de MAb17 a sC4 sin una unión anterior de CD4 a la proteína de la envoltura es una indicación de una proteína de la envoltura parcialmente desplegada o en un estado previo al inicio del cambio conformacional (es decir, una Env inestable que adopta la conformación "abierta" en ausencia de unión de CD4). Los resultados del ensayo ELISA se muestran en las FIGS. 3A y 3B.

Como se muestra en la FIG. 3B, sC4 muestra una fuerte unión a MAb 17b con una unión anterior a CD4, pero no hay una unión detectable a MAb 17b sin una unión anterior a CD4. Por el contrario, como se muestra en la FIG. 3A, sC1 mostró una unión a MAb 17 mucho menor tanto con unión anterior a CD4 como sin ella. Los resultados sugieren que sC4 tiene un patrón de plegamiento correcto, sin una exposición del sitio de unión del correceptor antes de la unión a CD4.

El plegamiento de las proteínas también se analizó mediante electroforesis en gel de poliacrilamida (PAGE) en condiciones nativas de sC1 y sC4 para evaluar la estructura cuaternaria de las variantes solubles de la proteína y la posible formación de puentes disulfuro incorrectos entre los protómeros. Después de la electroforesis en un gel en condiciones nativas, se detectó la proteína en el gel mediante un análisis de transferencia Western. Tal como muestran los resultados en la FIG. 4, la mayoría de sC4 está presente en un estado trimérico, que es la estructura cuaternaria correcta.

En conjunto, los resultados de los experimentos del plegamiento de proteínas demuestran que la proteína de la envoltura del VIH sintética soluble sC4 tiene el perfil de plegamiento deseado, que ha mejorado en comparación con el perfil de plegamiento del antígeno mos2Env existente (representado por sC1).

#### Expresión en la superficie celular

También se estudió la expresión en la superficie celular de las variantes unidas a la membrana de las proteínas de la envoltura del VIH C1 (completa), C4 (completa, véase la FIG. 1A), C1D7 y C4D7. Las células HEK293T se transfectaron de manera transitoria solamente con un plásmido que codifica eGFP (control negativo, CN) o con un plásmido que codifica eGFP junto con una construcción de expresión que codifica una variante de una proteína de la envoltura del VIH. Dos días después de la transfección, las células se sometieron a un análisis por clasificación celular activada por fluorescencia (FACS) tras la exposición a varios anticuerpos policlonales y monoclonales dirigidos contra gp120 y anticuerpos secundarios, y después se examinaron para determinar los niveles de expresión en la superficie celular de la proteína de la envoltura. Se evaluó la calidad de las variantes de la envoltura determinando los niveles de expresión globales utilizando un anticuerpo policlonal anti-gp120 y evaluando la unión relativa de los anticuerpos ampliamente neutralizantes PG9 y PG16, que presentan una dependencia de la estructura cuaternaria, y se unen preferiblemente al trimero de la envoltura plegado correctamente.

Los resultados de los experimentos de expresión en la superficie celular se muestran en la FIG. 5. Los niveles de expresión en la superficie de las variantes truncadas C1D7 y C4D7 según se midieron utilizando un anticuerpo anti-gp120, son mucho más elevados que los niveles de expresión en la superficie de sus homólogos completos, C1 y C4, respectivamente. Esto confirma que la delección de 144 restos del extremo carboxi de Env incrementa los niveles de expresión en la superficie de la envoltura. La construcción C4 completa de la invención también mostró una mejor unión a PG9 y PG16 en comparación con C1 completa, lo que sugiere que la secuencia de la envoltura C4 está plegada adecuadamente (es decir, un trímero) en la superficie celular.

Los resultados también demuestran que la variante C1D7, que es esencialmente Mos2Env con un dominio transmembrana añadido y 7 aminoácidos del dominio citoplasmático, se puede expresar en la superficie de células HEK293T. Esto contrasta con la construcción soluble en Ad26.mos2Env, que no se puede expresar con niveles detectables en la superficie cuando se transfecta a células A549. Sin embargo, la unión relativa a PG9 y PG16 es apenas detectable por encima del fondo, lo que sugiere que la secuencia de la envoltura C1D7 está plegada de manera deficiente y, probablemente, no está presente como un trímero intacto en la superficie celular.

En general, la variante de la envoltura C4D7 tiene el perfil de unión al anticuerpo más óptimo, con una expresión de gp120 mayor que la de su homólogo completo C4 y con un aumento en la unión a PG9 y PG16 de más de 15 veces en comparación con C1 y C1D7 (FIG. 5).

#### Ejemplo 3: Estabilidad de los vectores que codifican las secuencias de la envoltura del VIH

Trabajos previos en los laboratorios de los autores (no publicados) indicaron que los vectores de adenovirus 26 (Ad26) que codifican la secuencia del antígeno mos2Env mostraban unas proporciones PV/UI relativamente altas (lo que indica una calidad menor de los lotes de productos de adenovirus) y además que tales vectores mostraban problemas de estabilidad. En consecuencia, era importante examinar la estabilidad de las construcciones de las proteínas de la envoltura del VIH sintéticas de la invención en un contexto de adenovirus.

Se generaron vectores de Ad26 recombinantes (rAd26) que codifican las secuencias de antígenos del VIH de la invención C4, C4D7 y sC4 tal como se ha descrito anteriormente en el Ejemplo 1 en células PER.C6 (denominados respectivamente "rAd26.C4", "rAd26.C4D7", y "rAd26.sC4", respectivamente). Se recogieron clones del vector (placas) y se aumentó de escala para generar los lotes de investigación. Se aumentaron de escala como máximo 5 clones víricos (placas) hasta un formato T25 y se realizaron pases en serie durante 10 pases en un formato T25 (siendo los pases 1-3 las etapas de transfección y purificación de la placa, seguido de 10 pases en un formato T25, lo que dio como resultado un total de 13 pases). Se evaluó la estabilidad genética con un número de pases víricos (npv) de 3, 5, 10 y 13 mediante un ensayo de PCR del casete del transgén E1, seguido de secuenciación en npv 13. Los resultados se muestran en la FIG. 6.

Los vectores de rAd26 que codifican C4 completa (rAd26.C4) mostraron unas características de crecimiento deficientes, según se determinó por la ausencia de un efecto citopatógeno completo (CPE) en 2-3 días; inestabilidad genética, según se determina mediante las delecciones de la región del casete del transgén E1; o una combinación de estas (FIG. 6). Debido a las características de crecimiento deficientes y la inestabilidad genética observada, se dejó de considerar este vector que codifica C4 completa.

Por el contrario, para los vectores de rAd26 que codifican C4D7 (rAd26.C4D7) y sC4 (rAd26.sC4), todas las placas propagadas siguieron siendo genéticamente estables durante el transcurso del experimento (FIG. 6). Así pues, las construcciones de sC4 y C4D7 novedosas tuvieron un comportamiento mejor que la construcción de mos2Env original en lo que respecta a la estabilidad en un contexto de vector adenovírico. La estabilidad genética examinada hasta un npv 13 representa una propagación varios pases más allá de la utilizada en la preparación a escala industrial de los vectores.

#### Ejemplo 4: Expresión y antigenicidad *in vivo* de las secuencias de la envoltura del VIH en vectores de adenovirus

Se evaluaron la expresión y la antigenicidad de rAd26.C4D7 y rAd26.sC4 por separado o combinados con un vector de Ad26 recombinante que codifica mos1Env (SEQ ID NO: 5) (en lo sucesivo "rAd26.mos1 Env") en células A549 (línea celular humana) transducidas con un vector *in vitro* (datos no mostrados). Los análisis por citometría de flujo demostraron que todos los antígenos se expresaron en los cultivos celulares transducidos con  $2 \times 10^4$  partículas víricas (pv) de los antígenos de la envoltura solos como controles o con  $1 \times 10^4$  pv de 2 antígenos Env combinados mediante transducción de adenovirus. Todas las transducciones contuvieron adicionalmente dosis únicas ( $1 \times 10^4$  pv) de vectores de adenovirus que codifican mos1GagPol ("rAd26.mos1GagPol") y mos2GagPol ("rAd26.mos2GagPol") (Barouch et al, Nat Med 2010, 16:319-323), de manera que las combinaciones de los vectores evaluadas mostraron las mismas proporciones relativas de los diferentes vectores adenovíricos como se pretendía para su uso preclínico y clínico. Preferiblemente, los vectores que codifican las proteínas de la envoltura del VIH sintéticas de la invención se combinan con los vectores que codifican los antígenos mos1GagPol y mos2GagPol para su uso clínico.

La combinación de rAd26.mos1 Env y rAd26.C4D7 generaba una cobertura máxima de los epítomos evaluados según se determina mediante la unión de anticuerpos monoclonales. En particular, la exposición del epítomo PG16,

que fue debida a la transformación con Ad26.C4D7, resulta prometedor para su uso como una vacuna ya que PG16 representa un anticuerpo monoclonal ampliamente neutralizante que reconoce la región bucle V1/V2 de Env del VIH-1 (Walker et al., Science. 2009). Por lo tanto, la proteína de la envoltura del VIH sintética de la invención obtenida a partir de la secuencia de C4 aumenta la amplitud de la respuesta inmunitaria contra la proteína de la envoltura del VIH en comparación con la respuesta inmunitaria generada únicamente con mos1Env. En el estudio RV144, se ha demostrado que las respuestas de anticuerpos inducidas por la vacuna dirigidos hacia la región de la proteína de la envoltura se correlacionan con la protección frente a la infección por VIH-1 (Haynes et al., N Engl J Med. 2012) y, por lo tanto, la proteína de la envoltura del VIH sintética de la invención es una candidata prometedora para incluirla en las pautas de vacunación contra el VIH.

#### Ejemplo 5: Inmunogenicidad de los vectores que codifican proteínas de la envoltura del VIH sintéticas

Se examinaron en conejos las secuencias de proteínas de la envoltura del VIH sintéticas de la invención en un contexto de vector de Ad26 para determinar si estas construcciones representaban una alternativa inmunogénica a la construcción rAd26.mos2Env.

La inmunogenicidad de un vector de adenovirus que codifica mos1Env (rAd26.mos1Env; SEQ ID NO: 5) se examinó solo y combinado con vectores de adenovirus que codifican las proteínas de la envoltura del VIH sintéticas de la invención (rAd26.C4D7 y rAd26.sC4; que comprenden la SEQ ID NO: 8, en particular las SEQ ID NO: 18 y 19, respectivamente). En todos los casos, también se administraron vectores de adenovirus 26 que codifican los antígenos mos1GagPol y mos2GagPol (rAd26.mos1GagPol [SEQ ID NO: 28] y rAd26.mos2GagPol [SEQ ID NO: 29], respectivamente). Más específicamente, se comparó la inmunogenicidad de rAd26.mos1Env solo (vacuna trivalente: rAd26.mos1GagPol, rAd26.mos2GagPol y rAd26.mos1Env) con la inmunogenicidad de rAd26.mos1Env en combinación con uno de rAd26.C4D7 o rAd26.sC4 (vacuna tetravalente: la administración de rAd26.mos1GagPol, rAd26.mos2GagPol, rAd26.mos1Env y rAd26.C4D7; o la administración de rAd26.mos1GagPol, rAd26.mos2GagPol, rAd26.mos1Env y rAd26.sC4). Esta comparación de la vacuna trivalente, que carece de cualquier vector que codifica las proteínas de la envoltura del VIH sintéticas de la invención, con la vacuna tetravalente, que contiene vectores que codifican las proteínas de la envoltura del VIH sintéticas de la invención, permite determinar si las proteínas de la envoltura del VIH de la invención contribuyen a la amplitud de la protección.

La administración se realizó en pautas de vacunación, en las que se administraron estos vectores de Ad26 las semanas 0 y 6 como una doble sensibilización y una proteína gp140 del clado C (como una proteína Env gp140 trivalente que tiene la SEQ ID NO: 7 sin la secuencia del péptido señal de los restos 1-29, véase también el documento WO 2010/042942) las semanas 12 y 18 como un doble refuerzo (véase p. ej., Barouch et al., 2015, Science 349: 320-324). La Tabla 1 describe las pautas de vacunación utilizadas para el presente estudio. rAd26.Vacío se refiere a un vector de control que carece de cualquier gen que codifica una secuencia para una proteína antigénica del VIH. Cada grupo contuvo seis conejos.

**Tabla 1:** Pautas de vacunación examinadas en el estudio de la inmunogenicidad en conejos

Grupo	Primera y segunda inmunizaciones			Tercera y cuarta inmunizaciones			N=
	adenovectores	Dosis (pv)	Dosis total (pv)	Refuerzo con proteína	Dosis (ug)	Adyuvante	
1	rAd26.Mos1Env	2,5x10 <sup>10</sup>	5x10 <sup>10</sup>	GP140 (clado C)	10	AdjuPhos 250 µg	6
	rAd26.Mos1GagPol	1,25x10 <sup>10</sup>					
	rAd26.Mos2Gagpol	1,25x10 <sup>10</sup>					
2	rAd26.Mos1Env	1,25x10 <sup>10</sup>	5x10 <sup>10</sup>	GP140 (clado C)	10	AdjuPhos 250 µg	6
	rAd26.C4D7	1,25x10 <sup>10</sup>					
	rAd26.Mos1GagPol	1,25x10 <sup>10</sup>					
	rAd26.Mos2Gagpol	1,25x10 <sup>10</sup>					
3	rAd26.Mos1Env	1,25x10 <sup>10</sup>	5x10 <sup>10</sup>	GP140 (clado C)	10	AdjuPhos 250 µg	6
	rAd26.sC4	1,25x10 <sup>10</sup>					
	rAd26.Mos1GagPol	1,25x10 <sup>10</sup>					
	rAd26.Mos2Gagpol	1,25x10 <sup>10</sup>					
control	rAd26.Vacío	5x10 <sup>10</sup>	5x10 <sup>10</sup>	NA	0	AdjuPhos 250 µg	6

La comparación de la vacuna de Ad26 trivalente (que carece de los antígenos Env novedosos de la invención) con la vacuna de Ad26 tetravalente (que comprende los antígenos sC4 o C4D7 Env novedosos) permite examinar si los antígenos novedosos contribuyen a la amplitud de la protección. Se utilizó un ensayo de neutralización con células TZM-bl establecido [Montefiori DC. Methods Mol Biol 2009,485:395-405; Sarzotti-Kelsoe M et al., J Immunol Methods 2014,409:131-146] para medir la actividad neutralizante de los candidatos para la vacuna.

Los resultados se muestran en la Fig. 7 y se analizaron estadísticamente utilizando la vacuna trivalente (grupo 1 de la Tabla 1) como grupo de control y comparándola con cada una de las vacunas cuadrivalentes novedosas (grupos 2 y 3 de la Tabla 1).

5 En general, las adenoconstrucciones obtenidas a partir de C4 novedosas (es decir, que codifican las proteínas Env que comprenden la SEQ ID NO: 8, y que son una alternativa a mos2Env) fueron inmunogénicas después de dos inmunizaciones intramusculares homologas en conejos.

10 No se observó capacidad de neutralización de los sueros inmunológicos de conejo contra los pseudovirus del nivel 1B (no se muestran los datos), lo que no es inesperado ya que se sabe que tales virus resultan más difíciles de neutralizar.

15 La capacidad de neutralización de los pseudovirus de los sueros inmunológicos de conejo contra los virus de nivel 1A del clado B no se vio afectada por la adición de nuevos componentes (datos que no se muestran). Esto demuestra que el antígeno novedoso no interfirió de manera negativa con la inmunogenicidad del antígeno del clado B existente presente en la vacuna (aunque los nuevos componentes estuvieron dirigidos al clado C, no pudiendo excluirse a priori tal interferencia no deseable antes de haberla examinado).

20 La capacidad de neutralización de los pseudovirus de los sueros inmunológicos de conejo contra un virus del nivel 1A del clado C se vio significativamente potenciada en la inmunización con adeno que contiene C4D7 novedosa cuadrivalente (cuadrivalente, grupo 2) en comparación con la trivalente (que tiene únicamente mos1 Env) sola (grupo 1) (Fig 7 cuadro B). Además, la capacidad de neutralización de los pseudovirus de los sueros inmunológicos de conejo contra un virus del nivel 1A del clado C en la semana 8 se vio significativamente potenciada en la inmunización con adeno que contiene sC4 novedosa cuadrivalente (cuadrivalente, grupo 3) en comparación con la trivalente (que tiene únicamente mos1Env) sola (grupo 1) (Fig 7 cuadro B).

25 En conclusión, las construcciones C4D7 y sC4 codificadas en Ad26 fueron inmunogénicas y su adición aumentó la capacidad de unión y de neutralización de una vacuna que tiene mos1Env (principalmente clado B) como único componente Env codificado por Ad26, hacia las cepas del clado C (Fig 7B).

30 Ejemplo 6: Inmunogenicidad de las pautas de vacunación que incluyen los vectores que codifican proteínas de la envoltura del VIH sintéticas de la invención

35 Un estudio en conejos adicional evaluó la combinación del vector tetravalente Ad26.Mos4.HIV (que consiste en cuatro vectores adenovíricos: Ad26.Mos1GagPol [que codifica la SEQ ID NO: 28], Ad26.Mos2GagPol [que codifica la SEQ ID NO: 29], Ad26.Mos1Env [que codifica la SEQ ID NO: 5] y Ad26.Mos2SEnv [el nombre "C4D7" tal como se ha utilizado anteriormente también se refiere como "Mos2S"; este vector codifica la SEQ ID NO: 18 novedosa de acuerdo con la invención], en una mezcla 1:1:1:1 con una dosis total de  $5 \times 10^9$  pv), aplicada por vía intramuscular como inmunizaciones de doble sensibilización en las semanas 0 y 6, en combinación con refuerzos con la proteína Env del VIH-1 recombinante utilizando gp140 del clado C [que tiene la secuencia de los restos de aminoácidos 30-708 de la SEQ ID NO: 7], gp140 mosaico [que tiene la secuencia de los restos de aminoácidos 30-724 de la SEQ ID NO: 36] o la combinación de gp140 del clado C y gp140 mosaico, en las semanas 13 y 19. Estos refuerzos con proteína se aplicaron por vía intramuscular con una dosis total de 10 o 50 microgramos de proteína combinados con 250 mcg de adyuvante de tipo fosfato de aluminio formulado el día de la inmunización.

45 Los resultados indican que todas las pautas examinadas fueron inmunogénicas en todos los animales, induciendo títulos de anticuerpos elevados y una actividad de neutralización moderada contra los virus pseudotipados con Env del nivel 1. Si se utilizaba gp140 mosaico como antígeno vacuna, ya sea solo o combinado con gp140 del clado C, los títulos de ELISA específicos de gp140 mosaico y el reconocimiento de pseudovirus del Clado B aumentaron significativamente en la semana 15 en comparación con el grupo de referencia reforzado únicamente con gp140 del clado C. La magnitud del efecto global de la mejora fue moderada y mayor para el grupo reforzado con la combinación bivalente gp140 del clado C-gp140 mosaico en comparación con únicamente gp140 mosaico. En la semana 21 del estudio, estas diferencias desaparecieron y las respuestas inmunitarias medidas para las cohortes que recibieron los refuerzos bivalentes con gp140 del clado C-gp140 mosaico o los refuerzos monovalentes con gp140 del clado C fueron estadísticamente indistinguibles.

50 La pauta con la proteína bivalente mostró una inducción de títulos de ELISA y reconocimiento de pseudovirus del clado C comparables a la pauta de refuerzo solo con gp140 del clado C, lo que indica que la inclusión del inmunógeno relacionado con el clado B gp140 mosaico no tuvo un efecto negativo en la cobertura de los antígenos del clado C, a la vez que potenció significativamente la cobertura del clado B en la semana 15 del estudio.

60 Los datos confirman que el vector Ad26.Mos2SEnv que codifica un antígeno Env sintético de acuerdo con la invención se puede utilizar con éxito en las pautas de vacunación.

65

Referencias

1. Barouch et al, Nat Med 2010, 16: 319-323
2. WO 2010/059732
- 5 3. Schiernle et al., PNAS 94: 8640-8645, 1997
4. Abrahamyan et al., J Virol 79: 106-115, 2005
5. US20120076812
6. Barouch et al., Cell 155:1-9, 2013
7. Havenga, et al., 2006, J Gen Virol 87: 2135-43;
- 10 8. WO 03/104467
9. WO 2004/001032
10. WO 2007/104792
11. Abbink et al., (2007) Virol 81(9): 4654-63
12. Patente de los Estados Unidos n.º 7.270.811
- 15 13. Vogels et al., (2003) J Virol 77(15): 8263-71
14. WO 00/70071
15. WO2012/082918
16. Walker LM, Phogat SK, Chan-Hui PY, Wagner D, Phung P, Goss JL, et al. Broad and potent neutralizing antibodies from an African donor reveal a new HIV-1 vaccine target. Science 2009,326:285-289.
- 20 17. Haynes BF, Gilbert PB, McElrath MJ, Zolla-Pazner S, Tomaras GD, Alam SM, et al. Immune-correlates analysis of an HIV-1 vaccine efficacy trial. N Engl J Med 2012,366:1275-1286.
18. Barouch et al. (2015) Science 349: 320-324
19. Montefiori DC. Measuring HIV neutralization in a luciferase reporter gene assay. Methods Mol Biol 2009,485:395-405.
- 25 20. Sarzotti-Kelsoe M, Bailer RT, Turk E, Lin CL, Bilska M, Greene KM, et al. Optimization and validation of the TZM-bl assay for standardized assessments of neutralizing antibodies against HIV-1. J Immunol Methods 2014,409:131-146.
21. Edwards et al., J. Virology, 2002, 76:2683-2691.

30 LISTADO DE SECUENCIAS

- <110> Janssen Vaccines & Prevention B.V.
- 35 <120> Antígenos del virus de la inmunodeficiencia humana, vectores, composiciones y métodos de uso de los mismos
- <130> 0265 EP P01 PRI
- <160> 37
- 40 <170> PatentIn versión 3.5
- <210> 1
- <211> 500
- 45 <212> PRT
- <213> Secuencia artificial
- <220>
- <223> secuencia del antígeno mos1Gag mosaico
- 50 <400> 1

# ES 2 898 271 T3

Met	Gly	Ala	Arg	Ala	Ser	Val	Leu	Ser	Gly	Gly	Glu	Leu	Asp	Arg	Trp	1	5	10	15
Glu	Lys	Ile	Arg	Leu	Arg	Pro	Gly	Gly	Lys	Lys	Lys	Tyr	Arg	Leu	Lys	20	25	30	
His	Ile	Val	Trp	Ala	Ser	Arg	Glu	Leu	Glu	Arg	Phe	Ala	Val	Asn	Pro	35	40	45	
Gly	Leu	Leu	Glu	Thr	Ser	Glu	Gly	Cys	Arg	Gln	Ile	Leu	Gly	Gln	Leu	50	55	60	
Gln	Pro	Ser	Leu	Gln	Thr	Gly	Ser	Glu	Glu	Leu	Arg	Ser	Leu	Tyr	Asn	65	70	75	80
Thr	Val	Ala	Thr	Leu	Tyr	Cys	Val	His	Gln	Arg	Ile	Glu	Ile	Lys	Asp	85	90	95	
Thr	Lys	Glu	Ala	Leu	Glu	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Lys	100	105	110	
Lys	Lys	Ala	Gln	Gln	Ala	Ala	Ala	Asp	Thr	Gly	Asn	Ser	Ser	Gln	Val	115	120	125	
Ser	Gln	Asn	Tyr	Pro	Ile	Val	Gln	Asn	Ile	Gln	Gly	Gln	Met	Val	His	130	135	140	
Gln	Ala	Ile	Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Val	Glu	145	150	155	160



# ES 2 898 271 T3

Glu	Lys	Ala	Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Ser	Ala	Leu	Ser	165	170	175
Glu	Gly	Ala	Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	180	185	190
Gly	His	Gln	Ala	Ala	Met	Gln	Met	Leu	Lys	Glu	Thr	Ile	Asn	Glu	Glu	195	200	205
Ala	Ala	Glu	Trp	Asp	Arg	Val	His	Pro	Val	His	Ala	Gly	Pro	Ile	Ala	210	215	220
Pro	Gly	Gln	Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	225	230	235
Ser	Thr	Leu	Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro	Pro	Ile	245	250	255
Pro	Val	Gly	Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	260	265	270
Ile	Val	Arg	Met	Tyr	Ser	Pro	Val	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly	275	280	285
Pro	Lys	Glu	Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Tyr	Lys	Thr	Leu	290	295	300
Arg	Ala	Glu	Gln	Ala	Ser	Gln	Asp	Val	Lys	Asn	Trp	Met	Thr	Glu	Thr	305	310	315
Leu	Leu	Val	Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Lys	Ala	325	330	335
Leu	Gly	Pro	Ala	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly	340	345	350
Val	Gly	Gly	Pro	Gly	His	Lys	Ala	Arg	Val	Leu	Ala	Glu	Ala	Met	Ser	355	360	365
Gln	Val	Thr	Asn	Ser	Ala	Thr	Ile	Met	Met	Gln	Arg	Gly	Asn	Phe	Arg	370	375	380
Asn	Gln	Arg	Lys	Thr	Val	Lys	Cys	Phe	Asn	Cys	Gly	Lys	Glu	Gly	His	385	390	395
Ile	Ala	Lys	Asn	Cys	Arg	Ala	Pro	Arg	Lys	Lys	Gly	Cys	Trp	Lys	Cys	405	410	415

# ES 2 898 271 T3

Gly Lys Glu Gly His Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn  
420 425 430

Phe Leu Gly Lys Ile Trp Pro Ser Asn Lys Gly Arg Pro Gly Asn Phe  
435 440 445

Leu Gln Asn Arg Pro Glu Pro Thr Ala Pro Pro Glu Glu Ser Phe Arg  
450 455 460

Phe Gly Glu Glu Thr Thr Thr Pro Ser Gln Lys Gln Glu Pro Ile Asp  
465 470 475 480

Lys Glu Met Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp  
485 490 495

Pro Ser Ser Gln  
500

<210> 2  
<211> 491  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> secuencia del antígeno mos2Gag mosaico  
<400> 2

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Lys Trp  
1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys  
20 25 30

His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro  
35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Lys Gln Leu  
50 55 60

Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Phe Asn  
65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Ala Glu Ile Glu Val Arg Asp  
85 90 95

Thr Lys Glu Ala Leu Asp Lys Ile Glu Glu Glu Gln Asn Lys Ser Gln  
100 105 110

# ES 2 898 271 T3

Gln Lys Thr Gln Gln Ala Lys Glu Ala Asp Gly Lys Val Ser Gln Asn  
 115 120 125  
 Tyr Pro Ile Val Gln Asn Leu Gln Gly Gln Met Val His Gln Pro Ile  
 130 135 140  
 Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Ile Glu Glu Lys Ala  
 145 150 155 160  
 Phe Ser Pro Glu Val Ile Pro Met Phe Thr Ala Leu Ser Glu Gly Ala  
 165 170 175  
 Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly Gly His Gln  
 180 185 190  
 Ala Ala Met Gln Met Leu Lys Asp Thr Ile Asn Glu Glu Ala Ala Glu  
 195 200 205  
 Trp Asp Arg Leu His Pro Val His Ala Gly Pro Val Ala Pro Gly Gln  
 210 215 220  
 Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr Ser Asn Leu  
 225 230 235 240  
 Gln Glu Gln Ile Ala Trp Met Thr Ser Asn Pro Pro Ile Pro Val Gly  
 245 250 255  
 Asp Ile Tyr Lys Arg Trp Ile Ile Leu Gly Leu Asn Lys Ile Val Arg  
 260 265 270  
 Met Tyr Ser Pro Thr Ser Ile Leu Asp Ile Lys Gln Gly Pro Lys Glu  
 275 280 285  
 Pro Phe Arg Asp Tyr Val Asp Arg Phe Phe Lys Thr Leu Arg Ala Glu  
 290 295 300  
 Gln Ala Thr Gln Asp Val Lys Asn Trp Met Thr Asp Thr Leu Leu Val  
 305 310 315 320  
 Gln Asn Ala Asn Pro Asp Cys Lys Thr Ile Leu Arg Ala Leu Gly Pro  
 325 330 335  
 Gly Ala Thr Leu Glu Glu Met Met Thr Ala Cys Gln Gly Val Gly Gly  
 340 345 350  
 Pro Ser His Lys Ala Arg Val Leu Ala Glu Ala Met Ser Gln Thr Asn  
 355 360 365

# ES 2 898 271 T3

Ser Thr Ile Leu Met Gln Arg Ser Asn Phe Lys Gly Ser Lys Arg Ile  
370 375 380

Val Lys Cys Phe Asn Cys Gly Lys Glu Gly His Ile Ala Arg Asn Cys  
385 390 395 400

Arg Ala Pro Arg Lys Lys Gly Cys Trp Lys Cys Gly Lys Glu Gly His  
405 410 415

Gln Met Lys Asp Cys Thr Glu Arg Gln Ala Asn Phe Leu Gly Lys Ile  
420 425 430

Trp Pro Ser His Lys Gly Arg Pro Gly Asn Phe Leu Gln Ser Arg Pro  
435 440 445

Glu Pro Thr Ala Pro Pro Ala Glu Ser Phe Arg Phe Glu Glu Thr Thr  
450 455 460

Pro Ala Pro Lys Gln Glu Pro Lys Asp Arg Glu Pro Leu Thr Ser Leu  
465 470 475 480

Arg Ser Leu Phe Gly Ser Asp Pro Leu Ser Gln  
485 490

<210> 3

<211> 850

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia del antígeno mos1Pol mosaico

<400> 3

Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val Lys Leu Lys Pro  
1 5 10 15

Gly Met Asp Gly Pro Arg Val Lys Gln Trp Pro Leu Thr Glu Glu Lys  
20 25 30

Ile Lys Ala Leu Thr Ala Ile Cys Glu Glu Met Glu Lys Glu Gly Lys  
35 40 45

Ile Thr Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr Pro Val Phe Ala  
50 55 60

Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu Val Asp Phe Arg  
65 70 75 80

# ES 2 898 271 T3

Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val Gln Leu Gly Ile  
 85 90 95  
 Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val Thr Val Leu Ala  
 100 105 110  
 Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu Gly Phe Arg Lys  
 115 120 125  
 Tyr Thr Ala Phe Thr Ile Pro Ser Thr Asn Asn Glu Thr Pro Gly Ile  
 130 135 140  
 Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys Gly Ser Pro Ala  
 145 150 155 160  
 Ile Phe Gln Cys Ser Met Thr Arg Ile Leu Glu Pro Phe Arg Ala Lys  
 165 170 175  
 Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Ala Ala Leu Tyr Val Gly  
 180 185 190  
 Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys Ile Glu Glu Leu Arg  
 195 200 205  
 Glu His Leu Leu Lys Trp Gly Phe Thr Thr Pro Asp Lys Lys His Gln  
 210 215 220  
 Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys  
 225 230 235 240  
 Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys Asp Ser Trp Thr Val  
 245 250 255  
 Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile  
 260 265 270  
 Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu Leu Arg Gly Ala  
 275 280 285  
 Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu Glu Ala Glu Leu Glu  
 290 295 300  
 Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr  
 305 310 315 320  
 Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly His  
 325 330 335

# ES 2 898 271 T3

Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys  
 340 345 350  
 Thr Gly Lys Tyr Ala Lys Met Arg Thr Ala His Thr Asn Asp Val Lys  
 355 360 365  
 Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Met Glu Ser Ile Val Ile  
 370 375 380  
 Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Ile Gln Lys Glu Thr Trp  
 385 390 395 400  
 Glu Thr Trp Trp Thr Asp Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp  
 405 410 415  
 Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu  
 420 425 430  
 Lys Asp Pro Ile Ala Gly Val Glu Thr Phe Tyr Val Ala Gly Ala Ala  
 435 440 445  
 Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asp Arg Gly  
 450 455 460  
 Arg Gln Lys Ile Val Ser Leu Thr Glu Thr Thr Asn Gln Lys Thr Ala  
 465 470 475 480  
 Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly Ser Glu Val Asn  
 485 490 495  
 Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro  
 500 505 510  
 Asp Lys Ser Glu Ser Glu Leu Val Asn Gln Ile Ile Glu Gln Leu Ile  
 515 520 525  
 Lys Lys Glu Arg Val Tyr Leu Ser Trp Val Pro Ala His Lys Gly Ile  
 530 535 540  
 Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Ser Gly Ile Arg Lys  
 545 550 555 560  
 Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His Glu Lys  
 565 570 575  
 Tyr His Ser Asn Trp Arg Ala Met Ala Ser Asp Phe Asn Leu Pro Pro  
 580 585 590

# ES 2 898 271 T3

Val	Val	Ala	Lys	Glu	Ile	Val	Ala	Ser	Cys	Asp	Gln	Cys	Gln	Leu	Lys	595	600	605
Gly	Glu	Ala	Met	His	Gly	Gln	Val	Asp	Cys	Ser	Pro	Gly	Ile	Trp	Gln	610	615	620
Leu	Ala	Cys	Thr	His	Leu	Glu	Gly	Lys	Ile	Ile	Leu	Val	Ala	Val	His	625	630	635
Val	Ala	Ser	Gly	Tyr	Ile	Glu	Ala	Glu	Val	Ile	Pro	Ala	Glu	Thr	Gly	645	650	655
Gln	Glu	Thr	Ala	Tyr	Phe	Ile	Leu	Lys	Leu	Ala	Gly	Arg	Trp	Pro	Val	660	665	670
Lys	Val	Ile	His	Thr	Ala	Asn	Gly	Ser	Asn	Phe	Thr	Ser	Ala	Ala	Val	675	680	685
Lys	Ala	Ala	Cys	Trp	Trp	Ala	Gly	Ile	Gln	Gln	Glu	Phe	Gly	Ile	Pro	690	695	700
Tyr	Asn	Pro	Gln	Ser	Gln	Gly	Val	Val	Ala	Ser	Met	Asn	Lys	Glu	Leu	705	710	715
Lys	Lys	Ile	Ile	Gly	Gln	Val	Arg	Asp	Gln	Ala	Glu	His	Leu	Lys	Thr	725	730	735
Ala	Val	Gln	Met	Ala	Val	Phe	Ile	His	Asn	Phe	Lys	Arg	Lys	Gly	Gly	740	745	750
Ile	Gly	Gly	Tyr	Ser	Ala	Gly	Glu	Arg	Ile	Ile	Asp	Ile	Ile	Ala	Thr	755	760	765
Asp	Ile	Gln	Thr	Lys	Glu	Leu	Gln	Lys	Gln	Ile	Ile	Lys	Ile	Gln	Asn	770	775	780
Phe	Arg	Val	Tyr	Tyr	Arg	Asp	Ser	Arg	Asp	Pro	Ile	Trp	Lys	Gly	Pro	785	790	795
Ala	Lys	Leu	Leu	Trp	Lys	Gly	Glu	Gly	Ala	Val	Val	Ile	Gln	Asp	Asn	805	810	815
Ser	Asp	Ile	Lys	Val	Val	Pro	Arg	Arg	Lys	Val	Lys	Ile	Ile	Lys	Asp	820	825	830
Tyr	Gly	Lys	Gln	Met	Ala	Gly	Ala	Asp	Cys	Val	Ala	Gly	Arg	Gln	Asp	835	840	845
Glu	Asp															850		

<210> 4  
 <211> 850  
 <212> PRT  
 <213> Secuencia artificial

5

<220>  
 <223> secuencia del antígeno mos2Pol mosaico

<400> 4

10

```

Met  Ala  Pro  Ile  Ser  Pro  Ile  Glu  Thr  Val  Pro  Val  Lys  Leu  Lys  Pro
1      5      10      15

Gly  Met  Asp  Gly  Pro  Lys  Val  Lys  Gln  Trp  Pro  Leu  Thr  Glu  Glu  Lys
20      25      30

Ile  Lys  Ala  Leu  Val  Glu  Ile  Cys  Thr  Glu  Met  Glu  Lys  Glu  Gly  Lys
35      40      45

Ile  Ser  Lys  Ile  Gly  Pro  Glu  Asn  Pro  Tyr  Asn  Thr  Pro  Ile  Phe  Ala
50      55      60

Ile  Lys  Lys  Lys  Asp  Ser  Thr  Lys  Trp  Arg  Lys  Leu  Val  Asp  Phe  Arg
65      70      75      80

Glu  Leu  Asn  Lys  Arg  Thr  Gln  Asp  Phe  Trp  Glu  Val  Gln  Leu  Gly  Ile
85      90      95

Pro  His  Pro  Ala  Gly  Leu  Lys  Lys  Lys  Lys  Ser  Val  Thr  Val  Leu  Ala
100     105     110

Val  Gly  Asp  Ala  Tyr  Phe  Ser  Val  Pro  Leu  Asp  Glu  Asp  Phe  Arg  Lys
115     120     125

Tyr  Thr  Ala  Phe  Thr  Ile  Pro  Ser  Ile  Asn  Asn  Glu  Thr  Pro  Gly  Ile
130     135     140

Arg  Tyr  Gln  Tyr  Asn  Val  Leu  Pro  Gln  Gly  Trp  Lys  Gly  Ser  Pro  Ala
145     150     155     160

Ile  Phe  Gln  Ser  Ser  Met  Thr  Lys  Ile  Leu  Glu  Pro  Phe  Arg  Lys  Gln
165     170     175

Asn  Pro  Asp  Ile  Val  Ile  Tyr  Gln  Tyr  Met  Ala  Ala  Leu  Tyr  Val  Gly
180     185     190

```



Ser Asp Leu Glu Ile Gly Gln His Arg Thr Lys Ile Glu Glu Leu Arg  
 195 200 205

Gln His Leu Leu Arg Trp Gly Phe Thr Thr Pro Asp Lys Lys His Gln  
 210 215 220

Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu His Pro Asp Lys  
 225 230 235 240

Trp Thr Val Gln Pro Ile Val Leu Pro Glu Lys Asp Ser Trp Thr Val  
 245 250 255

Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp Ala Ser Gln Ile  
 260 265 270

Tyr Ala Gly Ile Lys Val Lys Gln Leu Cys Lys Leu Leu Arg Gly Thr  
 275 280 285

Lys Ala Leu Thr Glu Val Val Pro Leu Thr Glu Glu Ala Glu Leu Glu  
 290 295 300

Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val His Gly Val Tyr  
 305 310 315 320

Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln Lys Gln Gly Gln  
 325 330 335

Gly Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe Lys Asn Leu Lys  
 340 345 350

Thr Gly Lys Tyr Ala Arg Met Arg Gly Ala His Thr Asn Asp Val Lys  
 355 360 365

Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Thr Glu Ser Ile Val Ile  
 370 375 380

Trp Gly Lys Thr Pro Lys Phe Lys Leu Pro Ile Gln Lys Glu Thr Trp  
 385 390 395 400

Glu Ala Trp Trp Thr Glu Tyr Trp Gln Ala Thr Trp Ile Pro Glu Trp  
 405 410 415

Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp Tyr Gln Leu Glu  
 420 425 430

Lys Glu Pro Ile Val Gly Ala Glu Thr Phe Tyr Val Ala Gly Ala Ala

# ES 2 898 271 T3

435		440		445
Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val Thr Asp Arg Gly				
450		455		460
Arg Gln Lys Val Val Ser Leu Thr Asp Thr Thr Asn Gln Lys Thr Ala				
465		470		475 480
Leu Gln Ala Ile His Leu Ala Leu Gln Asp Ser Gly Leu Glu Val Asn				
		485		490 495
Ile Val Thr Ala Ser Gln Tyr Ala Leu Gly Ile Ile Gln Ala Gln Pro				
		500		505 510
Asp Lys Ser Glu Ser Glu Leu Val Ser Gln Ile Ile Glu Gln Leu Ile				
		515		520 525
Lys Lys Glu Lys Val Tyr Leu Ala Trp Val Pro Ala His Lys Gly Ile				
		530		535 540
Gly Gly Asn Glu Gln Val Asp Lys Leu Val Ser Arg Gly Ile Arg Lys				
		545		550 555 560
Val Leu Phe Leu Asp Gly Ile Asp Lys Ala Gln Glu Glu His Glu Lys				
		565		570 575
Tyr His Ser Asn Trp Arg Ala Met Ala Ser Glu Phe Asn Leu Pro Pro				
		580		585 590
Ile Val Ala Lys Glu Ile Val Ala Ser Cys Asp Lys Cys Gln Leu Lys				
		595		600 605
Gly Glu Ala Ile His Gly Gln Val Asp Cys Ser Pro Gly Ile Trp Gln				
		610		615 620
Leu Ala Cys Thr His Leu Glu Gly Lys Val Ile Leu Val Ala Val His				
		625		630 635 640
Val Ala Ser Gly Tyr Ile Glu Ala Glu Val Ile Pro Ala Glu Thr Gly				
		645		650 655
Gln Glu Thr Ala Tyr Phe Leu Leu Lys Leu Ala Gly Arg Trp Pro Val				
		660		665 670
Lys Thr Ile His Thr Ala Asn Gly Ser Asn Phe Thr Ser Ala Thr Val				
		675		680 685

# ES 2 898 271 T3

Lys Ala Ala Cys Trp Trp Ala Gly Ile Lys Gln Glu Phe Gly Ile Pro  
690 695 700

Tyr Asn Pro Gln Ser Gln Gly Val Val Ala Ser Ile Asn Lys Glu Leu  
705 710 715 720

Lys Lys Ile Ile Gly Gln Val Arg Asp Gln Ala Glu His Leu Lys Thr  
725 730 735

Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys Arg Lys Gly Gly  
740 745 750

Ile Gly Glu Tyr Ser Ala Gly Glu Arg Ile Val Asp Ile Ile Ala Ser  
755 760 765

Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln Ile Thr Lys Ile Gln Asn  
770 775 780

Phe Arg Val Tyr Tyr Arg Asp Ser Arg Asp Pro Leu Trp Lys Gly Pro  
785 790 795 800

Ala Lys Leu Leu Trp Lys Gly Glu Gly Ala Val Val Ile Gln Asp Asn  
805 810 815

Ser Asp Ile Lys Val Val Pro Arg Arg Lys Ala Lys Ile Ile Arg Asp  
820 825 830

Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln Asp  
835 840 845

Glu Asp  
850

<210> 5  
<211> 685  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> secuencia del antígeno mos1Env mosaico

<400> 5

Met Arg Val Thr Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp  
1 5 10 15

Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala Ala Gly Lys  
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr

# ES 2 898 271 T3

35					40					45					
Thr	Thr	Leu	Phe	Cys	Ala	Ser	Asp	Ala	Lys	Ala	Tyr	Asp	Thr	Glu	Val
50					55					60					
His	Asn	Val	Trp	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn	Pro
65					70					75					80
Gln	Glu	Val	Val	Leu	Glu	Asn	Val	Thr	Glu	Asn	Phe	Asn	Met	Trp	Lys
				85					90					95	
Asn	Asn	Met	Val	Glu	Gln	Met	His	Glu	Asp	Ile	Ile	Ser	Leu	Trp	Asp
			100					105					110		
Gln	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu
		115					120					125			
Asn	Cys	Thr	Asp	Asp	Val	Arg	Asn	Val	Thr	Asn	Asn	Ala	Thr	Asn	Thr
	130					135					140				
Asn	Ser	Ser	Trp	Gly	Glu	Pro	Met	Glu	Lys	Gly	Glu	Ile	Lys	Asn	Cys
145					150					155					160
Ser	Phe	Asn	Ile	Thr	Thr	Ser	Ile	Arg	Asn	Lys	Val	Gln	Lys	Gln	Tyr
				165					170					175	
Ala	Leu	Phe	Tyr	Lys	Leu	Asp	Val	Val	Pro	Ile	Asp	Asn	Asp	Ser	Asn
			180					185					190		
Asn	Thr	Asn	Tyr	Arg	Leu	Ile	Ser	Cys	Asn	Thr	Ser	Val	Ile	Thr	Gln
		195					200					205			
Ala	Cys	Pro	Lys	Val	Ser	Phe	Glu	Pro	Ile	Pro	Ile	His	Tyr	Cys	Ala
	210					215					220				
Pro	Ala	Gly	Phe	Ala	Ile	Leu	Lys	Cys	Asn	Asp	Lys	Lys	Phe	Asn	Gly
225					230					235					240
Thr	Gly	Pro	Cys	Thr	Asn	Val	Ser	Thr	Val	Gln	Cys	Thr	His	Gly	Ile
				245					250					255	
Arg	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu
			260					265					270		
Glu	Glu	Val	Val	Ile	Arg	Ser	Glu	Asn	Phe	Thr	Asn	Asn	Ala	Lys	Thr
		275					280					285			

## ES 2 898 271 T3

Ile	Met	Val	Gln	Leu	Asn	Val	Ser	Val	Glu	Ile	Asn	Cys	Thr	Arg	Pro
290					295					300					
Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	His	Ile	Gly	Pro	Gly	Arg	Ala	Phe
305					310					315					320
Tyr	Thr	Ala	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn
				325					330					335	
Ile	Ser	Arg	Ala	Asn	Trp	Asn	Asn	Thr	Leu	Arg	Gln	Ile	Val	Glu	Lys
			340					345					350		
Leu	Gly	Lys	Gln	Phe	Gly	Asn	Asn	Lys	Thr	Ile	Val	Phe	Asn	His	Ser
		355					360					365			
Ser	Gly	Gly	Asp	Pro	Glu	Ile	Val	Met	His	Ser	Phe	Asn	Cys	Gly	Gly
	370					375					380				
Glu	Phe	Phe	Tyr	Cys	Asn	Ser	Thr	Lys	Leu	Phe	Asn	Ser	Thr	Trp	Thr
385					390				395						400
Trp	Asn	Asn	Ser	Thr	Trp	Asn	Asn	Thr	Lys	Arg	Ser	Asn	Asp	Thr	Glu
				405					410					415	
Glu	His	Ile	Thr	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp
			420					425					430		
Gln	Glu	Val	Gly	Lys	Ala	Met	Tyr	Ala	Pro	Pro	Ile	Arg	Gly	Gln	Ile
		435					440					445			
Arg	Cys	Ser	Ser	Asn	Ile	Thr	Gly	Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly
	450					455					460				
Asn	Asp	Thr	Ser	Gly	Thr	Glu	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met
465					470					475					480
Arg	Asp	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Lys	Ile
				485					490					495	
Glu	Pro	Leu	Gly	Val	Ala	Pro	Thr	Lys	Ala	Lys	Arg	Arg	Val	Val	Gln
			500					505					510		
Ser	Glu	Lys	Ser	Ala	Val	Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu
		515					520					525			
Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Met	Thr	Leu	Thr	Val
	530					535					540				

# ES 2 898 271 T3

Gln Ala Arg Leu Leu Leu Ser Gly Ile Val Gln Gln Gln Asn Asn Leu  
545 550 555 560

Leu Arg Ala Ile Glu Ala Gln Gln His Leu Leu Gln Leu Thr Val Trp  
565 570 575

Gly Ile Lys Gln Leu Gln Ala Arg Val Leu Ala Val Glu Arg Tyr Leu  
580 585 590

Lys Asp Gln Gln Leu Leu Gly Ile Trp Gly Cys Ser Gly Lys Leu Ile  
595 600 605

Cys Thr Thr Thr Val Pro Trp Asn Ala Ser Trp Ser Asn Lys Ser Leu  
610 615 620

Asp Lys Ile Trp Asn Asn Met Thr Trp Met Glu Trp Glu Arg Glu Ile  
625 630 635 640

Asn Asn Tyr Thr Ser Leu Ile Tyr Thr Leu Ile Glu Glu Ser Gln Asn  
645 650 655

Gln Gln Glu Lys Asn Glu Gln Glu Leu Leu Glu Leu Asp Lys Trp Ala  
660 665 670

Ser Leu Trp Asn Trp Phe Asp Ile Ser Asn Trp Leu Trp  
675 680 685

<210> 6

<211> 684

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia del antígeno mos2Env mosaico

<400> 6

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp  
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val Met Gly Asn  
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys  
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val  
50 55 60

# ES 2 898 271 T3

His	Asn	Val	Trp	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn	Pro	65	70	75	80
Gln	Glu	Met	Val	Leu	Glu	Asn	Val	Thr	Glu	Asn	Phe	Asn	Met	Trp	Lys	85	90	95	
Asn	Asp	Met	Val	Asp	Gln	Met	His	Glu	Asp	Ile	Ile	Arg	Leu	Trp	Asp	100	105	110	
Gln	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu	115	120	125	
Glu	Cys	Arg	Asn	Val	Arg	Asn	Val	Ser	Ser	Asn	Gly	Thr	Tyr	Asn	Ile	130	135	140	
Ile	His	Asn	Glu	Thr	Tyr	Lys	Glu	Met	Lys	Asn	Cys	Ser	Phe	Asn	Ala	145	150	155	160
Thr	Thr	Val	Val	Glu	Asp	Arg	Lys	Gln	Lys	Val	His	Ala	Leu	Phe	Tyr	165	170	175	
Arg	Leu	Asp	Ile	Val	Pro	Leu	Asp	Glu	Asn	Asn	Ser	Ser	Glu	Lys	Ser	180	185	190	
Ser	Glu	Asn	Ser	Ser	Glu	Tyr	Tyr	Arg	Leu	Ile	Asn	Cys	Asn	Thr	Ser	195	200	205	
Ala	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe	Asp	Pro	Ile	Pro	Ile	210	215	220	
His	Tyr	Cys	Ala	Pro	Ala	Gly	Tyr	Ala	Ile	Leu	Lys	Cys	Asn	Asn	Lys	225	230	235	240
Thr	Phe	Asn	Gly	Thr	Gly	Pro	Cys	Asn	Asn	Val	Ser	Thr	Val	Gln	Cys	245	250	255	
Thr	His	Gly	Ile	Lys	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	260	265	270	
Ser	Leu	Ala	Glu	Glu	Glu	Ile	Ile	Ile	Arg	Ser	Glu	Asn	Leu	Thr	Asn	275	280	285	
Asn	Ala	Lys	Thr	Ile	Ile	Val	His	Leu	Asn	Glu	Thr	Val	Asn	Ile	Thr	290	295	300	
Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	Arg	Ile	Gly	Pro	305	310	315	320

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln  
 325 330 335

Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly  
 340 345 350

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Asn Phe  
 355 360 365

Thr Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn  
 370 375 380

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Gly  
 385 390 395 400

Thr Tyr Met Pro Asn Gly Thr Asn Ser Asn Ser Ser Ser Asn Ile Thr  
 405 410 415

Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly  
 420 425 430

Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser  
 435 440 445

Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly  
 450 455 460

Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg  
 465 470 475 480

Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys  
 485 490 495

Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Ser  
 500 505 510

Glu Lys Ser Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly  
 515 520 525

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln  
 530 535 540

Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu  
 545 550 555 560

Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly  
 565 570 575



Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln  
580 585 590

Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys  
595 600 605

Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr  
610 615 620

Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly  
625 630 635 640

Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln  
645 650 655

Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn  
660 665 670

Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp  
675 680

<210> 7

<211> 708

<212> PRT

<213> Secuencia artificial

<220>

<223> trímero de gp140 del clado C estabilizado: C97ZA012-gp140-foldon con mutaciones de escisión

<400> 7

Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp  
1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Ala Glu Asn  
20 25 30

Leu Trp Val Gly Asn Met Trp Val Thr Val Tyr Tyr Gly Val Pro Val  
35 40 45

Trp Thr Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Thr Lys Ala  
50 55 60

Tyr Asp Arg Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro  
65 70 75 80

Thr Asp Pro Asn Pro Gln Glu Ile Val Leu Glu Asn Val Thr Glu Asn  
85 90 95

# ES 2 898 271 T3

Phe Asn Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile  
 100 105 110  
 Ile Ser Leu Trp Asp Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro  
 115 120 125  
 Leu Cys Val Thr Leu His Cys Thr Asn Ala Thr Phe Lys Asn Asn Val  
 130 135 140  
 Thr Asn Asp Met Asn Lys Glu Ile Arg Asn Cys Ser Phe Asn Thr Thr  
 145 150 155 160  
 Thr Glu Ile Arg Asp Lys Lys Gln Gln Gly Tyr Ala Leu Phe Tyr Arg  
 165 170 175  
 Pro Asp Ile Val Leu Leu Lys Glu Asn Arg Asn Asn Ser Asn Asn Ser  
 180 185 190  
 Glu Tyr Ile Leu Ile Asn Cys Asn Ala Ser Thr Ile Thr Gln Ala Cys  
 195 200 205  
 Pro Lys Val Asn Phe Asp Pro Ile Pro Ile His Tyr Cys Ala Pro Ala  
 210 215 220  
 Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys Thr Phe Ser Gly Lys Gly  
 225 230 235 240  
 Pro Cys Asn Asn Val Ser Thr Val Gln Cys Thr His Gly Ile Lys Pro  
 245 250 255  
 Val Val Ser Thr Gln Leu Leu Leu Asn Gly Ser Leu Ala Glu Lys Glu  
 260 265 270  
 Ile Ile Ile Arg Ser Glu Asn Leu Thr Asp Asn Val Lys Thr Ile Ile  
 275 280 285  
 Val His Leu Asn Lys Ser Val Glu Ile Val Cys Thr Arg Pro Asn Asn  
 290 295 300  
 Asn Thr Arg Lys Ser Met Arg Ile Gly Pro Gly Gln Thr Phe Tyr Ala  
 305 310 315 320  
 Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln Ala Tyr Cys Asn Ile Ser  
 325 330 335  
 Gly Ser Lys Trp Asn Glu Thr Leu Lys Arg Val Lys Glu Lys Leu Gln

# ES 2 898 271 T3

340										345					350				
Glu	Asn	Tyr	Asn	Asn	Asn	Lys	Thr	Ile	Lys	Phe	Ala	Pro	Ser	Ser	Gly				
		355					360					365							
Gly	Asp	Leu	Glu	Ile	Thr	Thr	His	Ser	Phe	Asn	Cys	Arg	Gly	Glu	Phe				
	370					375					380								
Phe	Tyr	Cys	Asn	Thr	Thr	Arg	Leu	Phe	Asn	Asn	Asn	Ala	Thr	Glu	Asp				
385					390					395					400				
Glu	Thr	Ile	Thr	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp				
				405					410					415					
Gln	Gly	Val	Gly	Arg	Ala	Met	Tyr	Ala	Pro	Pro	Ile	Ala	Gly	Asn	Ile				
			420					425					430						
Thr	Cys	Lys	Ser	Asn	Ile	Thr	Gly	Leu	Leu	Leu	Val	Arg	Asp	Gly	Gly				
		435					440					445							
Glu	Asp	Asn	Lys	Thr	Glu	Glu	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asn	Met				
	450					455					460								
Lys	Asp	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Ile	Glu	Leu				
465					470					475					480				
Lys	Pro	Leu	Gly	Ile	Ala	Pro	Thr	Gly	Ala	Lys	Glu	Arg	Val	Val	Glu				
			485						490					495					
Arg	Glu	Glu	Arg	Ala	Val	Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu				
			500					505					510						
Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Leu	Thr	Leu	Thr	Val				
	515						520					525							
Gln	Ala	Arg	Gln	Leu	Leu	Ser	Ser	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu				
	530					535					540								
Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Met	Leu	Gln	Leu	Thr	Val	Trp				
545					550					555					560				
Gly	Ile	Lys	Gln	Leu	Gln	Thr	Arg	Val	Leu	Ala	Ile	Glu	Arg	Tyr	Leu				
			565						570					575					
Lys	Asp	Gln	Gln	Leu	Leu	Gly	Ile	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile				
		580						585					590						

# ES 2 898 271 T3

Cys Thr Thr Asn Val Pro Trp Asn Ser Ser Trp Ser Asn Lys Ser Gln  
595 600 605

Thr Asp Ile Trp Asn Asn Met Thr Trp Met Glu Trp Asp Arg Glu Ile  
610 615 620

Ser Asn Tyr Thr Asp Thr Ile Tyr Arg Leu Leu Glu Asp Ser Gln Thr  
625 630 635 640

Gln Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys  
645 650 655

Asn Leu Trp Ser Trp Phe Asp Ile Ser Asn Trp Leu Trp Tyr Ile Lys  
660 665 670

Ser Arg Ile Glu Gly Arg Gly Ser Gly Gly Tyr Ile Pro Glu Ala Pro  
675 680 685

Arg Asp Gly Gln Ala Tyr Val Arg Lys Asp Gly Glu Trp Val Leu Leu  
690 695 700

Ser Thr Phe Leu  
705

<210> 8

<211> 625

<212> PRT

<213> Secuencia artificial

<220>

<223> fragmento de C4: gp120-gp41 truncada sin péptido señal y dominio transmembrana

<400> 8

Met Gly Asn Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys  
1 5 10 15

Asp Ala Lys Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu  
20 25 30

Lys Glu Val His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp  
35 40 45

Pro Asn Pro Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn  
50 55 60

Met Trp Lys Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser  
65 70 75 80

# ES 2 898 271 T3

Leu	Trp	Asp	Ala	Ser	Leu	Glu	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	85	90	95
Val	Thr	Leu	Asn	Cys	Arg	Asn	Val	Arg	Asn	Val	Ser	Ser	Asn	Gly	Thr	100	105	110
Tyr	Asn	Ile	Ile	His	Asn	Glu	Thr	Tyr	Lys	Glu	Met	Lys	Asn	Cys	Ser	115	120	125
Phe	Asn	Ala	Thr	Thr	Val	Val	Glu	Asp	Arg	Lys	Gln	Lys	Val	His	Ala	130	135	140
Leu	Phe	Tyr	Arg	Leu	Asp	Ile	Val	Pro	Leu	Asp	Glu	Asn	Asn	Ser	Ser	145	150	155
Glu	Lys	Ser	Ser	Glu	Asn	Ser	Ser	Glu	Tyr	Tyr	Arg	Leu	Ile	Asn	Cys	165	170	175
Asn	Thr	Ser	Ala	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe	Asp	Pro	180	185	190
Ile	Pro	Ile	His	Tyr	Cys	Ala	Pro	Ala	Gly	Tyr	Ala	Ile	Leu	Lys	Cys	195	200	205
Asn	Asn	Lys	Thr	Phe	Asn	Gly	Thr	Gly	Pro	Cys	Asn	Asn	Val	Ser	Thr	210	215	220
Val	Gln	Cys	Thr	His	Gly	Ile	Lys	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	225	230	235
Leu	Asn	Gly	Ser	Leu	Ala	Glu	Glu	Glu	Ile	Ile	Ile	Arg	Ser	Glu	Asn	245	250	255
Leu	Thr	Asn	Asn	Ala	Lys	Thr	Ile	Ile	Val	His	Leu	Asn	Glu	Thr	Val	260	265	270
Asn	Ile	Thr	Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	Arg	275	280	285
Ile	Gly	Pro	Gly	Gln	Thr	Phe	Tyr	Ala	Thr	Gly	Asp	Ile	Ile	Gly	Asp	290	295	300
Ile	Arg	Gln	Ala	His	Cys	Asn	Leu	Ser	Arg	Asp	Gly	Trp	Asn	Lys	Thr	305	310	315
Leu	Gln	Gly	Val	Lys	Lys	Lys	Leu	Ala	Glu	His	Phe	Pro	Asn	Lys	Thr	325	330	335

# ES 2 898 271 T3

Ile Lys Phe Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His  
 340 345 350  
 Thr Phe Asn Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu  
 355 360 365  
 Phe Asn Glu Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro  
 370 375 380  
 Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala  
 385 390 395 400  
 Ile Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile  
 405 410 415  
 Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro  
 420 425 430  
 Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn  
 435 440 445  
 Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu  
 450 455 460  
 Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Lys  
 465 470 475 480  
 Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala  
 485 490 495  
 Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg  
 500 505 510  
 Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala  
 515 520 525  
 Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys  
 530 535 540  
 Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln  
 545 550 555 560  
 Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr  
 565 570 575  
 Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile  
 580 585 590

# ES 2 898 271 T3

Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr  
595 600 605

Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu  
610 615 620

Lys  
625

<210> 9  
<211> 29  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> Secuencia señal

<400> 9

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp  
1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val  
20 25

<210> 10  
<211> 29  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> Secuencia señal

<400> 10

Met Arg Val Thr Gly Ile Arg Lys Asn Tyr Gln His Leu Trp Arg Trp  
1 5 10 15

Gly Thr Met Leu Leu Gly Ile Leu Met Ile Cys Ser Ala  
20 25

<210> 11  
<211> 29  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> Secuencia señal

<400> 11

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp  
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val  
20 25

<210> 12  
<211> 29  
<212> PRT  
<213> Secuencia artificial

<220>  
 <223> Secuencia señal

5 <400> 12

Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp  
 1 5 10 15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala  
 20 25

10 <210> 13  
 <211> 22  
 <212> PRT  
 <213> Secuencia artificial

15 <220>  
 <223> dominio transmembrana

<400> 13

Ile Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe  
 1 5 10 15

Ala Val Leu Ser Ile Val  
 20

20 <210> 14  
 <211> 7  
 <212> PRT  
 <213> Secuencia artificial

25 <220>  
 <223> región citoplasmática truncada

<400> 14

30 Asn Arg Val Arg Gln Gly Tyr  
 1 5

35 <210> 15  
 <211> 32  
 <212> PRT  
 <213> Secuencia artificial

40 <220>  
 <223> dominio de trimerización GCN4

<400> 15

Met Lys Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile Tyr  
 1 5 10 15

His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu Val  
 20 25 30

45 <210> 16  
 <211> 30  
 <212> PRT  
 <213> Secuencia artificial



<220>

<223> dominio de trimerización foldon

<400> 16

5

Gly Ser Gly Gly Tyr Ile Pro Glu Ala Pro Arg Asp Gly Gln Ala Tyr  
1 5 10 15

Val Arg Lys Asp Gly Glu Trp Val Leu Leu Ser Thr Phe Leu  
20 25 30

<210> 17

<211> 862

10

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia de C4

15

<400> 17

Met Arg Val Arg Gly Met Leu Arg Asn Trp Gln Gln Trp Trp Ile Trp  
1 5 10 15

Ser Ser Leu Gly Phe Trp Met Leu Met Ile Tyr Ser Val Met Gly Asn  
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Asp Ala Lys  
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val  
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro  
65 70 75 80

Gln Glu Ile Val Leu Gly Asn Val Thr Glu Asn Phe Asn Met Trp Lys  
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp

## ES 2 898 271 T3

100								105				110					
Ala	Ser	Leu	Glu	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu		
		115					120					125					
Asn	Cys	Arg	Asn	Val	Arg	Asn	Val	Ser	Ser	Asn	Gly	Thr	Tyr	Asn	Ile		
	130					135					140						
Ile	His	Asn	Glu	Thr	Tyr	Lys	Glu	Met	Lys	Asn	Cys	Ser	Phe	Asn	Ala		
145					150					155					160		
Thr	Thr	Val	Val	Glu	Asp	Arg	Lys	Gln	Lys	Val	His	Ala	Leu	Phe	Tyr		
				165					170					175			
Arg	Leu	Asp	Ile	Val	Pro	Leu	Asp	Glu	Asn	Asn	Ser	Ser	Glu	Lys	Ser		
			180					185					190				
Ser	Glu	Asn	Ser	Ser	Glu	Tyr	Tyr	Arg	Leu	Ile	Asn	Cys	Asn	Thr	Ser		
		195					200					205					
Ala	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe	Asp	Pro	Ile	Pro	Ile		
	210					215					220						
His	Tyr	Cys	Ala	Pro	Ala	Gly	Tyr	Ala	Ile	Leu	Lys	Cys	Asn	Asn	Lys		
225					230					235					240		
Thr	Phe	Asn	Gly	Thr	Gly	Pro	Cys	Asn	Asn	Val	Ser	Thr	Val	Gln	Cys		
				245					250					255			
Thr	His	Gly	Ile	Lys	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly		
			260					265					270				
Ser	Leu	Ala	Glu	Glu	Glu	Ile	Ile	Ile	Arg	Ser	Glu	Asn	Leu	Thr	Asn		
		275					280					285					
Asn	Ala	Lys	Thr	Ile	Ile	Val	His	Leu	Asn	Glu	Thr	Val	Asn	Ile	Thr		
	290					295					300						
Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	Arg	Ile	Gly	Pro		
305					310					315					320		
Gly	Gln	Thr	Phe	Tyr	Ala	Thr	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln		
				325					330					335			
Ala	His	Cys	Asn	Leu	Ser	Arg	Asp	Gly	Trp	Asn	Lys	Thr	Leu	Gln	Gly		
			340					345					350				

Val	Lys	Lys	Lys	Leu	Ala	Glu	His	Phe	Pro	Asn	Lys	Thr	Ile	Lys	Phe	355	360	365	
Ala	Pro	His	Ser	Gly	Gly	Asp	Leu	Glu	Ile	Thr	Thr	His	Thr	Phe	Asn	370	375	380	
Cys	Arg	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser	Asn	Leu	Phe	Asn	Glu	385	390	395	400
Ser	Asn	Ile	Glu	Arg	Asn	Asp	Ser	Ile	Ile	Thr	Leu	Pro	Cys	Arg	Ile	405	410	415	
Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	Glu	Val	Gly	Arg	Ala	Ile	Tyr	Ala	420	425	430	
Pro	Pro	Ile	Ala	Gly	Asn	Ile	Thr	Cys	Arg	Ser	Asn	Ile	Thr	Gly	Leu	435	440	445	
Leu	Leu	Thr	Arg	Asp	Gly	Gly	Ser	Asn	Asn	Gly	Val	Pro	Asn	Asp	Thr	450	455	460	
Glu	Thr	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asn	Asn	Trp	Arg	Ser	465	470	475	480
Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Glu	Val	Lys	Pro	Leu	Gly	Val	Ala	485	490	495	
Pro	Thr	Glu	Ala	Lys	Arg	Arg	Val	Val	Glu	Arg	Glu	Lys	Arg	Ala	Val	500	505	510	
Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Ile	Leu	Gly	Ala	Ala	Gly	Ser	Thr	515	520	525	
Met	Gly	Ala	Ala	Ser	Ile	Thr	Leu	Thr	Val	Gln	Ala	Arg	Gln	Leu	Leu	530	535	540	
Ser	Gly	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	545	550	555	560
Gln	Gln	His	Met	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	565	570	575	
Thr	Arg	Val	Leu	Ala	Ile	Glu	Arg	Tyr	Leu	Gln	Asp	Gln	Gln	Leu	Leu	580	585	590	
Gly	Leu	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	Ala	Val	Pro	595	600	605	

ES 2 898 271 T3

Trp	Asn	Thr	Ser	Trp	Ser	Asn	Lys	Ser	Gln	Thr	Asp	Ile	Trp	Asp	Asn
610						615								620	
Met	Thr	Trp	Met	Gln	Trp	Asp	Lys	Glu	Ile	Gly	Asn	Tyr	Thr	Gly	Glu
625					630					635					640
Ile	Tyr	Arg	Leu	Leu	Glu	Glu	Ser	Gln	Asn	Gln	Gln	Glu	Lys	Asn	Glu
				645					650					655	
Lys	Asp	Leu	Leu	Ala	Leu	Asp	Ser	Trp	Asn	Asn	Leu	Trp	Asn	Trp	Phe
			660					665					670		
Ser	Ile	Ser	Lys	Trp	Leu	Trp	Tyr	Ile	Lys	Ile	Phe	Ile	Met	Ile	Val
		675					680					685			
Gly	Gly	Leu	Ile	Gly	Leu	Arg	Ile	Ile	Phe	Ala	Val	Leu	Ser	Ile	Val
	690					695					700				
Asn	Arg	Val	Arg	Gln	Gly	Tyr	Ser	Pro	Leu	Ser	Leu	Gln	Thr	Leu	Thr
705					710					715					720
Gln	Asn	Pro	Gly	Gly	Leu	Asp	Arg	Leu	Gly	Arg	Ile	Glu	Glu	Glu	Gly
				725					730					735	
Gly	Glu	Gln	Asp	Lys	Asp	Arg	Ser	Ile	Arg	Leu	Val	Asn	Gly	Phe	Phe
			740					745					750		
Ala	Leu	Phe	Trp	Asp	Asp	Leu	Arg	Ser	Leu	Cys	Leu	Phe	Ser	Tyr	His
		755					760					765			
Arg	Leu	Arg	Asp	Phe	Ile	Leu	Ile	Val	Ala	Arg	Ala	Val	Glu	Leu	Leu
	770					775					780				
Gly	Arg	Ser	Ser	Leu	Arg	Gly	Leu	Gln	Arg	Gly	Trp	Glu	Ile	Leu	Lys
785					790					795					800
Tyr	Leu	Gly	Ser	Leu	Leu	Gln	Tyr	Trp	Gly	Leu	Glu	Leu	Lys	Lys	Ser
				805					810					815	
Ala	Ile	Asn	Leu	Leu	Asp	Thr	Ile	Ala	Ile	Ala	Val	Ala	Glu	Gly	Thr
			820					825					830		
Asp	Arg	Ile	Ile	Glu	Leu	Ile	Gln	Arg	Ile	Cys	Arg	Ala	Ile	Cys	Asn
		835					840					845			
Ile	Pro	Arg	Arg	Ile	Arg	Gln	Gly	Phe	Glu	Ala	Ala	Leu	Gln		
	850					855									

<210> 18  
<211> 711  
<212> PRT

&lt;213&gt; Secuencia artificial

&lt;220&gt;

&lt;223&gt; secuencia de C4D7

5

&lt;400&gt; 18

Met	Arg	Val	Arg	Gly	Met	Leu	Arg	Asn	Trp	Gln	Gln	Trp	Trp	Ile	Trp	1	5	10	15
Ser	Ser	Leu	Gly	Phe	Trp	Met	Leu	Met	Ile	Tyr	Ser	Val	Met	Gly	Asn	20	25	30	
Leu	Trp	Val	Thr	Val	Tyr	Tyr	Gly	Val	Pro	Val	Trp	Lys	Asp	Ala	Lys	35	40	45	
Thr	Thr	Leu	Phe	Cys	Ala	Ser	Asp	Ala	Lys	Ala	Tyr	Glu	Lys	Glu	Val	50	55	60	
His	Asn	Val	Trp	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn	Pro	65	70	75	80
Gln	Glu	Ile	Val	Leu	Gly	Asn	Val	Thr	Glu	Asn	Phe	Asn	Met	Trp	Lys	85	90	95	
Asn	Asp	Met	Val	Asp	Gln	Met	His	Glu	Asp	Ile	Ile	Ser	Leu	Trp	Asp	100	105	110	
Ala	Ser	Leu	Glu	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu	115	120	125	
Asn	Cys	Arg	Asn	Val	Arg	Asn	Val	Ser	Ser	Asn	Gly	Thr	Tyr	Asn	Ile	130	135	140	
Ile	His	Asn	Glu	Thr	Tyr	Lys	Glu	Met	Lys	Asn	Cys	Ser	Phe	Asn	Ala	145	150	155	160
Thr	Thr	Val	Val	Glu	Asp	Arg	Lys	Gln	Lys	Val	His	Ala	Leu	Phe	Tyr	165	170	175	
Arg	Leu	Asp	Ile	Val	Pro	Leu	Asp	Glu	Asn	Asn	Ser	Ser	Glu	Lys	Ser	180	185	190	
Ser	Glu	Asn	Ser	Ser	Glu	Tyr	Tyr	Arg	Leu	Ile	Asn	Cys	Asn	Thr	Ser	195	200	205	

# ES 2 898 271 T3

Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile  
 210 215 220  
 His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys  
 225 230 235 240  
 Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys  
 245 250 255  
 Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly  
 260 265 270  
 Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn  
 275 280 285  
 Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr  
 290 295 300  
 Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro  
 305 310 315 320  
 Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln  
 325 330 335  
 Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly  
 340 345 350  
 Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Lys Phe  
 355 360 365  
 Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His Thr Phe Asn  
 370 375 380  
 Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu Phe Asn Glu  
 385 390 395 400  
 Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro Cys Arg Ile  
 405 410 415  
 Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Ile Tyr Ala  
 420 425 430  
 Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu  
 435 440 445  
 Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro Asn Asp Thr  
 450 455 460

Glu	Thr	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	Asn	Asn	Trp	Arg	Ser	465	470	475	480
Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Glu	Val	Lys	Pro	Leu	Gly	Val	Ala	485	490	495	
Pro	Thr	Glu	Ala	Lys	Arg	Arg	Val	Val	Glu	Arg	Glu	Lys	Arg	Ala	Val	500	505	510	
Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Ile	Leu	Gly	Ala	Ala	Gly	Ser	Thr	515	520	525	
Met	Gly	Ala	Ala	Ser	Ile	Thr	Leu	Thr	Val	Gln	Ala	Arg	Gln	Leu	Leu	530	535	540	
Ser	Gly	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu	Arg	Ala	Ile	Glu	Ala	545	550	555	560
Gln	Gln	His	Met	Leu	Gln	Leu	Thr	Val	Trp	Gly	Ile	Lys	Gln	Leu	Gln	565	570	575	
Thr	Arg	Val	Leu	Ala	Ile	Glu	Arg	Tyr	Leu	Gln	Asp	Gln	Gln	Leu	Leu	580	585	590	
Gly	Leu	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys	Thr	Thr	Ala	Val	Pro	595	600	605	
Trp	Asn	Thr	Ser	Trp	Ser	Asn	Lys	Ser	Gln	Thr	Asp	Ile	Trp	Asp	Asn	610	615	620	
Met	Thr	Trp	Met	Gln	Trp	Asp	Lys	Glu	Ile	Gly	Asn	Tyr	Thr	Gly	Glu	625	630	635	640
Ile	Tyr	Arg	Leu	Leu	Glu	Glu	Ser	Gln	Asn	Gln	Gln	Glu	Lys	Asn	Glu	645	650	655	
Lys	Asp	Leu	Leu	Ala	Leu	Asp	Ser	Trp	Asn	Asn	Leu	Trp	Asn	Trp	Phe	660	665	670	
Ser	Ile	Ser	Lys	Trp	Leu	Trp	Tyr	Ile	Lys	Ile	Phe	Ile	Met	Ile	Val	675	680	685	
Gly	Gly	Leu	Ile	Gly	Leu	Arg	Ile	Ile	Phe	Ala	Val	Leu	Ser	Ile	Val	690	695	700	
Asn	Arg	Val	Arg	Gln	Gly	Tyr										705	710		

<210> 19  
<211> 704

# ES 2 898 271 T3

<212> PRT

<213> Secuencia artificial

<220>

5 <223> secuencia de sC4

<400> 19

Met	Arg	Val	Arg	Gly	Met	Leu	Arg	Asn	Trp	Gln	Gln	Trp	Trp	Ile	Trp	1	5	10	15
Ser	Ser	Leu	Gly	Phe	Trp	Met	Leu	Met	Ile	Tyr	Ser	Val	Met	Gly	Asn	20	25	30	
Leu	Trp	Val	Thr	Val	Tyr	Tyr	Gly	Val	Pro	Val	Trp	Lys	Asp	Ala	Lys	35	40	45	
Thr	Thr	Leu	Phe	Cys	Ala	Ser	Asp	Ala	Lys	Ala	Tyr	Glu	Lys	Glu	Val	50	55	60	
His	Asn	Val	Trp	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn	Pro	65	70	75	80
Gln	Glu	Ile	Val	Leu	Gly	Asn	Val	Thr	Glu	Asn	Phe	Asn	Met	Trp	Lys	85	90	95	
Asn	Asp	Met	Val	Asp	Gln	Met	His	Glu	Asp	Ile	Ile	Ser	Leu	Trp	Asp	100	105	110	
Ala	Ser	Leu	Glu	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu	115	120	125	
Asn	Cys	Arg	Asn	Val	Arg	Asn	Val	Ser	Ser	Asn	Gly	Thr	Tyr	Asn	Ile	130	135	140	
Ile	His	Asn	Glu	Thr	Tyr	Lys	Glu	Met	Lys	Asn	Cys	Ser	Phe	Asn	Ala	145	150	155	160
Thr	Thr	Val	Val	Glu	Asp	Arg	Lys	Gln	Lys	Val	His	Ala	Leu	Phe	Tyr	165	170	175	
Arg	Leu	Asp	Ile	Val	Pro	Leu	Asp	Glu	Asn	Asn	Ser	Ser	Glu	Lys	Ser	180	185	190	
Ser	Glu	Asn	Ser	Ser	Glu	Tyr	Tyr	Arg	Leu	Ile	Asn	Cys	Asn	Thr	Ser	195	200	205	



# ES 2 898 271 T3

Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile  
210 215 220

His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys  
225 230 235 240

Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys  
245 250 255

Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly  
260 265 270

Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn  
275 280 285

Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr  
290 295 300

Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro  
305 310 315 320

Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln  
325 330 335

Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly  
340 345 350

Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Lys Phe  
355 360 365

Ala Pro His Ser Gly Gly Asp Leu Glu Ile Thr Thr His Thr Phe Asn  
370 375 380

Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Asn Leu Phe Asn Glu  
385 390 395 400

Ser Asn Ile Glu Arg Asn Asp Ser Ile Ile Thr Leu Pro Cys Arg Ile  
405 410 415

Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly Arg Ala Ile Tyr Ala  
420 425 430

Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser Asn Ile Thr Gly Leu  
435 440 445

Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly Val Pro Asn Asp Thr

450		455		460
Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg Asn Asn Trp Arg Ser				
465		470		475
				480
Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys Pro Leu Gly Val Ala				
		485		490
				495
Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg Glu Glu Arg Ala Val				
		500		505
				510
Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly Ala Ala Gly Ser Thr				
		515		520
				525
Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln Ala Arg Gln Leu Leu				
		530		535
				540
Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu Arg Ala Ile Glu Ala				
		545		550
				555
				560
Gln Gln His Met Leu Gln Leu Thr Val Trp Gly Ile Lys Gln Leu Gln				
		565		570
				575
Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln Asp Gln Gln Leu Leu				
		580		585
				590
Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys Thr Thr Ala Val Pro				
		595		600
				605
Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr Asp Ile Trp Asp Asn				
		610		615
				620
Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly Asn Tyr Thr Gly Glu				
		625		630
				635
				640
Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln Gln Glu Lys Met Lys				
		645		650
				655
Gln Ile Glu Asp Lys Ile Glu Glu Ile Leu Ser Lys Ile Tyr His Ile				
		660		665
				670
Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile Gly Glu Val Gly Ser				
		675		680
				685
Gly Ala Pro Thr Lys Ala Lys Arg Arg Val Val Gln Arg Glu Lys Arg				
		690		695
				700

<210> 20  
 <211> 4050  
 <212> ADN

<213> Secuencia artificial

<220>

<223> secuencia de nucleótidos que codifica mos1GagPol

5

<400> 20

atgggagcca gagccagcgt gctgtccgga ggggagctgg accgctggga gaagatcagg	60
ctgaggcctg gagggaagaa gaagtacagg ctgaagcaca tcgtgtgggc cagcagagag	120
ctggaacggt ttgccgtgaa ccctggcctg ctggaaacca gcgagggctg taggcagatt	180
ctgggacagc tgcagcccag cctgcagaca ggcagcgagg aactgcggag cctgtacaac	240
accgtggcca ccctgtactg cgtgcaccag cggatcgaga tcaaggacac caaagaagcc	300
ctggaaaaga tcgaggaaga gcagaacaag agcaagaaga aagcccagca ggctgccgct	360
gacacaggca acagcagcca ggtgtcccag aactacccca tcgtgcagaa catccaggga	420
cagatggtgc accaggccat cagccctcgg accctgaacg cctgggtgaa ggtggtggag	480
gaaaaggcct tcagccctga ggtgatcccc atgttctctg ccctgagcga gggagccaca	540
ccccaggacc tgaacaccat gctgaacacc gtgggagggc accaggctgc catgcagatg	600
ctgaaagaga caatcaacga ggaagctgcc gagtgggaca gggccacccc agtgcacgct	660
ggacctatcg ctcttgacca gatgagagag cccagaggca gcgatattgc tggcaccacc	720
tccacactgc aggaacagat cggctggatg accaacaacc ctcccattcc tgtgggagag	780
atctacaagc ggtggatcat tctgggactg aacaagatcg tgcggatgta cagccctgtg	840
agcatcctgg acatcaggca gggacccaaa gagcccttca gggactacgt ggaccgggtc	900
tacaagaccc tgagagccga gcaggccagc caggacgtga agaactggat gaccgagaca	960
ctgctggtgc agaacgcaa ccctgactgc aagaccatcc tgaaagccct gggacctgct	1020
gccaccctgg aagagatgat gacagcctgc caggagtggt gaggacctgg ccacaaggcc	1080
aggggtgctg ccgaggccat gagccagggt accaactctg ccaccatcat gatgcagaga	1140
ggcaacttcc ggaaccagag aaagaccgtg aagtgcctca actgtggcaa agagggacac	1200
attgccaaaga actgcagggc tcccaggaag aaaggctgct ggaagtgcgg aaaagaaggc	1260
caccagatga aggactgcac cgagaggcag gccaaacttcc tgggcaagat ctggcctagc	1320
aacaagggca ggcctggcaa ctctctgcag aacagaccgc agcccaccgc tcctcccgag	1380
gaaagcttcc ggtttggcga ggaaaccacc acccctagcc agaagcagga acccatcgac	1440
aaagagatgt accctctggc cagcctgaag agcctgttcg gcaacgaccc cagcagccag	1500
atggctccca tcagcccaat cgagacagtg cctgtgaagc tgaagcctgg catggacgga	1560
cccagggtga agcagtggcc tctgaccgag gaaaagatca aagccctgac agccatctgc	1620

gaggaaatgg	aaaaagaggg	caagatcacc	aagatcggac	ccgagaaccc	ctacaacacc	1680
cctgtgttcg	ccatcaagaa	gaaagacagc	accaagtggg	ggaaactggt	ggacttcaga	1740
gagctgaaca	agcggaccca	ggacttctgg	gaggtgcagc	tgggcatccc	tcaccctgct	1800
ggcctgaaga	aaaagaaaag	cgtgaccgtg	ctggctgtgg	gagatgccta	cttcagcgtg	1860
cctctggacg	agggcttccg	gaagtacaca	gccttcacca	tccccagcac	caacaacgag	1920
acacctggca	tcagatacca	gtacaacgtg	ctgcctcagg	gctggaaagg	cagccctgcc	1980
atcttccagt	gcagcatgac	cagaatcctg	gaacccttca	gagccaagaa	ccctgagatc	2040
gtgatctacc	agtatatggc	tgccctctac	gtgggcagcg	acctggaaat	cggacagcac	2100
agagccaaaa	tcgaagaact	ccgcgagcac	ctgctgaagt	ggggattcac	caccctcgac	2160
aagaagcacc	agaaagagcc	tcccttcctg	tggatgggct	acgagctgca	ccctgacaag	2220
tggaccgtgc	agcccatcca	gctgccagag	aaggactcct	ggaccgtgaa	cgacatccag	2280
aaactggtcg	gcaagctgaa	ctgggccagc	cagatctacc	ctggcatcaa	agtcagacag	2340
ctgtgtaagc	tgctgagggg	agccaaagca	ctgaccgaca	tcgtgcctct	gacagaagaa	2400
gccgagctgg	aactggccga	gaacagagag	atcctgaaag	aaccctgca	cggagtgtac	2460
tacgaccct	ccaaggacct	gattgccgag	atccagaaac	agggacacga	ccagtggacc	2520
taccagatct	atcaggaacc	tttcaagaac	ctgaaaacag	gcaagtacgc	caagatgcgg	2580
acagcccaca	ccaacgacgt	gaagcagctg	accgaagccg	tgcaaaaaat	cgccatggaa	2640
agcatcgtga	tctggggaaa	gacacccaag	ttcaggctgc	ccatccagaa	agagacatgg	2700
gaaacctggt	ggaccgacta	ctggcaggcc	acctggattc	ccgagtggga	gttcgtgaac	2760
acccaccccc	tggatgaagct	gtggtatcag	ctggaaaagg	accctatcgc	tggcgtggag	2820
acattctacg	tggctggagc	tgccaacaga	gagacaaagc	tgggcaaggc	tggctacgtg	2880
accgacagag	gcagacagaa	aatcgtgagc	ctgaccgaaa	ccaccaacca	gaaaacagcc	2940
ctgcaggcca	tctatctggc	actgcaggac	agcggaagcg	aggtgaacat	cgtgacagcc	3000
agccagtatg	ccctgggcat	catccaggcc	cagcctgaca	agagcgagag	cgagctggtg	3060
aaccagatca	tcgagcagct	gatcaagaaa	gaacgggtgt	acctgagctg	ggtgccagcc	3120
cacaagggca	tcggagggaa	cgagcaggtg	gacaagctgg	tgtccagcgg	aatccggaag	3180
gtgctgttcc	tggacggcat	cgataaagcc	caggaagagc	acgagaagta	ccacagcaat	3240
tggagagcca	tggccagcga	cttcaacctg	cctcccgtag	tggccaaaga	aatcgtggcc	3300
agctgcgacc	agtgccagct	gaaaggcgag	gccatgcacg	gacaggtgga	ctgctccctt	3360
ggcatctggc	agctggcatg	caccacacctg	gaaggcaaga	tcattctggt	ggccgtgcac	3420
gtggccagcg	gatacatcga	agccgaagtg	atccctgccg	agacagggca	ggaacagcc	3480
tacttcatcc	tgaagctggc	tggcagatgg	cctgtgaagg	tgatccacac	agccaacggc	3540

agcaacttca cctctgctgc cgtgaaggct gcctgttggt gggctggcat tcagcaggaa 3600  
 tttggcatcc cctacaatcc ccagtctcag ggagtgggtg ccagcatgaa caaagagctg 3660  
 aagaagatca tcggacaggt cagggatcag gccgagcacc tgaaaactgc cgtccagatg 3720  
 gccgtgttca tccacaactt caagcgggaag ggagggatcg gaggggtactc tgctggcgag 3780  
 cggatcatcg acatcattgc caccgatatc cagaccaaag agctgcagaa acagatcatc 3840  
 aagatccaga acttcagggt gtactacagg gacagcaggg accccatctg gaagggacct 3900  
 gccaaagctgc tgtggaaagg cgaaggagcc gtcgtcatcc aggacaacag cgacatcaag 3960  
 gtggtgcccc gacggaagggt gaaaatcatc aaggactacg gcaaacagat ggctggagcc 4020  
 gactgtgtcg ctggcaggca ggacgaggac 4050

<210> 21

<211> 4023

<212> ADN

<213> Secuencia artificial

<220>

<223> secuencia de nucleótidos que codifica mos2GagPol

<400> 21

atgggagcca gagccagcat cctgcgagga gggaagctgg acaagtggga gaagatcagg 60  
 ctgaggcctg gaggggaagaa aactacatg ctgaagcacc tggctctgggc cagcagagag 120  
 ctggaacggt ttgccctcaa tcctggcctg ctggaaacca gcgagggctg caagcagatc 180  
 atcaagcagc tgcagcctgc cctgcagaca ggcaccgagg aactgcggag cctgttcaac 240  
 accgtggcca ccctgtactg cgtgcatgcc gagatcgaag tgagggacac caaagaagcc 300  
 ctggacaaga tcgaggaaga gcagaacaag agccagcaga aaaccagca ggccaaagaa 360  
 gccgacggca aggtctccca gaactacccc atcgtgcaga acctgcaggg acagatggtg 420  
 caccagccca tcagccctcg gacactgaat gcctgggtga aggtgatcga ggaaaaggcc 480  
 ttcagccctg aggtgatccc catgttcaca gccctgagcg agggagccac accccaggac 540  
 ctgaacacca tgctgaacac cgtgggaggg caccaggctg ccatgcagat gctgaaggac 600  
 accatcaacg aggaagctgc cgagtgggac aggtgcacc ctgtgcacgc tggacctgtg 660  
 gctcctggcc agatgagaga gccagaggc agcgatattg ctggcaccac ctccaatctg 720  
 caggaacaga tcgcctggat gaccagcaac cctcccatcc ctgtgggaga catctacaag 780  
 cgggtggatca tcctgggact gaacaagatc gtgcggatgt acagccctac ctccatcctg 840  
 gacatcaagc agggacccaa agagcctttc agggactacg tggaccggtt cttcaagacc 900  
 ctgagagccg agcaggccac ccaggacgtg aagaactgga tgaccgacac cctgctggtg 960  
 cagaacgcca accctgactg caagaccatc ctgagagccc tgggacctgg agccaccctg 1020

gaagagatga tgacagcctg ccagggagtg ggaggaccct ctcacaaggc taggggtgctg	1080
gccgaggcca tgagccagac caacagcacc atcctgatgc agcggagcaa cttcaagggc	1140
agcaagcgga tcgtgaagtg cttcaactgt ggcaaagagg gacacattgc cagaaactgt	1200
agggcaccca ggaagaaagg ctgctggaag tgcggaaaag aaggccacca gatgaaggac	1260
tgcaccgaga ggcaggccaa cttcctgggc aagatctggc ctagccacaa gggcagacct	1320
ggcaacttcc tgcagagcag acccgagccc accgctcctc cagccgagag cttccggttc	1380
gaggaaacca cccctgctcc caagcaggaa cctaaggaca gagagcctct gaccagcctg	1440
agaagcctgt tcggcagcga ccctctgagc cagatggctc ccatctcccc tatcgagaca	1500
gtgcctgtga agctgaagcc tggcatggac ggaccaagtg tgaacagtg gcctctgacc	1560
gagggaaaaga tcaaaagccct ggtggagatc tgtaccgaga tggaaaaaga gggcaagatc	1620
agcaagatcg gacccgagaa cccctacaac acccctatct tcgccatcaa gaagaaagac	1680
agcaccaagt ggaggaaact ggtggacttc agagagctga acaagcggac ccaggacttc	1740
tgggaggtgc agctgggcat ccctcaccct gctggcctga agaaaaagaa aagcgtgacc	1800
gtgctggccg tgggagatgc ctacttcagc gtgcctctgg acgaggactt cagaaagtac	1860
acagccttca ccatccccag catcaacaac gagacacctg gcatacagata ccagtacaac	1920
gtgctgcctc agggatggaa gggctctcct gcaatcttcc agagcagcat gaccaagatc	1980
ctggaaccct tccggaagca gaaccctgac atcgtgatct accagtacat ggcagccctg	2040
tacgtcggca gcgacctgga aatcggacag caccggacca agatcgaaga actcaggcag	2100
cacctgctgc ggtggggatt caccaccctc gacaagaagc accagaaaga gcctcccttc	2160
ctgtggatgg gctacgagct gcaccagac aagtggaccg tgcagcccat cgtgctgcct	2220
gagaaggact cctggaccgt gaacgacatc cagaaactgg tcggcaagct gaactgggcc	2280
agccagatct acgctggcat caaagtgaag cagctgtgta agctcctgag aggcacccaa	2340
gccctgaccg aggtggtgcc actgacagag gaagccgagc tggaaactggc cgagaacaga	2400
gagatcctga aagaaccctg gcacggagtg tactacgacc ccagcaagga cctgattgcc	2460
gagatccaga agcagggaca gggacagtgg acctaccaga tctaccagga acccttcaag	2520
aacctgaaaa caggcaagta cgccaggatg aggggagccc acaccaacga cgtcaaacag	2580
ctgaccgaag ccgtgcagaa gatcgccacc gagagcatcg tgatttgggg aaagacaccc	2640
aagttcaagc tgcccatcca gaaagagaca tgggaggcct ggtggaccga gtactggcag	2700
gccacctgga tccccgagtg ggagttcgtg aacaccccac ccctggtgaa gctgtggtat	2760
cagctggaaa aagaacccat cgtgggagcc gagacattct acgtggctgg agctgccaac	2820
agagagacaa agctgggcaa ggctggctac gtgaccgaca gaggcaggca gaaagtgggtg	2880
tccctgaccg ataccaccaa ccagaaaaca gccctgcagg ccatccacct ggctctgcag	2940

gactctggcc	tggaagtga	catcgtgaca	gccagccagt	atgccctggg	catcattcag	3000
gcacagcctg	acaagagcga	gagcgagctg	gtgtctcaga	tcattgagca	gctgatcaag	3060
aaagaaaagg	tgtacctggc	ctgggtgcc	gccacaaagg	ggatcggagg	gaacgagcag	3120
gtggacaagc	tggtgtccag	gggcatccg	aaggtgctgt	ttctggacgg	catcgacaaa	3180
gcccaggaag	agcacgagaa	gtaccacagc	aattggagag	ccatggccag	cgagttcaac	3240
ctgcctccca	tcgtggccaa	agaaatcgtg	gcctcttgcg	acaagtgcc	gctgaaaggc	3300
gaggccattc	acggacaggt	ggactgcagc	ccaggcatct	ggcagctggc	ctgcacccac	3360
ctggaaggca	aggtgatcct	ggtggccgtg	cacgtggcct	ctggatacat	cgaagccgaa	3420
gtgatccctg	ccgagacagg	ccaggaaaca	gcctacttcc	tgctgaagct	ggctggcagg	3480
tggcctgtga	aaaccatcca	cacagccaac	ggcagcaact	tcacctctgc	caccgtgaag	3540
gctgcctgtt	ggtgggctgg	cattaagcag	gaatttggca	tcccctacaa	ccctcagtct	3600
cagggaagtgg	tgccctccat	caacaaagag	ctgaagaaga	tcatcggaca	ggtcagggat	3660
caggccgagc	atctgaaaac	agccgtccag	atggccgtgt	tcattccaaa	cttcaagcgg	3720
aagggaaggga	tcggagagta	ctctgctggc	gagaggatcg	tggaatttat	cggcagcgat	3780
atccagacca	aagaactgca	gaagcagatc	acaaagatcc	agaacttcag	ggtgtactac	3840
agggaacagca	gagatcccct	gtggaaggga	cctgccaagc	tgctgtggaa	aggcgaaggga	3900
gccgtcgtca	tccaggacaa	cagcgacatc	aaggtggtgc	ccagacggaa	ggccaagatc	3960
atcagagact	acggcaaaca	gatggctggc	gacgactgcg	tcgcctctag	gcaggacgag	4020
gac						4023

<210> 22  
 <211> 2055  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> secuencia de nucleótidos que codifica mos1Env

<400> 22

atgcgggtga	ccggcatccg	gaagaactac	cagcacctgt	ggcgggtggg	caccatgctg	60
ctgggcatcc	tgatgatttg	ctctgccgcc	ggaaagctgt	gggtgaccgt	gtactacggc	120
gtgcccgtgt	ggaaagaggc	caccaccacc	ctgttctgcg	ccagcgacgc	caaggcctac	180
gacaccgagg	tgcacaacgt	gtggggccacc	cacgcctgcg	tgcccaccga	ccccaacccc	240
caggaagtgg	tcctggaaaa	cgtgaccgag	aacttcaaca	tgtggaagaa	caacatggtg	300
gagcagatgc	acgaggacat	catcagcctg	tgggaccaga	gcctgaagcc	ctgcgtgaag	360
ctgaccccc	tgtgcgtgac	cctgaactgc	accgacgacg	tgcggaacgt	gaccaacaac	420

# ES 2 898 271 T3

gccaccaaca ccaacagcag ctggggcgag cctatggaaa agggcgagat caagaactgc 480  
agcttcaaca tcaccacctc catccggaac aaggtgcaga agcagtacgc cctgttctac 540  
aagctggacg tgggtgcccac cgacaacgac agcaacaaca ccaactaccg gctgatcagc 600  
tgcaacacca gcgtgatcac ccaggcctgc cccaaggtgt ccttcgagcc catccccatc 660  
cactactgcg cccctgccgg cttcgccatc ctgaagtgca acgacaagaa gttcaacggc 720  
accggccccct gcaccaacgt gagcacctg cagtgcaccc acggcatccg gcccggtgtg 780  
tccacccagc tgctgctgaa cggcagcctg gccgaggaag aggtggtgat cagaagcgag 840  
aatttcacca acaatgccaa gaccatcatg gtgcagctga acgtgagcgt ggagatcaac 900  
tgcacccggc ccaacaacaa cacccggaag agcatccaca tcggccctgg cagggccttc 960  
tacacagccg gcgacatcat cggcgacatc cggcaggccc actgcaacat cagccggggc 1020  
aactggaaca acaccctgcg gcagatcgtg gagaagctgg gcaagcagtt cggcaacaac 1080  
aagaccatcg tgttcaacca cagcagcggc ggagaccccg agatcgtgat gcacagcttc 1140  
aactgtggcg gcgagttctt ctactgcaac agcaccaagc tgttcaacag cacctggacc 1200  
tggaacaact ccacctgga taacaccaag cggagcaacg acaccgaaga gcacatcacc 1260  
ctgccctgcc ggatcaagca gattatcaat atgtggcagg aggtcggcaa ggccatgtac 1320  
gccctccca tccggggcca gatccggtgc agcagcaaca tcaccggcct gctgctgacc 1380  
cgggacggcg gcaacgatac cagcggcacc gagatcttcc ggcctggcgg cggagatatg 1440  
cgggacaact ggcgagcgga gctgtacaag tacaaggtgg tgaagatcga gccctgggc 1500  
gtggctccca ccaaggccaa gcggcgggtg gtgcagagcg agaagagcgc cgtgggcac 1560  
ggcgccgtgt ttctgggctt cctgggagcc gccggaagca ccatgggagc cgccagcatg 1620  
accctgaccg tgcaggcccg gctgctgctg tccggcatcg tgcagcagca gaacaacctg 1680  
ctccgggcca tcgaggccca gcagcacctg ctgcagctga ccgtgtgggg catcaagcag 1740  
ctgcaggcca ggggtgctggc cgtggagaga tacctgaagg atcagcagct cctggggatc 1800  
tggggctgca gcggcaagct gatctgcacc accaccgtgc cctggaacgc cagctggtcc 1860  
aacaagagcc tggacaagat ctggaacaat atgacctgga tggaatggga gcgcgagatc 1920  
aacaattaca ccagcctgat ctacacctg atcgaggaaa gccagaacca gcaggaaaag 1980  
aacgagcagg aactgctgga actggacaag tgggccagcc tgtggaactg gttcgacatc 2040  
agcaactggc tgtgg 2055

<210> 23  
<211> 2052  
<212> ADN  
<213> Secuencia artificial

<220>  
<223> secuencia de nucleótidos que codifica mos2Env

<400> 23



atgagagtgc	ggggcatcca	gcggaactgg	ccccagtgg	ggatctgggg	catcctgggc	60
ttttggatga	tcatcatctg	cggggtgatg	ggcaacctgt	gggtgaccgt	gtactacggc	120
gtgcccgtgt	ggaaagaggc	caagaccacc	ctgttctgcg	ccagcgacgc	caaggcctac	180
gagaaagagg	tgcacaacgt	gtgggccacc	cacgcctgcg	tgccaccga	ccccaacccc	240
caggaaatgg	tcctggaaaa	cgtgaccgag	aacttcaaca	tgtggaagaa	cgacatggtg	300
gaccagatgc	acgaggacat	catccggctg	tgggaccaga	gcctgaagcc	ctgctggaag	360
ctgaccccc	tgtgctgac	cctggaatgc	cggaactga	gaaactgag	cagcaacggc	420
acctacaaca	tcatccacaa	cgagacctac	aaagagatga	agaactgcag	cttcaacgcc	480
accaccgtgg	tggaggaccg	gaagcagaag	gtgcacgccc	tgttctaccg	gctggacatc	540
gtgcccctgg	acgagaacaa	cagcagcgag	aagtccagcg	agaacagctc	cgagtactac	600
cggctgatca	actgcaacac	cagcgccatc	accaggcct	gccccaaagg	gtccttcgac	660
cccatcccca	tccactactg	cgccccctgc	ggctacgcca	tcctgaagtg	caacaacaag	720
accttcaacg	gcaccggccc	ctgcaacaac	gtgagcaccg	tgcagtgcac	ccacggcatc	780
aagcccgtgg	tgtccaccca	gctgctgctg	aacggcagcc	tggccgagga	agagatcatc	840
atccggtccg	agaacctgac	caacaacgcc	aagaccatca	tcgtgcacct	gaatgagacc	900
gtgaacatca	cctgcacccg	gccccacaac	aacaccggga	agagcatccg	gatcggccct	960
ggccagacct	tttacgccac	cggcgacatc	atcggcgaca	tccggcaggc	ccactgcaac	1020
ctgagccggg	acggctggaa	caagaccctg	cagggcgtga	agaagaagct	ggccgagcac	1080
ttccccaata	agaccatcaa	cttcaccagc	agcagcggcg	gagacctgga	aatcaccacc	1140
cacagcttca	actgcagggg	cgagttcttc	tactgcaata	cctccggcct	gttcaatggc	1200
acctacatgc	ccaacggcac	caacagcaac	agcagcagca	acatcacctc	gccctgccgg	1260
atcaagcaga	tcatcaatat	gtggcaggag	gtcggcaggg	ccatgtacgc	ccctcccatc	1320
gccggcaata	tcacctgccg	gtccaacatc	accggcctgc	tgtgaccag	ggacggcggc	1380
agcaacaacg	gcgtgcctaa	cgacaccgag	accttccggc	ctggcggcgg	agatatgcgg	1440
aacaactggc	ggagcgagct	gtacaagtac	aaggtggtgg	aggtgaagcc	cctgggcgtg	1500
gctcctaccg	aggccaagcg	gcgggtggtg	gagagcgaga	agagcgccgt	gggcatcggc	1560
gccgtgtttc	tgggcattct	gggagccgcc	ggaagcacca	tgggagccgc	cagcatcacc	1620
ctgaccgtgc	aggcccggca	gctgctgtcc	ggcatcgtgc	agcagcagag	caacctgctg	1680
agagccatcg	aggcccagca	gcacatgctg	cagctgaccg	tgtggggcat	caagcagctg	1740
cagacccggg	tgtgtggccat	cgagagatac	ctgcaggatc	agcagctcct	gggcctgtgg	1800

# ES 2 898 271 T3

	ggctgcagcg gcaagctgat ctgcaccacc gccgtgccct ggaacaccag ctggtccaac	1860
	aagagccaga cgcacatctg ggacaacatg acctggatgc agtgggacaa agagatcggc	1920
	aactacaccg gcgagatcta caggctgctg gaagagagcc agaaccagca ggaaaagaac	1980
	gagaaggacc tgctggccct ggacagctgg aagaacctgt ggaactgggt cgacatcacc	2040
	aactggctgt gg	2052
5	<210> 24 <211> 829 <212> ADN <213> Secuencia artificial	
10	<220> <223> promotor de CMV utilizado para la expresión de antígenos en vectores de Ad26  <400> 24	
	tcaatatttg ccattagcca tattattcat tggttatata gcataaatca atattggcta	60
	ttggccattg catacgttgt atccatatca taatatgtac atttatattg gctcatgtcc	120
	aacattaccg ccatgttgac attgattatt gactagttat taatagtaat caattacggg	180
	gtcattagtt catagcccat atatggagtt ccgcgttaca taacttacgg taaatggccc	240
	gcctggctga ccgccaacg acccccgccc attgacgtca ataatgacgt atgttcccat	300
	agtaacgcca atagggactt tccattgacg tcaatgggtg gagtatttac ggtaaactgc	360
	ccacttggca gtacatcaag tgtatcatat gccaagtacg cccctattg acgtcaatga	420
	cggtaaatgg cccgcctggc attatgccca gtacatgacc ttatgggact ttcctacttg	480
	gcagtacatc tacgtattag tcatcgctat taccatgggt atgcggtttt ggcagtacat	540
	caatgggctg ggatagcggg ttgactcacg gggatttcca agtctccacc ccattgacgt	600
	caatgggagt ttgttttggc accaaaatca acgggacttt ccaaaatgtc gtaacaactc	660
	cgccccattg acgcaaattg gcggtagggc tgtacgggtg gaggtctata taagcagagc	720
	tcgttttagtg aaccgtcaga tcgcctggag acgccatcca cgctgttttg acctccatag	780
	aagacaccgg gaccgatcca gcctccgcgg ccgggaacgg tgcattgga	829
15	<210> 25 <211> 2586 <212> ADN <213> Secuencia artificial	
20	<220> <223> secuencia de nucleótidos que codifica C4  <400> 25	
	atgagagtgc ggggcatgct gagaaactgg cagcagtggg ggatctgggt cagcctgggc	60
	ttctggatgc tgatgatcta cagcgtgatg ggcaacctgt gggtcaccgt gtactacggc	120
25	gtgcccgtgt ggaaggacgc caagaccacc ctgttttgcg cctccgatgc caaggcctac	180

gagaaagagg	tgcacaacgt	ctggggccacc	cacgcctgtg	tgcccaccga	ccccaatccc	240
caggaaatcg	tcctgggcaa	cgtgaccgag	aacttcaaca	tgtggaagaa	cgacatggtc	300
gatcagatgc	acgaggacat	catctccctg	tgggacgcct	ccctggaacc	ctgcgtgaag	360
ctgaccctc	tgtgcgtgac	cctgaactgc	cggaaactgc	gcaactgtc	cagcaacggc	420
acctacaaca	tcatccacaa	cgagacatac	aaagagatga	agaactgcag	cttcaacgct	480
accaccgtgg	tcgaggaccg	gaagcagaag	gtgcacgccc	tgttctaccg	gctggacatc	540
gtgcccctgg	acgagaacaa	cagcagcgag	aagtccctccg	agaacagctc	cgagtactac	600
agactgatca	actgcaacac	cagcgccatc	acccaggcct	gccccaaagg	gtccttcgac	660
cctatcccca	tccactactg	cgcccctgcc	ggctacgcca	tcctgaagtg	caacaacaag	720
accttcaatg	gcaccggccc	ctgcaacaat	gtgtccaccg	tgcaagtgcac	ccacggcatc	780
aagcccgtgg	tgtctacca	gctgctgctg	aacggcagcc	tggccgagga	agagatcatt	840
atcagaagcg	agaacctgac	caacaacgcc	aaaaccatca	tcgtccacct	gaacgaaacc	900
gtgaacatca	cctgtaccgg	gcctaacaac	aacacccgga	agtccatccg	gatcggccct	960
ggccagacct	tttacgccac	cggcgatatt	atcggcgaca	tccggcaggc	ccactgcaat	1020
ctgagccggg	acggctggaa	caagacactg	cagggcgctca	agaagaagct	ggccgaacac	1080
ttccctaaca	agactatcaa	gttcgcccct	cactctggcg	gcgacctgga	aatcaccacc	1140
cacaccttca	actgtcgggg	cgagttcttc	tactgcaata	cctccaacct	gttcaacgag	1200
agcaacatcg	agcgggaacga	cagcatcatc	acactgcctt	gccggatcaa	gcagattatc	1260
aatatgtggc	aggaagtggg	cagagccatc	tacgcccctc	caatcgccgg	caacatcaca	1320
tgccggtcca	atatcaccgg	cctgctgctc	accagagatg	gcggctccaa	caatggcgtg	1380
ccaaacgaca	ccgagacatt	cagacccggc	ggaggcgaca	tgcggaacaa	ttggcggagc	1440
gagctgtaca	agtacaagg	ggtggaagt	aagcccctgg	gcgtggcccc	taccgaggcc	1500
aagagaagag	tggctgaacg	cgagaagcgg	gccgtgggaa	tcggagccgt	gtttctggga	1560
atcctgggag	ccgctggctc	taccatgggc	gctgcctcta	tcacctgac	agtgcaggcc	1620
agacagctgc	tcagcggcat	cgtgcagcag	cagagcaacc	tgctgagagc	cattgaggcc	1680
cagcagcaca	tgctgcagct	gaccgtgtgg	ggcattaagc	agctccagac	acgggtgctg	1740
gccatcgaga	gatacctgca	ggatcagcag	ctcctggggc	tgtggggctg	tagcggcaag	1800
ctgatctgta	ccaccgccgt	gccctggaat	acctcttgga	gcaacaagag	ccagaccgac	1860
atctgggaca	acatgacctg	gatgcagtgg	gacaaagaaa	tcggcaacta	taccggcgag	1920
atctatagac	tgctggaaga	gtcccagaac	cagcaggaaa	agaacgagaa	ggacctgctg	1980
gccctggatt	cttggaaaca	tctgtggaac	tggttcagca	tctccaagt	gctgtggtac	2040

# ES 2 898 271 T3

atcaagatct tcatcatgat cgtgggcggc ctgatcggcc tgcggatcat ctttgccgtg	2100
ctgagcatcg tgaaccgcgt gcggcaggga tacagccctc tgagcctgca gaccctgact	2160
cagaaccctg gcggactgga cagactgggc cggattgagg aagaaggcgg cgagcaggac	2220
aaggatcgga gcatcaggct ggtcaacggc ttcttcgctc tgttttgga cgacctgcgg	2280
agcctgtgcc tgttcagcta ccacagactg cgggacttta tcctgattgt ggccagagcc	2340
gtcgaactgc tggggagaag ctctctgaga ggccctgcagc ggggctggga gattctgaag	2400
tacctgggct ccctgctgca gtactggggc ctggaactga agaagtctgc catcaatctg	2460
ctcgacacaa tcgctattgc cgtggccgaa ggcaccgata gaatcatcga gctgatccag	2520
cggatctgcc gggccatctg caacatcccc agacggatca gacagggctt cgaggccgct	2580
ctgcag	2586

<210> 26  
 <211> 2133  
 <212> ADN  
 <213> Secuencia artificial

<220>  
 <223> secuencia de nucleótidos que codifica C4D7

<400> 26

atgagagtgc ggggcatgct gagaaactgg cagcagtggg ggatctgggc cagcctgggc	60
ttctggatgc tgatgatcta cagcgtgatg ggcaacctgt gggtcaccgt gtactacggc	120
gtgcccgtgt ggaaggacgc caagaccacc ctgttttgcg cctccgatgc caaggcctac	180
gagaaagagg tgcacaacgt ctggggccacc cagcctgtg tggccaccga ccccaatccc	240
caggaaatcg tcctgggcaa cgtgaccgag aacttcaaca tgtggaagaa cgacatgggc	300
gatcagatgc acgaggacat catctccctg tgggacgcct ccctggaacc ctgctgaag	360
ctgacctctc tgtgctgac cctgaactgc cggaaactgc gcaacgtgtc cagcaacggc	420
acctacaaca tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct	480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttctaccg gctggacatc	540
gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac	600
agactgatca actgcaaacac cagcgccatc acccaggcct gcccgaagg gtccttcgac	660
cctatcccca tcactactg cggccctgcc ggctacgcca tcctgaagtg caacaacaag	720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcaagtgcac ccacggcatc	780
aagcccgtgg tgtctaccca gctgctgctg aacggcagcc tggccgagga agagatcatt	840
atcagaagcg agaactgac caacaacgcc aaaaccatca tcgtccacct gaacgaaacc	900
gtgaacatca cctgtacccg gcctaacaac aacacccgga agtccatccg gatcggccct	960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc ccactgcaat	1020

ctgagccggg	acggctggaa	caagacactg	cagggcgtca	agaagaagct	ggccgaacac	1080
ttccctaaca	agactatcaa	gttcgcccct	cactctggcg	gcgacctgga	aatcaccacc	1140
cacaccttca	actgtcgggg	cgagttcttc	tactgcaata	cctccaacct	gttcaacgag	1200
agcaacatcg	agcggaacga	cagcatcatc	acactgcctt	gccggatcaa	gcagattatc	1260
aatatgtggc	aggaagtggg	cagagccatc	tacgcccctc	caatcgccgg	caacatcaca	1320
tgccggtcca	atatcaccgg	cctgctgctc	accagagatg	gcggctccaa	caatggcgtg	1380
ccaaacgaca	ccgagacatt	cagacccggc	ggaggcgaca	tgcggaacaa	ttggcgagac	1440
gagctgtaca	agtacaaggt	ggtggaagtg	aagcccctgg	gcgtggcccc	taccgaggcc	1500
aagagaagag	tggtcgaacg	cgagaagcgg	gccgtgggaa	tcggagccgt	gtttctggga	1560
atcctgggag	ccgctggctc	taccatgggc	gctgcctcta	tcaccctgac	agtgcaggcc	1620
agacagctgc	tcagcggcat	cgtgcagcag	cagagcaacc	tgctgagagc	cattgaggcc	1680
cagcagcaca	tgctgcagct	gaccgtgtgg	ggcattaagc	agctccagac	acgggtgctg	1740
gccatcgaga	gatacctgca	ggatcagcag	ctcctggggc	tgtggggctg	tagcggcaag	1800
ctgatctgta	ccaccgccgt	gccctggaat	acctcttgga	gcaacaagag	ccagaccgac	1860
atctgggaca	acatgacctg	gatgcagtgg	gacaaagaaa	tcggcaacta	taccggcgag	1920
atctatagac	tgctggaaga	gtcccagaac	cagcaggaaa	agaacgagaa	ggacctgctg	1980
gccctggatt	cttggaaaca	tctgtggaac	tggttcagca	tctccaagtg	gctgtggtac	2040
atcaagatct	tcatcatgat	cgtgggcggc	ctgatcggcc	tgcgatcat	ctttgccgtg	2100
ctgagcatcg	tgaaccgcgt	gcggcagggc	tac			2133

<210> 27

<211> 2112

<212> ADN

<213> Secuencia artificial

<220>

<223> secuencia de nucleótidos que codifica sC4

<400> 27

atgagagtgc	ggggcatgct	gagaaactgg	cagcagtggc	ggatctggtc	cagcctgggc	60
ttctggatgc	tgatgatcta	cagcgtgatg	ggcaacctgt	gggtcaccgt	gtactacggc	120
gtgccctgtg	ggaaggacgc	caagaccacc	ctgttttgcg	cctccgatgc	caaggcctac	180
gagaaagagg	tgacacaacgt	ctgggcccacc	caogcctgtg	tgcccaccga	ccccaatccc	240
caggaaatcg	tcctgggcaa	cgtgaccgag	aacttcaaca	tgtggaagaa	cgacatggctc	300
gatcagatgc	acgaggacat	catctccctg	tgggacgcct	ccctggaacc	ctgcgtgaag	360
ctgaccccctc	tgtgcgtgac	cctgaactgc	cggaacgtgc	gcaacgtgtc	cagcaacggc	420

# ES 2 898 271 T3

```

acctacaaca tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct      480
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttctaccg gctggacatc      540
gtgcccctgg acgagaacaa cagcagcgag aagtccctccg agaacagctc cgagtactac      600
agactgatca actgcaacac cagcgccatc acccaggcct gcccgaaggt gtccttcgac      660
cctatcccca tcactactg cgcccctgcc ggctacgcca tcctgaagtg caacaacaag      720
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc      780
aagcccgtgg tgtctacca gctgctgctg aacggcagcc tggccgagga agagatcatt      840
atcagaagcg agaactgac caacaacgcc aaaaccatca tcgtccacct gaacgaaacc      900
gtgaacatca cctgtaccgg gcctaacaac aacacccgga agtccatccg gatcggccct      960
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc cactgcaat     1020
ctgagccggg acggctggaa caagacactg cagggcgctc agaagaagct ggccgaacac     1080
ttccctaaca agactatcaa gttogcccct cactctggcg gcgacctgga aatcaccacc     1140
cacaccttca actgtcgggg cgagttcttc tactgcaata cctccaacct gttcaacgag     1200
agcaacatcg agcggaacga cagcatcatc aactgcctt gccggatcaa gcagattatc     1260
aatatgtggc aggaagtggg cagagccatc tacgcccctc caatcgccgg caacatcaca     1320
tgccggtcca atatcaccgg cctgctgctc accagagatg gcggtccaa caatggcgtg     1380
ccaaacgaca ccgagacatt cagacccggc ggaggcgaca tgcggaacaa ttggcggagc     1440
gagctgtaca agtacaaggt ggtggaagtg aagcccctgg gcgtggcccc taccgaggcc     1500
aagagaagag tggtcgaacg cgaggaacgg gccgtgggaa tcggagccgt gtttctggga     1560
atcctgggag ccgctggctc taccatgggc gctgcctcta tcaccctgac agtgcaggcc     1620
agacagctgc tcagcgcat cgtgcagcag cagagcaacc tgctgagagc cattgaggcc     1680
cagcagcaca tgctgcagct gaccgtgtgg ggcattaagc agctccagac acgggtgctg     1740
gccatcgaga gatacctgca ggatcagcag ctccctgggc tgtggggctg tagcggcaag     1800
ctgatctgta ccaccgccgt gccctggaat acctcttgga gcaacaagag ccagaccgac     1860
atctgggaca acatgacctg gatgcagtgg gacaaagaaa tcggcaacta taccggcgag     1920
atctatagac tgctggaaga gtcccagaac cagcaggaaa agatgaagca gatcgaggac     1980
aagatcgaag agattctgag caagatctac cacatcgaga acgagatcg ccgcatcaag     2040
aaactgatcg gcgaagtggg atccggcgct cccacaaagg ccaaaagacg ggtggtgcag     2100
cgcgagaaac gc                                                                2112

```

<210> 28  
 <211> 1350  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> secuencia del antígeno mos1GagPol mosaico  
 <400> 28

# ES 2 898 271 T3

Met Gly Ala Arg Ala Ser Val Leu Ser Gly Gly Glu Leu Asp Arg Trp  
1 5 10 15

Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys Lys Tyr Arg Leu Lys  
20 25 30

His Ile Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Val Asn Pro  
35 40 45

Gly Leu Leu Glu Thr Ser Glu Gly Cys Arg Gln Ile Leu Gly Gln Leu  
50 55 60

Gln Pro Ser Leu Gln Thr Gly Ser Glu Glu Leu Arg Ser Leu Tyr Asn  
65 70 75 80

Thr Val Ala Thr Leu Tyr Cys Val His Gln Arg Ile Glu Ile Lys Asp  
85 90 95

Thr Lys Glu Ala Leu Glu Lys Ile Glu Glu Glu Gln Asn Lys Ser Lys  
100 105 110

Lys Lys Ala Gln Gln Ala Ala Ala Asp Thr Gly Asn Ser Ser Gln Val  
115 120 125

Ser Gln Asn Tyr Pro Ile Val Gln Asn Ile Gln Gly Gln Met Val His  
130 135 140

Gln Ala Ile Ser Pro Arg Thr Leu Asn Ala Trp Val Lys Val Val Glu  
145 150 155 160

Glu Lys Ala Phe Ser Pro Glu Val Ile Pro Met Phe Ser Ala Leu Ser  
165 170 175

Glu Gly Ala Thr Pro Gln Asp Leu Asn Thr Met Leu Asn Thr Val Gly  
180 185 190

Gly His Gln Ala Ala Met Gln Met Leu Lys Glu Thr Ile Asn Glu Glu  
195 200 205

Ala Ala Glu Trp Asp Arg Val His Pro Val His Ala Gly Pro Ile Ala  
210 215 220

Pro Gly Gln Met Arg Glu Pro Arg Gly Ser Asp Ile Ala Gly Thr Thr

# ES 2 898 271 T3

225						230										235														240
Ser	Thr	Leu	Gln	Glu	Gln	Ile	Gly	Trp	Met	Thr	Asn	Asn	Pro	Pro	Ile															
				245					250					255																
Pro	Val	Gly	Glu	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys															
			260					265					270																	
Ile	Val	Arg	Met	Tyr	Ser	Pro	Val	Ser	Ile	Leu	Asp	Ile	Arg	Gln	Gly															
		275					280					285																		
Pro	Lys	Glu	Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Tyr	Lys	Thr	Leu															
	290					295					300																			
Arg	Ala	Glu	Gln	Ala	Ser	Gln	Asp	Val	Lys	Asn	Trp	Met	Thr	Glu	Thr															
305						310				315																				
Leu	Leu	Val	Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Lys	Ala															
				325					330																					
Leu	Gly	Pro	Ala	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly															
			340					345					350																	
Val	Gly	Gly	Pro	Gly	His	Lys	Ala	Arg	Val	Leu	Ala	Glu	Ala	Met	Ser															
		355					360					365																		
Gln	Val	Thr	Asn	Ser	Ala	Thr	Ile	Met	Met	Gln	Arg	Gly	Asn	Phe	Arg															
	370						375				380																			
Asn	Gln	Arg	Lys	Thr	Val	Lys	Cys	Phe	Asn	Cys	Gly	Lys	Glu	Gly	His															
385						390				395					400															
Ile	Ala	Lys	Asn	Cys	Arg	Ala	Pro	Arg	Lys	Lys	Gly	Cys	Trp	Lys	Cys															
				405					410						415															
Gly	Lys	Glu	Gly	His	Gln	Met	Lys	Asp	Cys	Thr	Glu	Arg	Gln	Ala	Asn															
			420					425					430																	
Phe	Leu	Gly	Lys	Ile	Trp	Pro	Ser	Asn	Lys	Gly	Arg	Pro	Gly	Asn	Phe															
		435					440					445																		
Leu	Gln	Asn	Arg	Pro	Glu	Pro	Thr	Ala	Pro	Pro	Glu	Glu	Ser	Phe	Arg															
		450					455					460																		
Phe	Gly	Glu	Glu	Thr	Thr	Thr	Pro	Ser	Gln	Lys	Gln	Glu	Pro	Ile	Asp															
465						470				475					480															



Lys Glu Met Tyr Pro Leu Ala Ser Leu Lys Ser Leu Phe Gly Asn Asp  
 485 490 495

Pro Ser Ser Gln Met Ala Pro Ile Ser Pro Ile Glu Thr Val Pro Val  
 500 505 510

Lys Leu Lys Pro Gly Met Asp Gly Pro Arg Val Lys Gln Trp Pro Leu  
 515 520 525

Thr Glu Glu Lys Ile Lys Ala Leu Thr Ala Ile Cys Glu Glu Met Glu  
 530 535 540

Lys Glu Gly Lys Ile Thr Lys Ile Gly Pro Glu Asn Pro Tyr Asn Thr  
 545 550 555 560

Pro Val Phe Ala Ile Lys Lys Lys Asp Ser Thr Lys Trp Arg Lys Leu  
 565 570 575

Val Asp Phe Arg Glu Leu Asn Lys Arg Thr Gln Asp Phe Trp Glu Val  
 580 585 590

Gln Leu Gly Ile Pro His Pro Ala Gly Leu Lys Lys Lys Lys Ser Val  
 595 600 605

Thr Val Leu Ala Val Gly Asp Ala Tyr Phe Ser Val Pro Leu Asp Glu  
 610 615 620

Gly Phe Arg Lys Tyr Thr Ala Phe Thr Ile Pro Ser Thr Asn Asn Glu  
 625 630 635 640

Thr Pro Gly Ile Arg Tyr Gln Tyr Asn Val Leu Pro Gln Gly Trp Lys  
 645 650 655

Gly Ser Pro Ala Ile Phe Gln Cys Ser Met Thr Arg Ile Leu Glu Pro  
 660 665 670

Phe Arg Ala Lys Asn Pro Glu Ile Val Ile Tyr Gln Tyr Met Ala Ala  
 675 680 685

Leu Tyr Val Gly Ser Asp Leu Glu Ile Gly Gln His Arg Ala Lys Ile  
 690 695 700

Glu Glu Leu Arg Glu His Leu Leu Lys Trp Gly Phe Thr Thr Pro Asp  
 705 710 715 720

Lys Lys His Gln Lys Glu Pro Pro Phe Leu Trp Met Gly Tyr Glu Leu  
 725 730 735

# ES 2 898 271 T3

His Pro Asp Lys Trp Thr Val Gln Pro Ile Gln Leu Pro Glu Lys Asp  
 740 745 750  
 Ser Trp Thr Val Asn Asp Ile Gln Lys Leu Val Gly Lys Leu Asn Trp  
 755 760 765  
 Ala Ser Gln Ile Tyr Pro Gly Ile Lys Val Arg Gln Leu Cys Lys Leu  
 770 775 780  
 Leu Arg Gly Ala Lys Ala Leu Thr Asp Ile Val Pro Leu Thr Glu Glu  
 785 790 795 800  
 Ala Glu Leu Glu Leu Ala Glu Asn Arg Glu Ile Leu Lys Glu Pro Val  
 805 810 815  
 His Gly Val Tyr Tyr Asp Pro Ser Lys Asp Leu Ile Ala Glu Ile Gln  
 820 825 830  
 Lys Gln Gly His Asp Gln Trp Thr Tyr Gln Ile Tyr Gln Glu Pro Phe  
 835 840 845  
 Lys Asn Leu Lys Thr Gly Lys Tyr Ala Lys Met Arg Thr Ala His Thr  
 850 855 860  
 Asn Asp Val Lys Gln Leu Thr Glu Ala Val Gln Lys Ile Ala Met Glu  
 865 870 875 880  
 Ser Ile Val Ile Trp Gly Lys Thr Pro Lys Phe Arg Leu Pro Ile Gln  
 885 890 895  
 Lys Glu Thr Trp Glu Thr Trp Trp Thr Asp Tyr Trp Gln Ala Thr Trp  
 900 905 910  
 Ile Pro Glu Trp Glu Phe Val Asn Thr Pro Pro Leu Val Lys Leu Trp  
 915 920 925  
 Tyr Gln Leu Glu Lys Asp Pro Ile Ala Gly Val Glu Thr Phe Tyr Val  
 930 935 940  
 Ala Gly Ala Ala Asn Arg Glu Thr Lys Leu Gly Lys Ala Gly Tyr Val  
 945 950 955 960  
 Thr Asp Arg Gly Arg Gln Lys Ile Val Ser Leu Thr Glu Thr Thr Asn  
 965 970 975  
 Gln Lys Thr Ala Leu Gln Ala Ile Tyr Leu Ala Leu Gln Asp Ser Gly  
 980 985 990

# ES 2 898 271 T3

Ser	Glu	Val	Asn	Ile	Val	Thr	Ala	Ser	Gln	Tyr	Ala	Leu	Gly	Ile	Ile
		995					1000					1005			
Gln	Ala	Gln	Pro	Asp	Lys	Ser	Glu	Ser	Glu	Leu	Val	Asn	Gln	Ile	
	1010					1015					1020				
Ile	Glu	Gln	Leu	Ile	Lys	Lys	Glu	Arg	Val	Tyr	Leu	Ser	Trp	Val	
	1025					1030					1035				
Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn	Glu	Gln	Val	Asp	Lys	Leu	
	1040					1045					1050				
Val	Ser	Ser	Gly	Ile	Arg	Lys	Val	Leu	Phe	Leu	Asp	Gly	Ile	Asp	
	1055					1060					1065				
Lys	Ala	Gln	Glu	Glu	His	Glu	Lys	Tyr	His	Ser	Asn	Trp	Arg	Ala	
	1070					1075					1080				
Met	Ala	Ser	Asp	Phe	Asn	Leu	Pro	Pro	Val	Val	Ala	Lys	Glu	Ile	
	1085					1090					1095				
Val	Ala	Ser	Cys	Asp	Gln	Cys	Gln	Leu	Lys	Gly	Glu	Ala	Met	His	
	1100					1105					1110				
Gly	Gln	Val	Asp	Cys	Ser	Pro	Gly	Ile	Trp	Gln	Leu	Ala	Cys	Thr	
	1115					1120					1125				
His	Leu	Glu	Gly	Lys	Ile	Ile	Leu	Val	Ala	Val	His	Val	Ala	Ser	
	1130					1135					1140				
Gly	Tyr	Ile	Glu	Ala	Glu	Val	Ile	Pro	Ala	Glu	Thr	Gly	Gln	Glu	
	1145					1150					1155				
Thr	Ala	Tyr	Phe	Ile	Leu	Lys	Leu	Ala	Gly	Arg	Trp	Pro	Val	Lys	
	1160					1165					1170				
Val	Ile	His	Thr	Ala	Asn	Gly	Ser	Asn	Phe	Thr	Ser	Ala	Ala	Val	
	1175					1180					1185				
Lys	Ala	Ala	Cys	Trp	Trp	Ala	Gly	Ile	Gln	Gln	Glu	Phe	Gly	Ile	
	1190					1195					1200				
Pro	Tyr	Asn	Pro	Gln	Ser	Gln	Gly	Val	Val	Ala	Ser	Met	Asn	Lys	
	1205					1210					1215				
Glu	Leu	Lys	Lys	Ile	Ile	Gly	Gln	Val	Arg	Asp	Gln	Ala	Glu	His	

# ES 2 898 271 T3

1220	1225	1230
Leu Lys Thr Ala Val Gln Met Ala Val Phe Ile His Asn Phe Lys 1235 1240 1245		
Arg Lys Gly Gly Ile Gly Gly Tyr Ser Ala Gly Glu Arg Ile Ile 1250 1255 1260		
Asp Ile Ile Ala Thr Asp Ile Gln Thr Lys Glu Leu Gln Lys Gln 1265 1270 1275		
Ile Ile Lys Ile Gln Asn Phe Arg Val Tyr Tyr Arg Asp Ser Arg 1280 1285 1290		
Asp Pro Ile Trp Lys Gly Pro Ala Lys Leu Leu Trp Lys Gly Glu 1295 1300 1305		
Gly Ala Val Val Ile Gln Asp Asn Ser Asp Ile Lys Val Val Pro 1310 1315 1320		
Arg Arg Lys Val Lys Ile Ile Lys Asp Tyr Gly Lys Gln Met Ala 1325 1330 1335		
Gly Ala Asp Cys Val Ala Gly Arg Gln Asp Glu Asp 1340 1345 1350		

<210> 29

<211> 1341

<212> PRT

<213> Secuencia artificial

<220>

<223> secuencia del antígeno mos2GagPol mosaico

<400> 29

Met Gly Ala Arg Ala Ser Ile Leu Arg Gly Gly Lys Leu Asp Lys Trp 1 5 10 15
Glu Lys Ile Arg Leu Arg Pro Gly Gly Lys Lys His Tyr Met Leu Lys 20 25 30
His Leu Val Trp Ala Ser Arg Glu Leu Glu Arg Phe Ala Leu Asn Pro 35 40 45
Gly Leu Leu Glu Thr Ser Glu Gly Cys Lys Gln Ile Ile Lys Gln Leu 50 55 60
Gln Pro Ala Leu Gln Thr Gly Thr Glu Glu Leu Arg Ser Leu Phe Asn 65 70 75 80

Thr	Val	Ala	Thr	Leu	Tyr	Cys	Val	His	Ala	Glu	Ile	Glu	Val	Arg	Asp	85	90	95
Thr	Lys	Glu	Ala	Leu	Asp	Lys	Ile	Glu	Glu	Glu	Gln	Asn	Lys	Ser	Gln	100	105	110
Gln	Lys	Thr	Gln	Gln	Ala	Lys	Glu	Ala	Asp	Gly	Lys	Val	Ser	Gln	Asn	115	120	125
Tyr	Pro	Ile	Val	Gln	Asn	Leu	Gln	Gly	Gln	Met	Val	His	Gln	Pro	Ile	130	135	140
Ser	Pro	Arg	Thr	Leu	Asn	Ala	Trp	Val	Lys	Val	Ile	Glu	Glu	Lys	Ala	145	150	155
Phe	Ser	Pro	Glu	Val	Ile	Pro	Met	Phe	Thr	Ala	Leu	Ser	Glu	Gly	Ala	165	170	175
Thr	Pro	Gln	Asp	Leu	Asn	Thr	Met	Leu	Asn	Thr	Val	Gly	Gly	His	Gln	180	185	190
Ala	Ala	Met	Gln	Met	Leu	Lys	Asp	Thr	Ile	Asn	Glu	Glu	Ala	Ala	Glu	195	200	205
Trp	Asp	Arg	Leu	His	Pro	Val	His	Ala	Gly	Pro	Val	Ala	Pro	Gly	Gln	210	215	220
Met	Arg	Glu	Pro	Arg	Gly	Ser	Asp	Ile	Ala	Gly	Thr	Thr	Ser	Asn	Leu	225	230	235
Gln	Glu	Gln	Ile	Ala	Trp	Met	Thr	Ser	Asn	Pro	Pro	Ile	Pro	Val	Gly	245	250	255
Asp	Ile	Tyr	Lys	Arg	Trp	Ile	Ile	Leu	Gly	Leu	Asn	Lys	Ile	Val	Arg	260	265	270
Met	Tyr	Ser	Pro	Thr	Ser	Ile	Leu	Asp	Ile	Lys	Gln	Gly	Pro	Lys	Glu	275	280	285
Pro	Phe	Arg	Asp	Tyr	Val	Asp	Arg	Phe	Phe	Lys	Thr	Leu	Arg	Ala	Glu	290	295	300
Gln	Ala	Thr	Gln	Asp	Val	Lys	Asn	Trp	Met	Thr	Asp	Thr	Leu	Leu	Val	305	310	315
Gln	Asn	Ala	Asn	Pro	Asp	Cys	Lys	Thr	Ile	Leu	Arg	Ala	Leu	Gly	Pro			

				325						330					335				
Gly	Ala	Thr	Leu	Glu	Glu	Met	Met	Thr	Ala	Cys	Gln	Gly	Val	Gly	Gly				
			340					345					350						
Pro	Ser	His	Lys	Ala	Arg	Val	Leu	Ala	Glu	Ala	Met	Ser	Gln	Thr	Asn				
		355					360					365							
Ser	Thr	Ile	Leu	Met	Gln	Arg	Ser	Asn	Phe	Lys	Gly	Ser	Lys	Arg	Ile				
	370					375					380								
Val	Lys	Cys	Phe	Asn	Cys	Gly	Lys	Glu	Gly	His	Ile	Ala	Arg	Asn	Cys				
385					390					395					400				
Arg	Ala	Pro	Arg	Lys	Lys	Gly	Cys	Trp	Lys	Cys	Gly	Lys	Glu	Gly	His				
				405					410						415				
Gln	Met	Lys	Asp	Cys	Thr	Glu	Arg	Gln	Ala	Asn	Phe	Leu	Gly	Lys	Ile				
			420					425					430						
Trp	Pro	Ser	His	Lys	Gly	Arg	Pro	Gly	Asn	Phe	Leu	Gln	Ser	Arg	Pro				
		435					440					445							
Glu	Pro	Thr	Ala	Pro	Pro	Ala	Glu	Ser	Phe	Arg	Phe	Glu	Glu	Thr	Thr				
	450					455					460								
Pro	Ala	Pro	Lys	Gln	Glu	Pro	Lys	Asp	Arg	Glu	Pro	Leu	Thr	Ser	Leu				
465					470				475						480				
Arg	Ser	Leu	Phe	Gly	Ser	Asp	Pro	Leu	Ser	Gln	Met	Ala	Pro	Ile	Ser				
			485						490					495					
Pro	Ile	Glu	Thr	Val	Pro	Val	Lys	Leu	Lys	Pro	Gly	Met	Asp	Gly	Pro				
			500					505					510						
Lys	Val	Lys	Gln	Trp	Pro	Leu	Thr	Glu	Glu	Lys	Ile	Lys	Ala	Leu	Val				
		515					520					525							
Glu	Ile	Cys	Thr	Glu	Met	Glu	Lys	Glu	Gly	Lys	Ile	Ser	Lys	Ile	Gly				
	530					535					540								
Pro	Glu	Asn	Pro	Tyr	Asn	Thr	Pro	Ile	Phe	Ala	Ile	Lys	Lys	Lys	Asp				
545					550					555					560				
Ser	Thr	Lys	Trp	Arg	Lys	Leu	Val	Asp	Phe	Arg	Glu	Leu	Asn	Lys	Arg				
				565					570					575					

Thr	Gln	Asp	Phe	Trp	Glu	Val	Gln	Leu	Gly	Ile	Pro	His	Pro	Ala	Gly	
			580					585					590			
Leu	Lys	Lys	Lys	Lys	Ser	Val	Thr	Val	Leu	Ala	Val	Gly	Asp	Ala	Tyr	
		595					600					605				
Phe	Ser	Val	Pro	Leu	Asp	Glu	Asp	Phe	Arg	Lys	Tyr	Thr	Ala	Phe	Thr	
	610					615					620					
Ile	Pro	Ser	Ile	Asn	Asn	Glu	Thr	Pro	Gly	Ile	Arg	Tyr	Gln	Tyr	Asn	
625					630					635					640	
Val	Leu	Pro	Gln	Gly	Trp	Lys	Gly	Ser	Pro	Ala	Ile	Phe	Gln	Ser	Ser	
				645					650					655		
Met	Thr	Lys	Ile	Leu	Glu	Pro	Phe	Arg	Lys	Gln	Asn	Pro	Asp	Ile	Val	
			660					665					670			
Ile	Tyr	Gln	Tyr	Met	Ala	Ala	Leu	Tyr	Val	Gly	Ser	Asp	Leu	Glu	Ile	
		675					680					685				
Gly	Gln	His	Arg	Thr	Lys	Ile	Glu	Glu	Leu	Arg	Gln	His	Leu	Leu	Arg	
	690					695					700					
Trp	Gly	Phe	Thr	Thr	Pro	Asp	Lys	Lys	His	Gln	Lys	Glu	Pro	Pro	Phe	
705					710					715					720	
Leu	Trp	Met	Gly	Tyr	Glu	Leu	His	Pro	Asp	Lys	Trp	Thr	Val	Gln	Pro	
				725					730					735		
Ile	Val	Leu	Pro	Glu	Lys	Asp	Ser	Trp	Thr	Val	Asn	Asp	Ile	Gln	Lys	
			740					745					750			
Leu	Val	Gly	Lys	Leu	Asn	Trp	Ala	Ser	Gln	Ile	Tyr	Ala	Gly	Ile	Lys	
		755					760					765				
Val	Lys	Gln	Leu	Cys	Lys	Leu	Leu	Arg	Gly	Thr	Lys	Ala	Leu	Thr	Glu	
	770					775					780					
Val	Val	Pro	Leu	Thr	Glu	Glu	Ala	Glu	Leu	Glu	Leu	Ala	Glu	Asn	Arg	
785					790					795					800	
Glu	Ile	Leu	Lys	Glu	Pro	Val	His	Gly	Val	Tyr	Tyr	Asp	Pro	Ser	Lys	
				805					810					815		
Asp	Leu	Ile	Ala	Glu	Ile	Gln	Lys	Gln	Gly	Gln	Gly	Gln	Trp	Thr	Tyr	
			820					825					830			

Gln	Ile	Tyr	Gln	Glu	Pro	Phe	Lys	Asn	Leu	Lys	Thr	Gly	Lys	Tyr	Ala	835	840	845
Arg	Met	Arg	Gly	Ala	His	Thr	Asn	Asp	Val	Lys	Gln	Leu	Thr	Glu	Ala	850	855	860
Val	Gln	Lys	Ile	Ala	Thr	Glu	Ser	Ile	Val	Ile	Trp	Gly	Lys	Thr	Pro	865	870	875
Lys	Phe	Lys	Leu	Pro	Ile	Gln	Lys	Glu	Thr	Trp	Glu	Ala	Trp	Trp	Thr	885	890	895
Glu	Tyr	Trp	Gln	Ala	Thr	Trp	Ile	Pro	Glu	Trp	Glu	Phe	Val	Asn	Thr	900	905	910
Pro	Pro	Leu	Val	Lys	Leu	Trp	Tyr	Gln	Leu	Glu	Lys	Glu	Pro	Ile	Val	915	920	925
Gly	Ala	Glu	Thr	Phe	Tyr	Val	Ala	Gly	Ala	Ala	Asn	Arg	Glu	Thr	Lys	930	935	940
Leu	Gly	Lys	Ala	Gly	Tyr	Val	Thr	Asp	Arg	Gly	Arg	Gln	Lys	Val	Val	945	950	955
Ser	Leu	Thr	Asp	Thr	Thr	Asn	Gln	Lys	Thr	Ala	Leu	Gln	Ala	Ile	His	965	970	975
Leu	Ala	Leu	Gln	Asp	Ser	Gly	Leu	Glu	Val	Asn	Ile	Val	Thr	Ala	Ser	980	985	990
Gln	Tyr	Ala	Leu	Gly	Ile	Ile	Gln	Ala	Gln	Pro	Asp	Lys	Ser	Glu	Ser	995	1000	1005
Glu	Leu	Val	Ser	Gln	Ile	Ile	Glu	Gln	Leu	Ile	Lys	Lys	Glu	Lys		1010	1015	1020
Val	Tyr	Leu	Ala	Trp	Val	Pro	Ala	His	Lys	Gly	Ile	Gly	Gly	Asn		1025	1030	1035
Glu	Gln	Val	Asp	Lys	Leu	Val	Ser	Arg	Gly	Ile	Arg	Lys	Val	Leu		1040	1045	1050
Phe	Leu	Asp	Gly	Ile	Asp	Lys	Ala	Gln	Glu	Glu	His	Glu	Lys	Tyr		1055	1060	1065
His	Ser	Asn	Trp	Arg	Ala	Met	Ala	Ser	Glu	Phe	Asn	Leu	Pro	Pro		1070	1075	1080



# ES 2 898 271 T3

Ile Val	Ala Lys	Glu Ile	Val Val	Ala Ser	Cys Asp	Lys Cys	Gln Leu
1085			1090			1095	
Lys Gly	Glu Ala	Ile His	Gly Gln	Val Asp	Cys Ser	Pro Gly	Ile
1100			1105			1110	
Trp Gln	Leu Ala	Cys Thr	His Leu	Glu Gly	Lys Val	Ile Leu	Val
1115			1120			1125	
Ala Val	His Val	Ala Ser	Gly Tyr	Ile Glu	Ala Glu	Val Ile	Pro
1130			1135			1140	
Ala Glu	Thr Gly	Gln Glu	Thr Ala	Tyr Phe	Leu Leu	Lys Leu	Ala
1145			1150			1155	
Gly Arg	Trp Pro	Val Lys	Thr Ile	His Thr	Ala Asn	Gly Ser	Asn
1160			1165			1170	
Phe Thr	Ser Ala	Thr Val	Lys Ala	Ala Cys	Trp Trp	Ala Gly	Ile
1175			1180			1185	
Lys Gln	Glu Phe	Gly Ile	Pro Tyr	Asn Pro	Gln Ser	Gln Gly	Val
1190			1195			1200	
Val Ala	Ser Ile	Asn Lys	Glu Leu	Lys Lys	Ile Ile	Gly Gln	Val
1205			1210			1215	
Arg Asp	Gln Ala	Glu His	Leu Lys	Thr Ala	Val Gln	Met Ala	Val
1220			1225			1230	
Phe Ile	His Asn	Phe Lys	Arg Lys	Gly Gly	Ile Gly	Glu Tyr	Ser
1235			1240			1245	
Ala Gly	Glu Arg	Ile Val	Asp Ile	Ile Ala	Ser Asp	Ile Gln	Thr
1250			1255			1260	
Lys Glu	Leu Gln	Lys Gln	Ile Thr	Lys Ile	Gln Asn	Phe Arg	Val
1265			1270			1275	
Tyr Tyr	Arg Asp	Ser Arg	Asp Pro	Leu Trp	Lys Gly	Pro Ala	Lys
1280			1285			1290	
Leu Leu	Trp Lys	Gly Glu	Gly Ala	Val Val	Ile Gln	Asp Asn	Ser
1295			1300			1305	
Asp Ile	Lys Val	Val Pro	Arg Arg	Lys Ala	Lys Ile	Ile Arg	Asp

# ES 2 898 271 T3

1310

1315

1320

Tyr Gly Lys Gln Met Ala Gly Asp Asp Cys Val Ala Ser Arg Gln  
1325 1330 1335

Asp Glu Asp  
1340

<210> 30  
<211> 709  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> secuencia de sC1

<400> 30

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp  
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Cys Arg Val Met Gly Asn  
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys  
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val  
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro  
65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys  
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Arg Leu Trp Asp  
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu  
115 120 125

Glu Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile  
130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala  
145 150 155 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr  
165 170 175

# ES 2 898 271 T3

Arg Leu Asp Ile Val Pro Leu Asp Glu Asn Asn Ser Ser Glu Lys Ser  
 180 185 190  
 Ser Glu Asn Ser Ser Glu Tyr Tyr Arg Leu Ile Asn Cys Asn Thr Ser  
 195 200 205  
 Ala Ile Thr Gln Ala Cys Pro Lys Val Ser Phe Asp Pro Ile Pro Ile  
 210 215 220  
 His Tyr Cys Ala Pro Ala Gly Tyr Ala Ile Leu Lys Cys Asn Asn Lys  
 225 230 235 240  
 Thr Phe Asn Gly Thr Gly Pro Cys Asn Asn Val Ser Thr Val Gln Cys  
 245 250 255  
 Thr His Gly Ile Lys Pro Val Val Ser Thr Gln Leu Leu Leu Asn Gly  
 260 265 270  
 Ser Leu Ala Glu Glu Glu Ile Ile Ile Arg Ser Glu Asn Leu Thr Asn  
 275 280 285  
 Asn Ala Lys Thr Ile Ile Val His Leu Asn Glu Thr Val Asn Ile Thr  
 290 295 300  
 Cys Thr Arg Pro Asn Asn Asn Thr Arg Lys Ser Ile Arg Ile Gly Pro  
 305 310 315 320  
 Gly Gln Thr Phe Tyr Ala Thr Gly Asp Ile Ile Gly Asp Ile Arg Gln  
 325 330 335  
 Ala His Cys Asn Leu Ser Arg Asp Gly Trp Asn Lys Thr Leu Gln Gly  
 340 345 350  
 Val Lys Lys Lys Leu Ala Glu His Phe Pro Asn Lys Thr Ile Asn Phe  
 355 360 365  
 Thr Ser Ser Ser Gly Gly Asp Leu Glu Ile Thr Thr His Ser Phe Asn  
 370 375 380  
 Cys Arg Gly Glu Phe Phe Tyr Cys Asn Thr Ser Gly Leu Phe Asn Gly  
 385 390 395 400  
 Thr Tyr Met Pro Asn Gly Thr Asn Ser Asn Ser Ser Ser Asn Ile Thr  
 405 410 415  
 Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly

420							425							430						
Arg	Ala	Met	Tyr	Ala	Pro	Pro	Ile	Ala	Gly	Asn	Ile	Thr	Cys	Arg	Ser					
		435					440					445								
Asn	Ile	Thr	Gly	Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly	Ser	Asn	Asn	Gly					
	450					455					460									
Val	Pro	Asn	Asp	Thr	Glu	Thr	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg					
465					470					475					480					
Asn	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Glu	Val	Lys					
				485					490					495						
Pro	Leu	Gly	Val	Ala	Pro	Thr	Glu	Ala	Lys	Arg	Arg	Val	Val	Glu	Arg					
			500					505					510							
Glu	Glu	Arg	Ala	Val	Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Ile	Leu	Gly					
		515					520					525								
Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Ile	Thr	Leu	Thr	Val	Gln					
	530					535					540									
Ala	Arg	Gln	Leu	Leu	Ser	Gly	Ile	Val	Gln	Gln	Gln	Ser	Asn	Leu	Leu					
545					550					555					560					
Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Met	Leu	Gln	Leu	Thr	Val	Trp	Gly					
				565					570					575						
Ile	Lys	Gln	Leu	Gln	Thr	Arg	Val	Leu	Ala	Ile	Glu	Arg	Tyr	Leu	Gln					
		580						585					590							
Asp	Gln	Gln	Leu	Leu	Gly	Leu	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	Cys					
		595					600					605								
Thr	Thr	Ala	Val	Pro	Trp	Asn	Thr	Ser	Trp	Ser	Asn	Lys	Ser	Gln	Thr					
	610					615					620									
Asp	Ile	Trp	Asp	Asn	Met	Thr	Trp	Met	Gln	Trp	Asp	Lys	Glu	Ile	Gly					
625					630					635					640					
Asn	Tyr	Thr	Gly	Glu	Ile	Tyr	Arg	Leu	Leu	Glu	Glu	Ser	Gln	Asn	Gln					
				645					650					655						
Gln	Glu	Lys	Met	Lys	Gln	Ile	Glu	Asp	Lys	Ile	Glu	Glu	Ile	Leu	Ser					
			660					665					670							

Lys Ile Tyr His Ile Glu Asn Glu Ile Ala Arg Ile Lys Lys Leu Ile  
675 680 685

Gly Glu Val Gly Ser Gly Ala Pro Thr Lys Ala Lys Arg Arg Val Val  
690 695 700

Gln Arg Glu Lys Arg  
705

<210> 31  
<211> 867  
<212> PRT  
<213> Secuencia artificial

<220>  
<223> secuencia de C1

<400> 31

Met Arg Val Arg Gly Ile Gln Arg Asn Trp Pro Gln Trp Trp Ile Trp  
1 5 10 15

Gly Ile Leu Gly Phe Trp Met Ile Ile Ile Cys Arg Val Met Gly Asn  
20 25 30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Lys  
35 40 45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Glu Lys Glu Val  
50 55 60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro  
65 70 75 80

Gln Glu Met Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys  
85 90 95

Asn Asp Met Val Asp Gln Met His Glu Asp Ile Ile Arg Leu Trp Asp  
100 105 110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu  
115 120 125

Glu Cys Arg Asn Val Arg Asn Val Ser Ser Asn Gly Thr Tyr Asn Ile  
130 135 140

Ile His Asn Glu Thr Tyr Lys Glu Met Lys Asn Cys Ser Phe Asn Ala  
145 150 155 160

Thr Thr Val Val Glu Asp Arg Lys Gln Lys Val His Ala Leu Phe Tyr

165										170					175				
Arg	Leu	Asp	Ile	Val	Pro	Leu	Asp	Glu	Asn	Asn	Ser	Ser	Glu	Lys	Ser				
			180					185					190						
Ser	Glu	Asn	Ser	Ser	Glu	Tyr	Tyr	Arg	Leu	Ile	Asn	Cys	Asn	Thr	Ser				
		195					200					205							
Ala	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe	Asp	Pro	Ile	Pro	Ile				
	210					215					220								
His	Tyr	Cys	Ala	Pro	Ala	Gly	Tyr	Ala	Ile	Leu	Lys	Cys	Asn	Asn	Lys				
225					230					235					240				
Thr	Phe	Asn	Gly	Thr	Gly	Pro	Cys	Asn	Asn	Val	Ser	Thr	Val	Gln	Cys				
				245					250					255					
Thr	His	Gly	Ile	Lys	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly				
			260					265					270						
Ser	Leu	Ala	Glu	Glu	Glu	Ile	Ile	Ile	Arg	Ser	Glu	Asn	Leu	Thr	Asn				
		275					280					285							
Asn	Ala	Lys	Thr	Ile	Ile	Val	His	Leu	Asn	Glu	Thr	Val	Asn	Ile	Thr				
	290					295					300								
Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	Arg	Ile	Gly	Pro				
305					310					315				320					
Gly	Gln	Thr	Phe	Tyr	Ala	Thr	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln				
				325					330					335					
Ala	His	Cys	Asn	Leu	Ser	Arg	Asp	Gly	Trp	Asn	Lys	Thr	Leu	Gln	Gly				
			340					345					350						
Val	Lys	Lys	Lys	Leu	Ala	Glu	His	Phe	Pro	Asn	Lys	Thr	Ile	Asn	Phe				
		355					360					365							
Thr	Ser	Ser	Ser	Gly	Gly	Asp	Leu	Glu	Ile	Thr	Thr	His	Ser	Phe	Asn				
	370					375					380								
Cys	Arg	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser	Gly	Leu	Phe	Asn	Gly				
385					390					395					400				
Thr	Tyr	Met	Pro	Asn	Gly	Thr	Asn	Ser	Asn	Ser	Ser	Ser	Asn	Ile	Thr				
				405					410					415					

Leu Pro Cys Arg Ile Lys Gln Ile Ile Asn Met Trp Gln Glu Val Gly  
 420 425 430

Arg Ala Met Tyr Ala Pro Pro Ile Ala Gly Asn Ile Thr Cys Arg Ser  
 435 440 445

Asn Ile Thr Gly Leu Leu Leu Thr Arg Asp Gly Gly Ser Asn Asn Gly  
 450 455 460

Val Pro Asn Asp Thr Glu Thr Phe Arg Pro Gly Gly Gly Asp Met Arg  
 465 470 475 480

Asn Asn Trp Arg Ser Glu Leu Tyr Lys Tyr Lys Val Val Glu Val Lys  
 485 490 495

Pro Leu Gly Val Ala Pro Thr Glu Ala Lys Arg Arg Val Val Glu Arg  
 500 505 510

Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly  
 515 520 525

Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln  
 530 535 540

Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu  
 545 550 555 560

Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly  
 565 570 575

Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln  
 580 585 590

Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys  
 595 600 605

Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr  
 610 615 620

Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly  
 625 630 635 640

Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln  
 645 650 655

Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn  
 660 665 670

# ES 2 898 271 T3

Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile  
 675 680 685  
 Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala  
 690 695 700  
 Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr Ser Pro Leu Ser  
 705 710 715 720  
 Leu Gln Thr Leu Thr Gln Asn Pro Gly Gly Leu Asp Arg Leu Gly Arg  
 725 730 735  
 Ile Glu Glu Glu Gly Gly Glu Gln Asp Lys Asp Arg Ser Ile Arg Leu  
 740 745 750  
 Val Asn Gly Phe Phe Ala Leu Phe Trp Asp Asp Leu Arg Ser Leu Cys  
 755 760 765  
 Leu Phe Ser Tyr His Arg Leu Arg Asp Phe Ile Leu Ile Val Ala Arg  
 770 775 780  
 Ala Val Glu Leu Leu Gly Arg Ser Ser Leu Arg Gly Leu Gln Arg Gly  
 785 790 795 800  
 Trp Glu Ile Leu Lys Tyr Leu Gly Ser Leu Leu Gln Tyr Trp Gly Leu  
 805 810 815  
 Glu Leu Lys Lys Ser Ala Ile Asn Leu Leu Asp Thr Ile Ala Ile Ala  
 820 825 830  
 Val Ala Glu Gly Thr Asp Arg Ile Ile Glu Leu Ile Gln Arg Ile Cys  
 835 840 845  
 Arg Ala Ile Cys Asn Ile Pro Arg Arg Ile Arg Gln Gly Phe Glu Ala  
 850 855 860  
 Ala Leu Gln  
 865

<210> 32  
 <211> 716  
 <212> PRT  
 <213> Secuencia artificial

<220>  
 <223> secuencia de C1D7

<400> 32



# ES 2 898 271 T3

Met	Arg	Val	Arg	Gly	Ile	Gln	Arg	Asn	Trp	Pro	Gln	Trp	Trp	Ile	Trp	1	5	10	15
Gly	Ile	Leu	Gly	Phe	Trp	Met	Ile	Ile	Ile	Cys	Arg	Val	Met	Gly	Asn	20	25	30	
Leu	Trp	Val	Thr	Val	Tyr	Tyr	Gly	Val	Pro	Val	Trp	Lys	Glu	Ala	Lys	35	40	45	
Thr	Thr	Leu	Phe	Cys	Ala	Ser	Asp	Ala	Lys	Ala	Tyr	Glu	Lys	Glu	Val	50	55	60	
His	Asn	Val	Trp	Ala	Thr	His	Ala	Cys	Val	Pro	Thr	Asp	Pro	Asn	Pro	65	70	75	80
Gln	Glu	Met	Val	Leu	Glu	Asn	Val	Thr	Glu	Asn	Phe	Asn	Met	Trp	Lys	85	90	95	
Asn	Asp	Met	Val	Asp	Gln	Met	His	Glu	Asp	Ile	Ile	Arg	Leu	Trp	Asp	100	105	110	
Gln	Ser	Leu	Lys	Pro	Cys	Val	Lys	Leu	Thr	Pro	Leu	Cys	Val	Thr	Leu	115	120	125	
Glu	Cys	Arg	Asn	Val	Arg	Asn	Val	Ser	Ser	Asn	Gly	Thr	Tyr	Asn	Ile	130	135	140	
Ile	His	Asn	Glu	Thr	Tyr	Lys	Glu	Met	Lys	Asn	Cys	Ser	Phe	Asn	Ala	145	150	155	160
Thr	Thr	Val	Val	Glu	Asp	Arg	Lys	Gln	Lys	Val	His	Ala	Leu	Phe	Tyr	165	170	175	
Arg	Leu	Asp	Ile	Val	Pro	Leu	Asp	Glu	Asn	Asn	Ser	Ser	Glu	Lys	Ser	180	185	190	
Ser	Glu	Asn	Ser	Ser	Glu	Tyr	Tyr	Arg	Leu	Ile	Asn	Cys	Asn	Thr	Ser	195	200	205	
Ala	Ile	Thr	Gln	Ala	Cys	Pro	Lys	Val	Ser	Phe	Asp	Pro	Ile	Pro	Ile	210	215	220	
His	Tyr	Cys	Ala	Pro	Ala	Gly	Tyr	Ala	Ile	Leu	Lys	Cys	Asn	Asn	Lys	225	230	235	240
Thr	Phe	Asn	Gly	Thr	Gly	Pro	Cys	Asn	Asn	Val	Ser	Thr	Val	Gln	Cys	245	250	255	

Thr	His	Gly	Ile	Lys	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	260	265	270
Ser	Leu	Ala	Glu	Glu	Glu	Ile	Ile	Ile	Arg	Ser	Glu	Asn	Leu	Thr	Asn	275	280	285
Asn	Ala	Lys	Thr	Ile	Ile	Val	His	Leu	Asn	Glu	Thr	Val	Asn	Ile	Thr	290	295	300
Cys	Thr	Arg	Pro	Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	Arg	Ile	Gly	Pro	305	310	315
Gly	Gln	Thr	Phe	Tyr	Ala	Thr	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln	325	330	335
Ala	His	Cys	Asn	Leu	Ser	Arg	Asp	Gly	Trp	Asn	Lys	Thr	Leu	Gln	Gly	340	345	350
Val	Lys	Lys	Lys	Leu	Ala	Glu	His	Phe	Pro	Asn	Lys	Thr	Ile	Asn	Phe	355	360	365
Thr	Ser	Ser	Ser	Gly	Gly	Asp	Leu	Glu	Ile	Thr	Thr	His	Ser	Phe	Asn	370	375	380
Cys	Arg	Gly	Glu	Phe	Phe	Tyr	Cys	Asn	Thr	Ser	Gly	Leu	Phe	Asn	Gly	385	390	395
Thr	Tyr	Met	Pro	Asn	Gly	Thr	Asn	Ser	Asn	Ser	Ser	Ser	Asn	Ile	Thr	405	410	415
Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp	Gln	Glu	Val	Gly	420	425	430
Arg	Ala	Met	Tyr	Ala	Pro	Pro	Ile	Ala	Gly	Asn	Ile	Thr	Cys	Arg	Ser	435	440	445
Asn	Ile	Thr	Gly	Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly	Ser	Asn	Asn	Gly	450	455	460
Val	Pro	Asn	Asp	Thr	Glu	Thr	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	Arg	465	470	475
Asn	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Glu	Val	Lys	485	490	495
Pro	Leu	Gly	Val	Ala	Pro	Thr	Glu	Ala	Lys	Arg	Arg	Val	Val	Glu	Arg	500	505	510

# ES 2 898 271 T3

Glu Lys Arg Ala Val Gly Ile Gly Ala Val Phe Leu Gly Ile Leu Gly  
 515 520 525  
 Ala Ala Gly Ser Thr Met Gly Ala Ala Ser Ile Thr Leu Thr Val Gln  
 530 535 540  
 Ala Arg Gln Leu Leu Ser Gly Ile Val Gln Gln Gln Ser Asn Leu Leu  
 545 550 555 560  
 Arg Ala Ile Glu Ala Gln Gln His Met Leu Gln Leu Thr Val Trp Gly  
 565 570 575  
 Ile Lys Gln Leu Gln Thr Arg Val Leu Ala Ile Glu Arg Tyr Leu Gln  
 580 585 590  
 Asp Gln Gln Leu Leu Gly Leu Trp Gly Cys Ser Gly Lys Leu Ile Cys  
 595 600 605  
 Thr Thr Ala Val Pro Trp Asn Thr Ser Trp Ser Asn Lys Ser Gln Thr  
 610 615 620  
 Asp Ile Trp Asp Asn Met Thr Trp Met Gln Trp Asp Lys Glu Ile Gly  
 625 630 635 640  
 Asn Tyr Thr Gly Glu Ile Tyr Arg Leu Leu Glu Glu Ser Gln Asn Gln  
 645 650 655  
 Gln Glu Lys Asn Glu Lys Asp Leu Leu Ala Leu Asp Ser Trp Lys Asn  
 660 665 670  
 Leu Trp Asn Trp Phe Asp Ile Thr Asn Trp Leu Trp Tyr Ile Lys Ile  
 675 680 685  
 Phe Ile Met Ile Val Gly Gly Leu Ile Gly Leu Arg Ile Ile Phe Ala  
 690 695 700  
 Val Leu Ser Ile Val Asn Arg Val Arg Gln Gly Tyr  
 705 710 715

<210> 33

<211> 2127

<212> ADN

<213> Secuencia artificial

<220>

<223> secuencia de nucleótidos que codifica sC1

<400> 33

# ES 2 898 271 T3

atgagagtgc	ggggcattca	gagaaactgg	ccccagtggt	ggatctgggg	catcctgggc	60
ttttggatga	tcattatctg	ccgcgtgatg	ggcaacctgt	gggtcacctg	gtactacggc	120
gtgcccgtgt	ggaagagggc	caagaccacc	ctgttctgcg	ccagcgacgc	caaggcctac	180
gagaaagagg	tgcacaacgt	ctgggccacc	cacgcctgtg	tgccaccga	ccccaatccc	240
caggaaatgg	tcctggaaaa	cgtgaccgag	aacttcaaca	tgtggaagaa	cgacatgggtg	300
gaccagatgc	acgaggacat	catccggctg	tgggaccaga	gcctgaagcc	ctgcgtgaag	360
ctgacccttc	tgtgcgtgac	cctggaatgc	cggaacgtgc	gcaacgtgtc	cagcaacggc	420
acctacaata	tcattccaaa	cgagacatac	aaagagatga	agaactgcag	cttcaacgct	480
accaccgtgg	tcgaggaccg	gaagcagaag	gtgcacgccc	tgttttaccg	gctggacatc	540
gtgcccctgg	acgagaacaa	cagcagcgag	aagtcctccg	agaacagctc	cgagtactac	600
agactgatca	actgcaacac	cagcgccatc	acccaggcct	gccccaaagg	gtccttcgac	660
cctatcccca	tccactactg	cgcccctgcc	ggctacgcca	tcctgaagtg	caacaacaag	720
accttcaatg	gcaccggccc	ctgcaacaat	gtgtccaccg	tgcaagtgcac	ccacggcatc	780
aagcccgtgg	tgtctaccca	gctgctgctg	aacggcagcc	tggccgagga	agagatcatc	840
atcagaagcg	agaacctgac	caacaacgcc	aagacaatca	tcgtccacct	gaacgaaacc	900
gtgaacatca	cctgtaccgg	gcctaacaac	aacaccggga	agtccatccg	gatcggccct	960
ggccgagacct	tttacgccac	cgcgcatatt	atcggcgaca	tccggcaggc	ccactgcaat	1020
ctgagccggg	acggctggaa	caagacactg	cagggcgtca	agaagaagct	ggccgaacac	1080
ttccccaaca	aaaccatcaa	cttcaccagc	tcctctggcg	gcgacctgga	aatcaccacc	1140
cacagcttta	actgcagagg	cgagttcttc	tactgcaata	cctccggcct	gttcaatgga	1200
acctacatgc	ccaacggggac	caacagcaac	tccagcagca	atatcaccct	gccttgccgg	1260
atcaagcaga	ttatcaatat	gtggcaggaa	gtgggcagag	ctatgtacgc	ccctccaatc	1320
gccggcaaca	tcacatgcag	aagcaacatt	accggcctgc	tgctcaccag	ggacggcggc	1380
tctaacaatg	gcgtgccaaa	cgacaccgag	acattcagac	ccggcggagg	cgacatgcgg	1440
aacaattggc	ggagcgagct	gtacaagtac	aagtggtggg	aagtgaagcc	cctgggcgtg	1500
gccctaccg	aagccaagag	aagagtggtc	gaacgcgagg	aacgggccgt	gggcattgga	1560
gccgtgtttc	tgggaatcct	gggagccgct	ggcagcacca	tgggcgctgc	ctctatcaca	1620
ctgacagtgc	aggccagaca	gctcctgagc	ggcatcgtgc	agcagcagag	caacctgctg	1680
agagccatcg	aggcacagca	gcacatgctg	cagctgaccg	tgtggggcat	taagcagctc	1740
cagacacggg	tgttgcccat	tgagagatac	ctgcaggatc	agcagctgct	cggcctgtgg	1800
ggctgtagcg	gcaagctgat	ctgtaccacc	gccgtgcctt	ggaacacctc	ctggtccaac	1860
aagagccaga	ccgacatctg	ggacaacatg	acctggatgc	agtgggacaa	agaaatcggc	1920

# ES 2 898 271 T3

aactataccg gcgagatcta ccgactgctg gaagagtccc agaaccagca ggaaaagatg 1980  
aagcagatcg aggacaagat cgaagagatt ctgagcaaaa tctaccacat cgagaacgag 2040  
atcgcccgca tcaagaaact gatcggcgaa gtgggatccg gcgctccac aaaggccaaa 2100  
agacgggtgg tgcagcgca gaaacgc 2127

<210> 34  
<211> 2601  
<212> ADN  
<213> Secuencia artificial

<220>  
<223> secuencia de nucleótidos que codifica C1

<400> 34

atgagagtgc ggggcattca gagaaactgg cccagtggt ggatctgggg catcctgggc 60  
ttttggatga tcattatctg ccgcgtgatg ggcaacctgt gggtcaccgt gtactacggc 120  
gtgcccgtgt ggaaagaggc caagaccacc ctgttctgcg ccagcgacgc caaggcctac 180  
gagaaagagg tgcacaacgt ctggggccacc cagcctgtg tgcccaccga cccaatccc 240  
caggaaatgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg 300  
gaccagatgc acgaggacat catccggctg tgggaccaga gcctgaagcc ctgcgtgaag 360  
ctgaccctc tgtgctgac cctggaatgc cggaacgtgc gcaacgtgtc cagcaacggc 420  
acctacaata tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct 480  
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttttaccg gctggacatc 540  
gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac 600  
agactgatca actgcaacac cagcgccatc acccaggcct gcccgaagg gtccctcgac 660  
cctatcccca tcactactg cggccctgcc ggctacgcca tcctgaagtg caacaacaag 720  
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc 780  
aagcccgtgg tgtctacca gctgctgctg aacggcagcc tggccgagga agagatcatc 840  
atcagaagcg agaacctgac caacaacgcc aagacaatca tcgtccacct gaacgaaacc 900  
gtgaacatca cctgtacccg gcctaacaac aacacccgga agtccatccg gatcggccct 960  
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc ccaactgcaat 1020  
ctgagccggg acggctggaa caagacactg cagggcgtca agaagaagct ggccgaacac 1080  
ttccccaaca aaaccatcaa cttcaccagc tcctctggcg gcgacctgga aatcaccacc 1140  
cacagcttta actgcagagg cgagttcttc tactgcaata cctccggcct gttcaatgga 1200  
acctacatgc ccaacgggac caacagcaac tccagcagca atatcaccct gccttgccgg 1260  
atcaagcaga ttatcaatat gtggcaggaa gtgggcagag ctatgtacgc ccctccaatc 1320

gcccggcaaca tcacatgcag aagcaacatt accggcctgc tgctcaccag ggacggcggc 1380  
tctaacaatg gcgtgccaaa cgacaccgag acattcagac ccggcgaggg cgacatgcgg 1440  
aacaattggc ggagcgagct gtacaagtac aaggtggtgg aagtgaagcc cctgggcgtg 1500  
gcccctaccg aagccaagag aagagtggtc gaacgcgaga agcggggccgt gggcattgga 1560  
gccgtgtttc tgggaatcct gggagccgct ggcagcacca tgggcgctgc ctctatcaca 1620  
ctgacagtgc aggccagaca gctcctgagc ggcacgtgc agcagcagag caacctgctg 1680  
agagccatcg aggcacagca gcacatgctg cagctgaccg tgtggggcat taagcagctc 1740  
cagacacggg tgctggccat tgagagatac ctgcaggatc agcagctgct cggcctgtgg 1800  
ggctgtagcg gcaagctgat ctgtaccacc gccgtgcctt ggaacacctc ctggtccaac 1860  
aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agaaatcggc 1920  
aactataccg gcgagatcta ccgactgctg gaagagtccc agaaccagca ggaaaagaac 1980  
gagaaggacc tgctggccct ggacagctgg aaaaatctgt ggaattgggt cgacatcacc 2040  
aactggctgt ggtacatcaa gatcttcac atgatcgtgg gcggcctgat cggcctgcgg 2100  
atcatctttg ccgtgctgag catcgtgaac cgcgtgcggc agggatacag ccctctgagc 2160  
ctgcagaccc tgaccagaa tccaggcggg ctggatcggc tgggcccgat tgaggaagaa 2220  
ggcggcgagc aggacaagga ccgcagcatc agactcgtga acggcttctt cgctctgttt 2280  
tgggacgacc tgcgagacct gtgcctgttc tcctaccaca gactgcggga ctttatcctg 2340  
attgtggcca gagccgtcga gctgctgggc agatcttctc tgagaggcct gcagcggggc 2400  
tgggagattc tgaagtacct gggctccctg ctgcagtatt ggggcctgga actgaagaag 2460  
tccgccatca atctgctcga cacaatcgct attgccgtgg ccgaaggcac cgacagaatc 2520  
atcgagctga tccagcggat ctgccgggac atctgcaaca tcccagacg gatcagacag 2580  
ggctttgaag ccgccctcca g 2601

<210> 35

<211> 2148

<212> ADN

<213> Secuencia artificial

<220>

<223> secuencia de nucleótidos que codifica C1D7

<400> 35

atgagagtgc ggggcattca gagaaactgg cccagtggt ggatctgggg catcctgggc 60  
ttttggatga tcattatctg ccgcgtgatg ggcaacctgt gggtcaccgt gtactacggc 120  
gtgcccgtgt ggaaagaggc caagaccacc ctgttctgcg ccagcgacgc caaggcctac 180  
gagaaagagg tgcacaacgt ctgggccacc cagcctgtg tgcccaccga cccaatccc 240  
caggaaatgg tcctggaaaa cgtgaccgag aacttcaaca tgtggaagaa cgacatggtg 300

gaccagatgc acgaggacat catccggctg tgggaccaga gcctgaagcc ctgcgtgaag 360  
ctgaccctc tgtgcgtgac cctggaatgc cggacgtgc gcaacgtgtc cagcaacggc 420  
acctacaata tcatccacaa cgagacatac aaagagatga agaactgcag cttcaacgct 480  
accaccgtgg tcgaggaccg gaagcagaag gtgcacgccc tgttttaccg gctggacatc 540  
gtgcccctgg acgagaacaa cagcagcgag aagtcctccg agaacagctc cgagtactac 600  
agactgatca actgcaacac cagcgccatc acccaggcct gcccgaaggt gtccttcgac 660  
cctatcccca tccactactg cgcccctgcc ggctacgcca tcctgaagtg caacaacaag 720  
accttcaatg gcaccggccc ctgcaacaat gtgtccaccg tgcagtgcac ccacggcatc 780  
aagcccgtgg tgtctacca gctgctgctg aacggcagcc tggccgagga agagatcatc 840  
atcagaagcg agaacctgac caacaacgcc aagacaatca tcgtccacct gaacgaaacc 900  
gtgaacatca cctgtaccog gcctaacaac aacacccgga agtccatccg gatcggccct 960  
ggccagacct tttagccac cggcgatatt atcggcgaca tccggcaggc cactgcaat 1020  
ctgagccggg acggctggaa caagacactg cagggcgta agaagaagct ggccgaacac 1080  
ttccccaaca aaaccatcaa cttaccagc tcctctggcg gcgacctgga aatcaccacc 1140  
cacagcttta actgcagagg cgagttcttc tactgcaata cctccggcct gttcaatgga 1200  
acctacatgc ccaacgggac caacagcaac tccagcagca atatcaccct gccttgccgg 1260  
atcaagcaga ttatcaatat gtggcaggaa gtgggcagag ctatgtacgc ccctccaatc 1320  
gccggcaaca tcacatgcag aagcaacatt accggcctgc tgctcaccag ggacggcggc 1380  
tctaacaatg gcgtgccaaa cgacaccgag acattcagac ccggcgagg cgacatgcgg 1440  
aacaattggc ggagcgagct gtacaagtac aaggtggtgg aagtgaagcc cctggcgctg 1500  
gcccctaccg aagccaagag aagagtggtc gaacgcgaga agcgggccgt gggcattgga 1560  
gccgtgtttc tgggaatcct gggagccgct ggcagcacca tggcgctgc ctctatcaca 1620  
ctgacagtgc aggccagaca gtcctgagc ggcacgtgc agcagcagag caacctgctg 1680  
agagccatcg aggcacagca gcacatgctg cagctgaccg tgtggggcat taagcagctc 1740  
cagacacggg tgctggccat tgagagatac ctgcaggatc agcagctgct cggcctgtgg 1800  
ggctgtagcg gcaagctgat ctgtaccacc gccgtgcctt ggaacacctc ctggtccaac 1860  
aagagccaga ccgacatctg ggacaacatg acctggatgc agtgggacaa agaaatcggc 1920  
aactataccg gcgagatcta ccgactgctg gaagagtccc agaaccagca ggaaaagaac 1980  
gagaaggacc tgctggccct ggacagctgg aaaaatctgt ggaattggtt cgacatcacc 2040  
aactggctgt ggtacatcaa gatcttcac atgatcgtgg gcggcctgat cggcctgcgg 2100  
atcatctttg ccgtgctgag catcgtgaac cgcgtgcggc agggctac 2148

<210> 36  
<211> 724

&lt;212&gt; PRT

&lt;213&gt; Secuencia artificial

&lt;220&gt;

5 &lt;223&gt; secuencia del trímero de Env mosaico

&lt;400&gt; 36

```

Met Arg Val Arg Gly Ile Gln Arg Asn Cys Gln His Leu Trp Arg Trp
 1              5              10              15

Gly Thr Leu Ile Leu Gly Met Leu Met Ile Cys Ser Ala Ala Gly Lys
          20              25              30

Leu Trp Val Thr Val Tyr Tyr Gly Val Pro Val Trp Lys Glu Ala Thr
          35              40              45

Thr Thr Leu Phe Cys Ala Ser Asp Ala Lys Ala Tyr Asp Thr Glu Val
 50              55              60

His Asn Val Trp Ala Thr His Ala Cys Val Pro Thr Asp Pro Asn Pro
65              70              75              80

Gln Glu Val Val Leu Glu Asn Val Thr Glu Asn Phe Asn Met Trp Lys
          85              90              95

Asn Asn Met Val Glu Gln Met His Glu Asp Ile Ile Ser Leu Trp Asp
          100              105              110

Gln Ser Leu Lys Pro Cys Val Lys Leu Thr Pro Leu Cys Val Thr Leu
          115              120              125

Asn Cys Thr Asp Asp Val Arg Asn Val Thr Asn Asn Ala Thr Asn Thr
130              135              140

Asn Ser Ser Trp Gly Glu Pro Met Glu Lys Gly Glu Ile Lys Asn Cys
145              150              155              160

Ser Phe Asn Ile Thr Thr Ser Ile Arg Asn Lys Val Gln Lys Gln Tyr
          165              170              175

Ala Leu Phe Tyr Lys Leu Asp Val Val Pro Ile Asp Asn Asp Ser Asn
          180              185              190

Asn Thr Asn Tyr Arg Leu Ile Ser Cys Asn Thr Ser Val Ile Thr Gln
195              200              205

```



# ES 2 898 271 T3

Ala	Cys	Pro	Lys	Val	Ser	Phe	Glu	Pro	Ile	Pro	Ile	His	Tyr	Cys	Ala			
210						215					220							
Pro	Ala	Gly	Phe	Ala	Ile	Leu	Lys	Cys	Asn	Asp	Lys	Lys	Phe	Asn	Gly			
225					230					235					240			
Thr	Gly	Pro	Cys	Thr	Asn	Val	Ser	Thr	Val	Gln	Cys	Thr	His	Gly	Ile			
				245					250					255				
Arg	Pro	Val	Val	Ser	Thr	Gln	Leu	Leu	Leu	Asn	Gly	Ser	Leu	Ala	Glu			
			260					265					270					
Glu	Glu	Val	Val	Ile	Arg	Ser	Glu	Asn	Phe	Thr	Asn	Asn	Ala	Lys	Thr			
		275					280					285						
Ile	Met	Val	Gln	Leu	Asn	Val	Ser	Val	Glu	Ile	Asn	Cys	Thr	Arg	Pro			
290						295					300							
Asn	Asn	Asn	Thr	Arg	Lys	Ser	Ile	His	Ile	Gly	Pro	Gly	Arg	Ala	Phe			
305					310					315					320			
Tyr	Thr	Ala	Gly	Asp	Ile	Ile	Gly	Asp	Ile	Arg	Gln	Ala	His	Cys	Asn			
				325					330					335				
Ile	Ser	Arg	Ala	Asn	Trp	Asn	Asn	Thr	Leu	Arg	Gln	Ile	Val	Glu	Lys			
			340					345					350					
Leu	Gly	Lys	Gln	Phe	Gly	Asn	Asn	Lys	Thr	Ile	Val	Phe	Asn	His	Ser			
		355					360					365						
Ser	Gly	Gly	Asp	Pro	Glu	Ile	Val	Met	His	Ser	Phe	Asn	Cys	Gly	Gly			
370						375					380							
Glu	Phe	Phe	Tyr	Cys	Asn	Ser	Thr	Lys	Leu	Phe	Asn	Ser	Thr	Trp	Thr			
385					390					395					400			
Trp	Asn	Asn	Ser	Thr	Trp	Asn	Asn	Thr	Lys	Arg	Ser	Asn	Asp	Thr	Glu			
				405					410					415				
Glu	His	Ile	Thr	Leu	Pro	Cys	Arg	Ile	Lys	Gln	Ile	Ile	Asn	Met	Trp			
			420					425					430					
Gln	Glu	Val	Gly	Lys	Ala	Met	Tyr	Ala	Pro	Pro	Ile	Arg	Gly	Gln	Ile			
		435					440					445						
Arg	Cys	Ser	Ser	Asn	Ile	Thr	Gly	Leu	Leu	Leu	Thr	Arg	Asp	Gly	Gly			
450						455					460							

# ES 2 898 271 T3

Asn	Asp	Thr	Ser	Gly	Thr	Glu	Ile	Phe	Arg	Pro	Gly	Gly	Gly	Asp	Met	465	470	475	480
Arg	Asp	Asn	Trp	Arg	Ser	Glu	Leu	Tyr	Lys	Tyr	Lys	Val	Val	Lys	Ile	485	490	495	
Glu	Pro	Leu	Gly	Val	Ala	Pro	Thr	Lys	Ala	Lys	Glu	Arg	Val	Val	Gln	500	505	510	
Arg	Glu	Glu	Arg	Ala	Val	Gly	Ile	Gly	Ala	Val	Phe	Leu	Gly	Phe	Leu	515	520	525	
Gly	Ala	Ala	Gly	Ser	Thr	Met	Gly	Ala	Ala	Ser	Met	Thr	Leu	Thr	Val	530	535	540	
Gln	Ala	Arg	Leu	Leu	Leu	Ser	Gly	Ile	Val	Gln	Gln	Gln	Asn	Asn	Leu	545	550	555	560
Leu	Arg	Ala	Ile	Glu	Ala	Gln	Gln	His	Leu	Leu	Gln	Leu	Thr	Val	Trp	565	570	575	
Gly	Ile	Lys	Gln	Leu	Gln	Ala	Arg	Val	Leu	Ala	Val	Glu	Arg	Tyr	Leu	580	585	590	
Lys	Asp	Gln	Gln	Leu	Leu	Gly	Ile	Trp	Gly	Cys	Ser	Gly	Lys	Leu	Ile	595	600	605	
Cys	Thr	Thr	Thr	Val	Pro	Trp	Asn	Ala	Ser	Trp	Ser	Asn	Lys	Ser	Leu	610	615	620	
Asp	Lys	Ile	Trp	Asn	Asn	Met	Thr	Trp	Met	Glu	Trp	Glu	Arg	Glu	Ile	625	630	635	640
Asn	Asn	Tyr	Thr	Ser	Leu	Ile	Tyr	Thr	Leu	Ile	Glu	Glu	Ser	Gln	Asn	645	650	655	
Gln	Gln	Glu	Lys	Asn	Glu	Gln	Glu	Leu	Leu	Glu	Leu	Asp	Lys	Trp	Ala	660	665	670	
Ser	Leu	Trp	Asn	Trp	Phe	Asp	Ile	Ser	Asn	Trp	Leu	Trp	Tyr	Ile	Lys	675	680	685	
Ser	Arg	Ile	Glu	Gly	Arg	Gly	Ser	Gly	Gly	Tyr	Ile	Pro	Glu	Ala	Pro	690	695	700	
Arg	Asp	Gly	Gln	Ala	Tyr	Val	Arg	Lys	Asp	Gly	Glu	Trp	Val	Leu	Leu	705	710	715	720

Ser Thr Phe Leu

<212> PRT

<213> Secuencia artificial

<220>

5 <223> Restos 655-682 de la SEQ ID NO: 18

<400> 37

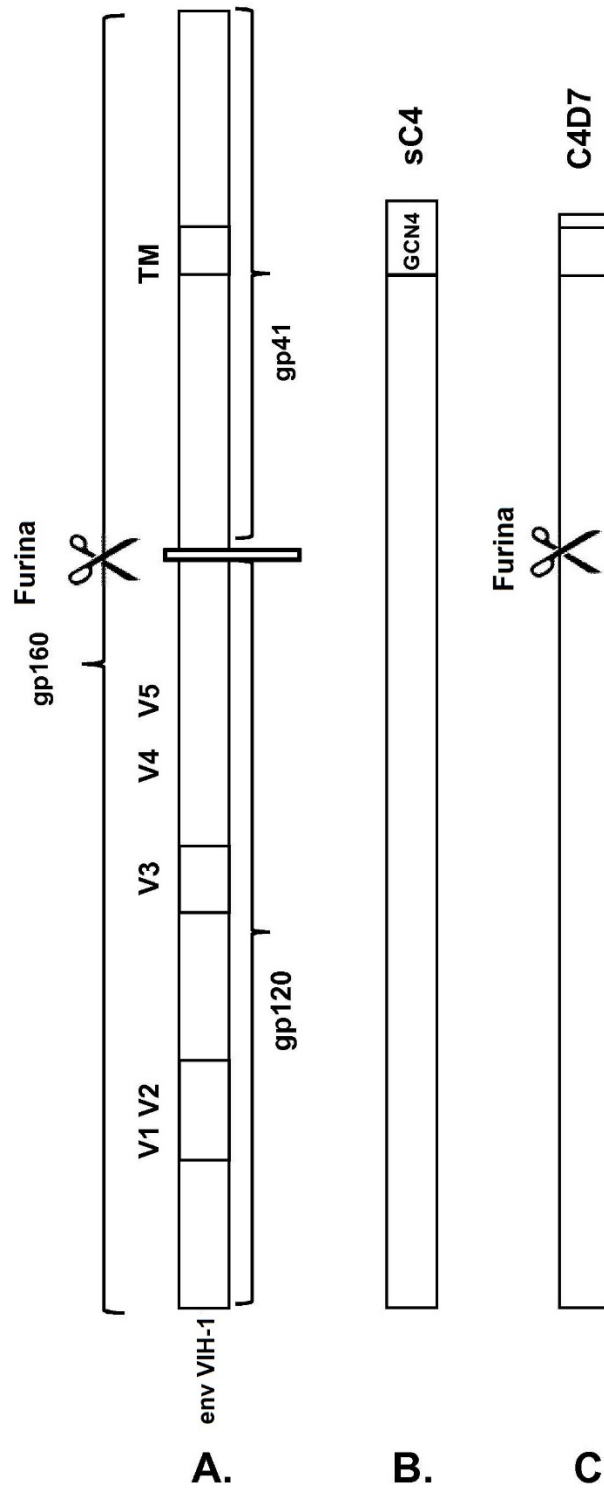
Asn	Glu	Lys	Asp	Leu	Leu	Ala	Leu	Asp	Ser	Trp	Asn	Asn	Leu	Trp	Asn
1				5				10					15		

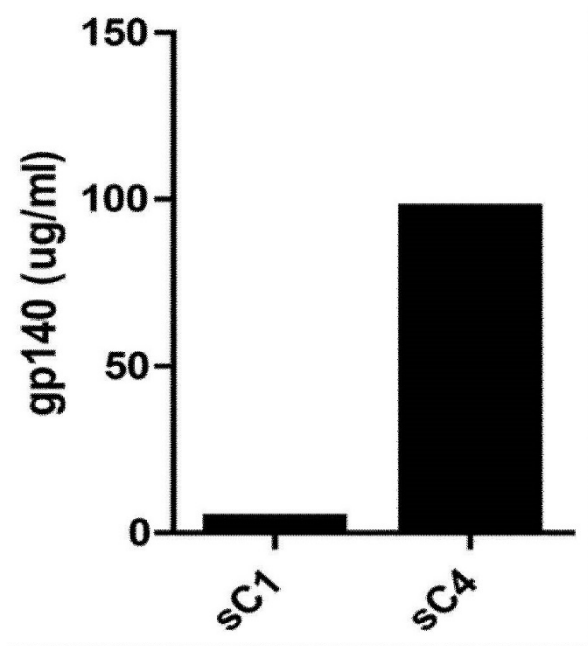
Trp	Phe	Ser	Ile	Ser	Lys	Trp	Leu	Trp	Tyr	Ile	Lys
			20					25			

## REIVINDICACIONES

1. Un vector que comprende un ácido nucleico que codifica una proteína de la envoltura del VIH sintética, en donde la proteína de la envoltura del VIH sintética comprende o consiste en la secuencia de aminoácidos de la SEQ ID NO: 18.
2. El vector de la reivindicación 1, en donde el vector es un vector vírico.
3. El vector de la reivindicación 1 o 2, en donde el vector es un vector adenovírico.
4. El vector de la reivindicación 3, en donde el vector adenovírico es un vector adenovírico humano o un vector adenovírico de simio.
5. El vector de la reivindicación 3, en donde el vector adenovírico es un vector de adenovirus del serotipo 26 (Ad26) humano.
6. El vector de la reivindicación 1 o 2, en donde el vector es un vector de virus adenoasociados, un vector del virus de la viruela, un vector MVA, un vector de virus entéricos, un vector del virus de la encefalitis equina de Venezuela, un vector del virus del bosque Semliki, un vector del virus del mosaico del tabaco, o un vector lentivírico.
7. El vector de la reivindicación 1 o 2, en donde el vector es un vector MVA.
8. El vector de la reivindicación 1 o 2, en donde el vector es un vector de tipo plásmido, cromosoma bacteriano artificial, cromosoma de levadura artificial, o bacteriófago.
9. Una composición que comprende una cantidad inmunológicamente eficaz de un vector según una cualquiera de las reivindicaciones 6 o 7 y un vehículo, en donde el ácido nucleico que codifica la proteína de la envoltura del VIH sintética está ligado operativamente a una secuencia promotora.
10. Un producto de tipo combinación de vacuna, que comprende:
  - (i) una primera composición que comprende una cantidad inmunológicamente eficaz de un vector de Ad26, que codifica una proteína de la envoltura del VIH sintética que tiene la secuencia de aminoácidos de la SEQ ID NO: 18; y
  - (ii) una segunda composición que comprende una cantidad inmunológicamente eficaz de un segundo vector de Ad26, que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5,
 en donde la primera y la segunda composición están presentes en la misma composición o en composiciones diferentes.
11. El producto de tipo combinación de vacuna de la reivindicación 10, que comprende además al menos una composición adicional que comprende:
  - una cantidad inmunológicamente eficaz de un vector que codifica al menos un polipéptido antigénico que tiene la secuencia de aminoácidos seleccionada del grupo que consiste en las SEQ ID NO: 1-4, 28 y 29, en donde la primera composición, la segunda composición y la composición adicional están presentes en la misma composición o en una o más composiciones diferentes.
12. El producto de tipo combinación de vacuna de la reivindicación 10 u 11, que además comprende al menos una composición adicional que comprende una cantidad inmunológicamente eficaz de uno o más polipéptidos antigénicos del VIH aislados, en donde la primera composición, segunda composición, composición adicional y otra composición están presentes en la misma composición o en una o más composiciones diferentes.
13. El producto de tipo combinación de vacuna de la reivindicación 12, en donde el uno o más polipéptidos antigénicos del VIH aislados incluyen uno o más polipéptidos antigénicos del VIH seleccionados del grupo que consiste en una proteína gp140 trimérica del clado C del VIH estabilizada que comprende los residuos 30-708 de la secuencia de aminoácidos de la SEQ ID NO: 7 y la proteína trimérica Env del VIH mosaico que comprende los residuos 30-724 de la SEQ ID NO: 36.
14. Un kit que comprende un producto de tipo combinación de vacuna de una cualquiera de las reivindicaciones 10-13.
15. La composición de la reivindicación 9, o el producto de tipo combinación de vacuna de una cualquiera de las reivindicaciones 10-13, para su uso en la inducción de una respuesta inmunitaria contra el VIH en un sujeto.
16. Una composición que comprende:

- (i) un primer vector de adenovirus 26 (Ad26) que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 18;
- 5 (ii) un segundo vector Ad26 que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5;
- (iii) un tercer vector Ad26 que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 28; y
- 10 (iv) un cuarto vector AD26 que codifica un polipéptido antigénico del VIH que comprende la secuencia aminoácidos de la SEQ ID NO: 29.
17. La composición de la reivindicación 16, en donde el primer, segundo, tercero y cuarto vectores Ad26 están presentes en una proporción de partículas víricas de 1:1:1:1
18. Una combinación de vacuna que comprende una composición de acuerdo con la reivindicación 16 y uno o más polipéptidos antigénicos del VIH aislados
- 15 19. La combinación de vacuna de la reivindicación 18, en donde el uno o más polipéptidos antigénicos del VIH aislados comprenden una proteína de la envoltura del VIH, por ejemplo, una proteína gp140 trimérica del clado C del VIH y/o una proteína trimérica de la envoltura mosaico.
- 20 20. La combinación de vacuna de la reivindicación 19, en donde la proteína gp140 trimérica del clado C comprende los residuos 30-708 de la secuencia de aminoácidos de la SEQ ID NO: 7, y/o en donde la proteína trimérica de la envoltura mosaico comprende los residuos 30-724 de la secuencia de aminoácidos de la SEQ ID NO: 36.
- 25 21. Un vector Ad26 que comprende un ácido nucleico que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 18, para su uso en la inducción de una respuesta inmunitaria contra el VIH en un sujeto.
- 30 22. Una vacuna con el vector adenovírico que comprende: (i) un primer vector de adenovirus 26 (Ad26) que codifica una proteína de la envoltura del VIH sintética que comprende la secuencia de aminoácidos de la SEQ ID NO: 18; (ii) un segundo vector Ad26 que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 5; (iii) un tercer vector Ad26 que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 28; y (iv) un cuarto vector Ad26 que codifica un polipéptido antigénico del VIH que comprende la secuencia de aminoácidos de la SEQ ID NO: 29, para su uso en una administración de sensibilización inicial,
- 35 y
- la proteína gp140 Env del VIH, por ejemplo, (a) la proteína gp140 del clado C que comprende los residuos 30-708 de la SEQ ID NO: 7, o (b) la proteína gp140 mosaico que comprende los residuos 30-724 de la SEQ ID NO: 36, o (c) una combinación de la proteína gp140 del clado C y una proteína gp140 mosaico,
- 40 para su uso junto con la vacuna con el vector adenovírico en una administración de refuerzo, en donde las administraciones de sensibilización inicial y de refuerzo son para generar una respuesta inmunitaria contra el VIH en un sujeto.

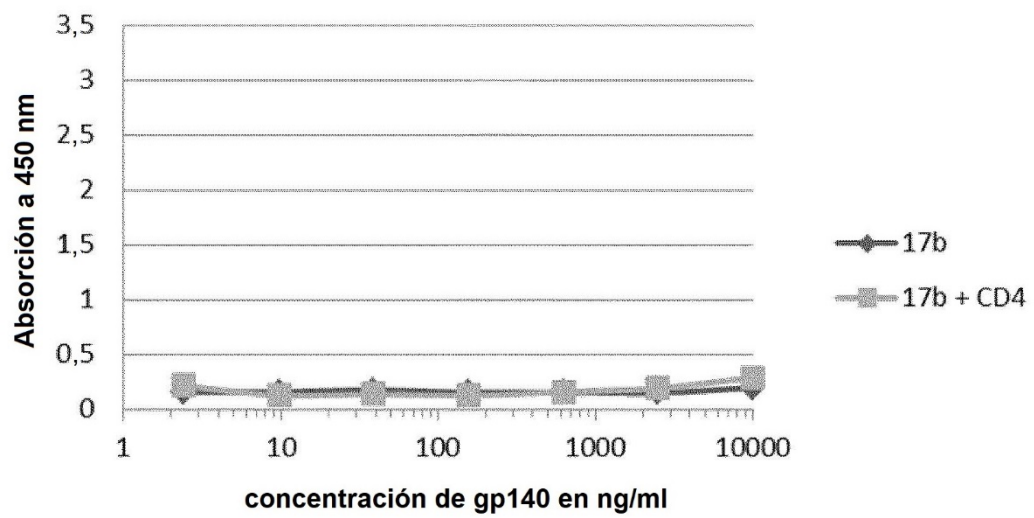




**Fig. 2**

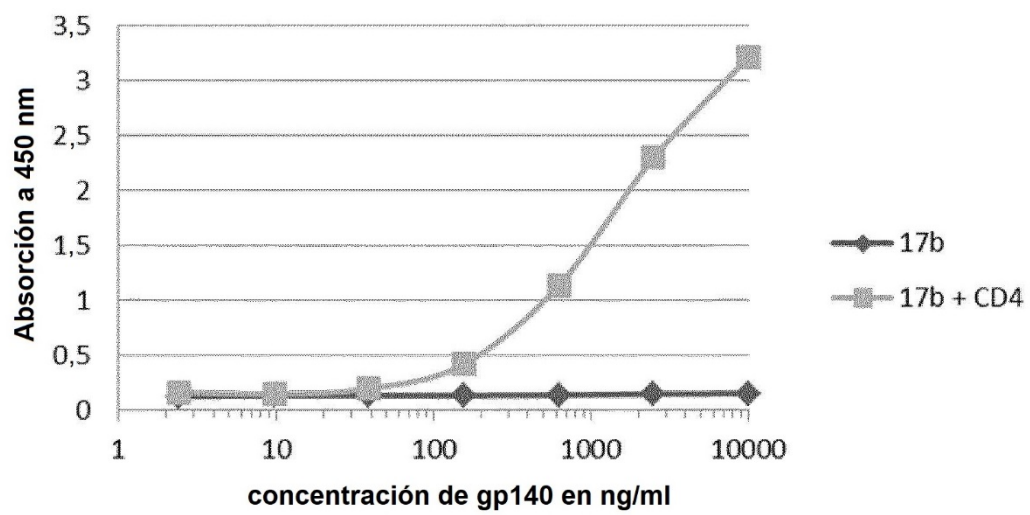
**A.**

**sC1**



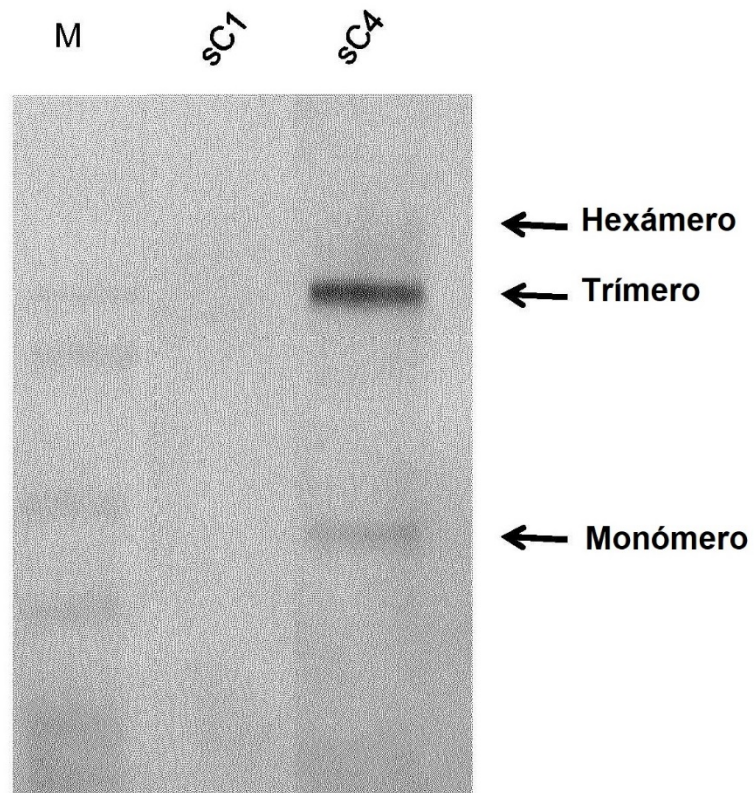
**B.**

**sC4**



**Fig. 3**





**Fig. 4**

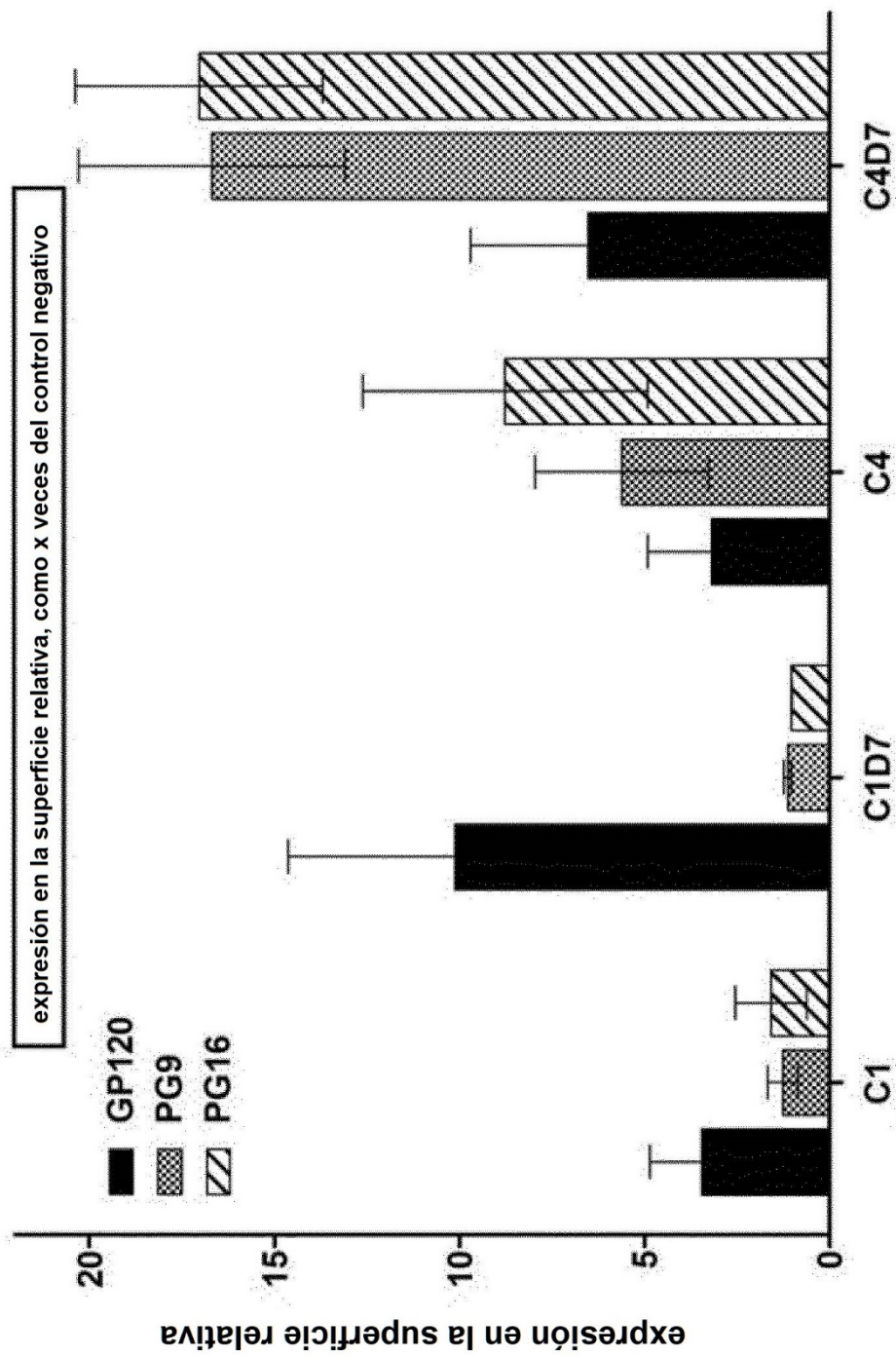
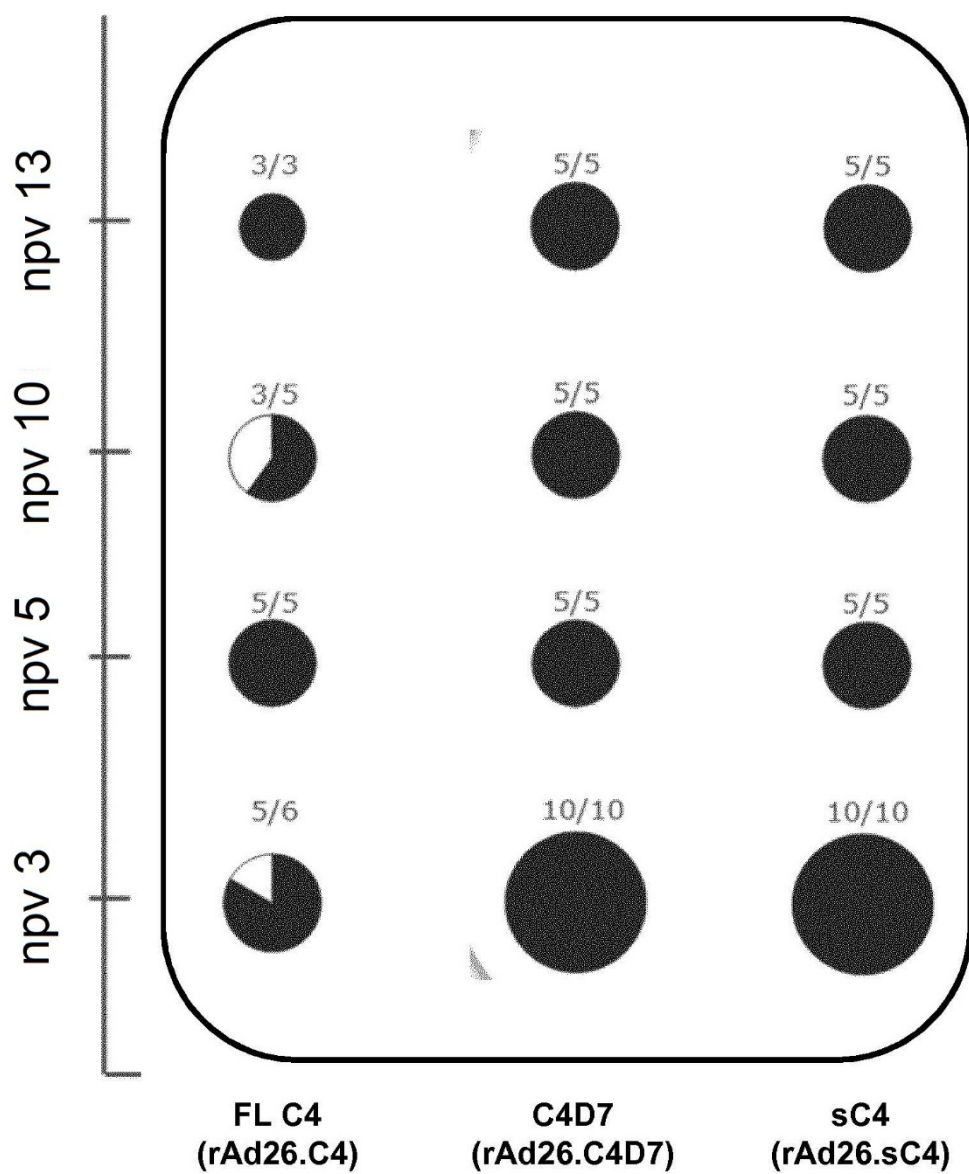
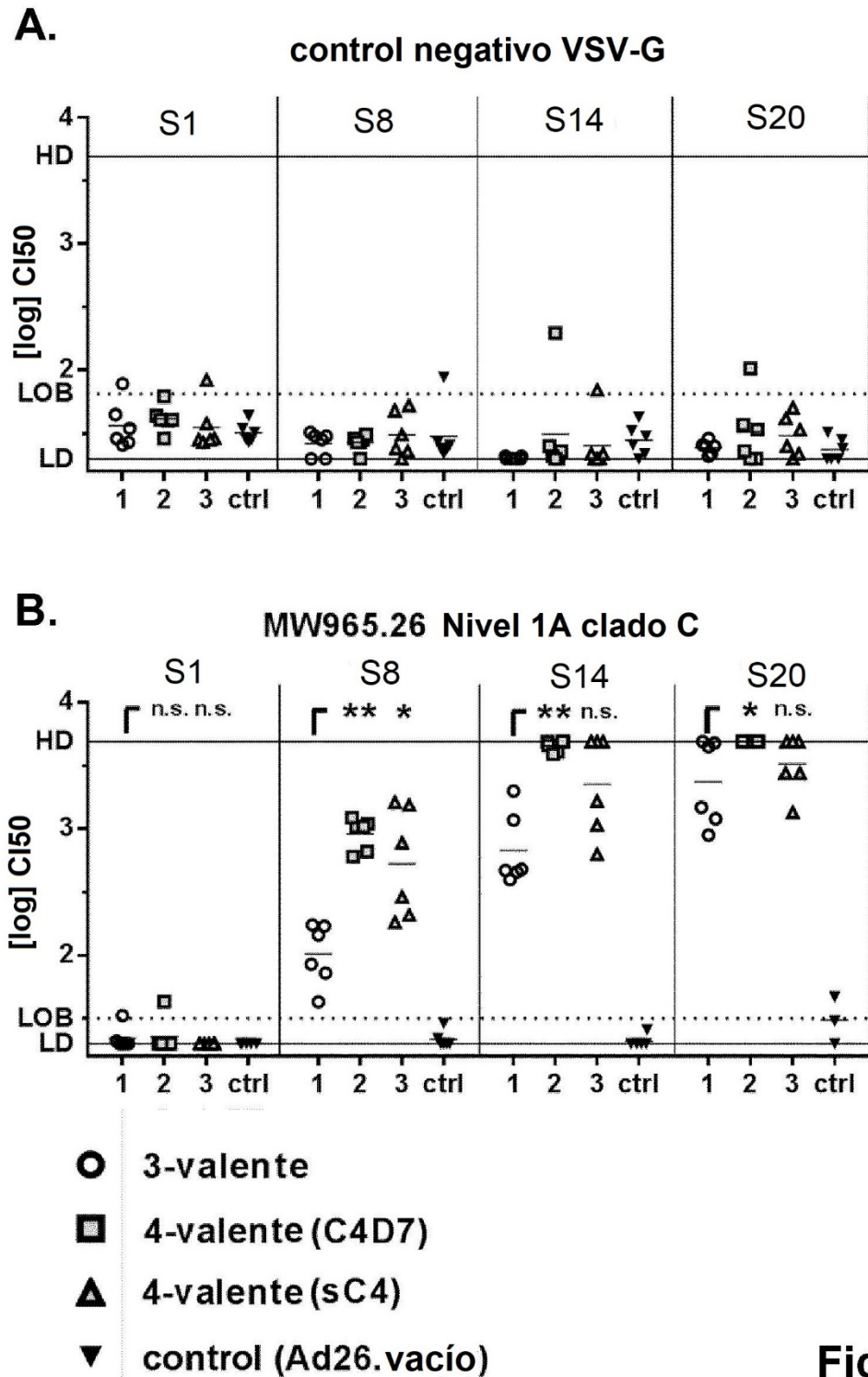


Fig. 5



**Fig. 6**



**Fig. 7**