

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 699 685**

51 Int. Cl.:

C07K 17/00 (2006.01)
C07K 14/005 (2006.01)
A61K 39/12 (2006.01)
A61K 39/235 (2006.01)
A61K 39/21 (2006.01)
A61K 39/275 (2006.01)
A61K 39/00 (2006.01)
C12N 7/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **18.11.2009 PCT/US2009/064999**
 87 Fecha y número de publicación internacional: **27.05.2010 WO10059732**
 96 Fecha de presentación y número de la solicitud europea: **18.11.2009 E 09828172 (8)**
 97 Fecha y número de publicación de la concesión europea: **12.09.2018 EP 2358757**

54 Título: **Vacunas antiviricas con inmunogenicidad celular mejorada**

30 Prioridad:

02.10.2009 US 248188 P
18.11.2008 US 115703 P
12.02.2009 US 152184 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
12.02.2019

73 Titular/es:

BETH ISRAEL DEACONESS MEDICAL CENTER
(50.0%)
330 Brookline Avenue
Boston, MA 02215, US y
LOS ALAMOS NATIONAL SECURITY, LLC
(50.0%)

72 Inventor/es:

BAROUCH, DAN, H.;
KORBER, BETTE, T. y
FISCHER, WILLIAM, M.

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 699 685 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

Vacunas antiviricas con inmunogenicidad celular mejorada

5 Campo de la invención

La invención proporciona una vacuna que comprende un polipéptido vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9, o un vector que codifica un polipéptido vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9, y kits para el tratamiento o la prevención de infecciones víricas. Las vacunas de la invención o como se describen en el presente documento incorporan polipéptidos víricos optimizados por ordenador que pueden aumentar la diversidad o amplitud y profundidad de la respuesta inmunitaria celular en sujetos vacunados.

15 Antecedentes de la invención

Las vacunas que inducen respuestas inmunitarias celulares contra virus deben reflejar la diversidad vírica global para tratar o prevenir eficazmente la infección vírica. Por ejemplo, el inicio de respuestas intensas y diversas de linfocitos T específicos del VIH-1 es probablemente crucial para una vacuna del VIH-1 eficaz. Las respuestas de linfocitos T citotóxicos (CTL) se correlacionan con progresión de enfermedad lenta en seres humanos, y la importancia de las respuestas de CTL en modelos de vacunación de primates no humanos está bien establecida. Aunque la envoltura (Env) altamente variable es la principal diana para anticuerpos neutralizantes contra el VIH y también será necesario adaptar antígenos de vacuna para inducir estas respuestas de anticuerpos, los componentes de vacuna de linfocitos T pueden dirigirse a proteínas más conservadas para desencadenar respuestas que sea más probable que reaccionen de forma cruzada. No obstante, incluso las proteínas del VIH-1 más conservadas son suficientemente diversas como para que la variación sea un problema. Los enfoques de vacuna de secuencia central artificial, tales como secuencias del VIH-1 consenso y ancestrales, esencialmente "dividen las diferencias" entre cepas, pueden estimular respuestas con reactividad cruzada potenciada en comparación con vacunas de cepas naturales. Los antígenos consenso representan secuencias antigénicas sintéticas que son el mejor "promedio" individual de todas las cepas en circulación. Aunque estos antígenos pueden inducir respuestas inmunitarias celulares dirigidas, la amplitud e intensidad de estas respuestas no mejoran sustancialmente con respecto a estrategias de vacunas previas.

Por ejemplo, Fisher *et al.* "Polyvalent Vaccine for Optimal Coverage of Potential T-Cell Epitopes in Global HIV-1 Variants," Nat. Med. 13(1):100-106 (2007) y Solicitud de Patente Internacional WO 2007/024941 desvelan una proteína mosaico denominada ENV-B.syn1.1 GP160 que comprende la parte GP41 completa de la proteína Env no escindida, que se puede usar en vacunaciones antiviricas.

El desarrollo de vacunas de próxima generación para tratar o prevenir infección vírica debe inducir una mayor amplitud de la inmunidad celular para permitir resultados de vacunación exitosos. La necesidad de dichas vacunas es particularmente urgente para el tratamiento o la prevención del VIH-1.

40 Sumario de la invención

La invención se define mediante las reivindicaciones.

45 En un primer aspecto de la invención, se proporciona una vacuna que comprende un polipéptido vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9, o un vector que codifica un polipéptido vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9.

50 En un segundo aspecto, se proporciona la vacuna para su uso en un método de tratamiento o reducción del riesgo de una infección por el virus de la inmunodeficiencia humana de tipo 1 (VIH-1) en un ser humano.

55 En un tercer aspecto de la invención, se proporciona un método de fabricación de una vacuna para el tratamiento o la reducción del riesgo de una infección por VIH-1 en un ser humano, comprendiendo dicho método sintetizar la vacuna de la invención.

60 En un cuarto aspecto de la invención, se proporciona un kit que comprende (a) la vacuna de la invención, (b) un vehículo, excipiente o diluyente farmacéuticamente aceptable, (c) instrucciones para el uso de los mismos y, opcionalmente, (d) un adyuvante. La invención también presenta una vacuna para el tratamiento o la reducción del riesgo de una infección vírica en un mamífero, tal como un ser humano, que incluye al menos dos polipéptidos víricos optimizados definidos (por ejemplo, 2, 3, 4, 5, o más polipéptidos víricos optimizados definidos), en donde los polipéptidos víricos optimizados corresponden al mismo producto génico vírico. La infección vírica está causada por el virus de la inmunodeficiencia humana de tipo 1 (VIH-1) y los productos génicos víricos pueden incluir Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr o Vpu. En una realización, la vacuna no incluye más de dos polipéptidos víricos optimizados correspondientes a uno de los productos génicos víricos de Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr o Vpu. En otra realización, la vacuna no incluye polipéptidos víricos optimizados correspondientes a Gag y Nef. En otra realización más, la vacuna incluye al menos dos polipéptidos víricos optimizados definidos (por ejemplo, 2, 3, 4, 5, o más

polipéptidos víricos optimizados definidos) para un primer producto génico vírico seleccionado de Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr y Vpu y uno o más polipéptidos víricos optimizados definidos (por ejemplo, 2, 3, 4, 5, o más polipéptidos víricos optimizados definidos) para un segundo producto génico vírico diferente del primer producto génico vírico seleccionado de Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr y Vpu.

5 También se desvela una vacuna para el tratamiento o a reducción del riesgo de una infección por el virus de la inmunodeficiencia humana de tipo 1 (VIH-1) en un ser mamífero, tal como un ser humano, que incluye un polipéptido vírico optimizado que tiene al menos siete aminoácidos contiguos (por ejemplo, al menos 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 50, 100, 150, 175, 200, 250, 300, 350, 400, 450, 500 o más aminoácidos contiguos de longitud) que tienen al menos 85 % de identidad de secuencia de aminoácidos con una cualquiera de las secuencias expuestas en las SEQ ID NO: 1-29. El polipéptido vírico optimizado puede tener al menos siete aminoácidos contiguos (por ejemplo, al menos 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 50, 100, 150, 175, 200, 250, 300, 350, 400, 450, 500 o más aminoácidos contiguos de longitud) que tienen identidad de secuencia de aminoácidos con una cualquiera de las secuencias expuestas en las SEQ ID NO: 1-29. El polipéptido vírico optimizado puede tener la secuencia de aminoácidos de una cualquiera de las secuencias expuestas en las SEQ NO: 1-29. La vacuna puede incluir al menos dos polipéptidos víricos optimizados seleccionados de uno cualquiera o más de los grupos a)-k): a) SEQ ID NO: 1 y 2; b) SEQ ID NO: 3, 4 y 5; c) SEQ ID NO: 6 y 7; d) SEQ ID NO: 8-12; e) SEQ ID NO: 13, 14 y 15; f) SEQ ID NO: 16, 17 y 18; g) SEQ ID NO: 19 y 20; h) SEQ ID NO: 21, 22 y 23; i) SEQ ID NO: 24 y 25; j) SEQ ID NO: 26 y 27; k) y SEQ ID NO: 21-22. La vacuna puede incluir un par de polipéptidos víricos optimizados seleccionados de uno cualquiera de los grupos a)-k) anteriores y uno o más polipéptidos víricos optimizados del mismo grupo o un grupo diferente a)-k). La vacuna puede incluir al menos tres o cuatro o más polipéptidos víricos optimizados de uno o más de los grupos a)-k).

25 La invención también presenta una vacuna para el tratamiento o la reducción del riesgo de una infección vírica en un mamífero, tal como un ser humano, que incluye al menos dos pares de polipéptidos víricos optimizados definidos, en donde cada par de polipéptidos víricos optimizados corresponde al mismo producto génico vírico, en donde no más de dos polipéptidos víricos optimizados incorporados en la vacuna corresponden al mismo producto génico vírico. En una realización, la vacuna incluye al menos tres pares de polipéptidos víricos optimizados definidos. En otra realización, la vacuna incluye al menos cuatro pares de polipéptidos víricos optimizados definidos. La infección vírica está causada por VIH-1 y los productos génicos víricos incluyen Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr o Vpu. En una realización, la vacuna no incluye más de dos polipéptidos víricos optimizados correspondientes a uno de los productos génicos víricos de Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr o Vpu. En otra realización, la vacuna no incluye polipéptidos víricos optimizados correspondientes a Gag y Nef. En una realización adicional, la vacuna incluye al menos tres pares de polipéptidos víricos optimizados definidos correspondientes a tres cualesquiera de los productos génicos víricos de Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr o Vpu. En otra realización, la vacuna incluye al menos cuatro pares de polipéptidos víricos optimizados definidos correspondientes a cuatro cualesquiera de los productos génicos víricos de Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr o Vpu.

40 En una realización, la vacuna induce una respuesta inmunitaria celular contra un producto génico vírico. En otra realización, la vacuna induce una respuesta inmunitaria celular contra el VIH-1. En una realización adicional, la secuencia de nucleótidos de al menos un polipéptido vírico optimizado definido está codificada por un ácido nucleico o vector. En una realización, el vector es un adenovirus recombinante, tal como adenovirus de serotipo 26 (Ad26), adenovirus de serotipo 34 (Ad34), adenovirus de serotipo 35 (Ad35), adenovirus de serotipo 48 (Ad48) o adenovirus de serotipo 5 HVR48 (Ad5HVR48). En una realización adicional, la vacuna está en combinación con un vehículo, excipiente o diluyente farmacéuticamente aceptable.

50 Se desvela un ácido nucleico que incluye la secuencia de nucleótidos de un polipéptido vírico optimizado que tiene al menos siete aminoácidos contiguos (por ejemplo, al menos 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 50, 100, 150, 175, 200, 250, 300, 350, 400, 450, 500 o más aminoácidos contiguos de longitud) que tienen al menos 85 % de identidad de secuencia de aminoácidos con una cualquiera de las secuencias de aminoácidos expuestas en las SEQ ID NO: 1-29. El polipéptido vírico optimizado puede tener al menos siete aminoácidos contiguos (por ejemplo, al menos 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 50, 100, 150, 175, 200, 250, 300, 350, 400, 450, 500 o más aminoácidos contiguos de longitud) que tienen identidad de secuencia con una cualquiera de las secuencias de aminoácidos expuestas en las SEQ ID NO: 1-29. El polipéptido vírico optimizado puede tener una cualquiera de las secuencias de aminoácidos expuestas en las SEQ ID NO: 1-29. El ácido nucleico puede incluir un vector. En una realización, el vector es un adenovirus recombinante, tal como adenovirus de serotipo 26 (Ad26), adenovirus de serotipo 34 (Ad34), adenovirus de serotipo 35 (Ad35), adenovirus de serotipo 48 (Ad48) o adenovirus de serotipo 5 HVR48 (Ad5HVR48).

60 Se desvela un polipéptido vírico optimizado que tiene al menos siete aminoácidos contiguos (por ejemplo, al menos 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 50, 100, 150, 175, 200, 250, 300, 350, 400, 450, 500 o más aminoácidos contiguos de longitud) que tienen al menos 85 % de identidad de secuencia de aminoácidos con una cualquiera de las secuencias de aminoácidos expuestas en las SEQ ID NO: 1-29. El polipéptido vírico optimizado puede tener al menos siete aminoácidos contiguos (por ejemplo, al menos 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 50, 100, 150, 175, 200, 250, 300, 350, 400, 450, 500 o más aminoácidos contiguos de longitud) que tienen identidad de secuencia con una cualquiera de las secuencias de

aminoácidos expuestas en las SEQ ID NO: 1-29. El polipéptido vírico optimizado puede tener una cualquiera de las secuencias de aminoácidos expuestas en las SEQ ID NO: 1-29.

Las vacunas de la invención son útiles en métodos para el tratamiento o la reducción del riesgo de infección por VIH-1 en un mamífero, tal como un ser humano. En una realización, la vacuna induce una respuesta inmunitaria celular contra un producto génico vírico.

Se desvela un método de fabricación de una vacuna para el tratamiento o la reducción del riesgo de una infección vírica en un mamífero, tal como un ser humano, poniendo en contacto un ácido nucleico de la invención con una célula y aislando un polipéptido vírico optimizado.

Por referencia al método del párrafo precedente, el polipéptido vírico optimizado induce una respuesta inmunitaria celular cuando se administra a un mamífero. La respuesta inmunitaria celular puede ser contra un producto génico vírico. En realizaciones de la invención, la infección vírica está causada por VIH-1 y los productos génicos víricos pueden incluir Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr o Vpu.

Se desvela un kit que incluye un ácido nucleico de la invención, un vehículo, excipiente o diluyente farmacéuticamente aceptable, e instrucciones para el uso de los mismos. El kit puede incluir también un adyuvante.

Como se desvela en el presente documento, el polipéptido vírico optimizado puede estar codificado por una secuencia de ácido nucleico que está optimizada para expresión en seres humanos (por ejemplo, una cualquiera de las SEQ ID NO: 5, 10, 11, 12, 15, 18 y 23).

Definiciones

Por "polipéptido vírico optimizado" o "polipéptido vírico optimizado por ordenador" se entiende un polipéptido inmunogénico que no es un péptido, polipéptido o proteína vírico de origen natural. Se general inicialmente secuencias polipeptídicas víricas optimizadas modificando la secuencia de aminoácidos de uno o más productos génicos víricos de origen natural (por ejemplo, péptidos, polipéptidos y proteínas) para aumentar la amplitud, intensidad, profundidad o longevidad de la respuesta inmunitaria antivírica (por ejemplo, respuestas inmunitarias celulares o humorales) generada tras la inmunización (por ejemplo, cuando se incorporan en una vacuna de la invención) de un mamífero (por ejemplo, un ser humano). Por tanto, el polipéptido vírico optimizado puede corresponder a una secuencia génica vírica "parental"; como alternativa, el polipéptido vírico optimizado puede no corresponder a una secuencia génica vírica "parental" específica, pero puede corresponder a secuencias análogas de diversas cepas o cuasiespecies de un virus. Las modificaciones de la secuencia génica vírica que pueden incluirse en un polipéptido vírico optimizado incluyen adiciones, sustituciones y supresiones de aminoácidos. En una realización de la invención, el polipéptido vírico optimizado es la secuencia de aminoácidos compuesta o fusionada de dos o más productos génicos víricos de origen natural (por ejemplo, aislados víricos naturales o clínicos) en la que cada epítipo potencial (por ejemplo, cada secuencia de aminoácidos contigua o solapante de 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 o más aminoácidos de longitud) se analiza y modifica para mejorar la inmunogenicidad del polipéptido vírico optimizado resultante. Los polipéptidos víricos optimizados que corresponden a productos génicos víricos diferentes también pueden fusionarse para facilitar la incorporación en una vacuna de la invención. Se describen métodos para generar un polipéptido vírico optimizado en, por ejemplo, Fisher *et al.* "Polyvalent Vaccine for Optimal Coverage of Potential T-Cell Epitopes in Global HIV-1 Variants," Nat. Med. 13(1):100-106 (2007) y la Publicación de Solicitud de Patente Internacional WO 2007/024941. Una vez que se ha generado la secuencia de polipéptido vírico optimizado, el polipéptido correspondiente puede producirse o administrarse mediante técnicas convencionales (por ejemplo, vectores víricos recombinantes, tales como los vectores adenovíricos divulgados en las publicaciones de solicitud de patente internacional 2006/040330 y WO 2007/104792).

Por "vehículo farmacéuticamente aceptable" se entiende un vehículo que es fisiológicamente aceptable para el mamífero tratado conservando al mismo tiempo las propiedades terapéuticas del compuesto con el que se administra. Un vehículo farmacéuticamente aceptable ejemplar es la solución salina fisiológica. Otros vehículos fisiológicamente aceptables y sus formulaciones son conocidos por un experto en la materia y se describen, por ejemplo, en Remington's Pharmaceutical Sciences (18ª edición, ed. A. Gennaro, 1990, Mack Publishing Company, Easton, PA).

Por "vector" se entiende una construcción de ADN que contiene un promotor unido operativamente a un gen o una región codificante cadena abajo (por ejemplo, un ADNc o fragmento de ADN genómico, que codifica un polipéptido o fragmento polipeptídico). La introducción del vector en una célula receptora (por ejemplo, una célula procariota o eucariota, por ejemplo, una bacteria, levadura, célula de insecto o célula de mamífero, dependiendo del promotor dentro del vector de expresión) u organismo (incluyendo, por ejemplo, un ser humano) permite que la célula exprese ARNm codificado por el vector, que después se traduce en el polipéptido vírico optimizado codificado de la invención. También se conocen bien en la técnica vectores para transcripción/traducción *in vitro* y se describen adicionalmente en el presente documento. Un vector puede ser un plásmido modificado genéticamente, virus o cromosoma artificial procedente de, por ejemplo, un bacteriófago, adenovirus, retrovirus, poxvirus o herpesvirus.

Por "producto génico vírico" se entiende cualquier péptido, polipéptido o proteína vírico de origen natural, o fragmento de los mismos. El producto génico vírico comprendido por la invención se obtiene del virus de la inmunodeficiencia humana de tipo 1 (VIH-1). Los productos génicos víricos del VIH-1 incluyen los polipéptidos Gag, Pol, Env, Nef, Tat, Rev, Vif y Vpr.

5 Breve descripción de los dibujos

10 La Figura 1 es un diagrama que ilustra la amplitud expandida de polipéptidos víricos Gag, Pol y Env del VIH-1 optimizados por ordenador frente a péptidos de epítomos de linfocitos T (PTE) potenciales globales en macacos Rhesus. Los animales inmunizados con los polipéptidos víricos optimizados (azul) reaccionaron con el mayor número de grupos de péptidos de recuerdo.

La Figura 2 es un diagrama que muestra que los polipéptidos víricos Gag, Pol y Env del VIH-1 modificados por ordenador expanden la amplitud de la respuesta inmunitaria celular específica de epítomo.

15 La Figura 3 ilustra la amplitud de respuestas inmunitarias celulares detectadas en macacos Rhesus después de la inmunización con los productos génicos víricos del VIH-1 Gag, Pol y Env procedentes de los polipéptidos víricos modificados por ordenador de la invención, así como animales inmunizados con antígenos del VIH-1 o antígenos aislados del VIH-1 de clado C. Los animales inmunizados con los polipéptidos víricos optimizados (azul) reaccionaron con el mayor número de grupos de péptidos de recuerdo. Ya que los animales son exogámicos, los grupos difieren entre un animal y otro. Gag, Pol y Env inducen cada uno muchas respuestas inmunitarias celulares y pueden tener patrones de reactividad compartidos.

20 Las Figuras 4A-4C son gráficos que muestran los epítomos potenciales compartidos entre las diferentes vacunas ensayadas (mosaico bivalente (Mos2), M consenso (Mcon) y clado C optimizado (OptC)) por polipéptido vírico (Pol (Fig. 4A), Gag (Fig. 4B) y Env (Fig. 4C)). Las Figuras 4A-C muestran la cobertura relativa del conjunto de genomas de longitud completa de la base de datos del VIH actual y los péptidos PTE por los diferentes candidatos a vacuna.

25 La Figura 5 es un gráfico que muestra que el número de respuestas de péptidos PTE (donde cada respuesta se considera un acontecimiento independiente con independencia del solapamiento) a la vacuna de mosaico bivalente (Mos2) es mayor que el número de respuestas a la vacuna consenso de grupo M (Mcon) y a la vacuna de cepa vírica natural (clado C optimizado (C Natural (óptimo))), que se ha seleccionado para proporcionar cobertura óptima de los antígenos de vacuna de colección del grupo M (OptC). La Figura 5 muestra el número de respuestas de péptidos PTE por animal por proteína, linfocito T CD8+ y linfocito T CD4+. Estadísticamente, Mos2 > Mcon ~ OptC (Mcon muestra una tendencia a más respuesta que OptC). El valor de p de Wilcoxon para Mos2 en comparación con Mcon: valor de p = 0,001058.

30 La Figura 6 es un diagrama que muestra el número de péptidos PTE que desencadena respuestas de linfocitos T. Un número mediano de 16 (intervalo: 12-29) péptidos PTE de la vacuna de mosaico bivalente (Mos2) desencadena una respuesta en linfocitos T CD8+, mientras que solamente un número mediano de 6 (intervalo: 0-7) péptidos Mcon y solamente un número mediano de 3 péptidos (intervalo: 0-3) de péptidos OptC desencadenan una respuesta en linfocitos T CD8+. Un número mediano de 4 (intervalo: 2-6) péptidos PTE de la vacuna de mosaico bivalente (Mos2) desencadena una respuesta en linfocitos T CD4+, mientras que solamente un número mediano de 1 (intervalo: 0-2) péptidos Mcon y solamente un número mediano de 0,5 péptidos (intervalo: 0-2) de péptidos OptC desencadenan una respuesta en linfocitos T CD4+. Por tanto, la tendencia para respuestas es Mos2 > Mcon > OptC.

35 La Figura 7 es un esquema que resume el mapeo de todos los péptidos PTE Gag de linfocitos T CD8+ que son reconocidos por linfocitos T de cada uno de los animales estudiados (véase Ejemplo 3 posteriormente). El número de animales, grupo de péptidos y número de péptidos marcan los límites de cada péptido reactivo. El símbolo indica el grupo: *, Mos2; ¥, ConM; ±, OptC. Gag se incluye aquí como un ejemplo. Tiende a haber agrupamiento de respuestas de CD8 incluso aunque los animales sean exogámicos. Los mosaicos tienen ventajas potenciales con respecto a las vacunas monovalentes. Los mosaicos tienen una mejor probabilidad de estimular una respuesta que reaccione con variantes más habituales. Los mosaicos también estimulan múltiples respuestas a las diferentes formas que están presentes en el cóctel. Por tanto, los mosaicos tienen el potencial de bloquear vías de escape habituales. En el estudio de los inventores, la vacuna de mosaico tendió a estimular respuestas de linfocitos T que reconocían más péptidos solapantes. Existen muchos puntos calientes de localización de péptidos reactivos. Los péptidos PTE se diseñan para maximizar la cobertura de epítomo (o nonúmero para un tramo de 9 aminoácidos contiguos) potencial del grupo M del VIH-1 en los reactivos peptídicos usados para evaluar vacunas. Inevitablemente, hay mucho solapamiento en péptidos PTE, pero debido al algoritmo, el solapamiento es habitualmente un solapamiento con algo de variación. La figura 7 desvela la SEQ ID NO: 42.

45 La Figura 8 es un esquema que resume el mapeo de los péptidos PTE Gag de linfocitos T CD4+ que son reconocidos por linfocitos T de cada uno de los animales estudiados. La figura 8 desvela la SEQ ID NO: 43.

50 La Figura 9 es un diagrama que ilustra patrones típicos de respuestas de PTE a la vacuna de ConM o a la vacuna natural óptima, alineando péptidos que inducen una respuesta con la región relevante de la vacuna. Son necesarias buenas coincidencias con tramos sólidos de identidad entre vacuna y péptido PTE diana para conseguir una reacción a estas vacunas. La figura desvela las SEQ ID NO: 44-57, respectivamente, en orden de aparición.

55 La Figura 10 es un diagrama que ilustra que las vacunas de mosaico generaron muchas respuestas que reconocieron múltiples péptidos solapantes variantes sin ninguna competición antigénica evidente y con respuestas locales amplias. En particular, se reconocieron cuatro péptidos PTE variables. Por otra parte, en la región de solapamiento se reconocieron ambas formas de mosaico, así como una combinación de las dos.

Finalmente, se reconoció una forma nueva (S). La Figura 10 desvela las SEQ ID NO: 58-63, respectivamente, en orden de aparición.

La Figura 11 es un diagrama que ilustra un patrón típico de respuestas de péptidos PTE CD8+ en animal vacunado con mosaico (361-07). Se ensayaron 22 péptidos PTE y se identificaron 8 regiones sensibles a CD8; 5 regiones incluían péptidos variables que coinciden con aminoácidos en uno u otro de los mosaicos. Se identificaron 5 regiones sensibles a CD4. Por tanto, las respuestas de linfocitos T a mosaicos muestran más péptidos variables en una región dada. Esto parece suceder en particular con respuestas de linfocitos T CD8. Esto podría ser el resultado de desencadenar múltiples clones de linfocitos T que reconocen variantes de epítomos, y estos pueden bloquear vías de escape adecuadas. No solamente hay más respuestas, sino que son más profundas y abarcan más variantes. La Figura 11 desvela respuestas de CD8 como SEQ ID NO: 64-101, respectivamente, en orden de aparición. Las respuestas de CD8 se desvelan como SEQ ID NO: 102-117, respectivamente, en orden de aparición. La Figura 12 es un gráfico que muestra el número de péptidos PTE variables solapantes que abarcan regiones diana de linfocitos T inducidos por vacuna.

La Figura 13 es un gráfico que muestra que la vacuna de 2 antígenos mosaico produce más respuestas de linfocitos T, en relación con las vacunas de Mcon y OptC, a regiones que contienen uno o más péptidos PTE solapantes. La Figura 13 es similar a la Figuras 5, monos mostrados en el mismo orden de derecha a izquierda, pero con la escala cambiada para reflejar el número de respuestas a regiones que contienen uno o más péptidos PTE solapantes en lugar de péptidos individuales.

La Figura 14 es un diagrama que muestra el número de respuestas de linfocitos T en animales después de la administración de vacunas de mosaico bivalente (Mos2), Mcon y OptC. La vacuna de mosaico bivalente (Mos2) desencadena un número mediano de 8 respuestas en linfocitos T CD8+, mientras que solamente un número mediano de 3 (intervalo: 0-6) y 1,5 péptidos (intervalo: 0-5) respuestas de linfocitos T CD8+ son desencadenadas por vacunas de Mcon y OptC, respectivamente. La vacuna de mosaico bivalente (Mos2) desencadena un número mediano de 3 (intervalo: 2-5) respuestas en linfocitos T CD4+, mientras que solamente un número mediano de 1 (intervalo: 0-2) y 0,5 (intervalo: 0-2) respuestas de linfocitos T CD4+ son desencadenadas por vacunas de Mcon y OptC, respectivamente. Por tanto, la tendencia para respuestas es Mos2 > Mcon > OptC.

La Figura 15 es un gráfico que muestra que las vacunas de mosaico pueden inducir más respuestas que reaccionan de forma cruzada con proteínas naturales del clado C que una vacuna natural del clado C: los péptidos agrupados GAG que representan 5 proteínas. Los animales vacunados con el consenso del grupo M o la proteína natural del clado C de cobertura óptima tuvieron 0-2 respuestas a los péptidos procedentes de estas proteínas, mientras que los animales vacunados con mosaico pudieron responder a grupos de 1-5 péptidos. La vacuna de mosaico induce más respuestas a cada una de las proteínas ensayadas que Mcon o el C óptimo. Las respuestas de linfocitos T inducidas por vacunas de mosaico también reconocieron más conjuntos de péptidos agrupados que abarcaban proteínas Gag *reales*. 10-12 subgrupos = péptidos 10x15meros (excepto 96ZM Gag, que es péptidos 5x20meros).

La Figura 16 es un gráfico que muestra que el diseño de mosaico es robusto para cambios en polipéptidos víricos a lo largo del tiempo (por ejemplo, Gag M).

La Figura 17 es un gráfico que muestra que la cobertura usando optimización de nonúmeros es robusta sobre longitudes casi de optimización (por ejemplo, 8-12meros) (se muestra Gag).

La Figura 18 es un gráfico que muestra que un aumento en el número de variantes aumenta la cobertura, pero tiene solamente rendimientos decrecientes (se muestra Gag).

Las Figuras 19A-19B son gráficos que muestran la amplitud y magnitud de respuestas de linfocitos T específicas de epítomo a péptidos PTE. La Figura 19A es un gráfico que muestra los números de respuestas de linfocitos T CD4+ (superior) y CD8+ (inferior) específicas de epítomo a péptidos PTE individuales después de una única inmunización de vectores rAd26 que expresan antígenos Gag, Pol y Env del VIH-1 mosaico (azul), M consenso (verde), clado B + clado C (morado) o clado C natural óptimo (rojo). Se representan monos individuales en el eje x. Los diferentes tonos de cada color reflejan respuestas a los diferentes antígenos (Gag, Pol, Env). La Figura 19B es un gráfico que muestra los números de regiones de respuesta de linfocitos T CD4+ (superior) y CD8+ (inferior).

Las Figuras 20A-20C muestran un esquema que muestra respuestas de linfocitos T CD8+ a péptidos PTE en la semana 4 después de inmunización mapeada en secuencias proteicas de Gag (Figura 20A) (SEQ ID NO: 118), Pol (Figura 20B) (SEQ ID NO: 119) y Env (Figura 20C) (SEQ ID NO: 120) de VIH-1. Los colores indican monos que recibieron los antígenos Gag, Pol y Env del VIH-1 mosaico (azul), M consenso (verde), clado B + clado C (morado) o clado C natural óptimo (rojo). Para cada epítomo, el número de monos, antígeno (G, Gag; P, Pol; E, Env), número de subgrupo y número de péptidos PTE individuales están indicados.

Las Figuras 21A-21C muestran un esquema que muestra respuestas de linfocitos T CD4+ a péptidos PTE en la semana 4 después de inmunización mapeada en secuencias proteicas de Gag (Figura 21A) (SEQ ID NO: 121), Pol (Figura 21B) (SEQ ID NO: 122) y Env (Figura 21C) (SEQ ID NO: 123) de VIH-1. Los colores indican monos que recibieron los antígenos Gag, Pol y Env del VIH-1 mosaico (azul), M consenso (verde), clado B + clado C (morado) o clado C natural óptimo (rojo). Para cada epítomo, el número de monos, antígeno (G, Gag; P, Pol; E, Env), número de subgrupo y número de péptidos PTE individuales están indicados.

La Figura 22 es un esquema que muestra el alineamiento de secuencias de vacunas con péptidos PTE reactivos en todos los monos en la semana 4 después de la inmunización con vectores rAd26 que expresan antígenos Gag, Pol y Env del VIH-1 mosaico, M consenso, clado B + clado C o clado C natural óptimo. Para cada mono, se muestran secuencias de vacunas en la parte superior, y se muestran péptidos PTE reactivos debajo de las secuencias de vacunas indicadas por los números de antígenos (G, Gag; P, Pol; E, Env) y péptidos PTE. La región de solapamiento mínima se muestra en negrita. Se muestran polimorfismos de secuencias entre los dos antígenos mosaico o los dos antígenos de clado B + clado C en azul. Las diferencias entre las secuencias de vacuna y los

péptidos PTE reactivos se muestran en rojo. La Figura 22 desvela las SEQ ID NO: 124-640, respectivamente, en orden de aparición.

Las regiones mínimas en los péptidos que probablemente contengan el epítipo de respuesta inmunitaria, basándose en el solapamiento entre péptidos reactivos cuando se produce, están en negrita en las vacunas. Si no hay ningún péptido solapante, se supone que el epítipo puede estar en cualquier parte del péptido, de modo que la región completa está en negrita. No se puede diferenciar entre diferentes respuestas de linfocitos T que se dirigen a epítopos con diferentes límites en un péptido, o respuestas de linfocitos T clonales promiscuos que pueden tolerar la variación cuando estén presentes variantes; cualquiera de los escenarios podría ser ventajoso en una respuesta inmunitaria de vacuna. El número de regiones diana corresponde al número mínimo de respuestas de linfocitos T necesario para explicar los datos.

Los aminoácidos en los que la vacuna y los péptidos no coinciden se escriben en rojo; si quedan en la región que probablemente porte el epítipo, están en rojo en negrita. Las diferencias de aminoácidos fuera de las regiones solapantes cuando múltiples péptidos solapan se marcan en rojo, pero en negrita.

Las vacunas siempre están en la parte superior. La letra para cada proteína (Gag es G, Pol es P, Envoltura es E) y el número de péptido se usan para marcar cada péptido PTE reactivo. Los números de proteínas y HXB2 están detrás de cada péptido.

Para las vacunas de mosaico y clado B+C, hay 2 antígenos de cada una y ambos se incluyen en el alineamiento; las diferencias de aminoácidos en las vacunas se indican en azul y, si el péptido reactivo porta el aminoácido variante en el segundo mosaico, también está en azul. En cada una de las posiciones donde los dos antígenos de vacuna difieren, los péptidos reactivos también se marcan en negrita para indicar las posiciones donde la inclusión de dos variantes puede haber influido en la respuesta inmunitaria a la vacuna y haber permitido mayor amplitud y profundidad.

Por ejemplo, la primera vacuna resumida es la vacuna de clado B+C, y el animal 287-95 es el primer animal para el que se enumeran respuestas. Hubo 3 respuestas de CD8 a péptidos PTE, 1 para CD4. Dos de los péptidos CD8 muestran solapamiento sustancial, E26 y E282, de modo que ambos pueden ser dianas para la misma respuesta de CTL; por lo tanto, también se observa que solo hay 2 regiones sensibles a CD8 y 1 región sensible a CD4. Para cada región sensible, se escribe el número de péptidos solapantes por región (por ejemplo, CD8: 12 CD4: 1) para evaluar la profundidad de las respuestas; el dos es rojo para indicar que la región de solapamiento es variable en los péptidos reactivos. Si la vacuna difiere, como el D/E en la segunda región reactiva, se marca en azul. Solamente la región de solapamiento está en negrita. La H en E282 no se encontró en ninguna de las vacunas de modo que está marcado en rojo; está dentro de la región de solapamiento de modo que está en negrita. Cada péptido reactivo tiene su proteína y numeración de HXB2 correspondiente indicado a la derecha.

Las Figuras 23A-23C son gráficos que muestran la magnitud de todas las respuestas de linfocitos CD8+ (Figura 23A y 23B) y CD4+ (Figura 23C) específicas de Gag, Pol y Env dispuestas de menor a mayor.

Las Figuras 24A-C muestran la profundidad de respuestas de linfocitos T específicas de epítipo a péptidos PTE. La Figura 24A es un esquema que muestra un ejemplo de respuestas de linfocitos T mapeadas en el mono 366 que recibieron los antígenos del clado C natural óptimo. La Figura 24B es un esquema que muestra un ejemplo de respuestas de linfocitos mapeados en el mono 361 que recibieron los antígenos mosaico bivalentes. En las Figuras 24A y 24B, se muestran secuencias de vacunas en la parte superior (OptC; Mos1, Mos2) y se muestran péptidos PTE reactivos bajo las secuencias de vacunas indicadas por el antígeno (G, Gag; P, Pol; E, Env) y los números de péptidos PTE. La región de solapamiento mínima se muestra en negrita. Se muestran polimorfismos de secuencias entre los dos antígenos mosaico en azul. Las diferencias entre las secuencias de vacuna y los péptidos PTE reactivos se muestran en rojo. Se muestran los alineamientos completos de todos los péptidos positivos organizados por regiones de respuesta en la Fig. 22. La Figura 24C es un gráfico que muestra la profundidad de respuestas de linfocitos T CD4+ (superior) y CD8+ (inferior) después de inmunización con vectores rAd26 que expresan antígenos M consenso, clado B + clado C o clado C natural óptimo. Se representan monos individuales en el eje x. Se muestra una variante de respuesta (tono claro) o >1 variantes de respuesta (tono oscuro) para cada región epitópica. La Figura 24A desvela las SEQ ID NO: 641-650, respectivamente, en orden de aparición. La Figura 24B desvela las SEQ ID NO: 651-685, respectivamente, en orden de aparición.

La Figura 25 es un gráfico que muestra la amplitud de respuestas de linfocitos T específicas de epítopos a péptidos Gag del VIH-1 de clados A, B y C. La amplitud de respuestas inmunitarias celulares se evaluó utilizando subgrupos de péptidos solapantes de las siguientes cepas de Gag de VIH-1: clado C DU422, clado C ZM651, consenso C, consenso A y consenso B. Se muestran números de subgrupos positivos después de una única inmunización de vectores rAd26 que expresan antígenos Gag, Pol y Env del VIH-1 mosaico (azul), M consenso (verde), clado B + clado C (morado) o clado C natural óptimo (rojo). Se representan monos individuales en el eje x.

La Figura 26A-D son gráficos que muestran las respuestas inmunitarias celulares y humorales después de la inmunización de refuerzo. Se muestran la magnitud (Fig. 26A) y amplitud (Fig. 26B) de respuestas de linfocitos T individuales en la semana 4 después de la sensibilización (lado izquierdo de cada panel) y en la semana 44 después

del refuerzo (lado derecho de cada panel) para cada mono. Los monos se sensibilizaron en la semana 0 con vectores rAd26 y se reforzaron en la semana 40 con vectores rAd5HVR48 que expresan antígenos Gag, Pol y Env del VIH-1 mosaico, M consenso o clado C natural óptimo. Se representan monos individuales en el eje x. En la Fig. 26A, el rojo indica respuestas de linfocitos T CD8+, el azul indica respuestas de linfocitos T CD4+, las líneas representan respuestas observadas en ambos puntos temporales y los puntos representan respuestas observados solamente en un punto temporal. En la Fig. 26B, los diferentes tonos de cada color reflejan respuestas a los diferentes antígenos (Gag, Pol, Env). La Figura 26C es un gráfico que muestra los títulos de puntos finales de ELISA específicos de Env en las semanas 0, 4 y 44. La Figura 26D es un gráfico que muestra los títulos de anticuerpos neutralizantes (NAb) para virus del clado A (DJ263.8), clado B (SF162.LS) y clado C (MW965.26) de nivel 1 en la semana 44. Los títulos de NAb para virus de leucemia murina como un control negativo fueron <20 para todas las muestras.

La Figura 27 es un gráfico que muestra la cobertura teórica de péptidos PTE por los diversos antígenos de vacuna. Se muestra el porcentaje de péptidos PTE de 9 aminoácidos que están abarcados por los antígenos Gag, Pol y Env del VIH-1 mosaico (azul), M consenso (verde), clado B + clado C (morado) o clado C natural óptimo (rojo).

Descripción detallada de la invención

La invención presenta polipéptidos víricos optimizados que se obtienen por ordenador de productos génicos víricos de origen natural. Los polipéptidos víricos optimizados de la invención permiten un aumento de la amplitud y profundidad de inmunidad específica de virus (por ejemplo, inmunidad celular, tal como respuestas inmunitarias basadas en linfocitos T) después de inmunización de un sujeto (por ejemplo, un ser humano) con uno o más polipéptidos víricos optimizados de la invención o vacunas (por ejemplo, un vector) que incorporan uno o más polipéptidos víricos optimizados de la invención. La invención proporciona vacunas que pueden administrarse a un sujeto (por ejemplo, un ser humano) infectado con o en riesgo de infectarse con una infección vírica. Las vacunas de la invención pueden incorporar al menos dos polipéptidos víricos optimizados definidos para cada producto génico vírico correspondiente representado. La incorporación de al menos dos polipéptidos víricos optimizados definidos permite aumento de la cobertura y representación de epítomos inmunogénicos en la vacuna, que los inventores han descubierto que dan como resultado un aumento del número total de respuestas inmunitarias específicas de virus después de la vacunación de un sujeto. También se desvelan métodos de administración y fabricación de vacunas, vectores y polipéptidos víricos optimizados a un sujeto (por ejemplo, un ser humano). Las composiciones, métodos y kits descritos en el presente documento pueden aumentar sustancialmente la diversidad, amplitud y/o profundidad de las respuestas inmunitarias celulares específicas de virus proporcionando al menos dos polipéptidos víricos optimizados definidos.

Polipéptidos víricos optimizados de la invención

En determinadas realizaciones, la presente invención proporciona vacunas polivalentes (por ejemplo, bivalentes) que incorporan polipéptidos víricos optimizados por ordenador que corresponden a y se obtienen de productos génicos víricos que circulan de forma natural. Se ensamblan proteínas mosaico polivalentes de secuencias naturales por recombinación informática y se optimizan para proporcionar máxima cobertura de epítomos de linfocitos T (PTE) potenciales para una valencia dada. Los antígenos mosaico son proteína de longitud completa que se diseñan para conservar la expresión y el procesamiento de antígenos naturales.

Los inventores han descubierto que la inmunización con dos polipéptidos víricos optimizados definidos correspondientes a y procedentes de un único producto génico vírico (es decir, una vacuna bivalente) induce un número sustancialmente mayor de respuestas inmunitarias celulares (por ejemplo, respuestas de linfocitos T) que vacunas monovalentes o polivalentes convencionales que incorporan polipéptidos de origen natural procedentes del mismo producto génico vírico (por ejemplo, secuencias basadas en aislados clínicos) o una secuencia consenso de dichos polipéptidos de origen natural procedentes del mismo producto génico vírico. En consecuencia, una vacuna que incorpora polipéptidos víricos optimizados por ordenador, cuyas secuencias proporcionan cobertura máxima de tramos cortos que no sean poco habituales de secuencias víricas en circulación, puede aumentar la amplitud y profundidad de la respuesta inmunitaria.

Se usa un algoritmo genético para crear conjuntos de polipéptidos víricos optimizados como mezclas "mosaico" de fragmentos de un conjunto arbitrario de secuencias de productos génicos víricos de origen natural proporcionadas como entradas. Esta estrategia de algoritmo genético usa secuencias proteicas desalineadas de una población vírica general como un conjunto de datos de entrada y por lo tanto tiene la virtud de ser "independiente de alineamiento". Crea polipéptidos víricos optimizados artificiales que se asemejan a proteínas víricas halladas en la naturaleza, pero no son de origen natural. El algoritmo genético puede ajustarse para optimizar polipéptidos víricos de diferentes longitudes, dependiendo de la diana pretendida o respuesta inmunitaria deseada. Ya que la mayoría de epítomos de linfocitos T son de nueve aminoácidos de longitud, el algoritmo genético utilizado para diseñar los polipéptidos víricos optimizados de la invención se basó en la optimización de cada secuencia de aminoácidos noamérica consecutiva de un producto génico vírico dado (por ejemplo, Gag de VIH-1). De acuerdo con este enfoque, los nonámeros (por ejemplo) que no existen en la naturaleza o que son muy poco habituales pueden excluirse - esta es una mejora en relación con estrategias de vacunas basadas en secuencia consenso ya que estas últimas pueden contener algunos nonámeros (por ejemplo) que aparecen poco o nada en la naturaleza. La definición de aptitud usada para el algoritmo

genético es que el cóctel polivalente más "apto" es la combinación de secuencias víricas de entrada que proporciona la mejor cobertura (mayor fracción de coincidencias perfectas) de todos los nonámeros en la población y está sometido a la restricción de que ningún nonámero está ausente o es poco habitual en la población. El algoritmo genético usado para generar los polipéptidos víricos optimizados de la invención se describe adicionalmente en la publicación de solicitud de patente internacional WO 2007/024941.

En una realización, la invención proporciona vacunas polivalentes (por ejemplo, bivalentes) del VIH-1 que incorporan el polipéptido del VIH-1 optimizado expuesto en la SEQ ID NO: 9). En determinadas realizaciones, la invención presenta una vacuna polivalente que incorpora dos o más polipéptidos del VIH-1 optimizados. En cada caso, los polipéptidos del VIH-1 optimizados se basan en todas las variantes del VIH-1 en circulación global, conocidas como el grupo principal (M) del VIH-1. Los inventores han generado un conjunto de polipéptidos del VIH-1 optimizados (SEQ ID NO: 1-29) que aumentan la amplitud y profundidad de la inmunidad celular basándose en genes de mosaico del grupo M que utilizan solamente dos variantes por gen (por ejemplo, dos secuencias polipeptídicas para cada uno de Gag, Pol, Env, Nef, Tat, Rev, Vif, Vpr y Vpu). Se ha obtenido el resultado nuevo y sorprendente en macacos Rhesus de que el uso de estos polipéptidos del VIH-1 en una vacuna polivalente (por ejemplo, bivalente) del grupo M del VIH-1 induce una amplitud y profundidad significativamente mayores de respuestas inmunitarias celulares específicas del VIH-1 en comparación con otras dos estrategias de antígenos de vacunas principales (antígenos de M consenso y antígenos de clado C naturales óptimos).

La invención también presenta la fusión de polipéptidos víricos optimizados que corresponden a diferentes productos génicos víricos. El algoritmo genético descrito anteriormente puede usarse para generar polipéptidos fusionados para su uso en una vacuna de la invención. Por ejemplo, las fusiones de polipéptidos del VIH-1 optimizadas de Gag/Nef (SEQ ID NO: 19-20), Gag/Pol (SEQ ID NO: 21-27) y Gag/Pol/Nef (SEQ ID NO: 28-29) pueden incorporarse en un vector de la invención para administración a un sujeto (por ejemplo, un ser humano) infectado con o en riesgo de infectarse con VIH-1. Las vacunas de la invención (bien en forma de polipéptido o de ácido nucleico) también pueden incluir uno o más polipéptidos no "mosaico" (o secuencias que los codifican, respectivamente), tales como, por ejemplo, las secuencias del clado C óptimas (SEQ ID NO: 30-36) o las secuencias consenso (SEQ ID NOS: 37-39).

Los polipéptidos víricos optimizados desvelados en la presente invención pueden prepararse convencionalmente por técnicas de síntesis química, tales como las descritas por Merrifield, J. Amer. Chem. Soc. 85:2149 (1963) (véase también, por ejemplo, Stemmer *et al.*, 164 Gene 49 (1995)). Por ejemplo, las vacunas pueden prepararse fácilmente usando síntesis peptídica de fase sólida (SPPS). Puede realizarse síntesis de fase sólida automática usando uno cualquiera de varios sintetizadores automáticos disponibles en el mercado, bien conocidos, tales como el sintetizador de péptidos Applied Biosystems ABI 433A. Como alternativa, los polipéptidos víricos optimizados de la invención pueden producirse de forma recombinante transfectando o transduciendo una célula o un organismo con un ácido nucleico o vector (por ejemplo, un vector vírico, como un adenovirus) que permite la expresión intracelular del polipéptido vírico optimizado. Pueden sintetizarse ácidos nucleicos y vectores que codifican la secuencia de nucleótidos de polipéptidos víricos optimizados de la invención mediante técnicas de ADN recombinante bien conocidas, incluyendo los descritos en el presente documento.

Vacunas de la invención

La invención también presenta vacunas que pueden administrarse a un paciente infectado con o en riesgo de infectarse con un virus (por ejemplo, VIH-1). Una vacuna de la invención contiene al menos uno de los polipéptidos víricos optimizados de la invención, como se analiza en el presente documento, incluyendo el polipéptido vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9. La vacuna de la invención puede ser un ácido nucleico que codifica la secuencia de nucleótidos de dos o más polipéptidos víricos optimizados de la invención (por ejemplo, el componente inmunogénico de un vector vírico recombinante (por ejemplo, subunidad) o de organismo completo (por ejemplo, virus completo)). Los ácidos nucleicos incluyen vectores (por ejemplo, vectores víricos, tales como adenovirus) que incorporan la secuencia de nucleótidos de dos o más polipéptidos víricos optimizados de la invención. Los polipéptidos víricos optimizados de la invención, así como vacunas, ácidos nucleicos y vectores que incorporan polipéptidos víricos optimizados, pueden expresarse de forma recombinante en una célula o un organismo o pueden administrarse directamente al sujeto (por ejemplo, un ser humano) infectado con, o en riesgo de infectarse con, un virus.

Vectores de la invención

La invención también presenta vectores que codifican las secuencias de nucleótidos (por ejemplo, ADN o ARN) de uno o más polipéptidos víricos optimizados de la invención, incluyendo un vector que codifica un polipéptido vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9. El vector puede ser un vehículo (por ejemplo, un liposoma), un plásmido, un cósmido, un cromosoma artificial de levadura o un virus que incluye una secuencia de nucleótidos que codifica uno o más polipéptidos víricos optimizados de la invención. El vector puede incluir secuencias de ácido nucleico adicionales de varias fuentes.

Pueden construirse vectores que codifican uno o más polipéptidos víricos optimizados de la invención usando cualquier técnica de biología molecular recombinante conocida en este campo. El vector, tras la transfección o

transducción de una célula u organismo diana, puede ser extracromosómico o puede estar integrado en el cromosoma de la célula hospedadora. El componente de ácido nucleico de un vector puede estar en un número de copias individual o múltiple por célula diana y puede ser lineal, circular o concatemerizado.

5 Los vectores de la invención también pueden incluir secuencias de sitios de entrada de ribosoma internos (IRES) para permitir la expresión de múltiples cadenas peptídicas o polipeptídicas de un único transcrito de ácido nucleico. Por ejemplo, un vector de la invención puede codificar uno o más polipéptidos víricos optimizados de la invención, así como otro polipéptido (por ejemplo, un marcador detectable, tal como proteína verde fluorescente (GFP)).

10 Los vectores de la invención incluyen además elementos de expresión génica que facilitan la expresión de polipéptidos víricos optimizados de la invención. Los elementos de expresión génica útiles para la expresión de un vector que codifica un polipéptido vírico optimizado de la invención incluyen, pero sin limitación, (a) secuencias reguladoras, tales como promotores de transcripción vírica y sus elementos potenciadores, tales como el promotor temprano de SV40, LTR del virus del sarcoma de Rous y LTR del virus de leucemia murina de Moloney; (b) regiones de corte y empalme y sitios de poliadenilación tales como los procedentes de la región tardía de SV40; y (c) sitios de poliadenilación tales como en SV40. También se incluyen orígenes de replicación plasmídicos, genes de resistencia a antibióticos o selección, múltiples sitios de clonación (por ejemplo, loci de escisión de enzimas de restricción) y otras secuencias génicas víricas (por ejemplo, secuencias que codifican elementos estructurales, funcionales o reguladores víricos, tales como la repetición terminal larga (LTR) del VIH).

20 Los vectores de la invención también pueden incluir polipéptidos víricos optimizados de la invención que se han optimizado para expresión en seres humanos, tales como, por ejemplo, una cualquiera de las SEQ ID NO: 11, 14-18 y 23.

25 Los vectores de la invención también pueden modificarse técnicamente para incluir un sitio de clonación múltiple (MCS) que tiene los siguientes sitios de escisión enzimática: XbaI-EcoRI-Kozak-Inicio...Detención-BamHI-NheI; y la siguiente secuencia: TCTAGA GAATTC GCCACC [ATG gen TAA TGA] GGATCC GCTAGC. Pueden usarse vectores que tienen este MCS con polipéptidos víricos optimizados que no tienen ningún sitio XbaI, EcoRI, BamHI, NheI interno y ningún tramo de 6 o más C o G.

30 *Administración in vivo*

35 Las vacunas de la invención son útiles en métodos para la administración *in vivo* de una o más vacunas de la invención (por ejemplo, un vector que codifica dos o más polipéptidos víricos optimizados de la invención) a un sujeto (por ejemplo, un ser humano) para facilitar la expresión de dos o más polipéptidos víricos optimizados de la invención. Tras administrar la vacuna al sujeto, se expresarán uno o más polipéptidos víricos optimizados de la invención que pueden inducir respuestas inmunitarias protectoras o terapéuticas (por ejemplo, respuestas inmunitarias celulares o humorales) dirigidas contra los inmunógenos víricos.

40 Pueden emplearse varios tipos de vectores para suministrar una secuencia de nucleótidos que codifica uno o más polipéptidos víricos optimizados de la invención directamente a un sujeto (por ejemplo, un ser humano). Los vectores de la invención incluyen virus, ADN desnudo, oligonucleótidos, lípidos catiónicos (por ejemplo, liposomas), polímeros catiónicos (por ejemplo, polisomas), virosomas y dendrímeros.

45 *Transfección y transducción ex vivo*

50 Las vacunas de la invención también pueden usarse en métodos para la transfección y transducción de células *ex vivo* (por ejemplo, células sanguíneas, tales como linfocitos), seguido de la administración de estas células a un sujeto (por ejemplo, un ser humano). Las células pueden ser autólogas para el sujeto tratado. Las células pueden transfectarse o transducirse *ex vivo* con uno o más vectores que codifican la secuencia de nucleótidos de uno o más polipéptidos víricos optimizados de la invención para permitir la expresión temporal o permanente de los polipéptidos víricos optimizados en el sujeto tratado. Tras administrar estas células modificadas al sujeto, se expresarán uno o más vectores víricos optimizados de la invención que pueden inducir respuestas inmunitarias protectoras o terapéuticas (por ejemplo, respuestas inmunitarias celulares o humorales) dirigidas contra los inmunógenos víricos.

55 Pueden emplearse varios tipos de vectores para suministrar una secuencia de nucleótidos que codifica uno o más polipéptidos víricos optimizados de la invención a una célula (por ejemplo, una célula sanguínea, tal como un linfocito). Los vectores de la invención incluyen virus, ADN desnudo, oligonucleótidos, lípidos catiónicos (por ejemplo, liposomas), polímeros catiónicos (por ejemplo, polisomas), virosomas y dendrímeros. Las vacunas de la presente invención también pueden usarse en métodos para la transfección o transducción *ex vivo* de células (por ejemplo, células sanguíneas) seguido de la administración de estas células de vuelta al sujeto donante para permitir la expresión de polipéptidos víricos optimizados de la invención que tienen propiedades inmunogénicas. Las células que pueden aislarse y transfectarse o transducirse *ex vivo* según estos métodos incluyen, pero sin limitación, células sanguíneas, células cutáneas, fibroblastos, células endoteliales, células de músculo esquelético, hepatocitos, células epiteliales de próstata y células endoteliales vasculares. Las células madre también son células apropiadas para transducción o transfección con un vector de la invención. Las células totipotentes, pluripotentes, multipotentes o unipotentes,

incluyendo células progenitoras de médula ósea y células madre hematopoyéticas (HSC), pueden aislarse y transfectarse o transducirse con un vector que codifica uno o más polipéptidos víricos optimizados de la invención y administrarse a un sujeto según estos métodos.

5 El método de transfección o transducción usado para expresar un vector vírico optimizado de la invención tiene una fuerte influencia en la fuerza y longevidad de la expresión proteica en la célula transfectada o transducida y, posteriormente, en el sujeto que recibe la célula. La presente invención proporciona vectores que son de naturaleza temporal (por ejemplo, vectores adenovíricos) o de larga duración (por ejemplo, vectores retrovíricos). Se conocen en la técnica secuencias reguladoras (por ejemplo, promotores y potenciadores) que pueden usarse para regular la expresión de proteínas. El tipo de célula que se transfecta o transduce también tiene una fuerte influencia en la fuerza y longevidad de la expresión de proteínas. Por ejemplo, puede esperarse que los tipos celulares con altas tasas de renovación tengan periodos más cortos de expresión de proteínas.

15 *Vectores víricos*

Pueden usarse vectores víricos que codifican la secuencia de nucleótidos de uno o más polipéptidos víricos optimizados de la invención como una vacuna de la invención. Por ejemplo, la secuencia de nucleótidos de uno o más polipéptidos víricos optimizados de la invención puede insertarse de forma recombinante en la de un genoma vírico natural o modificado (por ejemplo, atenuado) adecuado para la transducción de un sujeto (por ejemplo, administración *in vivo*) o células aisladas de un sujeto (por ejemplo, para transducción *ex vivo* seguida de administración de las células de nuevo al sujeto). Pueden realizarse modificaciones adicionales en el virus para mejorar la infecciosidad o el tropismo (por ejemplo, pseudotipado), reducir o eliminar la competencia replicativa o reducir la inmunogenicidad de los componentes víricos (por ejemplo, todos los componentes no relacionados con el agente de vacuna inmunogénica). Un vector de la invención puede ser expresado por la célula transducida y secretada al espacio extracelular o permanecer con la célula de expresión (por ejemplo, como una molécula intracelular o presentada en la superficie celular). También pueden usarse vectores víricos quiméricos o pseudotipados para transducir una célula para permitir la expresión de uno o más polipéptidos víricos optimizados de la invención. Se describen posteriormente vectores ejemplares.

30 *Adenovirus*

Los adenovirus recombinantes ofrecen varias ventajas significativas para su uso como vectores para la expresión de uno o más polipéptidos víricos optimizados de la invención. Los virus pueden prepararse hasta alta titulación, pueden infectar células no replicativas y pueden conferir transducción de alta eficacia de células diana *ex vivo* después de contacto con una población de células diana. Asimismo, los adenovirus no integran su ADN en el genoma del hospedador. Por tanto, su uso como vectores de expresión tiene un riesgo reducido de inducción de trastornos proliferativos espontáneos. En modelos animales, se ha descubierto en general que los vectores adenovíricos median en la expresión a alto nivel durante aproximadamente una semana. La duración de la expresión transgénica (expresión de un ácido nucleico que codifica un polipéptido vírico optimizado de la invención) puede prolongarse usando promotores específicos de célula o tejido. Otras mejoras en la ingeniería molecular del vector adenovírico en sí mismo han producido expresión transgénica más sostenida y menos inflamación. Esto se ve con vectores denominados de "segunda generación" que albergan mutaciones específicas en genes adenovíricos tempranos adicionales y vectores "vacíos" en los que prácticamente todos los genes víricos se suprimen utilizando una estrategia de Cre-Lox (Engelhardt *et al.*, Proc. Natl. Acad. Sci. USA 91:6196 (1994) y Kochanek *et al.*, Proc. Natl. Acad. Sci. USA 93:5731 (1996)).

Los vectores adenovíricos de serotipo poco habitual y quiméricos divulgados en las publicaciones de solicitud de patente internacional WO 2006/040330 y WO 2007/104792, son particularmente útiles como vectores de la invención. Por ejemplo, los adenovirus recombinantes rAd26, rAd34, rAd35, rAd48 y rAd5HVR48 pueden codificar uno o más polipéptidos víricos optimizados de la invención. Uno o más vectores víricos recombinantes que codifican polipéptidos víricos optimizados de la invención pueden administrarse a un sujeto para tratar o prevenir una infección vírica.

50 *Virus adenoasociados (VAA)*

Los virus adenoasociados (VAAr), procedentes de parvovirus no patógenos, también pueden usarse para expresar polipéptidos víricos optimizados de la invención ya que estos vectores no inducen casi ninguna respuesta inmunitaria celular antivector y producen expresión transgénica que dura meses en la mayoría de sistemas experimentales.

Retrovirus

60 Los retrovirus son útiles para la expresión de polipéptidos víricos optimizados de la invención. A diferencia de adenovirus, el genoma retrovírico se basa en ARN. Cuando un retrovirus infecta una célula, introducirá su ARN junto con varias enzimas en la célula. Las moléculas de ARN vírico del retrovirus producirán una copia de ADN bicatenario, denominada provirus, mediante un proceso denominado transcripción inversa. Después del transporte al núcleo celular, el ADN provírico se integra en un cromosoma de célula hospedadora, alterando de forma permanente el genoma de la célula transducida y cualquier célula descendiente que puede proceder de esta célula. La capacidad de introducir de forma permanente un gen en una célula u organismo es la característica definitoria de retrovirus usados

para terapia génica. Los retrovirus incluyen lentivirus, una familia de virus que incluye virus de la inmunodeficiencia humana (VIH) que incluye varias proteínas accesorias para facilitar la infección vírica e integración provírica. Los vectores lentivíricos, de "tercera generación", actuales, presentan incompetencia de replicación total, tropismo amplio y capacidad de transferencia génica aumentada para células de mamífero (véase, por ejemplo, Mangeat y Trono, *Human Gene Therapy* 16(8):913 (2005) y Wiznerowicz y Trono, *Trends Biotechnol.* 23(1):42 (2005)).

Otros vectores víricos

Además de vectores adenovíricos y retrovíricos, se conocen en este campo otros vectores víricos y técnicas que pueden usarse para expresar polipéptidos víricos optimizados de la invención en una célula (por ejemplo, una célula sanguínea, tal como un linfocito) o sujeto (por ejemplo, un ser humano). Estos virus incluyen poxvirus (por ejemplo, virus vaccinia y virus vaccinia modificado Ankara o (MVA); véase, por ejemplo, las patentes de los Estados Unidos n.º 4.603.112 y 5.762.938), herpesvirus, togavirus (por ejemplo, virus de la encefalitis equina venezolana; véase, por ejemplo, la patente de los Estados Unidos n.º 5.643.576), picornavirus (por ejemplo, poliovirus; véase, por ejemplo, la patente de los Estados Unidos n.º 5.639.649), baculovirus, y otros descritos por Wattanapitayakul y Bauer (*Biomed. Pharmacother.* 54:487 (2000)).

Otros vectores de expresión: ADN desnudo y oligonucleótidos

También pueden usarse ADN desnudo u oligonucleótidos que codifican uno o más polipéptidos víricos optimizados de la invención para expresar estos polipéptidos en una célula (por ejemplo, una célula sanguínea, tal como un linfocito) o sujeto (por ejemplo, un ser humano). Véase, por ejemplo, Cohen, *Science* 259:1691-1692 (1993); Fynan *et al.*, *Proc. Natl. Acad. Sci. USA*, 90:11478 (1993); y Wolff *et al.*, *BioTechniques* 11:474485 (1991). Este es el método más sencillo de transfección no vírica. Existen métodos eficaces para el suministro de ADN desnudo tales como electroporación y el uso de una "pistola génica", que dispara partículas de oro recubiertas con ADN en una célula usando gas de alta presión y partículas transportadoras (por ejemplo, oro).

Lipoplejos y poliplejos

Para mejorar el suministro de un ácido nucleico que codifica un polipéptido vírico optimizado de la invención en una célula o sujeto, pueden usarse lipoplejos (por ejemplo, liposomas) y poliplejos para proteger el ADN del vector de degradación indeseable durante el proceso de transfección. El ADN plasmídico puede cubrirse con lípidos en una estructura organizada como una micela o un liposoma. Cuando la estructura organizada forma un complejo con ADN, se denomina lipoplejo. Hay tres tipos de lípidos, aniónicos (con carga negativa), neutros o catiónicos (con carga positiva). Los lipoplejos que utilizan lípidos catiónicos han demostrado utilidad para la transferencia génica. Los lípidos catiónicos, debido a su carga positiva, forman complejo de forma natural con el ADN de carga negativa. Además, como resultado de su carga interaccionan con la membrana celular, se produce endocitosis del lipoplejo y el ADN se liberan al citoplasma. Los lípidos catiónicos también protegen contra la degradación del ADN por la célula.

Los complejos de polímeros con ADN se denominan poliplejos. La mayoría de los poliplejos consisten en polímeros catiónicos y su producción se regula por interacciones iónicas. Una gran diferencia entre los métodos de acción de poliplejos y lipoplejos es que los poliplejos no pueden liberar su carga de ADN al citoplasma, de modo que, con este fin, debe producirse cotransfección con agentes líticos de endosomas (para lisar el endosoma que se realiza durante la endocitosis) tales como adenovirus inactivados. Sin embargo, esto no siempre sucede; los polímeros tales como polietilimina tienen su propio método de alteración de endosomas, como también quitosano y trimetilquitosano.

Los lípidos catiónicos y polímeros ejemplares que pueden usarse en combinación con un ácido nucleico que codifica un polipéptido vírico optimizado de la invención para formar lipoplejos o poliplejos incluyen, pero sin limitación, polietilimina, lipofectina, lipofectamina, polilisina, quitosano, trimetilquitosano y alginato.

Métodos híbridos

Varios métodos híbridos de transferencia génica combinan dos o más técnicas. Los virosomas, por ejemplo, combinan lipoplejos (por ejemplo, liposomas) con un virus inactivado. Se ha mostrado que este enfoque da como resultado transferencia génica más eficaz en células epiteliales respiratorias que métodos víricos o liposómicos solos. Otros métodos implican la mezcla de otros vectores víricos con lípidos catiónicos o virus de hibridación. Cada uno de estos métodos puede usarse para facilitar la transferencia de un ácido nucleico que codifica polipéptidos víricos optimizados de la invención en una célula (por ejemplo, una célula sanguínea, tal como un linfocito) o sujeto (por ejemplo, un ser humano).

Dendrimeros

Los dendrimeros también pueden usarse para transferir un ácido nucleico que codifica un polipéptido vírico optimizado de la invención en una célula (por ejemplo, una célula sanguínea, tal como un linfocito) o sujeto (por ejemplo, un ser humano). Un dendrímero es una macromolécula altamente ramificada con una forma esférica. La superficie de la partícula puede funcionalizarse de muchas maneras y muchas de las propiedades de la construcción resultante se

determinan por su superficie. En particular es posible construir un dendrímero catiónico (es decir uno con una carga superficial positiva). Cuando está en presencia de material genómico tal como ADN o ARN, la complementariedad de la carga conduce a una asociación temporal del ácido nucleico con el dendrímero catiónico. Una vez que alcanza su destino, el complejo de dendrímero-ácido nucleico se toma después en la célula a través de endocitosis.

5 Administración *in vivo*

La invención también presenta vacunas útiles en métodos *in vivo* para inmunizar a un sujeto (por ejemplo, un ser humano). En una realización, una o más vacunas de la invención pueden ser para administración directa a un sujeto para inducir una respuesta inmunitaria protectora o terapéutica (por ejemplo, una respuesta inmunitaria celular o humoral) contra un virus (por ejemplo, VIH-1). Como alternativa, un vector que codifica uno o más polipéptidos víricos optimizados de la invención, como se han descrito anteriormente, puede administrarse directamente a un sujeto para prevenir o tratar una infección vírica. Un vector (por ejemplo, un vector vírico) que transfecta o transduce eficazmente una o más células *in vivo* puede inducir una respuesta inmunitaria amplia, duradera y potente en el sujeto tratado. Tras la transferencia del componente de ácido nucleico del vector de expresión en una célula hospedadora (por ejemplo, una célula sanguínea, tal como un linfocito), la célula hospedadora produce y presenta o secreta la vacuna de la invención, que después actúa para activar componentes del sistema inmunitario tales como células presentadoras de antígenos (APC), linfocitos T y linfocitos B, que dan como resultado el establecimiento de la inmunidad.

20 Composiciones farmacéuticas

La invención también presenta las vacunas, vectores y polipéptidos víricos optimizados de la invención en combinación con uno o más excipientes, diluyentes, tampones u otros vehículos aceptables farmacéuticamente aceptables. La formulación de una vacuna, vector o polipéptidos víricos optimizados empleará o permitirá la expresión de una cantidad eficaz de inmunógeno polipeptídico vírico optimizado. Es decir, se incluirá una cantidad de antígeno que provocará que el sujeto tratado (por ejemplo, un ser humano) produzca una respuesta inmunológica específica y suficiente para transmitir protección al sujeto para la exposición posterior a un virus (por ejemplo, VIH-1) o para tratar una infección vírica existente. Por ejemplo, una formulación de una vacuna de la invención puede permitir la expresión de una cantidad de antígeno que provocará que el sujeto produzca una respuesta inmunitaria celular amplia y específica. Un sujeto tratado con una vacuna, vector o polipéptido vírico optimizado de la invención también puede producir anticuerpos antivíricos (por ejemplo, anticuerpos neutralizantes) que pueden conferir un beneficio protector o terapéutico al sujeto. Una vacuna, vector o polipéptido vírico optimizado de la invención puede administrarse directamente a un sujeto, solo o en combinación con cualquier vehículo, sal o adyuvante farmacéuticamente aceptable conocido en la técnica.

Las sales farmacéuticamente aceptables pueden incluir sales de adición de ácidos no tóxicas o complejos metálicos que se usan habitualmente en la industria farmacéutica. Los ejemplos de sales de adición de ácidos incluyen ácidos orgánicos, tales como ácidos acético, láctico, pamoico, maleico, cítrico, málico, ascórbico, succínico, benzoico, palmítico, subérico, salicílico, tartárico, metanosulfónico, toluenosulfónico o trifluoroacético o similares; ácidos poliméricos tales como ácido tánico, carboximetilcelulosa o similares; y ácidos inorgánicos tales como ácido clorhídrico, ácido bromhídrico, ácido sulfúrico, ácido fosfórico o similares. Los complejos metálicos incluyen cinc, hierro, y similares. Un vehículo farmacéuticamente aceptable ejemplar es la solución salina fisiológica. Otros vehículos fisiológicamente aceptables y sus formulaciones son conocidos por un experto en la materia y se describen, por ejemplo, en Remington's Pharmaceutical Sciences, (18ª edición), ed. A. Gennaro, 1990, Mack Publishing Company, Easton, PA.

Las formulaciones farmacéuticas de una cantidad profiláctica o terapéuticamente eficaz de una vacuna, vector o polipéptido vírico optimizado de la invención pueden administrarse por vía oral, parenteral (por ejemplo, inyección intramuscular, intraperitoneal, intravenosa o subcutánea, inhalación, vía intradérmica, gotas ópticas o implante), nasal, vaginal, rectal, sublingual o tópica, en mezcla con un vehículo farmacéuticamente aceptable adaptado para la vía de administración. La concentración de una vacuna, vector o polipéptido vírico optimizado de la invención en la formulación puede variar de aproximadamente 0,1-100 % en peso.

Las formulaciones para administración parenteral de composiciones que contienen una vacuna, vector o polipéptido vírico optimizado de la invención incluyen soluciones, suspensiones o emulsiones acuosas o no acuosas estériles. Los ejemplos de vehículos adecuados incluyen propilenglicol, polietilenglicol, aceites vegetales, gelatina, nafaleno hidrogenados y ésteres orgánicos inyectables, tales como oleato de etilo. Dichas formulaciones también pueden contener adyuvantes, tales como agentes conservantes, humectantes, emulsionantes y dispersantes. Pueden usarse polímeros de lactida biocompatible, biodegradable, copolímero de lactida/glicólido o copolímeros de polioxietileno-polioxipropileno para controlar la liberación de los compuestos. Otros sistemas de suministro parenteral potencialmente útiles para composiciones que contienen una vacuna, vector o polipéptido vírico optimizado de la invención incluyen partículas de copolímeros de etileno-acetato de vinilo, bombas osmóticas, sistemas de infusión implantables y liposomas.

Las formulaciones líquidas pueden esterilizarse mediante, por ejemplo, filtración a través de un filtro que retenga bacterias, incorporando agentes esterilizantes en las composiciones o irradiando o calentando las composiciones.

Como alternativa, también pueden fabricarse en forma de composiciones sólidas, estériles, que se pueden disolver en agua estéril o en algún otro medio inyectable estéril inmediatamente antes de su uso.

Las composiciones que contienen vacuna, vector o polipéptido vírico optimizado de la invención para administración rectal o vaginal son preferentemente supositorios que pueden contener, además de sustancias activas, excipientes tales como manteca de cacao y una cera de supositorio. También se preparan composiciones para administración nasal o sublingual con excipientes convencionales conocidos en la técnica. Las formulaciones para inhalación pueden contener excipientes, por ejemplo, lactosa, o pueden ser soluciones acuosas que contienen, por ejemplo, éter de polioxietilen-9-laurilo, glicocolato y desoxicolato, o pueden ser soluciones oleosas para administración en forma de gotas nasales o nebulizador, o como un gel.

La cantidad del principio activo en las composiciones de la invención puede variarse. Un experto en la materia apreciará que las dosificaciones individuales exactas pueden ajustarse algo dependiendo de una diversidad de factores, incluyendo el péptido que se administra, el tiempo de administración, la vía de administración, la naturaleza de la formulación, la tasa de excreción, la naturaleza de las condiciones del sujeto y la edad, el peso, la salud y el sexo del paciente. Además, la gravedad de la afección tratada por la vacuna, vector o polipéptido vírico optimizado también tendrá una influencia en el nivel de dosificación. En general, se administran niveles de dosificación de entre 0,1 µg/kg y 100 mg/kg de peso corporal diariamente como una única dosis o dividida en múltiples dosis. Preferentemente, el intervalo de dosis general es de entre 250 µg/kg y 5,0 mg/kg de peso corporal al día. Son esperables amplias variaciones de la dosificación necesaria a la vista de las diferentes eficacias de las diversas vías de administración. Por ejemplo, se esperaría en general que la administración oral requiriera mayores niveles de dosificación que la administración por inyección intravenosa. Las variaciones en estos niveles de dosificación pueden ajustarse usando rutinas empíricas convencionales para optimización, que son bien conocidos en la técnica. En general, la dosificación profiláctica o terapéuticamente eficaz precisa puede ser determinada por el especialista clínico a cargo considerando los factores anteriormente identificados.

La cantidad de una vacuna, vector o polipéptido vírico optimizado de la invención presente en cada dosis proporcionada a un paciente se selecciona teniendo en consideración la edad, el peso, el sexo, la condición física general del paciente y similares. La cantidad de una vacuna, vector o polipéptido vírico optimizado necesaria para inducir una respuesta inmunitaria (por ejemplo, una respuesta inmunitaria celular) o producir un efecto exógeno en el paciente sin efectos secundarios adversos significativos varía dependiendo de la composición farmacéutica y la presencia opcional de un adyuvante. Las dosis iniciales pueden seguirse opcionalmente de refuerzos repetidos, cuando sea deseable. El método puede implicar administrar de forma crónica la vacuna, el vector o el polipéptido vírico optimizado de la invención. Para uso terapéutico o uso profiláctico, pueden ser deseables dosificaciones repetidas de la vacuna de inmunización, el vector o el polipéptido vírico optimizado, tales como un refuerzo anual o un refuerzo en otros intervalos. La dosificación administrada variará, por supuesto, dependiendo de factores conocidos tales como las características farmacodinámicas de la vacuna, el vector o el polipéptido vírico optimizado particular, y su modo y vía de administración; la edad, la salud y el peso del receptor; la naturaleza y el alcance de los síntomas, el tipo de tratamiento simultáneo, la frecuencia del tratamiento y la naturaleza del efecto deseado. Una vacuna, vector o polipéptido vírico optimizado de la invención puede administrarse en tratamientos crónicos para sujetos en riesgo de infección aguda debido a pinchazos o infección materna. Una frecuencia de dosificación para dichas infecciones "agudas" puede variar de dosificaciones diarias a una o dos veces a la semana i.v. o i.m., durante una duración de aproximadamente 6 semanas. La vacuna, el vector o polipéptido vírico optimizado también puede emplearse en tratamientos crónicos para pacientes infectados, o pacientes con infección avanzada con un virus (por ejemplo, VIH-1). En pacientes infectados, la frecuencia de administración crónica puede variar de dosificaciones diarias a una o dos veces a la semana i.m. y puede depender de la semivida del inmunógeno presente en la vacuna, el vector o el polipéptido vírico optimizado de la invención.

Adyuvantes

Una vacuna de la invención usada para vacunar a un mamífero (por ejemplo, un ser humano) que lo necesite contra un virus puede administrarse simultáneamente con uno o más adyuvantes farmacéuticamente aceptables para aumentar la inmunogenicidad de la vacuna. Los adyuvantes aprobados para uso humano incluyen sales de aluminio (alumbre). Estos adyuvantes han sido útiles para algunas vacunas incluyendo hepatitis B, difteria, polio, rabia y gripe. Otros adyuvantes útiles incluyen adyuvante completo de Freund (CFA), adyuvante incompleto de Freund (IFA), dipéptido de muramilo (MDP), análogos sintéticos de MDP, N-acetilmuramil-L-alanil-D-isoglutamil-L-alanina-2-[1,2-dipalmitoil-s-glicero-3-(hidroxifosforiloxi)]etilamida (MTP-PE) y composiciones que contienen un aceite metabolizable y un agente emulsionante, en donde el aceite y agente emulsionante están presentes en la forma de una emulsión de aceite en agua que tiene gotas de aceite siendo sustancialmente todos ellos de menos de un micrómetro de diámetro.

Kits

La invención proporciona kits que incluyen una composición farmacéutica que contiene una vacuna, vector o polipéptido vírico optimizado de la invención y un vehículo farmacéuticamente aceptable, en una cantidad terapéuticamente eficaz para prevenir o tratar una infección vírica. Los kits incluyen instrucciones para permitir que un especialista clínico (por ejemplo, un médico o enfermero) para administrar la composición contenida en los mismos.

Preferentemente, los kits incluyen múltiples envases de la composición o las composiciones farmacéuticas de una única dosis que contienen una cantidad eficaz de una vacuna, vector o polipéptido vírico optimizado de la invención. Opcionalmente, pueden incluirse en los kits instrumentos o dispositivos necesarios para administrar la o las composiciones farmacéuticas. Por ejemplo, un kit de la presente invención puede proporcionar una o más jeringas precargadas que contienen una cantidad eficaz de una vacuna, vector o polipéptido vírico optimizado de la invención. Asimismo, los kits también pueden incluir componentes adicionales tales como instrucciones o programas de administración para un paciente infectado con o en riesgo de infectarse con un virus para usar la o las composiciones farmacéuticas que contienen una vacuna, vector o polipéptido vírico optimizado de la invención.

Resultará evidente para los expertos en la materia que pueden hacerse diversas modificaciones y variaciones en las composiciones, métodos y kits de la presente invención sin alejarse del alcance de la invención. Por tanto, se pretende que la presente invención abarque las modificaciones y variaciones de la presente invención siempre que queden dentro del alcance de las reivindicaciones adjuntas y sus equivalentes.

Ejemplos

La presente invención se ilustra mediante los siguientes ejemplos, que no se pretenden en ningún modo que sean limitantes de la invención.

Ejemplo 1

Las secuencias de antígenos de mosaico Gag, Pol, Nef y Env (SEQ ID NO: 1-8) se construyeron usando el algoritmo genético analizado anteriormente. Estas secuencias se modificaron después para hacerlas prácticas para el desarrollo de vacunas eliminando la actividad de escisión/fusión en Env (SEQ ID NO: 9-11), eliminando la actividad catalítica en Pol (SEQ ID NO: 12-14), eliminando sitios de miristilación en Nef (SEQ ID NO: 16-18) y construyendo construcciones de fusión incluyendo GagNef, GagPol o GagPolNef (SEQ ID NO: 19-29). También se representan los genes de clado C naturales óptimos comparadores (SEQ ID NO: 30-36).

Ejemplo 2

Se inmunizaron monos rhesus con 3×10^{10} vectores rAd26 vp que expresan genes Gag, Pol y Env de secuencias de M consenso (Grupo 1), M mosaico bivalente (Grupo 2) o clado C natural óptimo (Grupo 3). Las secuencias de M consenso representan secuencias sintéticas que representan el mejor "promedio" de virus en circulación en todo el mundo. Las secuencias de M mosaico bivalentes se han descrito anteriormente. Las secuencias del clado C naturales óptimas son secuencias de origen natural de virus VIH-1 del clado C reales que tienen el carácter más de "tipo consenso". Se evaluó la amplitud inmunitaria celular evaluando el número de péptidos de respuesta del conjunto de péptidos de epítomos de linfocitos T (PTE) potenciales globales. Los péptidos PTE representan >85 % de secuencias del VIH-1 globales y están libremente disponibles del NIH.

Los resultados muestran que las nuevas secuencias de M mosaico bivalentes rindieron drásticamente mejor que estos otros dos conceptos antigénicos principales. Como se muestra en la Tabla 1, los antígenos de M mosaico bivalentes indujeron un aumento significativo de la amplitud de respuestas de linfocitos T específicas de Gag, específicas de Env, específicas de Pol y totales en comparación con antígenos de M consenso y antígenos de clado C naturales óptimos. (La media representa el n.º promedio de epítomos en cada grupo de monos; ETM representa el error típico de la media).

Tabla 1: Los antígenos Gag/Pol/Env del VIH-1 mosaico expanden la amplitud contra péptidos PTE globales en monos rhesus

Amplitud	Grupo I: M Consenso		Grupo II: M Mosaico bivalente		Grupo III: Clado C natural	
	Media	ETM	Media	ETM	Media	ETM
Gag	2,0	0,4	7,7	0,9	2,2	0,5
Env	2,0	0,4	4,0	0,6	1,6	0,5
Pol	2,7	0,5	8,1	1,4	2,4	0,5
Total	6,7	0,7	19,9	1,9	6,1	1,1

Ejemplo 3

Se inmunizaron macacos IM con 3×10^{10} vectores rAd26 vp que expresan genes Gag, Pol y Env de secuencias de M consenso (Grupo 1; n=7), M Mosaico bivalente (Grupo 2; n=7) o clado C natural óptimo (Grupo 3; n=6) descritas en el Ejemplo 2. Se evaluó la amplitud inmunitaria celular evaluando el número de péptidos de respuesta del conjunto de péptidos de epítomos de linfocitos T (PTE) potenciales globales.

Como una lectura, se evaluaron las respuestas de Elispot de IFN γ CD4/CD8 a péptidos PTE agrupados (magnitud). Los epítomos se mapearon exhaustivamente usando péptidos PTE de 15 unidades para evaluar el número de positivos (los positivos se definieron como 55 células formadoras de puntos (CFP) por cada 10^6 PBMC y 4x fondo). También se

ensayaron conjuntos agrupados de péptidos solapantes que abarcan 5 proteínas Gag para comparar las respuestas a un conjunto de proteínas completas.

5 Los resultados muestran que las secuencias de M mosaico bivalentes rindieron drásticamente mejor que los otros dos conceptos antigénicos principales (Mcon y OptC).

Ejemplo 4

10 Se usó modelización para validar la observación de los inventores de que las respuestas de linfocitos T aumentan como resultado de la vacuna de mosaico. Los inventores ajustaron modelos de regresión de Poisson que predijeron el número de péptidos reactivos en función de la vacuna, el polipéptido y el tipo de linfocito T y después realizaron una eliminación por etapas de interacciones. Los inventores observaron que, aunque la vacuna de mosaico produjo una potenciación altamente significativa en el número de respuestas de PTE positivas, lo hizo de forma más o menos uniforme en todas las poliproteínas y tipos de linfocitos T. Por tanto, se puede predecir el número de péptidos que tiene un efecto positivo en un animal combinando contribuciones que dependen, por separado, del tipo de linfocito T, el polipéptido y la vacuna que el animal recibió.

20 Estos modelos también incluyeron efectos aleatorios para explicar la variación entre animales. Esta es una precaución diseñada para permitir valores de p más creíbles, asignando apropiadamente el potencial predictivo del modelo.

Los inventores observaron los siguientes efectos:

- 25 a) Hay muchas más respuestas de CD8 que respuestas de CD4, por un factor de 4,37, $p < 2 \times 10^{-16}$;
- b) Hay menos respuestas en gp160 que en gag o pol, por un factor de 0,54, $p = 0,000830$, y ninguna diferencia significativa entre gag y pol (incluso cuando se normaliza por longitud de secuencia ya que pol es dos veces más largo que Gag y por lo tanto tiene más oportunidad de reaccionar); y
- c) La vacuna de mosaico genera significativamente más respuestas positivas que Mcon (por un factor de 3,6, $p = 6,26 \times 10^{-11}$) mientras que OptC genera menos, aunque la diferencia de Mcon-OptC no es significativa.

30 *Ejemplo 5*

Si se considera solamente el número mínimo de respuestas inducidas por una vacuna y detectadas por péptidos PTE, de modo que todos los péptidos que solapan en ≥ 8 aminoácidos independientemente de la variación se cuentan solamente 1 vez, las vacunas de mosaico aún generan un mayor número de respuestas a regiones definidas.

35 Para el CD8, contando cada conjunto de péptidos solapantes solamente una vez:

Sumario estadístico:

40 Mos2 > Mcon - OptC (Mcon muestra una tendencia a más respuesta que OptC)
 El valor de p de Wilcoxon para Mos2 en comparación con Mcon: valor de p = 0,0009992
 El valor de p de Wilcoxon para Mcon en comparación con C óptimo: valor de p = 0,2351

Sumario de los grupos:						
Vacuna	Mín.	1 ^{er} Cu.	Mediana	Media	3 ^{er} Cu.	Máx.
Mos2.cd8	7	7,5	8	9,4	11	14
Mcon.cd8	0	3	3	3,3	4	6
OptC.cd8	0	1	1,5	2	4,25	5

45 Para el CD4, contando cada conjunto de péptidos solapantes solamente una vez (hay muy poco solapamiento en CD4, de modo que este es casi igual que el primer recuento).

Sumario estadístico:

50 Mos2 >> Mcon ~ OptC (Mcon muestra una tendencia a más respuesta que OptC)
 El valor de p de Wilcoxon para Mos2 en comparación con Mcon: valor de p = 0,00198
 El valor de p de Wilcoxon para Mcon en comparación con C óptimo: 0,099

55

Sumario de los grupos:						
Vacuna	Mín.	1 ^{er} Cu.	Mediana	Media	3 ^{er} Cu.	Máx.
Mos2.cd4	2	2,5	3	3,4	4,5	5
Mcon.cd4	0	1	1	1,3	2	2
OptC.cd4	0	0	0,5	0,67	1	2

Ejemplo 6: Regresión de Poisson contando cada conjunto de péptidos solapantes solamente una vez

Usando péptidos PTE solapantes, los inventores determinaron lo siguiente, que están en amplio acuerdo con los resultados analizados en el Ejemplo 4 anterior, donde cada respuesta de PTE positiva se contó por separado:

- a) Hay muchas más respuestas de CD8 que respuestas de CD4, por un factor de 2,8, $p \approx 1 \times 10^{-7}$;
- b) La vacuna de mosaico genera significativamente más respuestas positivas que Mcon (por un factor de 2,84, $p \approx 4,3 \times 10^{-7}$), mientras que OptC genera menos, aunque la diferencia de Mcon-OptC no es significativa; y
- c) Hay más respuestas a Pol que a Gag y más a Gag que gp160, pero solo la diferencia de Pol-gp160, un factor de aproximadamente 2, fue significativa, $p < 0,001$.

Ejemplo 7:

La siguiente tabla es un recuento de las respuestas totales a Gag, Pol y respuestas de Env a las tres vacunas en los 7 animales vacunados con 2 mosaicos (Mos2) o Mcon y los 6 animales vacunados con el clado C natural óptimo (OptC):

	CD8			CD4		
	Env	Gag	Pol	Env	Gag	Pol
2Mos	13	20	33	3	10	11
ConM	8	7	8	2	3	4
OptC	4	5	5	1	2	1

La vacuna de OptC produjo una respuesta promedio en todos los monos que era ligeramente menor que la respuesta de linfocitos T CD8+ por proteína. La vacuna de Mcon mostró ~1 respuesta por proteína. Solamente con Mos2 se observa una diferencia en las proteínas, donde Env típicamente tiene menos respuestas que Gag o Pol.

Cada una de las proteínas en la vacuna de Mos2 indujo muchas respuestas y contribuyó a la respuesta general. La longitud relativa de las proteínas consenso después de las modificaciones para inactivar pol y la supresión del dominio de escisión y fusión en Env fue: 671 aminoácidos de Env, 851 de Pol, 498 de Gag (1,35 : 1,7 : 1).

Sumario

- Amplitud: Las 2 vacunas de mosaico inducen respuestas de linfocitos T que son capaces de reconocer muchas más regiones epitópicas que el M consenso o una única cepa natural óptima.
- Profundidad: La diversidad de los péptidos PTE reconocidos sugiere que ambas formas en los 2 mosaicos inducen diferentes respuestas de linfocitos T a los péptidos variantes, aumentando el potencial de reactividad cruzada.

Ejemplo 8:

Las vacunas del VIH-1 de mosaico de la invención expanden la amplitud y profundidad de respuestas inmunitarias celulares en monos Rhesus. Los inventores construyeron antígenos de Gag, Pol y Env del VIH-1 de mosaico que optimizaron la cobertura de PTE de secuencias del grupo M del VIH-1, que incluyen todos los clados del VIH-1 principales y linajes recombinantes en la base de datos de secuencia del VIH-1 de Los Alamos. Se utilizó una estrategia de mosaico bivalente para equilibrar los problemas de competición de cobertura teórica y utilidad práctica. Los antígenos de Gag, Pol y Env del VIH-1 de mosaico bivalentes expandieron sustancialmente la amplitud y magnitud (profundidad) de respuestas de linfocitos T CD8+ y CD4+ específicas de epítomos en monos rhesus, en relación con la respuesta inmunitaria observada usando antígenos del VIH-1 de secuencia natural y consenso en monos rhesus.

Los inventores inmunizaron 27 monos rhesus exogámicos con una única inyección de vectores de adenovirus de serotipo 26 recombinantes (rAd26) que expresaban los siguientes antígenos: antígenos de Gag, Pol y Env del VIH-1 de (i) mosaico bivalente (N=7), (ii) M consenso (N=7), (iii) clado B y clado C combinado bivalente (N=7) o (iv) clado C natural óptimo (N=6). Se administró una dosis total de 3×10^{10} partículas víricas de vectores rAd26 que expresaban estos antígenos una vez i.m. a cada animal. Los antígenos de clado C óptimos fueron las secuencias de cepa natural seleccionadas para proporcionar cobertura de PTE máxima de secuencias de clado C en la base de datos de secuencia del VIH-1 de Los Alamos (analizada en los Materiales y métodos posteriormente). Los inventores evaluaron la amplitud y magnitud (profundidad) de respuestas de linfocitos T específicas del VIH-1 inducidas por vacuna mediante ensayos ELISPOT de IFN- γ en la semana 4 después de inmunización utilizando grupos y subgrupos de péptidos que incluyeron todos los PTE hallados en al menos 15 % de las secuencias del grupo M del VIH-1. Todas las respuestas de péptidos individuales se resolvieron y se realizaron ensayos ELISPOT de IFN- γ pobres en células para determinar si los péptidos reactivos representaban epítomos de linfocitos T CD8+ o CD4+.

El número total de respuestas inmunitarias celulares específicas de Gag, Pol y Env a péptidos PTE inducidos por los antígenos de mosaico fue 3,8 veces mayor que el número de respuestas inducidas por los antígenos de secuencia consenso o natural (Fig. 19A; $P = 1 \times 10^{-11}$, comparando el mosaico con los antígenos consenso, el siguiente grupo más alto, basándose en un modelo de regresión de Poisson). Hubo 4,4 veces más respuestas de linfocitos T CD8+

que CD4+ ($P < 10^{-11}$) y menos respuestas a Env que a Gag o Pol ($P < 0,0007$). El número mediano de respuestas de linfocitos T CD8+ fue mayor para la vacuna de mosaico, seguido de las vacunas consenso, la combinada B+C y el clado C natural (medianas de 16, 5, 3 y 2 respuestas por animal en cada grupo, respectivamente). Aunque hay menos respuestas de linfocitos T CD4+ en general, el mismo patrón relativo surgió con el mayor número de respuestas de linfocitos T CD4+ a la vacuna de mosaico, seguido de las vacunas consenso, la combinada B+C y el clado C natural (medianas de 4, 1, 1 y 0,5 respuestas por animal en cada grupo, respectivamente). Los números de respuestas de linfocitos T CD8+ y CD4+ inducidas por las vacunas consenso, el B+C combinada y el clado C natural no fueron estadísticamente distinguibles.

Los péptidos PTE incluyen múltiples secuencias solapantes que reflejan de forma natural polimorfismos de secuencias del VIH-1 de origen natural y por lo tanto las respuestas de péptidos PTE abarcan tanto el reconocimiento de un epítipo particular (amplitud) como el reconocimiento cruzado de variantes de ese epítipo (profundidad). Se realizó un análisis conservativo de la amplitud evaluando el número de regiones epitópicas reactivas por mono en el que todos los péptidos PTE reactivos que solaparon en 8 o más aminoácidos se contaron como un acontecimiento. En este análisis conservativo, los inventores aún observaron que los antígenos de mosaico indujeron números 3,1 veces mayores de regiones epitópicas reactivas de Gag, Pol y Env en comparación con los antígenos consenso o antígenos de secuencia natural (Fig. 19B; $P = 1,6 \times 10^{-7}$, regresión de Poisson). Las regiones epitópicas mostraron algo de agrupamiento entre animales, como se demuestra por regiones de alta densidad epitópica (Figs. 20A-20C y Figs. 21A-21C). Se muestran los alineamientos completos de todos los péptidos positivos organizados por regiones de respuesta en la Fig. 22.

Estos datos muestran que los antígenos de mosaico aumentaron sustancialmente la amplitud de respuestas inmunitarias celulares en comparación con antígenos de consenso M y clado C natural. También se demostró que los antígenos de mosaico bivalentes eran superiores a la combinación bivalente de antígenos de clado B y clado C (Fig. 19A y 19B), lo que indica que la amplitud potenciada se debía al diseño de secuencia de mosaico y no reflejaba solamente el uso de dos secuencias antigénicas distintas por proteína. Para determinar si la amplitud aumentada inducida por antígenos de mosaico comprometía la potencia de las respuestas, los inventores evaluaron la magnitud de todas las respuestas de linfocitos T CD8+ y CD4+ individuales. La magnitud de estas respuestas resultó comparable entre todos los grupos (Fig. 23; $P = 0,58$ y $P = 0,99$, respectivamente, ensayos de Kolmogorov-Smirnov de dos lados). Por tanto, los antígenos de mosaico expandieron la amplitud inmunitaria celular sin comprometer la magnitud de respuestas específicas de epítopos individuales, lo que indica que las restricciones de inmunodominancia y competición antigénica no limitaron la inmunogenicidad de los antígenos de mosaico en este estudio.

A continuación, los inventores caracterizaron la profundidad de las respuestas inmunitarias celulares inducidas por los diversos regímenes de vacunas. Los inventores definieron la profundidad como el número de péptidos PTE variantes inducidos simultáneamente para una región epitópica particular. La inducción de respuestas a múltiples variantes epitópicas comunes puede aumentar la cobertura inmunológica de secuencias de virus infecciosos, bloquear las vías de escape *in vivo* comunes u obligar al virus a pasar por vías de escape terciarias que provocan altos costes de ajuste. Los antígenos de secuencia consenso y natural indujeron respuestas que se caracterizaron por un alto grado de identidad de secuencia entre las secuencias de vacuna y los péptidos PTE reactivos, ejemplificado por las respuestas en mono 366 que recibió los antígenos de clado C natural (Fig. 24A; véase también la Fig. 22). Por el contrario, los antígenos de mosaico indujeron respuestas que se caracterizaban por múltiples péptidos PTE reactivos en regiones epitópicas particulares. Estos péptidos representaban variantes comunes y con frecuencia reflejaban los polimorfismos contenidos en las secuencias de vacuna de mosaico, ejemplificado por las respuestas en el mono 361 (Fig. 24B; véase también la Fig. 22). Un sumario de todas las respuestas específicas de epítipo en estos animales demuestra que los antígenos de mosaico aumentaron la frecuencia de respuestas inmunitarias celulares a péptidos con dos o más variantes diana en comparación con los antígenos de secuencia consenso o natural (Fig. 24C; $P = 0,001$, ensayo de suma de rangos de Wilcoxon que compara el mosaico con los antígenos consenso, el siguiente grupo más alto).

Para complementar el análisis utilizando péptidos PTE, los inventores también evaluaron la amplitud de respuestas inmunitarias celulares en los monos vacunados con péptidos solapantes tradicionales que abarcaban 5 secuencias de Gag diferentes: clado C DU422, clado C ZM651, consenso C, consenso A y consenso B. La amplitud inmunitaria celular se determinó evaluando la reactividad a subgrupos de 10 péptidos solapantes que abarcaban cada secuencia de Gag. Los antígenos de mosaico indujeron mayor amplitud de respuestas de linfocitos T en comparación con los antígenos de secuencia consenso o natural frente a todas las secuencias de Gag que se ensayaron (Fig. 25; $P = 1 \times 10^{-7}$, regresión binomial). Por tanto, los antígenos de mosaico aumentaron la amplitud inmunitaria celular no solamente para péptidos PTE sino también para péptidos Gag reales de los clados A, B y C. Los antígenos de mosaico resultaron incluso superiores a los antígenos del clado C natural óptimos para inducir respuestas con péptidos Gag del clado C. Por otra parte, los antígenos de mosaico indujeron respuestas comparables a péptidos Gag de múltiples clados, mientras que los antígenos del clado C natural mostraron respuestas reducidas a péptidos Gag del clado A y clado B (Fig. 25).

Para evaluar la durabilidad de estas observaciones, los inventores reforzaron los monos que recibieron los antígenos de mosaico, consenso y clado C natural óptimo en la semana 40 con una dosis total de 3×10^{10} partículas víricas del vector heterólogo rAd5HVR48 que expresa antígenos Gag, Pol y Env del VIH-1 que coincidieron con las secuencias utilizadas en la inmunización inicial. La amplitud inmunitaria se determinó evaluando la reactividad a subgrupos de 10

péptidos PTE en la semana 4 (después de sensibilización) y en la semana 44 (después del refuerzo). La mayoría de respuestas de linfocitos T CD8+ y CD4+ que se observaron después de la inmunización de sensibilización se expandieron después del refuerzo (Fig. 26A, líneas rojas y azules), y también se detectaron varias nuevas respuestas (Fig. 26A, puntos rojos y azules). En la semana 44, la magnitud de respuestas inmunitarias celulares individuales resultaron comparables entre grupos (Fig. 26A). El número de respuestas de subgrupos inducidas por los antígenos de mosaico (mediana de 27 respuestas por animal), sin embargo, permaneció sustancialmente mayor que el número de respuestas de subgrupos inducidas por los antígenos consenso (mediana de 11 respuestas por animal) o los antígenos del clado C natural óptimo (mediana de 10 respuestas por animal) después de la inmunización de refuerzo (Fig. 26B). Tanto antes como después del refuerzo, hubo más respuestas por animal inducidas por la vacuna de mosaico que por las vacunas del clado C natural o consenso ($P < 0,001$, ensayos de suma de rangos de Wilcoxon para todas las comparaciones por pares).

Los inventores también midieron respuestas inmunitarias humorales específicas de Env después de inmunización de refuerzo mediante ELISA (Fig. 26C) y ensayos de neutralización de pseudovirus basados en luciferasa (Fig. 26D). Todos los grupos mostraron títulos de ELISA comparables con respecto a gp140 del clado C y respuestas de anticuerpo neutralizante (NAb) comparables al virus de clado C de nivel 1 MW965.26. Los antígenos de mosaico indujeron respuestas de Nab ligeramente mayores al virus del clado B de nivel 1 SF162.LS en comparación con los antígenos del clado C consenso o natural ($P = 0,02$, ensayo de suma de rangos de Wilcoxon), aunque los inventores no detectaron ninguna respuesta de NAb a virus de nivel 2 en ningún grupo.

Los datos de los inventores demuestran que los antígenos de Gag, Pol y Env del VIH-1 de mosaico aumentaron tanto la amplitud como la profundidad de respuestas inmunitarias celulares específicas de epítipo en comparación con antígenos de secuencia consenso o natural en monos rhesus, bastante coincidentes con las predicciones teóricas (Fig. 27). Los sorprendentes resultados con antígenos de mosaico en este estudio pueden haber reflejado el hecho de que los vectores rAd26 son particularmente eficaces en la inducción de respuestas de linfocitos T CD8+ así como el hecho de que los antígenos de mosaico parecen particularmente eficaces en el aumento de la amplitud de linfocitos T CD8+ (Fig. 19A y 19B). Los inventores también han observado amplitud de linfocitos T CD4+ potenciada con antígenos de mosaico, que aunque hubo números sustancialmente menores de estas respuestas.

Se ha mostrado que la amplitud de respuestas inmunitarias celulares específicas de Gag es crítica para el control del VIS en monos rhesus y para el control del VIH-1 en seres humanos. Por otra parte, en el estudio STEP de fase 2b, el candidato a vacuna del VIH-1 basado en rAd5 que expresaba antígenos de Gag, Pol y Nef del clado B natural indujo solamente una amplitud limitada de respuestas inmunitarias celulares específicas del VIH-1 y no observó ningún beneficio de la vacuna. Los vacunados en el estudio STEP desarrollaron una mediana de solamente 2-3 respuestas de linfocitos T específicas de epítipo, incluyendo una mediana de solamente 1 respuesta específica de epítipo a Gag, y esta amplitud muy estrecha de respuestas inmunitarias celulares probablemente proporcionó cobertura inmunológica insuficiente de la diversidad de virus infecciosos. También se ha indicado que se produce escape vírico de linfocitos T CD8+ rápidamente durante la infección por VIH-1 aguda, y por lo tanto las respuestas inmunitarias celulares inducidas por vacuna contra variantes de epítopos habituales también pueden resultar críticas. En conjunto, estos estudios enfatizan la necesidad de desarrollar estrategias de vacuna del VIH-1 que aumentan la amplitud y profundidad inmunitaria celular.

Ya que los inventores evaluaron antígenos del VIH-1 de mosaico en el presente estudio, no pudieron evaluar la eficacia protectora de estos regímenes de vacunas contra exposiciones al VIS. Sin embargo, los inventores han informado previamente de que la amplitud de respuestas inmunitarias celulares específicas del VIS inducidas por vectores rAd se correlacionaban con eficacia protectora contra exposiciones al VIS en monos rhesus (Liu *et al.*, Nature 457:87, 2009). Los inventores también han mostrado que las respuestas inmunitarias contra epítopos variantes pueden bloquear la evolución mutacional del VIS en monos rhesus *in vivo* (Barouch *et al.*, Nat. Immunol. 6:247, 2005), lo que sugiere la relevancia biológica de la expansión de la profundidad inmunitaria celular. La modelización de la eficacia protectora de vacunas de mosaico contra exposiciones al VIS en primates no humanos tiene limitaciones intrínsecas, ya que la diversidad observada de secuencias del grupo M del VIS y VIH-1 difiere sustancialmente y se ve influida por biología subyacente diferente. Por ejemplo, la presión de selección de linfocitos T CD8+ en hospedadores naturales tales como mangabey gris parece sustancialmente menor que en seres humanos. Por tanto, la evaluación adicional de antígenos de mosaico como vacunas del VIH-1 candidatas puede verse beneficiada por los ensayos clínicos.

En resumen, los inventores demuestran que antígenos de Gag, Pol y Env del VIH-1 de mosaico bivalente expandieron sustancialmente la amplitud y profundidad inmunitaria celular en monos rhesus. Estos hallazgos tienen implicaciones importantes para el desarrollo de vacuna del VIH-1, ya que la diversidad de virus global y el escape vírico de respuestas inmunitarias celulares representan obstáculos críticos en el desarrollo de una vacuna del VIH-1 basado en linfocitos T. Un cóctel bivalente de antígenos de mosaico también es práctico y potencialmente factible para desarrollo clínico. El aumento de la valencia de antígenos de mosaico puede aumentar la cobertura. Finalmente, la estrategia de antígeno de mosaico es generalizable y podría utilizarse para otros patógenos genéticamente diversos además del VIH-1.

Materiales y procedimientos

Diseño de antígenos y producción de vectores. Se construyeron antígenos de Gag, Pol y Env de mosaico bivalente para proporcionar cobertura óptima de secuencias del grupo M del VIH-1 en la base de datos de secuencia del VIH-1 de Los Alamos esencialmente como se describe (1, 2). Se seleccionaron antígenos del clado C natural óptimo para ser las secuencias que proporcionan cobertura de PTE óptima de secuencias del clado C en la base de datos de secuencia del VIH-1 de Los Alamos (C.IN.-70177 Gag, C.ZA.04.04ZASK208B1 Pol, C.SN.90.90SE_364 Env). Se seleccionaron antígenos del clado B para ser secuencias casi consenso o consenso (B.CAM-1 Gag, B.IIIB Pol, B.Con Env) y se usaron para complementar los antígenos del clado C óptimos para el enfoque de vacuna del clado B + C bivalente. Los antígenos de Pol contenían RT y IN sin PR e incluyeron mutaciones puntuales para eliminar la actividad catalítica como se describe (Priddy *et al.*, *Clinical infectious diseases* 46:1769, 2008). Los antígenos de Env gp140 contenían mutaciones puntuales para eliminar la actividad de escisión y fusión. Las secuencias de vacuna se representan en la Fig. 27. Se cultivaron vectores de adenovirus de serotipo 26 incompetente para replicación (rAd26) y de hexón quimérico rAd5HVR48 que expresan estos antígenos en células PER.55K y se purificaron por sedimentación de gradiente CsCl doble esencialmente como se describe (Abbink *et al.*, *J. Virol.* 81:4654, 2007 y Roberts *et al.*, *Nature* 441:239, 2006).

Animales e inmunizaciones. 27 monos rhesus exogámicos que no expresaban el alelo del MHC de clase I *Mamu-A*01* se alojaron en el New England Primate Research Center (NEPRC), Southborough, MA. Las inmunizaciones implicaron 3×10^{10} partículas víricas de vectores rAd26 o rAd5HVR48 que expresan antígenos de Gag, Pol y Env del VIH-1 de mosaico, M consenso, clado B + clado C o clado C natural óptimo suministradas como inyecciones de 1 ml i.m. en ambos músculos cuádriceps en las semanas 0 y 40. Todos los estudios animales fueron aprobados por el comité institucional de cuidado y uso de animales (IACUC).

Ensayos ELISPOT de IFN- γ . Las respuestas inmunitarias celulares específicas del VIH-1 en monos vacunados se evaluaron mediante ensayos ELISPOT de interferón- γ (IFN- γ) esencialmente como se describe (Roberts *et al.*, *Nature* 441:239, 2006, y Liu *et al.*, *Nature* 457:87, 2009). Se obtuvieron péptidos de epítomos de linfocitos T potenciales (PTE) Gag, Pol y Env del VIH-1 que incluían todos los PTE encontrados en al menos 15 % de secuencias del grupo M del VIH-1 así como péptidos de Gag del VIH-1 de las cepas de clado C DU422, clado C ZM651, consenso C, consenso A, y consenso B del NIH AIDS Research and Reference Reagent Program. Se recubrieron placas multiscreen de 96 pocillos (Millipore) durante una noche con 100 μ l/pocillo de 10 μ g/ml de anti-IFN- γ humano (BD Biosciences) en PBS de Dulbecco sin endotoxina (D-PBS). Las placas se lavaron después tres veces con D-PBS que contenía Tween-20 0,25 % (D-PBS/Tween), se bloquearon durante 2 h con D-PBS que contenía FBS 5 % a 37 °C, se lavaron tres veces con D-PBS/Tween, se aclararon con RPMI 1640 que contenía FBS 10 % para retirar el Tween-20, y se incubaron con 2 μ g/ml de cada péptido y 2 $\times 10^5$ PBMC por triplicado en volúmenes de reacción de 100 μ l. Después de una incubación de 18 h a 37°C, las placas se lavaron nueve veces con PBS/Tween y una vez con agua destilada. Después las placas se incubaron con anti-IFN- γ humano biotinilado 2 μ g/ml (BD Biosciences) durante 2 h a temperatura ambiente, se lavaron seis veces con PBS/Tween, y se incubaron durante 2 h con una dilución 1:500 de estreptavidina-fosfatasa alcalina (Southern Biotechnology Associates). Después de cinco lavados con PBS/Tween y uno con PBS, las placas se revelaron con cromógeno de nitro azul de tetrazolio/5-bromo-4-cloro-3-indolil-fosfato (Pierce), se detuvieron con lavado con agua del grifo, se secaron al aire y se leyeron usando un lector de ELISPOT (Cellular Technology Ltd). Se calcularon las células formadoras de puntos (CFP) por cada 10^6 PBMC. Los fondos del medio fueron típicamente < 15 CFP por cada 10^6 PBMC. Las respuestas positivas se definieron como > 55 CFP por cada 10^6 PBMC y > 4 veces el fondo.

Mapeo de epítomos. Se realizó mapeo de epítomos de linfocitos T CD8+ y CD4+ exhaustivo utilizando péptidos PTE Gag, Pol y Env que se obtuvieron del NIH AIDS Research and Reference Reagent Program. Se realizaron ensayos ELISPOT de IFN- γ en la semana 4 después de inmunización inicialmente con grupos de péptidos completos, así como con subgrupos que contienen 10 péptidos PTE. Todos los subgrupos de péptidos con respuestas positivas se desconvolucionaron y los epítomos se confirmaron con péptidos PTE de 15 aminoácidos individuales. Se realizaron después ensayos ELISPOT de IFN- γ pobres en células para determinar si los péptidos reactivos representaban epítomos de linfocitos T CD8+ o CD4+. También se realizó mapeo de epítomos parcial utilizando subgrupos de PTE 4 semanas después de la inmunización de refuerzo en la semana 44. Todas las respuestas límite se volvieron a ensayar y solo se consideraron positivas si se confirmaron. También se realizó mapeo de epítomos parcial utilizando subgrupos que contenían 10 péptidos de Gag solapantes para evaluar la amplitud frente a Gag del VIH-1 de diversos clados.

Ensayos inmunitarios humorales. Se evaluaron respuestas inmunitarias humorales específicas de Env mediante ELISA directos utilizando gp140 de Env del clado C del VIH-1 y ensayos de neutralización de pseudovirus basados en luciferasa esencialmente como se describe (Montefiori, Evaluating neutralizing antibodies against HIV, SIV and SHIV in luciferase reporter gene assays. *Current Protocols in Immunology*, Coligan, Kruisbeek, Margulies, Shevach, Strober y Coico, Ed. (John Wiley & Sons, 2004, págs. 1-15).

Análisis estadísticos. Todos los análisis estadísticos se realizaron usando el paquete R (Equipo, *Foundation for Statistical Computing, Viena, Austria*, 2009). Para analizar la amplitud de las respuestas inmunitarias celulares a péptidos PTE mapeados (Fig. 19A), los inventores ajustaron modelos de regresión de Poisson que predijeron el número de péptidos reactivos en función del grupo de vacuna, antígeno (Gag, Pol, Env) y subpoblación de linfocitos

(CD4, CD8). Los modelos de los inventores incluyeron efectos aleatorios para adaptarse a la variación entre animales y se ajustaron con la biblioteca de lme4 (Pinheiro, *Springer, Nueva York* (2000)) del paquete R. Los datos se ajustan bien a los modelos (parámetro de dispersión 1,0) y no hubo ninguna interacción significativa entre los tres factores explicativos. Por ejemplo, la potenciación 3,8 veces del número de péptidos PTE reconocidos por monos que recibieron los antígenos de mosaico en comparación con los que recibieron los antígenos de secuencia consenso o natural (Fig. 19A) aplicados igualmente a PTE de Gag, Pol y Env y mantenidos para respuestas por linfocitos T CD8+ así como CD4+. El análisis del número de regiones epitópicas reactivas (Fig. 19B) también incluyó modelos de regresión de Poisson con efectos aleatorios y de nuevo se ajustó bien (parámetro de dispersión 0,87) sin ninguna interacción significativa. Se realizaron comparaciones de la magnitud de respuestas de linfocitos T CD8+ y CD4+ (Fig. 23) utilizando ensayos de Kolmogorov-Smirnov de dos caras. También se realizaron ensayos no paramétricos para comparar la amplitud y profundidad de respuestas por mono entre diferentes vacunas (Figs. 19A y 24C). Inicialmente los inventores emplearon ensayos de Kruskal-Wallis para determinar si hubo una diferencia entre los 4 grupos de vacuna. En cada caso esto fue altamente significativo y los inventores evaluaron después todas las comparaciones por pares entre los 4 grupos de vacuna usando ensayos de suma de rangos de Wilcoxon. En cada una de estas comparaciones, la vacuna de mosaico indujo significativamente más respuestas por mono que las otras 3 vacunas. Para analizar la amplitud de respuestas a Gag del VIH-1 de diversos clados (Fig. 25), los inventores ajustaron los datos a modelos de regresión binomiales. Estos modelos usaron el grupo de vacuna como una variable explicativa e incluyeron efectos aleatorios para explicar la variación entre animales y entre cepas. Los datos tenían una dispersión un poco baja, pero los animales que recibieron la vacuna de mosaico aún indujeron un número significativamente mayor de respuestas. Se realizó evaluación de la cobertura de PTE usando herramientas disponibles en la base de datos de secuencia del VIH-1 de Los Alamos.

APÉNDICE DE SECUENCIAS

I. SECUENCIAS DE ENV GP160, GAG, POL, NEF DE MOSAICO M BIVALENTES

ENV1 GP160 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 1

MRVVTGIRKNYQHLLWRWGTMLLGILMICSAAAGKLWVTVVYGVVPVWKEATTTLFCASDA
 KAYDTEVHNVWATHACVPTDPNPQEVVLENVTFENFMWKNMVEQMHEDIISLWDQS
 LKPCVKLTPLCVTLNCTDDVRNVNTNATNTNSSWGEPMEKGEIKNCSFNITTSIRNK
 VQKQYALFYKLDVVPIDNDSNNTNYRLISCNTSVITQACPKVSFEPIPIHYCAPAGF
 AILKCNCKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEVIRSENFTNN
 AKTIMVQLNVSVEINCTRPNNNTRKSIHIGPGRAFYTAGDIIGDIRQAHCNISRANW
 NNTLRQIVEKLGKQFGNKTIVFNHSSGGDPEIVMHSFNCGGEFFYCNSTKLFNSTW
 TWNSTWNNTKRSNDTEEHITLPCRICKQIINMWQEVGKAMYAPP IRGQIRCSSNITG
 LLLTRDGGNDTSGTEIFRPGGGMRDNWRSELYKYKVKIEPLGVAPTKAKRRVVQR
EKRAVGIGAVFLGFLGAAGSTMGAASMTLTVQARLLLSGIVQQNNLLRAIEAQOHL
 LQLTVWGIKQLQARVLAVERYLKDQQLLGIWGC SGKLICTTTVPWNASWSNKSLDKI
 WNNMTWMEWEREINNYTSLIYTLIEESQNQEKNEQELLELDK WASLWNWFDISNWL
 WYIKIFIMIVGGLVGLRIVFAVLSIVNRVRQGYSPLSFQTRLPA PRGPDRPEGIEEE
 GGERDRDRSVRLVDGFLVLWDDLQSLCLFSYHRLRDL LLLIVELLGRRGWEALKYWW
 NLLQYWSQELKNSAISLLNATAVAVAEGTDRVIEALQRACRAILHIPRRIRQGLERL
 LL

ENV2 GP160 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 2

MRVREGIQRNWPQWWIWGILGFWMI IICRVMGNLWVTVYYGVPVWKEAKTTLFCASDA
KAYEKEVHNVWATHACVPTDPNPQEMVLENVTENFNMWKNMVDQMHEDI IRLWDQS
LKPCVKLTPLCVTLECRNVRNVSSNGTYNI IHNETYKEMKNCSFNATTVVEDRKQKV
HALFYRLDIVPLDENNSSEKSSSENSSEYYRLINCNTSAITQACPKVSFDPIPIHYCA
PAGYAILKCNKTFNGTGPCNNVSTVQCTHGKPVVSTQLLLNGSLAEIIIIRSEN
LTNNAKTIIVHLNETVNITCTRPNNTRKSRIRIGPGQTFYATGDIIGDIRQAHCNLS
RDGWNKTLQGVKKLAEHFPNKTINFTSSSGDLEITTHSFNCRGEFFYCNTSGLFN
GTYPNGTNSNSSNITLPCRIKQI INMWQEVGRAMYAPPIAGNITCRSNITGLLLT
RDGGSNNGVPNDTETFRPGGDMRNNWRSELYKYKVVEVKPLGVAPTEAKRRVVERE
KRAVGIGAVFLGILGAAGSTMGAASITLTVQARQLLSGIVQQSNLLRAIEAQQHML
QLTVWGILKQLQTRVLAIERYLQDQQLLGLWGCSGKLICTTAVPWNTSWSNKSQTDIW
DNMTWMQWDKEIGNYTGEIYRLLEESQNQQEKNEKDLLALDSWKNLWNWFDITNWLW
YIKIFIMIVGGLIGLR IILGVLSIVRRVRQGYSPLSFQTLTPNPRGLDRLGRIEEEG
GEQDRDRSIRLVNGFLALAWDDLRSCLFSYHQLRDFILIVARAVELLGRSSLRGLQ
RGWEALKYLGNLVQYWGLELKKGAISLLDTIATAVAEGTDRI IELIQSICRAIRNIP
RRIRQGFASLL

GAG1 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 3

5

MGARASVLSGGELDRWEKIRLRPGGKKYRLKHI V WASRELERFAVNPGLLETSEGC
RQILGQLQPSLQGTGSEELRSLYNTVATLYCVHQRIEIKDTKEALEKIEEEQNKSKKK
AQQAADTGNSSQVSONYPIVQNIQGMVHQAI SPRTLNAWVKVVEEKAFSPEVIM
FSALSEGATPQDLNMTLNTVGGHQAMQMLKETINEEAAEWDRVHPVHAGPIAPGQM
REPRGSDIAGTTS TLQEQIGWMTNPPPIPVGEIYKRWI IILGLNKIVRMYS PVSILDI
RQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQANANPDCKTILKALGPAAT
LEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKCFNCGKEG
HIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKG RPNFLQNRPEPT
APPEESFRFGEETTTSPSQKQEPIDKEMYPLASLKS LFGNDPSSQ

GAG2 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 4

10

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLV WASRELERFALNPGLLETSEGC
KQIIKQLQPALQGTGTEELRSLFNTVATLYCVHAEIEVRDTKEALDKIEEEQNKSQK
TQQAKEADGKVSQNYPIVQNLQGMVHQPI SPRTLNAWVKVIEEKAFSPEVIMFTA
LSEGATPQDLNMTLNTVGGHQAMQMLKDTINEEAAEWDR LHPVHAGPVAPGQMRP
RGSDIAGTTSNLQEQIAWMTSNPPIPVGDIYKRWI IILGLNKIVRMYSPTSILDIKQG
PKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEE
MMTACQGVGGPSHKARVLAEAMSQTNSTILMQRSNFKGSKRIVKCFNCGKEGHIARN
CRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKG RPNFLQSRPEPTAPPAE
SFRFEETTPAPKQEPKDREPLTSLRSLFGSDPLSQ

POL1 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 5

15

FFRENLAFFQGEAREFPSEQTRANSPTSRELQVRGDNPHSEAGAERQGTLNFPQITL
 WQRPLVSIKVGQIREALLDTGADDTVLEDINLPGKWKPKMIGGIGGFIVRQYDQI
 LIEICGKKAIGTVLVGPTPVNI IGRNMLTQLGCTLNFPISPIETVPVCLKPGMDGPR
 VKQWPLTEEKIKALTAICEEMEKEGKITKIGPENPYNTPVFAIKKKDSTKWRKLVDF
 RELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEGFRKYTAFTIPS
 TNNETPGIRYQYNVLPQGWKGSFAIFQCSMTRILEPFRAKNPEIIVYQYMDDLVGS
 DLEIGQHRAKIEELREHLLKWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPE
 KDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEALELAE
 NREILKEPVGHVYYDPSKDLIAEIQKQGHQWTYQIYQEPFKNLKTGKYAKMRTAHT
 NDVKQLTEAVQKIAMESIVIWGKTPKFRLP IQETWETWWTDYWQATWIPWEFVNT
 PPLVKLWYQLEKDP IAGVETFYVDGAANRETKLGKAGYVTDGRGRQKIVSLTETTNOK
 TELQAIYLALQDSGSEVNIVTDSQYALGIIQAQPKSESELVNIIEQLIKKERVYL
 SWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASDFNLPP
 VVAKEIVASCDQCQLKGEAMHGQVDCSPGIWQLDCTHLEGKIIILVAVHVASGYIEAE
 VIPAETGQETAYFILKLAGRWPVKVIHTDNGSNFTSAAVKAACWWAGIQQEFGIPYN
 PQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERII
 DIIATDIQTKELQKQI I KIQNFRVYYRDSRDP IWKGPAKLLWKGEGAVVIQDNSDIK
 VVPRRKVKI IKDYGKQ MAGADCVAGRQDED

POL2 DE MOSAICO (SECUENCIA DE AA)

5 SEQ ID NO: 6

FFRENLAFFPQKAREFSSEQTRANSPTRRELQVWGRDNNSLSEAGADRQGTVSFSFP
 QITLWQRPLVTIKIGGQLKEALLDTGADDTVLEEMNLPGRWKPKMIGGIGGFIVRQ
 YDQIPIEICGHKAIGTVLVGPTPVNI IGRNLLTQIGCTLNFPISPIETVPVCLKPGM
 DGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPIFAIKKKDSTKWRK
 LVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTA
 TIPSINNENETPGIRYQYNVLPQGWKGSFAIFQSSMTKILEPFKQNPDIVIYQYMDDL
 YVGS DLEIGQHRTKIEELRQHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPI
 VLPEKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVVPLTEEAEL
 ELAENREILKEPVGHVYYDPSKDLIAEIQKQGGQWTYQIYQEPFKNLKTGKYARMR
 GAHTNDVKQLTEAVQKIATESIVIWGKTPKFKLP IQETWEAWWTEYWQATWIPWE
 FVNT PPLVKLWYQLEKEPIVGAETFYVDGAANRETKLGKAGYVTDGRGRQKVSLTDT
 TNQKTELQAIHLALQDSGLEVNIVTDSQYALGIIQAQPKSESELVSIIEQLIKKE
 KVYLA WVP AHKGIGGNEQVDKLVSRGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASEF
 NLPP I VAKEIVASCDKQQLKGEA IHGQVDCSPGIWQLDCTHLEGKVIILVAVHVASGY
 IEAEVIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTSATVKAACWWAGIKQEFG
 IPYNPQSQGVVESINKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGEYSAG
 ERIVDIIASDIQTKELQKQITKIQNFRVYYRDSRDP LWKGPAKLLWKGEGAVVIQDN
 SDIKVVPRRKAKI IRDYGKQ MAGDDCVASRQDED

NEF1 DE MOSAICO (SECUENCIA DE AA)

10 SEQ ID NO: 7

MGGKWSKSSVVGWPAIRERMRAEPAADGVGAVSRDLEKHGAI TSSNTAANNADCAW
 LEAQEEEEVGFPPVRPQVPLRPMTYKGALDLSHFLKEKGGLEGLIYSQKRQDILDWV
 YHTQGYFPDQWNYTPGPGIRYPLTFGWCFKLVPEPEKIEEANEGENNSLLHPMSQH
 GMDDPEKEVLMWKFDSRLAFHHMARELHPEYYKDC

15 NEF2 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 8

MGGKWSKSSIVGWPAVRERIRRAEPAAEGVGAASQDLDDKYGALTSSNTAATNADCAW
LEAQEDEEVGFVPKQVPLRPMTYKAAFDSLFFLKEKGGLDGLIYSKKRQEIIDLWV
YNTQGGFFPDWQNYTPGPGVRYPLTFGWCFKLVVDPREVEEANKGENNCLLHPMNLH
GMDDPEREVLVWRFDSRLAFHHMAREKHPEYYKNC

5 II. SECUENCIAS DE ENV GP140 DE MOSAICO M BIVALENTES (DEFECTUOSAS EN ESCISIÓN/FUSIÓN)
ENV1 GP140 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 9

10

MRVTGIRKQNYQHLWRWGTMLLGLIMICSAAGKLWVTVYYGVPVWKEATTLFCASDA
KAYDTEVHNVWATHACVPTDPNPQEVVLENTENFNMWKNMVEQMHEDIISLWDQS
LKPCVKLTPLCVTLNCTDDVRNVTNNATNTNSSWGEPMEKGEIKNCSFNITTSIRNK
VQKQYALFYKLDVVPIDNDSNNTNYRLISCNSTSVITQACPKVSFEPPIPIHYCAPAGF
AILKCNDKKFNGTGPCTNVSTVQCTHGIRPVVSTQLLLNGSLAEEVVIRSENFNTNN
AKTIMVQLNVSVEINCTRPNNNTRKSIHIGPGRAFYTAGDIIGDIRQAHCNISRANW
NNTLRQIVEKLGKQFGNKTIVFNHSSGGDPEIVMHSFNCGGEFFFCNSTKLFNSTW
TWNNTWNNTKRSNDTEEHI TLPCRIKQI INMWQEVGKAMYAPP IRGQIRCSSNITG
LLLTRDGGNDTSGTEIFRPGGGDMRDNRSELYKYKVVKIEPLGVAPTAKRRVVQS
EKSAVGIGAVFLGFLGAAGSTMGAASMTLTVQARLLL SGIVQQQNNLLRAIEAQOHL
LQLTVWGIKQLQARVLAVERYLKDQQLLGIWGC SGKLICTTTVPWNASWSNKS LDKI
WNNMTWMEWEREINNYTSLIYTLIEESQNQOQEKNEQELLELDKWASLWNWFDI SNWL
W

ENV2 GP140 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 10

15

MRVGIQRNWPQWWIWGILGFWMI IICRVMGNLWVTVYYGVPVWKEAKTTLFCASDA
KAYEKEVHNVWATHACVPTDPNPQEMVLENTENFNMWKNMVDQMHEDIIRLWDQS
LKPCVKLTPLCVTLECRNVRNVSSNGTYNI IHNETYKEMKNCSFNATTVVEDRKQKV
HALFYRLDIVPLDENNSSEKSSSENSSEYYRLINCNTSAITQACPKVSFDPIPIHYCA
PAGYAILKCNKTFNGTGPCNNVSTVQCTHGKIPVVSTQLLLNGSLAEEIIIRSEN
LTNNAKTIIVHLNETVNI TCTRPNNNTRKSI RIGPGQTFYATGDIIGDIRQAHCNLS
RDGWNKTLQGVKKLAEHFPNKTINF TSSSGDLEITTHSFNCRGEFFFCNTSGLFN
GTYPNGTNSNSSNITLPCRIKQI INMWQEVGRAMYAPPIAGNITCRSNI TGLLLT
RDGGSNNGVPNDTETFRPGGGDMRNNNRSELYKYKVVVEVKPLGVAPTEAKRRVVESE
KSAVGIGAVFLGILGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQOHL
LQLTVWGIKQLQTRVLAIERYLQDQQLLGLWGC SGKLICTTAVPWNTSWSNKSQTDI W
DNMTWMOQWKEIGNYTGETYRLLEESQNQOQEKNEKDLLALDSWKNLWNWFDITNWLW

ENV GP140 DE MOS3 (SECUENCIA DE AA)

20 678 AA

SEQ ID NO: 11

MRVKGIRKNYQHLWKWGTMLLGMMLICSAAEQLWVTVVYGVVPVWRDAET
 TLFASDAKAYEREVHNIWATHACVPTDPNPQEIVLENVTEEFNMWKNDMV
 EQMHTDIISLWDESLKPCVKLAPLCVTLNCTNANLNCTNDNCNRTVDKMREE
 IKNCSFNMTTELDRDKKQKVYALFYKLDIVPIEKNSSEYRLINCNTSTITQACPK
 VTFEPIPIHYCTPAGFAILKCKDKKFNGTGPKKNVSTVQCTHGKIPVISTQLLL
 NGLAEGEIIIRSENITNNAKTIIVQLNESVVINCTRPGNNTRKSVRIGPGQAFY
 ATGEIIGDIRQAYCNISRAKWNNTLQIVTKLKEQFKNKTIVFNQSSGGDEPIT
 THSFNCGGEFFYCNTTQLFNSTWNSNSTWNDTTGSVTEGNDTITLPCRIVQIV
 NMWQRVGQAMYAPPIEGNITCKSNITGLLLVRDGGNINRTNETFRPGGGNMK
 DNWRSELYKYKVVEIKPLGVAPTRAKRRVVESEKSAVGLGAVFLGFLGTAG
 STMGAASLTLTVQARQVLSGIVQQSNLLKAIEAQHLLKLTWVGKQLQAR
 ILAVERYLRDQQLGIWGC SGKLICTTNPWNSSWSNKSQEEIWNMTWMQ
 WDREISNYTDIYRLLLEDSONQKEKNEQDLLALDKWASLWNWFSITNWLW

III. SECUENCIAS DE POL DE MOSAICO M BIVALENTES (EXTENSAMENTE INACTIVADAS, CON SUPRESIÓN DE PR, MUTACIONES DE 9 A PARA ELIMINAR LA ACTIVIDAD CATALÍTICA)

5

POL1 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 12

MAPISPIETVPVKLKPGMDGPRVKQWPLTEEKIKALTAICEEMEKEGKITKIGPENP
 YNTPVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLAVG
 DAYFSVPLDEGFRKYTAFTIPSTNNETPGIRYQYNVLPQGWKGS PAIFQCSMTRILE
 PFRAKNPEIIVYQYMAALYVGS DLEIGQHRAKIEELREHLLKKGFTTTPDKKHQKEPP
 FLWMGYELHDPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGKVRQLCKLL
 RGAKALTDIVPLTEEALELAENREILKEPVHGVYDPSKDLIAEIQKQGHQWTYQ
 IYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAMESIIVIWGKTPKFRLP IQKET
 WETWWTDYWQATWIPEWFEVNTPLVKLWYQLEKDP IAGVETFYVAGAANRETKL GK
 AGYVTDGRGRQKIVSLTETTNQKTALQAIYLALQDSGSEVNIVTASQYALGI IQAQP
 KSESELVNQIEQLIKKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKA
 QEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDQCQLKGEAMHGQVDCSPGIWQLAC
 THLEGKIIILVAVHVASGYIEAEVI PAETGQETAYFILKLAGRWPVKVIHTANGSNFT
 SAAVKAACWWAGIQQEFGI PYNPQS QGVVASMNKELKKIIGQVRDQAEHLKTAVQMA
 VFIHNFKRKGGIGGYSAGERIIDIIATDIQTKELOKQIITKIQNFRVYYRDSRDP IWK
 GPAKLLWKGEGAVVIQDNSDIKVVPRRKVKI IKDYGKQ MAGADCVAGRQDED

10

POL2 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 13

MAPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENP
 YNTPIFAIAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLAVG
 DAYFSVPLDEDFRKYTAFTIP SINNETPGIRYQYNVLPQGWKGS PAIFQSSMTKILE
 PFRKQNPDIIVYQYMAALYVGS DLEIGQHRTKIEELRQHLLRWGFTTTPDKKHQKEPP
 FLWMGYELHDPDKWTVQPIVLPKDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLL
 RGTALTEVVPLTEEALELAENREILKEPVHGVYDPSKDLIAEIQKQGGQWTYQ
 IYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKIATESIIVIWGKTPKFKLPIQKET
 WEAWWTEYWQATWIPEWFEVNTPLVKLWYQLEKEPIVGAETFYVAGAANRETKL GK
 AGYVTDGRGRQKVSLTDTTNQKTALQAIHLALQDSGLEVNIVTASQYALGI IQAQP
 KSESELVSQIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSRGIRKVLFLDGDIDKA
 QEEHEKYHSNWRAMASEFNLPPIVAKEIVASCDKQKGEA IHGQVDCSPGIWQLAC
 THLEGKVILVAVHVASGYIEAEVI PAETGQETAYFLLKLAGRWPVKTIHTANGSNFT
 SATVKAACWWAGIKQEFGI PYNPQS QGVVAS INKELKKIIGQVRDQAEHLKTAVQMA
 VFIHNFKRKGGIGEYSAGERIVDIIASDIQTKELOKQITKIQNFRVYYRDSRDP LWK
 GPAKLLWKGEGAVVIQDNSDIKVVPRRKAKI IRDYGKQ MAGDDCVASRQDED

15

POL V3 DE MOS3 (SECUENCIA DE AA)

851 AA

5 SEQ ID NO: 14

MAPISPIDTVPVTLKPGMDGPKIKQWPLTEEEKIKALTEICTEMEKEGKISRIGPENP
YNTPVFAIKKKNSTRWRKLVDFRELNKKTQDFWEVQLGIPHPAGLKKKRSVTVLAVG
DAYFSVPLDKDFRKYTAFTIIPSVNNETPGVRYQYNVLPQGWKGS PAIFQCSMTKILE
PFRAQNPEIVIIYQYVAALYVGS DLEIEQHRTKIEELRAHLLSWGFTTPDKKHOREPP
FLWMGYELHPDRWTVQPIELPEKESWTVNDIQKLVGKLNWASQIYPGIKVKQLCRL
RGAKALTEVIPLTKEAELELAENREILREPVHGVVYDPSKDLVAEIQKQGQDQWTYQ
IYQEPYKNLKTGKYARKRSAHTNDVRQLTEAVQKIALESIVIWGKIPKFRLPIQRET
WETWWTEYWQATWIPDWEFVNTPLVVKLWYQLEKEPIAGAETFYVAGASNRETKIGK
AGYVTDKGRQKVVS LTETTNQKAALQAIQLALQDSGPEVNIIVTASQYVLGIIQAQPD
RSESELVNQIIIEELIKKEKVYLSWVPAHKGIGGNEQVDKLVSAGIRKILFLDGIDKA
QEEHERYHSNWR TMASDFNLPPIVAKEIVANC DKCQLKGEAMHGQVDCSPGMWQLAC
THLEGKIIIVAVHVASGYMEAEVIPAETGQETAYYILKLAGRWPVKVHTANGSNFT
STTVKAACWWANVTQEFGI PYNPQSQGV IASMNKELKKIIGQVREQAEHLKTAVQMA
VLIHNFKRRGGIGGYSAGERIVDI IATDIQTRELOKQIIKIQNFRVYFRDSRDPVWK
GPAKLLWKGE GAVVIQDNSEIKVVPRR KVKIIRDY GKQMAGDDCVAGRODEDQ

IV. SECUENCIA DE GAG DE MOSAICO M BIVALENTE

10

GAG DE MOS3 (SECUENCIA DE AA)

508 AA

15 SEQ ID NO: 15

MGARASVLSGGKLD AWEKIRLRPGGKKKYK LKHIVWASRELD RFALNPGLLETAEGC
QQIIIEQLQPALQTGSEELKSLYNTVAVLYCVHQRIDVKDTKEALDKIEEIQNKSKQK
TQQAADTGSSSKVSONYPIVQNAQGQMVHQALSPRTLNAWVKVVEEKGFNPEVIM
FSALAE GATPQDLNMMMLNIVGGHQAMQILKDTINEEAADWDR LHPVHAGPIPPGQM
REPRGSDIAGTTSTPQEQIGWMTSNPPVPVGEIYKRWIIMGLNKIVRMYS PVSILDI
KQGPKESFRDYVDRFFKVLRAEQATQEVKNWMTETLLIQNANPDCKSILRALGPGAS
LEEMMTACQGVGGPSHKARILAEAMSQANNTNIMMQRGNFKGQKR IKCFNCGKEGHL
ARNCRAPRKRGCWKCGREGHQMKDCNERQANFLGKIWPSSKGRPGNFPQSRPEPTAP
LEPTAPPAEPTAPPAESFGFGEEITPSPKQE QKDREPLTSLKSLFGSDPLLQ

V. SECUENCIAS DE NEF DE MOSAICO M BIVALENTES (POSICIÓN 2 G A A PARA SUPRIMIR UN SITIO DE MIRISTILACIÓN

20

NEF DE MOS1

(206 AA)

25

SEQ ID NO: 16

MAGKWSKSSVVGWPAIRERMRR AEPAADGVGAVSRDLEKHGAI TSSNTAANNADCAW
LEAQEEEEVGFVVRPQVPLRPMTYK GALDLSHFLKEKGGLEGLIYSQKRQDILD L WV
YHTQGYFPDWQNYTPGPGIRYPLTFGWCFKLV PVEPEKIEEANE GENNSLLHPMSQH
GMDDPEKEVLMWKFDSRLAFHHMARELHPEYYKDC

30

NEF DE MOS2

(206 AA) - POSICIÓN 2 G A A PARA SUPRIMIR EL SITIO DE MIRISTILACIÓN

5 SEQ ID NO: 17

MAGKWSKSSIVGWPAVRERIRRAEPAAEGVGAASQDLDKYGALTSSNTAATNADCAW
LEAQEDEEVGFVPKQVPLRPMTYKAAFDSLFFLKEKGGLDGLIYSKKRQEIILDLWV
YNTQGGFFPDWQNYTPGPGVRYPLTFGWCFKLVVDPREVVEEANKGENNCLLHPMNLH
GMDDPEREVLVWRFDSRLAFHHMAREKHPEYYKNC

NEF DE MOS3

10

(208 AA)

SEQ ID NO: 18

15

MAGKWSKRSVVGWPAVRERMRRTEPAAEGVGAVSQDLDKHGALTSSNTAHNNADCAW
LQAQEEEEVEGFPVRPQVPRPMTYKAAVDLSHFLKEKGGLEGLIHSQKRQEIILDLW
VYHTQGGFFPDWHNYTPGPGTRFPLTFGWICYKLVVDPKEVEEANEENCLLHPMSQ
HGMEDEDREVLKWKFDSSLARRHMARELHPEFYKDC

VI. SECUENCIAS DE FUSIÓN DE GAGNEF DE MOSAICO M BIVALENTES

GAGNEF1 DE MOSAICO (SECUENCIA DE AA)

20

SEQ ID NO: 19

MGARASVLSGGELDRWEKIRLRPGGKKKYRLKHIWASRELERFAVNPGLLETSEGC
RQILGQLQPSLQGTGSEELRSLYNTVATLYCVHQRIEIKDTKEALEKIEEEQNKSKKK
AQQAADTGNSSQVSQNYPIVQNIQGGQMVHQAI SPRTLNAWVKVVEEKAFSPEVIPM
FSALSEGATPQDLNMLNTVGGHQAAMQMLKETINEEAAEWDRVHPVHAGPIAPGQM
REPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSPVSILDI

RQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQANANPDCKTILKALGPAAT
LEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTKVCFNCGKEG
HIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGRPGNLFQNRPEPT
APPEESFRFGEETTTSPSQKQEPIDKEMYPLASLKSFLGNDPSSQAGKWSKSSVVGWP
AIRERMRAEPAADGVGAVSRDLEKHGAI TSSNTAANNADCAWLEAQEEEEVEGFPVR
PQVPLRPMTYKGAALDLSHFLKEKGGLEGLIYSQKRQDIILDLWVYHTQGYFPDWQNYT
PGPGIRYPLTFGWCFKLVVPEPEKIEEANEENNSLLHPMSQHGMDDEPEKVLWVKF
DSRLAFHHMARELHPEYYKDC

25 GAGNEF2 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 20

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLEETSEGC
KQIIKQLQPALQGTGTEELRSLFNTVATLYCVHAEIEVRDTKEALDKIEEEQNKSOQK
TQOAKEADGKVSQNYPIVQNLQGMVHQPI SPRTLNAWVKVIEEKAFSPEVIMFTA
LSEGATPQDLNTMLNTVGGHQAMQMLKDTINEEAAEWDRLHPVHAGPVAPGQMRP
RGSIDIAGTTSNLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRMYSPTSILDIKQG
PKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEE
MMTACQGVGGPSHKARVLAEMSQTNSTILMQRSNFKGSKRIVKCFNCGKEGHIARN
CRAPRKKGCWKCQKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAE
SFRFEETTPAPKQEPKDREPLTSLRSLFGSDPLSQAGKWSKSSIVGWPVAVRERIRRA
EPAAEGVGAASQDLKYGALTSNTAATNADCAWLEAQEEDVEVGFVPKQVPLRPMT
YKAAFDSLFFLKEKGLDGLIYSKKRQEIIDLWVYNTQGGFFPDWQNYTPGPGVRYPL
TFGWCFLVFPVDPREVEEANKGENNCLLHPMNLHGMDDPEREVLVWRFD SRLAFHMM
AREKHPEYYKNC

5 VII. SECUENCIAS DE FUSIÓN DE GAGPOL DE MOSAICO M BIVALENTES (VERSIÓN 3; POL EXTENSIVAMENTE
INACTIVADO, CON SUPRESIÓN DE PR, MUTACIONES DE INACTIVACIÓN DE 9 A PARA ELIMINAR LA
ACTIVIDAD CATALÍTICA)

GAGPOL1 V3 DE MOSAICO (SECUENCIA DE AA)

10 SEQ ID NO: 21

MGARASVLSGGELDRWEKIRLRPGGKKKYRLKHIVWASRELERFAVNPGLLEETSEGC
RQILGQLQPSLQGTGSEELRSLYNTVATLYCVHQRIEIKDTKEALEKIEEEQNKSKKK
AQQAADTGNSSQVVSQNYPIVQNIQGMVHQAI SPRTLNAWVKVVEEKAFSPEVIM
FSALSEGATPQDLNTMLNTVGGHQAMQMLKETINEEAAEWDRVHPVHAGPIAPGQM
REPRGSDIAGTTSTLQEQIGWMTNPNPPIPVGEIYKRWIILGLNKIVRMYSPVSI
RQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTTETLLVQANANPDCKTILKALGPAAT
LEEMMTACQGVGGPGHKARVLAEMSQVTNSATIMMQRGNFRNQRKTVKCFNCGKEG
HIAKNCRAPRKKGCWKCQKEGHQMKDCTERQANFLGKIWPSNKGPRGNFLQNRPEPT
APPEESFRFGEETTTSPQKQEPIDKEMYPLASLKSLEFGNDPSSQMAPI SPIETVPVK
LKPMDGPRVKQWPLTEEKIKALTAICEEMEKEGKITKIGPENPYNTPVFAIKKSDS
TKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLA VGDAYFSVPLDEGFR
KYTAFTIPSTNNETPGIRYQYNVLPQGWKGS PAIFQCSMTRILEPFRAKNPEIYIQ

YMAALYVGS DLEIGQHRAKIEELREHLLKWGFTTPDKKHQKEPPFLWMGYELHPDKW
TVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLT
EEAELELAENREILKEPVHGVYDPSKDLIAEIQKQGHQDQWYQIYQEPFKNLKTGK
YAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKTPKFRLPIQKETWETWWTDYWQATW
IPEWEFVNTPPLVKLWYQLEKDP IAGVETFYVAGAANRETKLGKAGYVTDGRGRQKIV
SLTETTNTQKTALQAIYLAALQDSGSEVNIVTASQYALGI IQA QPKSESELVNQII EQ
LIKKERVYLSWVPAHKGIGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRA
MASDFNLPPVVAKEIVASCDQCQLKGEAMHGQVDCSPGIWQLACTHLEGKII LVAHV
VASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIHTANGSNFTSAAVKAACWWAGI
QQEFGIPYNPQSQGVVASMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIG
GYSAGERIIDIIATDIQTKE LQKQI IKIQNFRVYYRDSRDP IWKGPAKLLWKGE GAV
VIQDNSDIKVVPRRKVKI IKDYGKQ MAGADCVAGRQDED

GAGPOL2 V3 DE MOSAICO (SECUENCIA DE AA)

15

SEQ ID NO: 22

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGC
 KQIIKQLQPALQGTGTEELRSLFNTVATLYCVHAEIEVRDTKEALDKIEEEQNKSQK
 TQQAKEADGKVSQNYPIVQNLQGMVHQPISPRTLNAWVKVIEEKAFSPEVIPMFTA
 LSEGATPQDLNNTMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPVAPGQMRP
 RGSDIAGTTSNLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRMYSPSILDIKQG
 PKEPFRDYVDRFFKTLRAEQATQDVKNWMTDPLLQANANPDCKTILRALGPGATLEE
 MMTACQGVGGPSHKARVLAEAMSQTNSTILMQRSNFKGSKRIVKCFNCGKEGHIARN
 CRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAE
 SFRFEETTPAPKQEPKDRPLTSLRSLFGSDPLSQMAPISPIETVPVKLPGMDGPK
 VKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPIFAIKKDKSTKWRKLVDF
 RELNKRTQDFWEVQLGIPHPAGLKKKSVTVLAVGDAYFSVPLDEDFRKYTAFTIPS
 INNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRKQNPDIVIYQYMAALYVGS
 DLEIGQHRTKIEELRQHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPE
 KDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTKALTEVVPLTEEALELAE
 NREILKEPVHGVYYDPSKDLIAEIQKQGQGWTYQIYQEPFKNLKTGKYARMRGAHT
 NDVKQLTEAVQKIATESIVIWGKTPKFKLPIQKETWEAWWTEYWQATWIPWEFVNT
 PPLVKLWYQLEKEPIVGAETFYVAGAANRETKLGKAGYVTDGRGRQKVSLTDTTNQK
 TALQAIHLALQDSGLEVNIVTASQYALGI IQAOPDKSESELVSQIEQLIKKEKYYL
 AWVPAHKGIGGNEQVDKLVSRGIRKVLFLDGIDKAQEEHEKYHSNWRAMASEFNLP
 IVAKEIVASCDKCQLKGEAIHGQVDCSPGIWQLACTHLEGKVLVAVHVASGYIEAE
 VIPAETGQETAYFLLKLAGRWPVKTIHTANGSNFTSATVKAACWWAGIKQEBGIPYN
 PQSQGVVASINKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGEYSAGERIV
 DIIASDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPALLWKGEGAVVIQDNSDIK
 VVPRRKAKIIRDYGKQMGDDCVASRQDED

5 GAG-POL V3 DE MOS3 (SECUENCIAS DE AA)

FUSIÓN DE GAG-POL DE 1359 aa CON GAG COMPLETO Y POL MODIFICADO

SEQ ID NO: 23

10

MGARASVLSGGKLDWEKIRLRPGGKKKYKHLKHIWASRELDRLFALNPGLLETAEGC
 QQIIEQLQPALQGTGSEELKSLYNTVAVLYCVHQRIDVKDTKEALDKIEEIQNKSKQK

TQQAADTGS SSKVSNYP I VQNAQGQMVHQALSPRTLNAWVKVVEEKGFNPEV I PM
FSALAE GATPQDLNMM LNI VGGHQAA MQILKDTINEEAADWDR LHVPVHAGPIPPGQM
REPRGSDIAGTTSTPQE QIGWMTSNPPVPVGEIYKRWIIMGLNKIVRMYSPVSILDI
KQGPKE SFRDYVDRFFKVLRAEQATQEVKNWMTETLLIQNANPDCKSILRALGPGAS
LEEMMTACQGVGGP SHKARILAEAMSQANNTNIMMQRGNFKGQKRIKCFNCGKEGHL
ARNCRAPRKRGCWKCGREGHQMKDCNERQANFLGKIWPSSKGRPGNFPQSRPEPTAP
LEPTAPPAEPTAPPAESFGFGEEITPSPKQEQKDREPLTSLKSLFGSDPLLQMAPIS
PIDTVPVTLKPGMDGPKIKQWPLTEBKIKALTEICTEMEKEGKISRIGPENPYNTPV
FAIKKNSTRWRKLVDFRELNKKTQDFWEVQLGIPHPAGLKKKRSVTVLAVGDAYFS
VPLDKDFRKYTAFTIPSVNNETPGVRYQYNVLPQGWKGS PAIFQCSMTKILEPFRAQ
NPEIV IYQYVAALYVGS DLEIEQHRTKIEELRAHLLSWGFTTPDKKHQREPPFLWMG
YELHPDRWTVQPIELPEKESWTVNDIQKLVGKLNWASQIYPGIKVKQLCRLLRGAKA
LTEVIPLTKEAELELAENREILREP VHGVYDPSKDLVAEIQKQGQDQWTYQIYQEP
YKNLKTGKYARKRSAHTNDVRQLTEAVQKIALESIVIWGKIPKFRLP IQRETWETWW
TEYWQATWIPDWEFVNT PPLVKLWYOLEKEPIAGAETFYVAGASNRETKIGKAGYVT
DKGRQKVVSLTETTNOKAALQAIQLALQDSGPEVNIVTASQYV LGGIIQAQPDRSESE
LVNQIIEELIKKEKVYLSWVPAHKGIGGNEQVDKLVSAGIRKILFLDGIDKAQEEHE
RYHSNWR TMASDFNL PPIVAKEIVANCDKCQLKGEAMHGQVDCSPGMWQLACTHLEG
KIIIVAVHVASGYMEAEVIPAETGQETAYYILKLAGRWPVKV VHTANGSNFTSTTVK
AACWWANVTQEFGI PYNPQSQGV IASMNKELKKIIGQVREQAEHLKTAVQMAVLIHN
FKRRGGIGGYSAGERIVDIATDIQTRELQKQIKIQNFRVYFRDSRDPVWKGP AKL
LWKGEGAVVIQDNSEIKVVP RRKVKIIRDY GKQ MAGDDCVAGRQDEDQ

VIII. SECUENCIAS DE FUSIÓN DE GAGPOL DE MOSAICO M BIVALENTES (VERSIÓN 4; POL MÍNIMAMENTE INACTIVADO, PR-RT-IN COMPLETO)

5

GAGPOL1 V4 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 24

MGARASVLSGGELDRWEKIRLRPGGKKKYRLKHIVWASRELERFAVNPGLLETSEGC
RQILGQLQPSLQ TGSEELRSLYNTVATLYCVHQRIEIKDTKEALEKIEEEQNKSKKK
AQQAADTGNSSQVSNYP I VQNIQGQMVHQALSPRTLNAWVKVVEEKAFSPEV I PM
FSALSEGATPQDLN TMLNTVGGHQAA MQMLKETINEEAAEWDRVHPVHAGPIAPGQM
REPRGSDIAGTTSTLQE QIGWMTNPPPIPVGEIYKRWIILGLNKIVRMYSPVSILDI
RQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQNANPDCKTILKALGPAAT
LEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKCFNCGKEG
HIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWP SNKGRPGNFLQNRPEPT
APPEESFRFGEETTTSPSQKQEPIDKEMYPLASLKSLFGNDPSSQRENLAFOQGEARE
FPSEQTRANSPTSRELQVRGDNPHSEAGAERQGT LNFPQITLWQRPLVSIKVGQIR
EALLATGADDTVLEDINLPGKWKPKMIGGIGGF I KVGQYDQILIEICGKKAIGTVLV
GPTPVNIIGRNMLTQLGCTLNFPISPIETVPVKLPGMDGPRVKQWPLTEEKIKALT
AICEEMEKEGKITKIGPENPYNTPVFAIKK KDKTKWRKLVDFRELNKR TQDFWEVQL
GIPHPAGLKKKKSVTVLDVGDAYFSVPLDEGFRKYTAFTIPSTNNETPGIRYQYNV L
PQGWKGS PAIFQCSMTRILEPFRAKNPEIV IYQYMDHLYVGS DLEIGQHRAKIEELR
EHLLKWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGK
LNWASQIYPGIKVRQLCKLLRGAKALTDIVPLTEEAELELAENREILKEPVHGVYD

10

PSKDLIAEIQKQGHQDQWYQIYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIAM
 ESIVIWGKTPKFRPLPIQKETWETWWTDYWQATWIPEWEFVNTPPLVKLWYQLEKDPI
 AGVETFFYVDGAANRETKLGKAGYVTDGRGRQKIVSLTETTNOKTELOAIYLLALQDSGS
 EVNIVTDSQYALGIIQAQPKSESELVNOIIEQLIKKERVYLSWVPAHKGIGGNEQV
 DKLVS SGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASDFNLPPVVAKEIVASCDQCQL
 KGEAMHGQVDCSPGIWQLACTHLEGKIIILVAVHVASGYIEAEVIPAETGQETAYFLL
 KLAGRWPVKVIHTDNGSNFTSAAVKAACWWAGIQQEFGI PYNPQSQGVVESMNKELK
 KIIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIIDIIATDIQTKELQKQ
 IIKIQNFRVYYRDSRDP IWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKVKI IKDYGK
 QMAGADCVAGRQDED

GAGPOL2 V4 DE MOSAICO (SECUENCIA DE AA)

5 SEQ ID NO: 25

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGC
 KQIIKQLQPALQGTTEELRSLFNTVATLYCVHAEIEVRDTKEALDKIEEQNKSOQK
 TQQAKEADGKVSQNYPIVQNLQGMVHQPI SPRTLNAWVKVIEEKAFSPEVIMFTA
 LSEGATPQDLNMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPVAPGQMRP
 RGSDIAGTTSNLQEQIAWMTSNPPIPVGDIYKRWIIILGLNKIVRMYSPTSILDIKQG
 PKPEFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEE
 MMTACQGVGGPSHKARVLAEMSQTNSTILMQRSNFKGSKRIVKCFNCGKEGHIARN
 CRAPRKKGCWCKGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAE
 SFRFEETTPAPKQEPKDREPLTSLRSLFGSDPLSQRENLAFFQGGKAREFSSEQTRAN
 SPTRRELQVWGRDNNLSLSEAGADRQGTVSFSFPQITLWQRPLVTIKIGGOLKEALLA
 TGADDTVLEEMNLGRWPKMIGGIGGFIKVGQYDQIPIEICGHKAIGTVLVGPTPV
 NIIGRNLLTQIGCTLNFPISPIETVPVKLPGMDGPKVKQWPLTEEKIKALVEICTE
 MEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHP
 AGLKKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIIPSINNETPGIRYQYNVLPQGWK
 GSPAIFQSSMTKILEPFRKQNPDIYIYQYMDHLYVGS DLEIGQHRTKIEELRQHLLR
 WGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPKDSWTVNDIQKLVGKLNWAS
 QIYAGIKVKQLCKLLRGTKALTEVVPLTEEAELELAENREILKEPVHGVYYDPSKDL
 IAEIQKQGGQWYQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKIATESIVI
 WGKTPKFKLPIQKETWEAWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAET
 FYVDGAANRETKLGKAGYVTDGRGRQKVVSLTDTTNQKTELOAIHLALQDSGLEVNIV
 TDSQYALGIIQAQPKSESELVSQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVS
 RGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASEFNLPPIVAKEIVASCDKQKLGKAI
 HGQVDCSPGIWQLACTHLEGKVIILVAVHVASGYIEAEVIPAETGQETAYFLLKLAGR
 WPVKTIHTDNGSNFTSATVKAACWWAGIKQEFGI PYNPQSQGVVESINKELKKIIGQ
 VRDQAEHLKTAVQMAVFIHNFKRKGGIGEYSAGERIVDIIASDIQTKELQKQITKIQ
 NFRVYYRDSRDP LWKGPAKLLWKGEGAVVIQDNSDIKVVPRRKAKIIRDYGKQ MAGD
 DCVASRQDED

10 IX. SECUENCIAS DE FUSIÓN DE GAGPOL DE MOSAICO M BIVALENTES (VERSIÓN 5; POL MÍNIMAMENTE INACTIVADO, CON SUPRESIÓN DE PR)

GAGPOL1 V5 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 26

MGARASVLSGGELDRWEKIRLRPGGKKKYRLKHIVWASRELERFAVNPGLLETSEGC
 RQILGQLQPSLQGTGSEELRSLYNTVATLYCVHQRIEIKDTKEALEKIEEEQNKSKKK
 AQQAAADTGNSSQVSQNYPIVQNIQGQMVHQAI SPRTLNAWVKVVEEKAFSPEVIM
 FSALSEGATPQDLNNTMLNTVGGHQAMQMLKETINEEAAEWDRVHPVHAGPIAPGQM
 REPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSVSI LDI
 RQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQANANPDCKTILKALGPAAT
 LEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKCFNCGKEG
 HIAKNCRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSNKGRPGNFLQNRPEPT
 APPEESFRFGEETTTTPSQKQEPIDKEMYPLASLKS LFGNDPSSQMAPISPIETVPVK
 LKPGMDGPRVKQWPLTEEKIKALTAICEEMEKEGKITKIGPENPYNTPVFAIKK KDS
 TKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLVDVGDAYFSVPLDEGFR
 KYTAFTIPSTNNETPGIRYQYNVLPQGWKGSPAIFQCSMTRILEPFRAKNPEI VIYQ
 YMDHLYVGS DLEIGQHRAKIEELREHLLKWGF TTPDKKHQKEPPFLWMGYELHPDKW
 TVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLT
 EEAELELAENREILKEPVHGVYDPSKDLIAEIQKQGHQWTYQIYQEPFKNLKTGK
 YAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKTPKFRLPIQKETWETWWTDYWQATW
 IPEWEFVNTPLV KLWYQLEKDP IAGVETFYVDGAANRETKLGKAGYVTD RGRQKIV
 SLTETT NQKTELQAIY LALQDSGSEVNI VTD SQYALGIIQAQPKSESELVNQII EQ
 LIKKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEEHEKYHSNWRA
 MASDFNLPPVVAKEIVASCDQCQLKGEAMHGQVDCSPGIWQLACTHLEGKIILVAVH
 VASGYIEAEVIPAETGQETAYFILKLAGRWPVKVIHTDNGSNFTSAAVKAACWWAGI
 QQEFGIPYNPQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIG
 GYSAGERIIDIIATDIQTKELQKQIIKIQNFRVYYRDSRDP IWKGP AKLLWKGEGAV
 VIQDNSDIKVVPRRKVKIKDYGKQ MAGADCVAGRQDED

5 GAGPOL2 V5 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 27

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGC
 KQIIKQLQPALQGTGTEELRSLFNTVATLYCVHAEIEVRDTKEALDKIEEEQNKSOQK
 TQOAKEADGKVSQNYPIVQNLQGMVHQPI SPRTLNAWVKVIEEKAFSPEVIMFTA
 LSEGATPQDLNMLNTVGGHQAAMQMLKDTINEEAAEWDRHLHPVHAGPVAPGQMRP
 RGSDIAGTTSNLQEQIAWMTSNPPIPVGDIYKRWII LGLNKIVRMYSPSILDIKQG
 PKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEE
 MMTACQGVGGPSHKARVLAEAMSQTNSTILMQRSNFKGSKRIVKCFNCGKEGHIARN
 CRAPRKKGCWKCQKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAE
 SFRFEETTPAPKQEPKDRPLTSLRSLFGSDPLSQMAPISPIETVPVKLKPMDGPK
 VKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPIFAIKKDKSTKWRKLVDF
 RELNKRTQDFWEVQLGIPHPAGLKKKSVTVLDVGDAYFSVPLDEDFRKYTAFTIPS
 INNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILEPFRKQNPDIVIYQYMDHLYVGS
 DLEIGQHRTKIEELRQHLLRWGFTTPDKKHQKEPPFLWMGYELHDPKWTVPQIVLPE
 KDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTALTEVVPLTEEALELAE
 NREILKEPVGHVYDPSKDLIAEIQKQGQWQTYQIYQEPFKNLKTGKYARMGAHT
 NDVKQLTEAVQKIATESIVIWGKTPKFKLPIQKETWEAWWTEYWQATWIPWEFVNT
 PPLVKLWYQLEKEPIVGAETFYVDGAANRETCLGKAGYVTDGRGRQVVS LTDTTNOK
 TELQAIHLALQDSGLEVNIVTDSQYALGI IQAQPDKSESELVSQIIEQLIKKEKVYL
 AWVPAHKGIGGNEQVDKLVSRGIRKVLFLDGIDKAQEEHEKYHSNWRAMASEFNLPP
 IVAKEIVASCDCQKGEAIHGQVDCSPGIWQLACTHLEGKVI LVAHVHVASGYIEAE

VIPAETGQETAYFLLKLAGRWPVKTIHTDNGSNFTSATVKAACWWAGIKQEFGIPYN
 PQSQGVVESINKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGEYSAGERIV
 DIIASDIQTKELQKQITKIQNFRVYYRDSRDPLWKGPALLWKGEGAVVIQDNSDIK
 VVPRRKAKIIRDYGKQMGDDCVASRQDED

X. SECUENCIAS DE FUSIÓN DE GAGPOLNEF DE MOSAICO M BIVALENTES (POL EXTENSIVAMENTE INACTIVADO, CON SUPRESIÓN DE PR)

5

GAGPOLNEF1 DE MOSAICO (SECUENCIA DE AA)

SEQ ID NO: 28

MGARASVLSGGELDRWEKIRLRPGGKKKYRLKHIVWASRELERFAVNPGLLETSEGC
 RQILGQLQPSLQGTGSEELRSLYNTVATLYCVHQRIEIKDTKEALEKIEEEQNKSKKK
 AQQAAADTGNSSQVSONYPIVQNIQGQMVHQAI SPRTLNAWVKVVEEKAFSPEVIM
 FSALSEGATPQDLNMLNTVGGHQAMQMLKETINEEAAEWDRVHPVHAGPIAPGQM
 REPRGSDIAGTTSTLQEQIGWMTNNPPIPVGEIYKRWIILGLNKIVRMYSPVSILDI
 RQGPKEPFRDYVDRFYKTLRAEQASQDVKNWMTETLLVQANANPDCKTILKALGPAAT
 LEEMMTACQGVGGPGHKARVLAEAMSQVTNSATIMMQRGNFRNQRKTVKCFNCGKEG
 HIAKNCRAPRKKGCWKCCKEGHQMKDCTERQANFLGKIWPSNKGKRPNGFLQNRPEPT
 APPEESFRFGEETTTSPSQKQEPIDKEMYPLASLKSFLGNDPSSQMAPI SPIETVPVK
 LKPGMDGPRVKQWPLTEEKIKALTAICEEMEKEGKITKIGPENPYNTPVFAIKKKDS
 TKWRKLVDFRELNRKRTQDFWEVQLGIPHPAGLKKKKSVTVLAVGDAYFSVPLDEGFR
 KYTAFTIPSTNNETPGIRYQYNVLPQGWKGS PAIFQCSMTRILEPFRAKNPEI VIYQ
 YMAALYVGS DLEIGQHRAKIEELREHLLKWGF TTPDKKHQKEPPFLWGMGYELHPDKW
 TVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKLLRGAKALTDIVPLT
 EEAELELAENREILKEPVHGVYDPSKDLIAEIQKQGHQDQWYQIYQEPFKNLKTGK
 YAKMRTAHTNDVKQLTEAVQKIAMESIVIWGKTPKFRLP IQKETWETWWTDYQATW
 IPEWEFVNT PPLVKLWYQLEKDP IAGVETFYVAGAANRETKLGKAGYVTDGRQKIV
 SLTETTQKTAALQAIYLA LQDSGSEVNI VTA S QYALGIIQAQPKSESELVNQII EQ
 LIKKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRA
 MASDFNLPPVVAKEIVASCDQCQLKGEAMHGQVDCSPGIWQLACTHLEGKIIILVAVH
 VASGYIEAEVI PAETGQETAYFILKLAGRWPVKVIHTANGSNFTSAAVKAACWWAGI
 QQEFGIPYNPQS QGVVASMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIG
 GYSAGERIIDIIATDIQTKELQKQIIKIQNFRVYYRDSRDP IWKGPAKLLWKGEGAV
 VIQDNSDIKVVPRRKVKI IKDYGKQ MAGADCVAGRQEDMAGKWSKSSVVGWPAIRE
 RMRAEPAADGVGAVSRDLEKHGAITSSNTAANNADCAWLEAQEEEEVGFPPVRPQVP
 LRPMTYKGALDLSHFLKEKGGLEGLIYSQKRQDILDLWVYHTQGYFPDWQNYTPGPG
 IRYPLTFGWCFKLVPEPEKIEEANEGENNSLLHPMSQHGMDDPEKEVLMWKFD SRL
 AFHHMARELHPEYYKDC

GAGPOLNEF2 DE MOSAICO (SECUENCIA DE AA)

5 SEQ ID NO: 29

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGC
 KQIIKQLQPALQGTGTEELRSLFNTVATLYCVHAEIEVRDTKEALDKIEEEQNKSQK
 TQQAKEADGKVSQNYPIVQNLQGQMVHQPI SPRTLNAWVKVIEEKAFSPEVIMFTA

LSEGATPQDLNMLNTVGGHQAAAMQMLKDTINEEAAEWDRRLHPVHAGPVAPGQMPREP
 RGSDIAGTTSNLQEQIAWMTSNPPIPVGDIYKRWIILGLNKIVRMYSPSILDIKQG
 PKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEE
 MMTACQGVGGPSHKARVLAEMSQTNSTILMQRSNFKGSKRIVKCFNCGKEGHIARN
 CRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAE
 SFRFEETTPAPKQEPKDREPLTSLRSLFGSDPLSQMAPISPIETVPVKLKPMDGPK
 VKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNTPIFAIKKKDSTKWRKLVDF
 RELNKRTQDFWEVQLGIPHPAGLKKKSVTVLAVGDAYFSVPLDEDFRKYTAFTIPS
 INNETPGIRYQYNVLPQGWKGSPIFQSSMTKILEPFRKQNPDIVIYQYMAALYVGS
 DLEIGQHRTKIEELRQHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIVLPE
 KDSWTVNDIQKLVGKLNWASQIYAGIKVKQLCKLLRGTALTEVVPLTEEALELAE
 NREILKEPVHGVYDPSKDLIAEIQKQGQGWTYQIYQEPFKNLKTGKYARMGAHT
 NDVKQLTEAVQKIATESIVIWGKTPKFKLPIQKETWEAWWTEYWQATWIPWEFVNT
 PPLVKLWYQLEKEPIVGAETFYVAGAANRETGLGKAGYVTDGRGRQKVSLTDTTNQK
 TALQAIHLALQDSGLEVNIVTASQYALGIIOAQPKSESELVSQIIEQLIKKEKVYL
 AWVPAHKGIGGNEQVDKLVSRGIRKVLFLDGIDKAQEEHEKYHSNWRAMASEFNLP
 IVAKEIVASCDCQKQKGEA IHGQVDCSPGIWQLACTHLEGKVIILVAVHVASGYIEAE
 VIPAETGQETAYFLLKLAGRWPVKTIHTANGSNFTSATVKAACWWAGIKQEFGPIPN
 PQSQGVVASINKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGEYSAGERIV
 DIIASDIQTKELQKQITKIQNFRVYRDSRDPLWKGPAKLLWKGEGAVVIQDNSDIK
 VVPRRKAKIIRDYGKQMGDDCVASRQDEDMAGKWSKSSIIVGWPAVRERIRRAEPAA
 EGVGAASQDLDKYGALTSNTAATNADCAWLEAQEEDVEVGFVVKPQVPLRPMTYKAA
 FDLSEFLKEKGLDGLIYSKKRQEIIDLWVYNTQGFPPDWQNYTPGPGVRYPLTFGW
 CFKLVPVDPREVEEANKGENNCLLHPMNLHGMDPPEREVLVWRFDSRLAFHHMAREK
 HPEYYKNC

XI. SECUENCIAS DE ENV GP160, GAG, POL, NEF DE CLADO C ÓPTIMO

5 ENV GP160 DE CLADO C ÓPTIMO (SN90.90.SE364) (SECUENCIA DE AA)

SEQ ID NO: 30

MRVTGMLRNCQPWWIWGILGFWMLLIYNVGGNLWVTVYYGVPVWKEAKTTLFCASDA
 KAYEKEVHNWATHACVPTDPNPQEMVLENVTEYFNMWKNDMVDQMHEDIISLWDQS
 LKPCVKLTPLCVTLNCRNVTTSNNATSNNDNPNGEIKNCSFNITTELDRKRRNEYALF
 YRLDIVPLSGSKNSSNSSEYRLINCNTSAITQACPKVSFDPPIPIHYCAPAGYAILKC
 NNKTFNGTGPCNNVSTVQCTHGKIPVVSTQLLLNGSLAEGEIIIRSENLTNNAKTII
 VHLNESIEIVCARPNNNTRKSMRIGPGQTFYATGDIIGDIRQAHCNISGNWNATLEK
 VKGKLQEHFPGKNISFEPSSGGDLEITTHSFNCRGEFFYCDTSKLFNGTTHANSI
 TIQCRIKQIINMWQGVGRAIYAPPIAGNITCKSNITGLLLTRDGGTLNNDTEKFRPG
 GGDNRDNWRSELYKYKVEIKPLGIAPTAKRRVVEREKRAVGIGAVFLGLGAAGS
 TMGAASITLTVQARQLLSGIVQQSNLLRAIEAQQHMLQLTVWGIKQLQTRVLAIER
 YLKDQQLLGIWGCSGKIICTTAVPWNTSWSNKSLEDIWDNMTWMQWDREINNYTSII
 YSLLEESQNQQEKNEKDLLALDSWNNLWNWFNITKWLWYIKIFIMIVGGLIGLRIIF
 AVLSIVNRVRQGYSPLSFQTLIPNPRGPDRLGRIIEEGGEQDRDRSIRLVNGFLAIA
 WDDLRLSLCLFSYRRLRDFILIVARAVELLIQRGWETLKYLGSL?QYWGLELKKSAIS
 LLDTIAITVAEGTDRIIELVQRICRAISNIPRRIRQGFEEALQ

10

GAG DE CLADO C ÓPTIMO (IN.70177) (SECUENCIA DE AA)

SEQ ID NO: 31

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGC
KQILKQLQPALQTGTEELRSLYNTVATLYCVHAGIEVRDTKEALDKIEEEQNKGOQK
TQQAAGADGKVSQNYPIVQNLQGMVHQAI SPRTLNNAWVKVIEEKAFSPEVPMFTA
LSEGATPQDLNMLNTVGGHQAAMQMLKDTINEEAAEWDRLHPVHAGPIAPGQMRP
RGSDIAGTTSTLQEQIAWMTNPPVPGDIYKRWIILGLNKIVRMYSVPSILDIKQG
PKEFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEE
MMTACQGVGGPSHKARVLAEAMSQTGSTIMMQRSNFKGSKRIVKCFNCGKEGHIARN
CRAPRKKGCWKCQKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAE
SFRFEETTPAPKQELKDREPLTSLKSLFGSDPLSQ

5 POL DE CLADO C ÓPTIMO (ZA.04.04ZASK208B1) (SECUENCIA DE AA)

SEQ ID NO: 32

FFRENLAFFQQGEAREFPSEQARANSPTSREFQVRGDNPCSEAGVKGQGTNLNFPQITL
WQRPLVSIKVGQVKEALLDTGADDTVLEEINLPGKWKPKMIGGIGGFIKVRQYDQI
LIEICGKKAIGTVLVGPTPVNIIGRNMLTQLGCTLNFPISPIETVPVCLKKPGMDGPK
IKQWPLTEEKIKALMAICEEMEKEGKITKIGPENPYNTPIFAIKKDKSTKWRKLVDF
RELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDESEFRKYTAFTIPS
INNETPGIRYQYNVLPQGWKGSFAIFQSSMTKILEPFRAKNPEIIVYQYMDLTVGS
DLEIGQHRAKIEELREHLLRWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIQLPE
KDSWTVNDIQKLVGKLNWASQIYSGIKVRQLCKLLRGAKALTDIVPLTEEALELAE
NREILKEPVHGVYDPSKDLIAEIQKQGYDQWYQIYQEPFKNLKTGKYAKMRTAHT
NDVKQLTEAVQKIALESIVIWGKTPKFRLPKIQKETWEIWWTDYWQATWIPEWEFVNT
PPLVKLWYQLEKEPIAGAETFYVDGAANRETKIGKAGYVTDKGRQKIVTLTETTNOK
TELQAIQLALQDSGSEVNIVTDSQYALGI IQAQPKSESELVNOIIEQLINKERVYL
SWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASEFNLPP
VVAKEIVASCDKQCQLKGEATHGQVDCSPGIWQLDCTHLEGKVILVAHVHVASGYMEAE
VIPAETGQETAYYILKLAGRWPVKVIHTDNGSNFTSAAVKAACWWAGIQQEFGIPYN
PQSQGVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERII
DIIATDIQTKELQKQIIKIQNFRVYRDSRDP IWKGPALLWKGEAVVIQDNNDIK
VPPRRVKIKDYGKQMGADCVAGRQDED

10 NEF DE CLADO C ÓPTIMO (ZA00.1170MB) (SECUENCIA DE AA)

SEQ ID NO: 33

MGGKWSKSSIVGWPDVRRERMRRTEPAAGVGAASQDLDDKYGALTSNTTHNNADCAW
LEAQEEGEGVGFVVRPQVPLRPMTYKGAFDLSFFLKEKGLDGLIYSKKRQEILDWV
YHTQGFPPDWQNYTPGPGVRYPLTFGWCFKLVVDPREVEEANKGENNCLLHPMSLH
GMEDEEREVLKWEFDSSLARRHLARELHPEYYKDC

15 XII. SECUENCIA DE ENV GP140 DE CLADO C ÓPTIMO (DEFECTUOSA EN ESCISIÓN/FUSIÓN)

ENV GP140 DE CLADO C ÓPTIMO (SN90.90.SE364) (SECUENCIA DE AA)

20 SEQ ID NO: 34

MRVTGMLRNCQPWWIWIWILGFWMLLIYNVGGNLWVTVYYGVPVWKEAKTTLFCASDA
KAYEKEVHNWVATHACVPTDPNPQEMVLENVTEYFNMWKNDMVDQMHEDEIISLWDQS
LKPCVKLTPLCVTLNCRNVTTSNATSNDNPNGEIKNCSFNITTELDRKRRNEYALF
YRLDIVPLSGSKNSSNSSEYRLINCNTSAITQACPKVSFDPIPIHYCAPAGYAILKC
NNKTFNGTGPCNNVSTVQCTHGIKPVVSTQLLLNGSLAEGEIIIRSENLTNNAKTII
VHLNESIEIVCARPNNTRKSMRIGPGQTFYATGDIIGDIRQAHCNISGNWNATLEK
VKGKLOEHFPGKNISFEPSSGGDLEITTHSFNCRGEFFYCDTSKLFNGTHTANSI
TIQCRIKQIINMWQGVGRAIYAPPIAGNITCKSNITGLLLTRDGGTLNNDTEKFRPG
GGDMRDNRSELYKYKVVEIKPLGIAPTAKRRVVESEKSAVGIGAVFLGLGAAGS
TMGAASITLTVQARQLLSGIVQQSNLLRAIEAQQHMLQLTVWGIKQLQTRVLAIER
YLKDQQLLGIWGCSGKIICTTAVPWNTSWSNKSLEDIWDNMTWMQWDREINNYTSII
YSLLEESQNQQEKNEKDLLALDSWNNLWNWFNITKWLW

XIII. SECUENCIA DE POL DE CLADO C ÓPTIMO (EXTENSIVAMENTE INACTIVADO, CON SUPRESIÓN DE PR)

5 POL DE CLADO C ÓPTIMO (ZA.04.04ZASK208B1) (SECUENCIA DE AA)

SEQ ID NO: 35

MAPISPIETVPVVKLKPMDGPKIKQWPLTEEKIKALMAICEEMEKEGKITKIGPENP
YNTPIFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKSVTVLAVG
DAYFSVPLDESFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPIAFQSSMTKILE
PFRAKNPEIUIYQYMAALYVGSLEIGQHRAKIEELREHLLRWGFTTPDKKHQKEPP
FLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYSGIKVRQLCKLL
RGAKALTDIVPLTEEALELAENREILKEPVHGVYDPSKDLIAEIQKQGYDQWYQ
IYQEPFKNLKTGKYAKMRTAHTNDVKQLTEAVQKIALESIVIWGKTPKFRPLPIQKET
WEIWWTDYWQATWIPEWFEVNTPLVKLWYQLEKEPIAGAETFYVAGAANRETKIGK
AGYVTDKGRQKIVTLTETTQKTALQAIQLALQDSGSEVNIIVTASQYALGIIQAQPD
KSESELVNQIIEQLINKERVYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKA
QEEHEKYHSNWRAMASEFNLPPVVAKEIVASCDKQKGEAIIHGQVDCSPGIWQLAC
THLEGKVILVAVHVASGYMEAIEVIPAETGQETAYYILKLAGRWPVKVIHTANGSNFT
SAAVKAACWWAGIQQEFGIPYNPQSQGVVASMNKELKKIIGQVRDQAEHLKTAVQMA
VFIHNFKRKGGIGGYSAGERIIDIIATDIQTKELQKQIIKIQNFRVYYRDSRDPKWK
GPAKLLWKGEAVVIQDNSDIKVVPRRKVKI IKDYGKQ MAGADCVAGRQDED

10 XIV. SECUENCIA DE FUSIÓN DE GAGNEF DE CLADO C ÓPTIMO

GAGNEF DE CLADO C ÓPTIMO (IN.70177-ZA00.1170MB) (SECUENCIA DE AA)

15 SEQ ID NO: 36

MGARASILRGGKLDKWEKIRLRPGGKKHYMLKHLVWASRELERFALNPGLLETSEGC
 KQILKQLQPALQGTGTEELRSLYNTVATLYCVHAGIEVRDTKEALDKIEEEQNKGQOK
 TQQAKGADGKVSQNYPIVQNLQGQMVHQAI SPRTLNAWVKVIEEKAFSPEVIMFTA
 LSEGATPQDLNMLNTVGGHQAAMQMLKDTINEEAAEWDRHLHPVHAGPIAPGQMRP
 RGSIDIAGTTSTLQEQIAWMTNPPVPGVDIYKRWIILGLNKIVRMYSPVSIILDIKQG
 PKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQANANPDCKTILRALGPGATLEE
 MMTACQGVGGPSHKARVLAEAMSQTGSTIMMQRSNFKGSKRIVKCFNCGKEGHIARN
 CRAPRKKGCWKCGKEGHQMKDCTERQANFLGKIWPSHKGRPGNFLQSRPEPTAPPAE
 SFRFEETTPAPKQELKDREPLTSLKSLFGSDPLSQAGKWSKSSIVGWPDPVRRMRRT
 EPAAEGVGAASQDLDKYGALTSSTNTHNNADCAWLEAQEEGEGVGFVPRPQVPLRPMT
 YKGAFDLSFFLKEKGLDGLIYSKKRQEILDWVYHTQGFPPDWQNYTPGPGVRYPL
 TFGWCFKLVDPREVVEANKGENNCLLHPSMLHGMEDEEREVLKWEFDSSSLARRHL
 ARELHPEYYKDC

XV. SECUENCIAS CONSENSO

5 ENV DE M CONSENSO

SEQ ID NO: 37

MRVRGIQRNCQHLWRWGTLILGMLMICSAAENLWVTVYYGVPVWKEANTTLF
 CASDAKAYDTEVHNWATHACVPTDPNPQEIIVLENVTENFNMMWKNMVEQM
 HEDIISLWDQSLKPCVKLTPLCVTLNCTNVNVTNNTNNTTEEKGEIKNCSFNITTEI
 RDKKQKVYALFYRLDVVPIDDNNNSSNYRLINCNTSAITQACPKVSFEPIPIHYC
 APAGFAILKCNDDKFNGTGPKNVSTVQCTHGIKPVVSTQLLLNGSLAEEEEIIRS
 ENITNNAKTIIVQLNESVEINCTRPNNNTRKSIRIGPGQAFYATGDIIGDIRQAHCN
 I
 SGTKWNKTLOQVAKKLRHFNNKTIIFKPSSGGDLEITTHSFNCRGEFFYCNTSG
 LFNSTWINGNKTNNNTNDTITLPCRICKIINMWQGVGQAMYAPPIEGKITCKSNI
 TGLLLTRDGGNNNTNETEIFRPGGGMDRDNWRSELYKYKVVKIEPLGVAPTAK
 RRVVESEKSAVGIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQSNLLR
 AIEAQOHLQLTVWGIKQLQARVLAVERYLKDQQLLGIWGC SGKLICTTTVPWN
 SWSNKSQDEIWDNMTWMEWEREINNYTDIIYSLIEESQNQQEKNEQELLALDK
 WASLWNWFDITNWLW

10

GAG DE M CONSENSO

SEQ ID NO: 38

MGARASVLSGGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPGLLET
 SEGCKQIIGQLQPALQGTGSEELRSLYNTVATLYCVHQRIEVKDTKEALEKIEEEQN
 KSQOKTQQAADKGNSSKVSQNYPIVQNLQGQMVHQAI SPRTLNAWVKVIEEK
 AFSPEVIMFSA LSEGATPQDLNMLNTVGGHQAAMQMLKDTINEEAAEWDRL
 HPVHAGPIPPGQMRPGRSIDIAGTTSTLQEQIAWMTSNPPIPVGEIYKRWIILGLN
 KIVRMYSPVSIILDIRQGPKEPFRDYVDRFFKTLRAEQATQDVKNWMTDTLLVQAN
 ANPDCKTILKALGPGATLEEMMTACQGVGGPGHARVLAEAMSQVTNAAIMM
 QRGNFKGQRRIKCFNCGKEGHIARNCRAPRKKGCWKCGKEGHQMKDCTERQA
 NFLGKIWPSNKGRPGNFLQSRPEPTAPPAESFGFGEEITPSPKQEPKDEPPLTSLK
 SLFGNDPLSQ

15

POL DE M CONSENSO

SEQ ID NO: 39

MAPISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALTEICTEMEKEGKISKIGPEN
PYNTPIFAIKKDKSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSXTV
LDVGDAYFSVPLDEDFRKYTAFTTIP SINNETPGIRYQYNVLPQGWKGS PAIFQSSM
TKILEPFRTQNP EIVYQYMDHLYVGS DLEIGQHRAKIEELREHLLRWGFTTPDKK
HQKEPPFLWMGYELHPDKWTVQPIQLPEKDSWTVNDIQKLVGKLNWASQIYPGI
KVKQLCKLLRGAKALTDIVPLTEEALELAENREILKEPVHGVVYDPSKDLIAEIQ
KQGQDQWYQIYQEPFKNLKTGKYAKMRS AHTNDVKQLTEAVQKIATESIVIW
GKTPKFRLP IQKETWETWWTEYWQATWIP EWEFVNT PPLVKLWYQLEKEPIAG
AETFYVDGAANRETKLGKAGYVTD RGRQKV VSLTETT NQKTELQAIHLALQDS
GSEVNIIVTDSQYALGIIQAQPKSESELVNQIIEQLIKKEKVYLSWVPAHKGIGGN
EQVDKLVSTGIRKVLFLDGDIDKAQEEHEKYHSNWRAMASDFNLPPIVAKEIVASC
DKCQLKGEAMHGQVDCSPGIWQLACTHLEGKIIILVAVHVASGYIEAEVIPAETG
QETAYFILKLAGRWPVKVIHTDNGSNFTSAAVKAACWWAGIQQEFGIPYNPQSQ
GVVESMNKELKKIIGQVRDQAEHLKTAVQMAVFIHNFKRKGGIGGYSAGERIID I
IATDIQTKELQKQITKIQNFRVYYRDSRDP IWKGPAKLLWKGE GAVVIQDNSDIK
VPPRRKAKIIRDYGKQ MAGDDCVAGRQDED

5

REIVINDICACIONES

1. Una vacuna que comprende:

- 5 (a) un polipéptido vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9, o
(b), un vector que codifica un polipéptido vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9.

10 2. La vacuna según la reivindicación 1, en donde dicha vacuna comprende dicho polipéptido vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9 y en donde dicha vacuna comprende además:

- a) al menos dos polipéptidos de gag optimizados definidos seleccionados de uno cualquiera o más de los grupos:
(i) SEQ ID NO: 3 y 4, (ii) SEQ ID NO: 3 y 15, y (iii) SEQ ID NO: 4 y 15; y/o
15 b) al menos dos polipéptidos de pol optimizados definidos seleccionados de uno cualquiera o más de los grupos:
(i) SEQ ID NO: 12 y 13, (ii) SEQ ID NO: 12 y 14, y (iii) SEQ ID NO: 13 y 14.

3. La vacuna según la reivindicación 2,
en donde dichos al menos dos polipéptidos de gag optimizados definidos comprenden las secuencias de aminoácidos
expuestas en las SEQ ID NO: 3 y 4, respectivamente, y/o
20 en donde dichos al menos dos polipéptidos de pol optimizados definidos comprenden las secuencias de aminoácidos
expuestas en las SEQ ID NO: 12 y 13, respectivamente.

4. La vacuna según la reivindicación 1, en donde dicha vacuna comprende dicho vector que codifica un polipéptido
vírico optimizado que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 9 y en donde dicha vacuna
25 comprende además uno o más vectores que codifican:

- al menos dos polipéptidos de gag optimizados definidos seleccionados de uno cualquiera o más de los grupos: (i)
SEQ ID NO: 3 y 4, (ii) SEQ ID NO: 3 y 15, y (iii) SEQ ID NO: 4 y 15, y/o
30 al menos dos polipéptidos de pol optimizados definidos seleccionados de uno cualquiera o más de los grupos: (i)
SEQ ID NO: 12 y 13, (ii) SEQ ID NO: 12 y 14, y (iii) SEQ ID NO: 13 y 14.

5. La vacuna según la reivindicación 4, en donde dicha vacuna comprende uno o más vectores que codifican:

- 35 al menos dos polipéptidos de gag optimizados definidos que comprenden las secuencias de aminoácidos
expuestas en las SEQ ID NO: 3 y 4, respectivamente, y/o
al menos dos polipéptidos de pol optimizados definidos que comprenden las secuencias de aminoácidos expuestas
en las SEQ ID NO: 12 y 13, respectivamente.

6. La vacuna según la reivindicación 5, en donde dicha vacuna comprende uno o más vectores que codifican:

- 40 al menos dos polipéptidos de gag optimizados definidos que comprenden las secuencias de aminoácidos
expuestas en las SEQ ID NO: 3 y 4, respectivamente, y
al menos dos polipéptidos de pol optimizados definidos que comprenden las secuencias de aminoácidos expuestas
45 en las SEQ ID NO: 12 y 13, respectivamente.

7. La vacuna según la reivindicación 5 o 6, en donde dicha vacuna comprende además un vector que codifica un
polipéptido de env optimizado que tiene la secuencia expuesta en la SEQ ID NO: 10.

8. La vacuna según una cualquiera de las reivindicaciones 4 a 7, en donde dichos polipéptidos de gag, pol y/o env
50 optimizados están codificados por (i) un único vector o (ii) múltiples vectores.

9. La vacuna de una cualquiera de las reivindicaciones 1 y 4 a 8, en donde dicha vacuna comprende un vector vírico
seleccionado del grupo que consiste en adenovirus de serotipo 26 (Ad26), adenovirus de serotipo 34 (Ad34),
adenovirus de serotipo 35 (Ad35), adenovirus de serotipo 48 (Ad48) o adenovirus de serotipo 5 HVR48 (Ad5HVR48),
55 poxvirus y virus vaccinia Ankara modificado (MVA).

10. La vacuna de la reivindicación 9, en donde dicho vector vírico es un adenovirus de serotipo 26 (Ad26).

11. La vacuna de la reivindicación 9, en donde dicho vector vírico es un virus vaccinia Ankara modificado (MVA).

60 12. La vacuna de una cualquiera de las reivindicaciones 1 a 11, en donde dicha vacuna comprende adicionalmente
un vehículo, excipiente o diluyente farmacéuticamente aceptable.

65 13. La vacuna de una cualquiera de las reivindicaciones 1 a 12 para su uso en un método de tratamiento o reducción
del riesgo de una infección por el virus de la inmunodeficiencia humana de tipo 1 (VIH-1) en un ser humano.

14. Un método de fabricación de una vacuna para el tratamiento o la reducción del riesgo de una infección por el VIH-1 en un ser humano, comprendiendo dicho método sintetizar la vacuna de una cualquiera de las reivindicaciones 1 a 12.

5 15. Un kit que comprende:

- a) la vacuna de una cualquiera de las reivindicaciones 1 a 11;
- b) un vehículo, excipiente o diluyente farmacéuticamente aceptable;
- c) instrucciones para el uso de los mismos; y, opcionalmente,
- d) un adyuvante.

10

Figura 1

Número de Grupos Positivos: 4x fondo

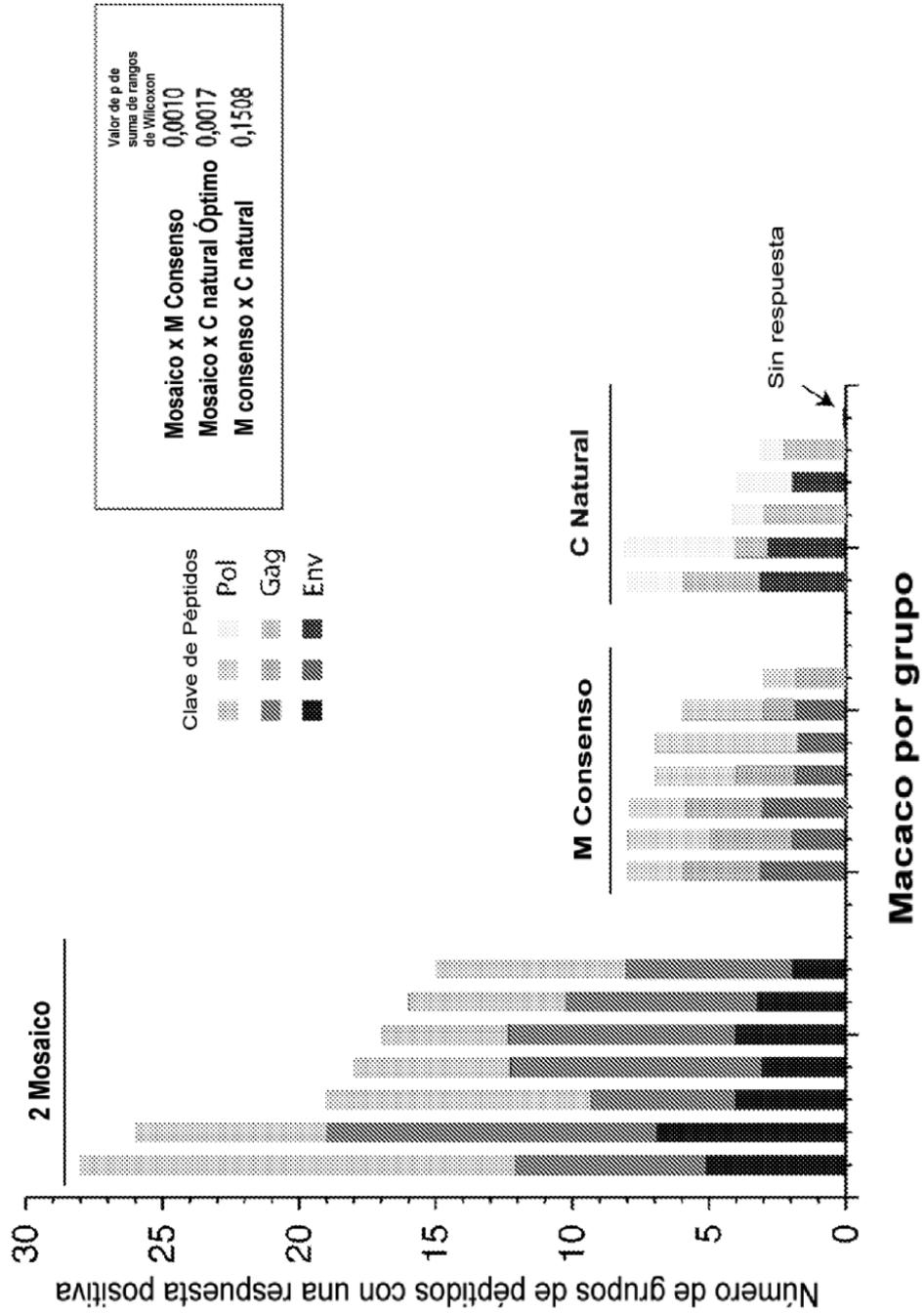


Figura 2

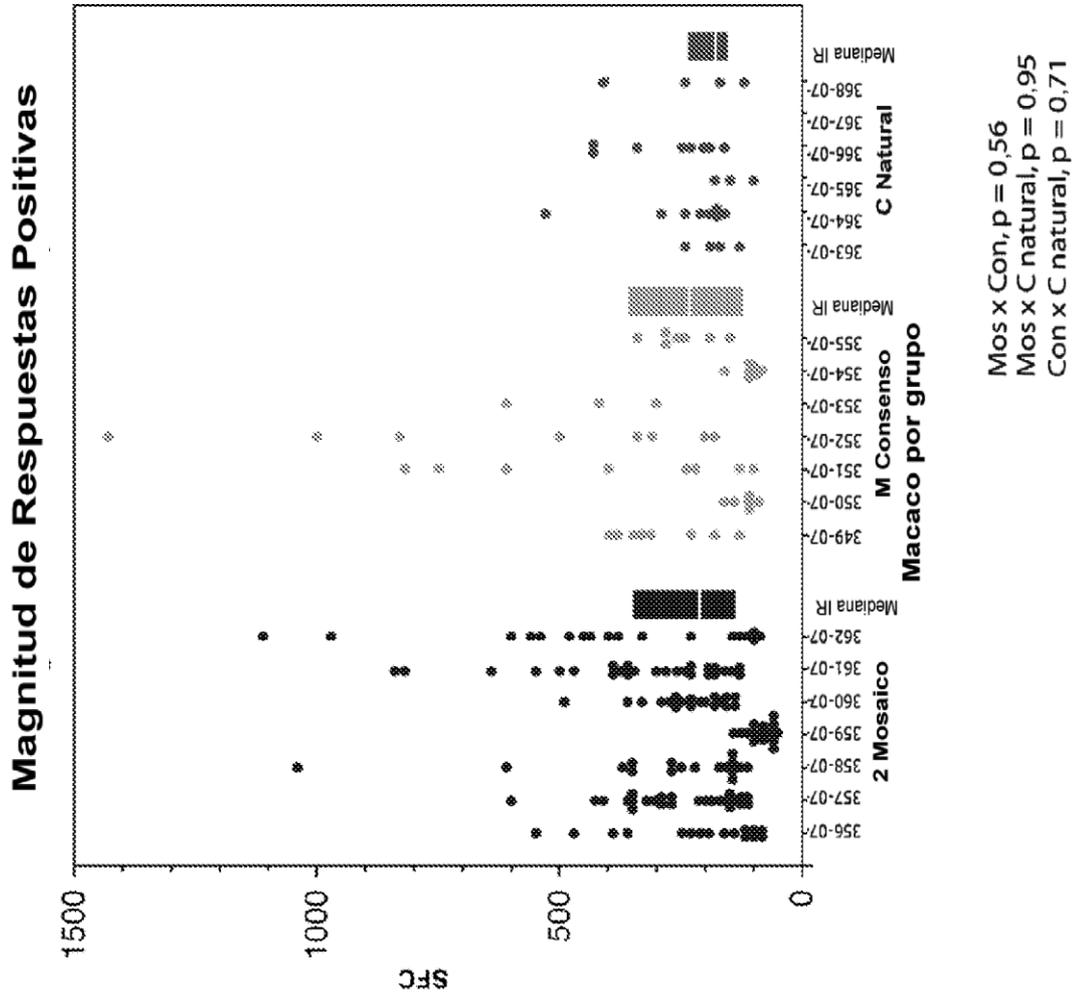


Figura 3

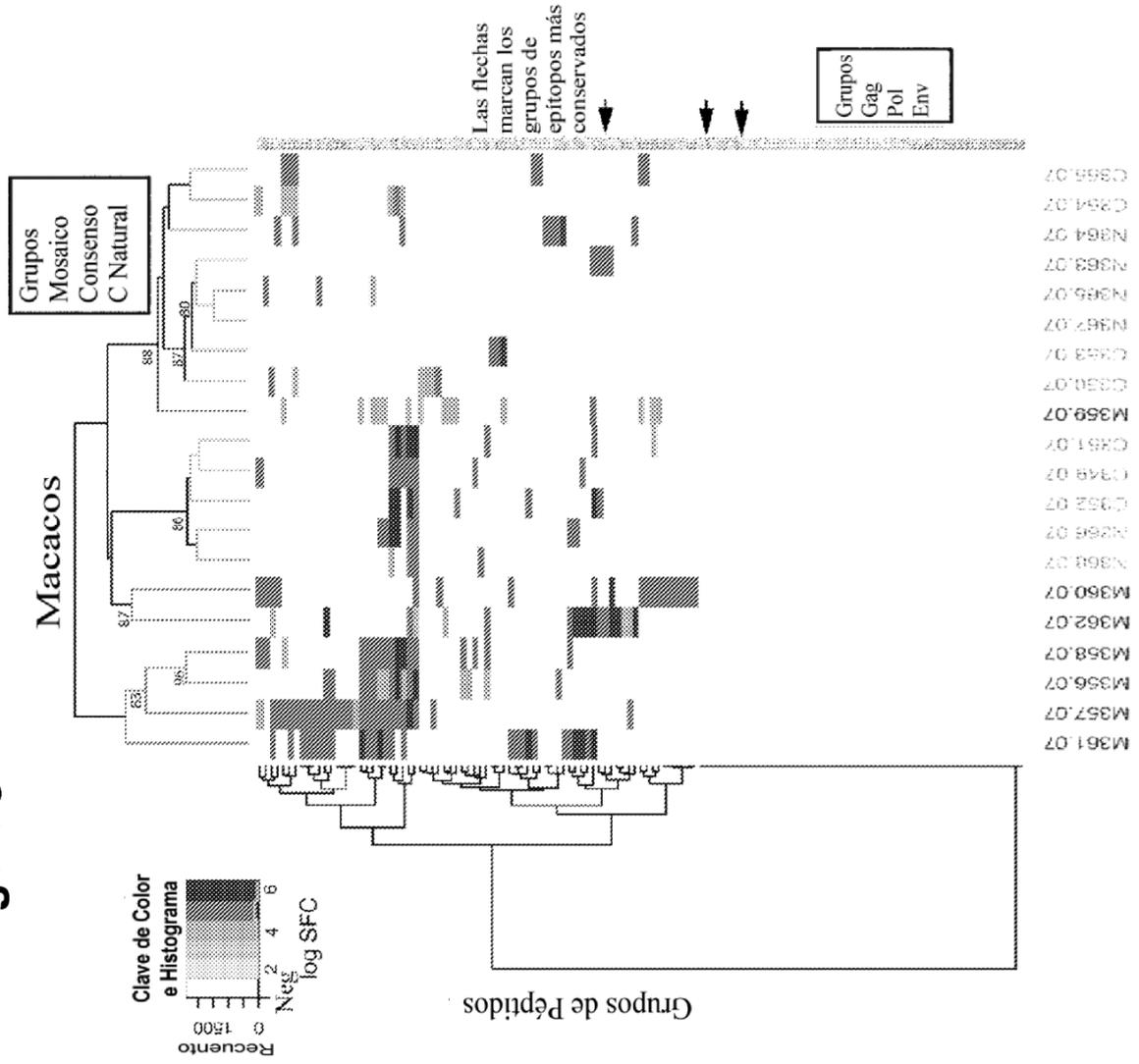


Figura 4
Cobertura noamérica de PTE y grupo M de diferentes
candidatos a vacuna

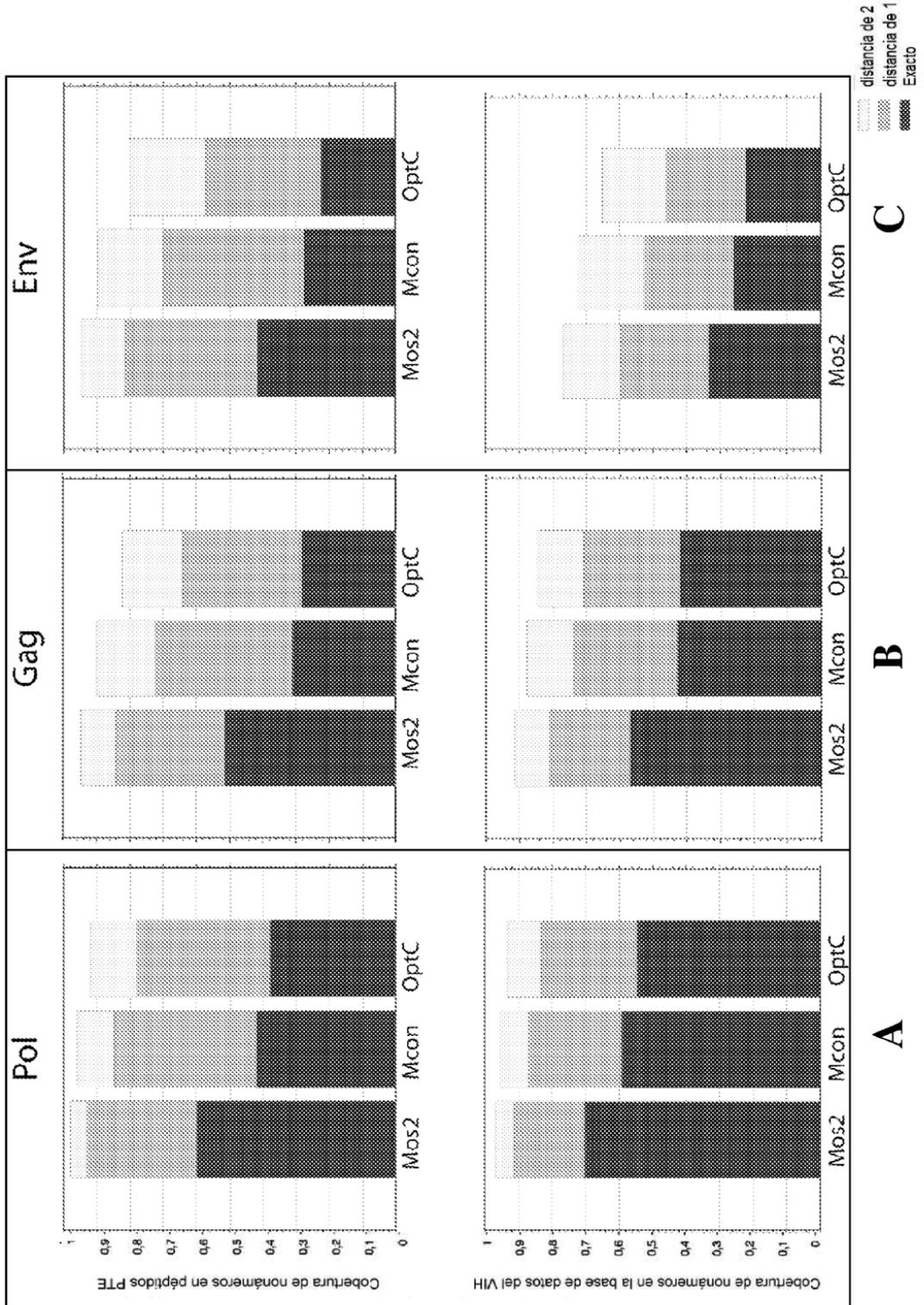


Figura 5

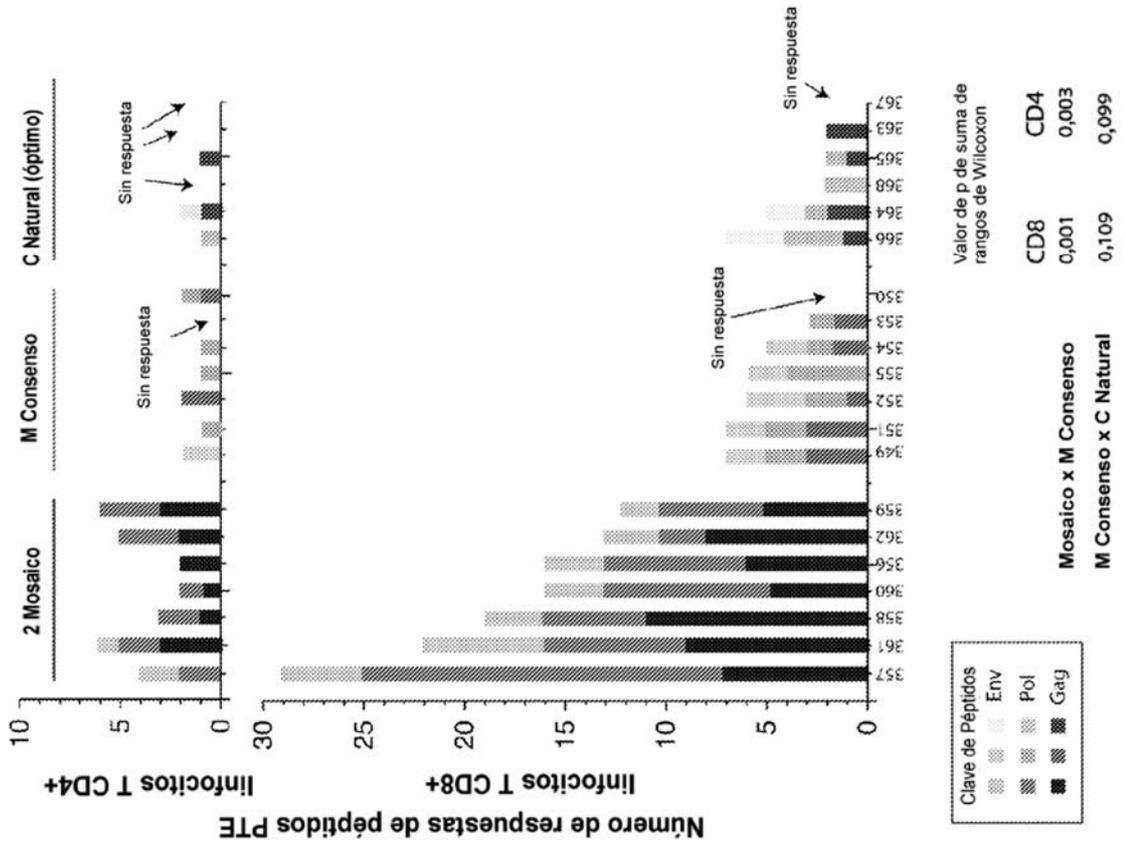


Figura 6

- Linfocitos T CD8: mediana (intervalo)
 - 2 Mosaico: 16 (12-29)
 - Mcon: 6 (0-7)
 - OptC: 3 (0-7)
- Linfocitos T CD4:
 - 2 Mosaico: 4 (2-6)
 - Mcon: 1 (0-2)
 - OptC: 0,5 (0-2)

Figura 7

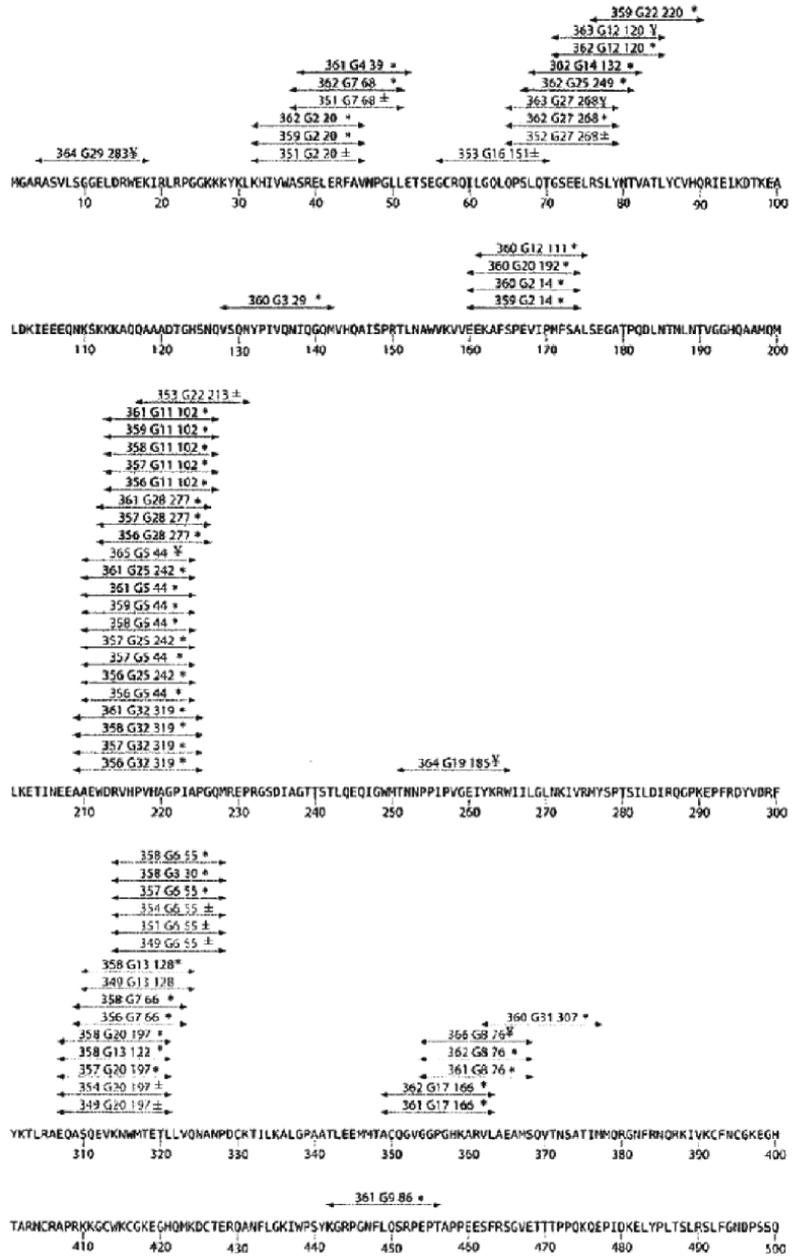


Figura 8

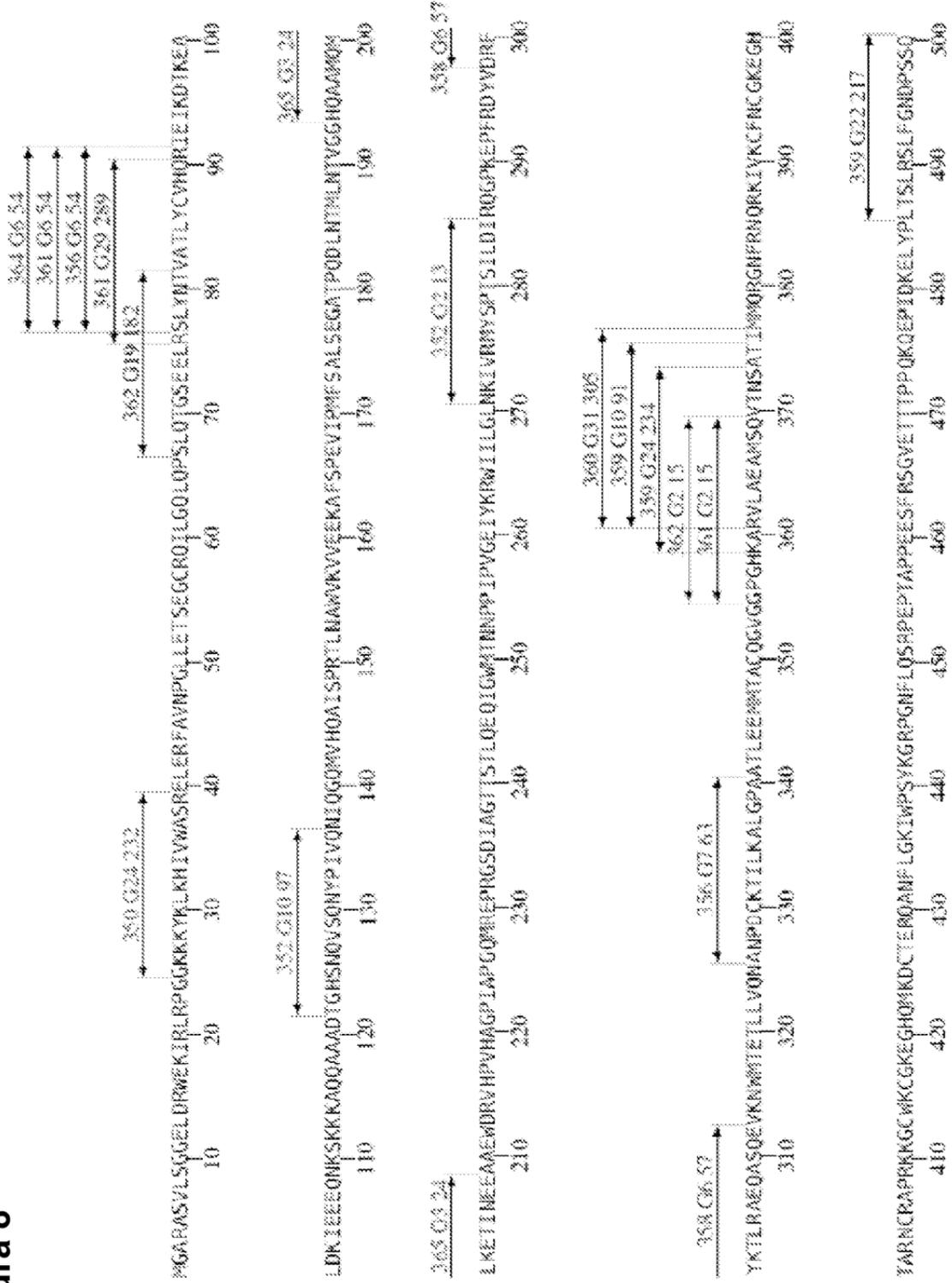


Figura 9

OptC 366-07:

5 respuestas de CD8:

OptC **IVQQQSNLIRAIETAAQQ**
 E54 VVQQQSNLIRAIETAAQ Env 548 562
 E72 -VQQQNNLIRAIETAAQH Env 549 563

OptC **AVFIHNEFKRKGIGGY**
 P22 AVFIHNEFKRKGIGG Pol 894 908
 P236 -VLIHNEFKRKGIGGY Pol 895 909

OptC **MAICEEMEKEGKITK**
 P224 TAICEEMEKEGKITK Pol 190 204

OptC **CTHGIKPVVSTQLLL**
 E15 CTHGIKPVVSTQLLL Env 247 261

OptC **GGPSHKARVIAEAMG**
 G76 GGPSHKARVIAEAMG Gag 354 368

1 respuesta de CD4:

OptC **IIGQVRDQAEHLKTA**
 P86 IIGQVRDQAEHLKTA Pol 876 890

Figura 10

Mos1	ICTTTPWNASW SNK SL	T...A
Mos2	ICTTAVPWNT SW SNKSQ	T...S
E334	ICTTTPWNASW SNR	A...A
E214	-CTTTPWNSW SNKT	A...T
E158	--TTAVPWNASW SNK SL	
E290	--TTAVPWNT SW SNK SL	

Figura 12

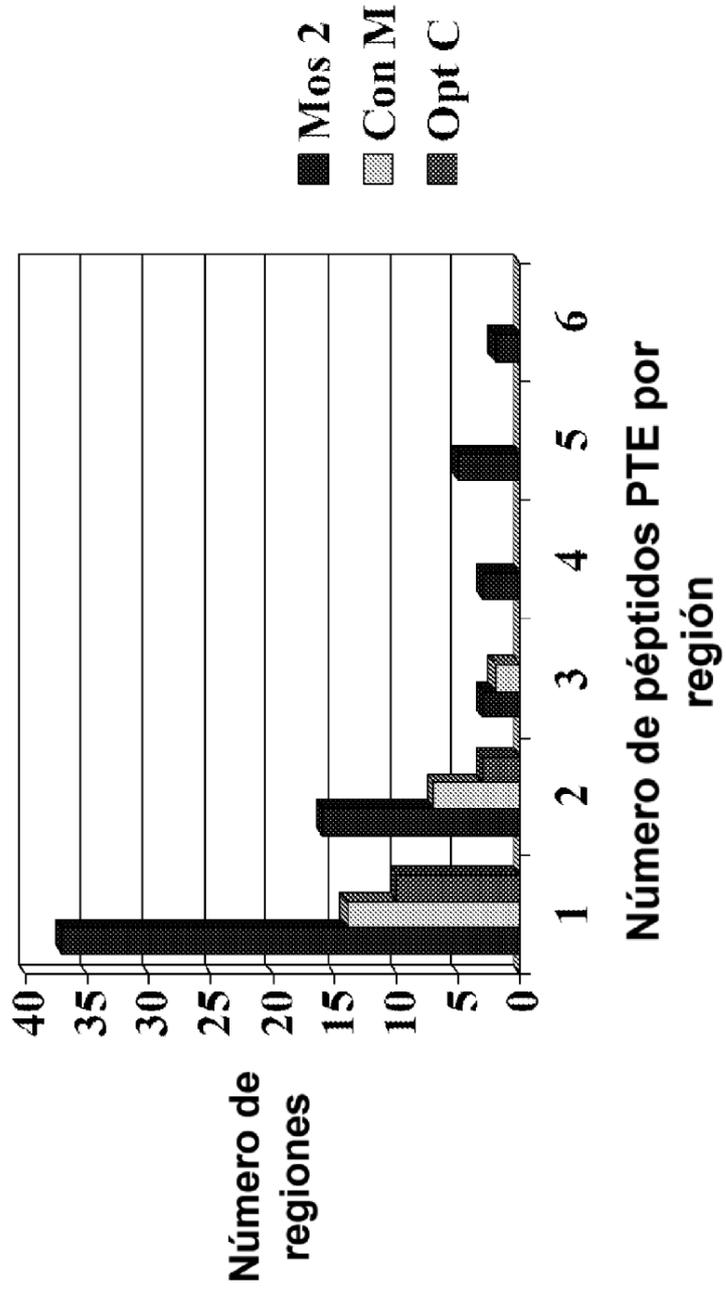


Figura 13

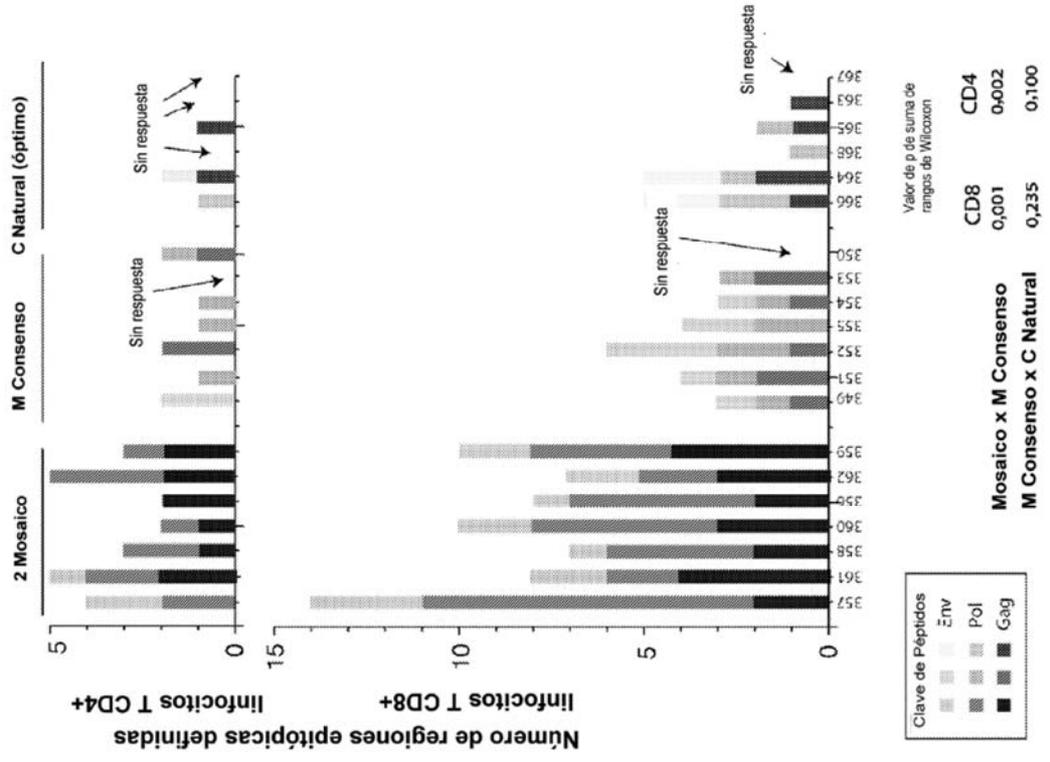


Figura 14

- Linfocitos T CD8: mediana (intervalo)
 - 2 Mosaico 8 (7-14)
 - Mcon: 3 (0-6)
 - OptC: 1,5 (0-5)
- Linfocitos T CD4
 - 2 Mosaico 3 (2-5)
 - Mcon: 1 (0-2)
 - OptC: 0,5 (0-2)

Figura 15

Respuestas de ensayo en grupos de vacunas contra Grupos de Péptidos Positivos basándose en 5 Gag diferentes

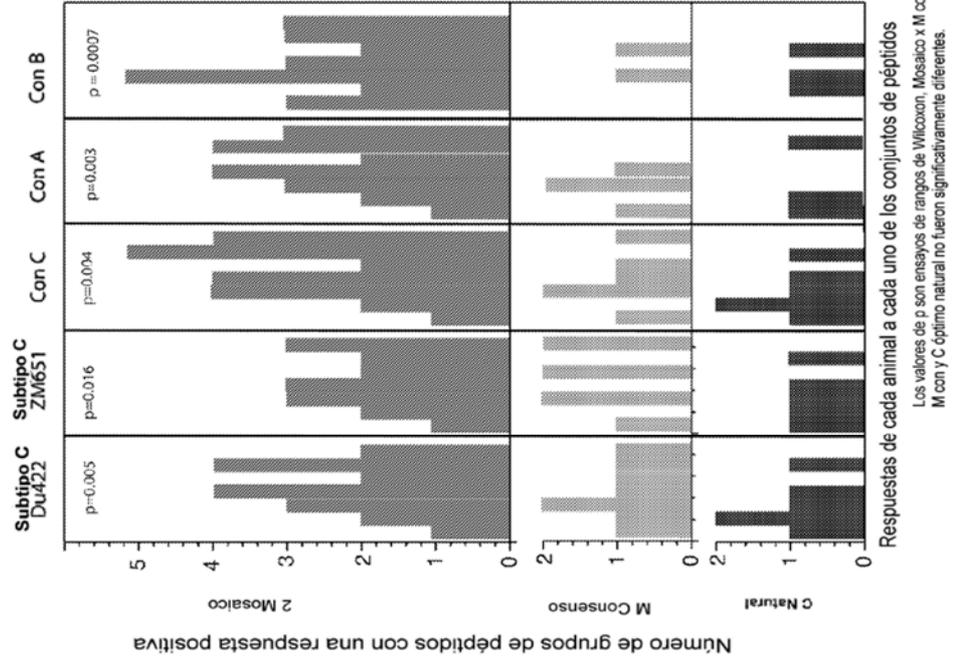
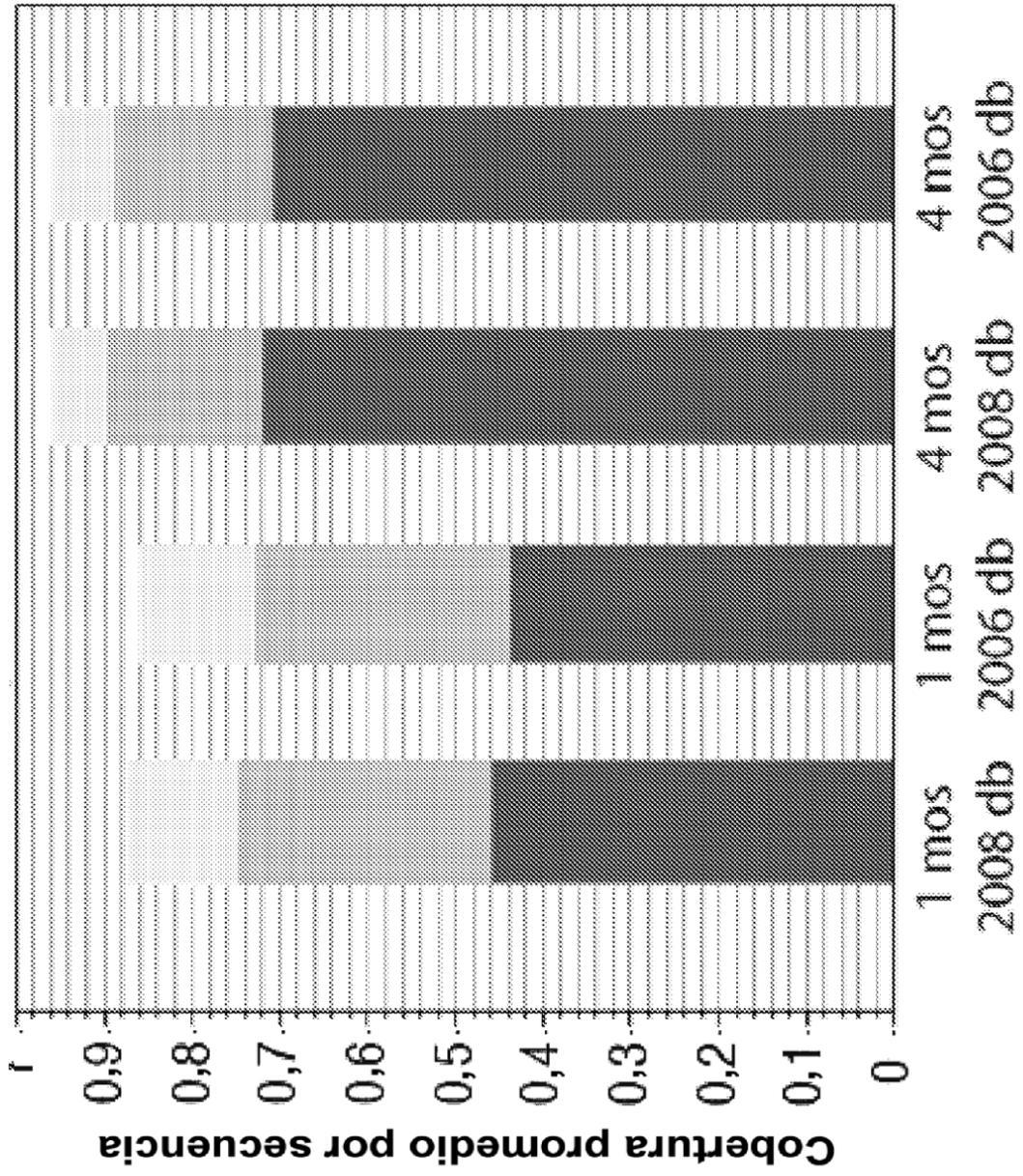


Figura 16



Número de Gag

2006 ~ 550
2008 ~ 1208

Figura 17

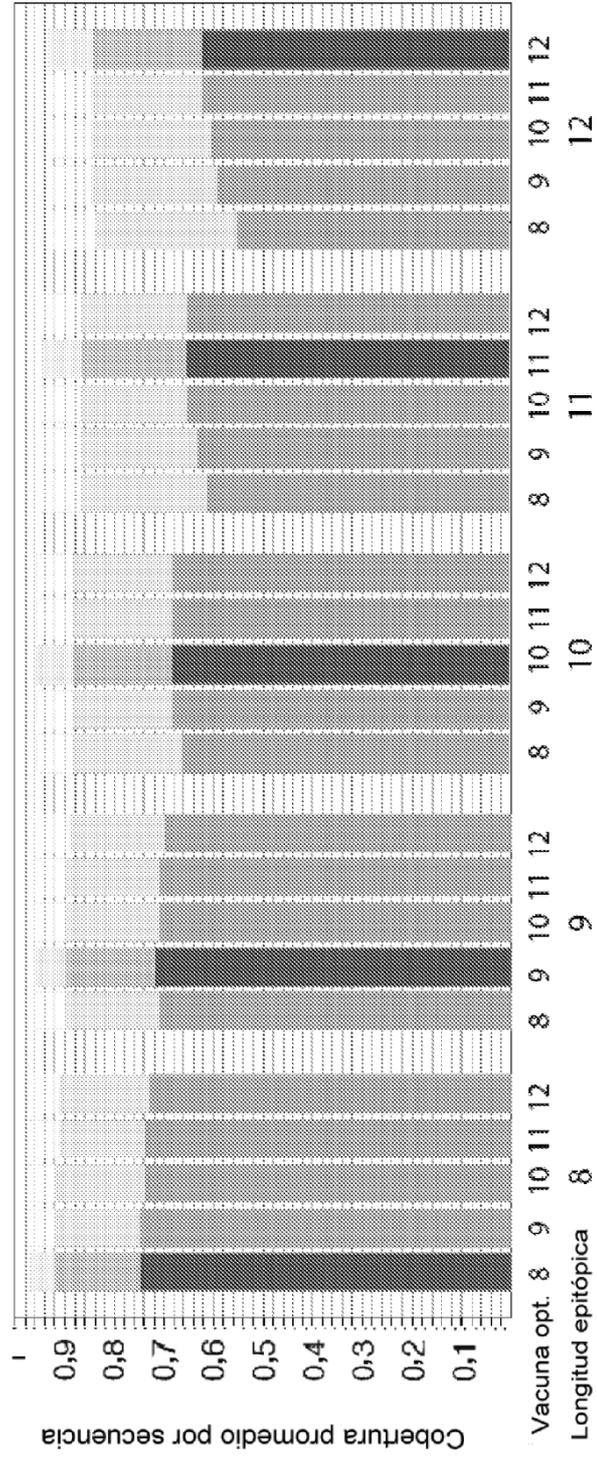
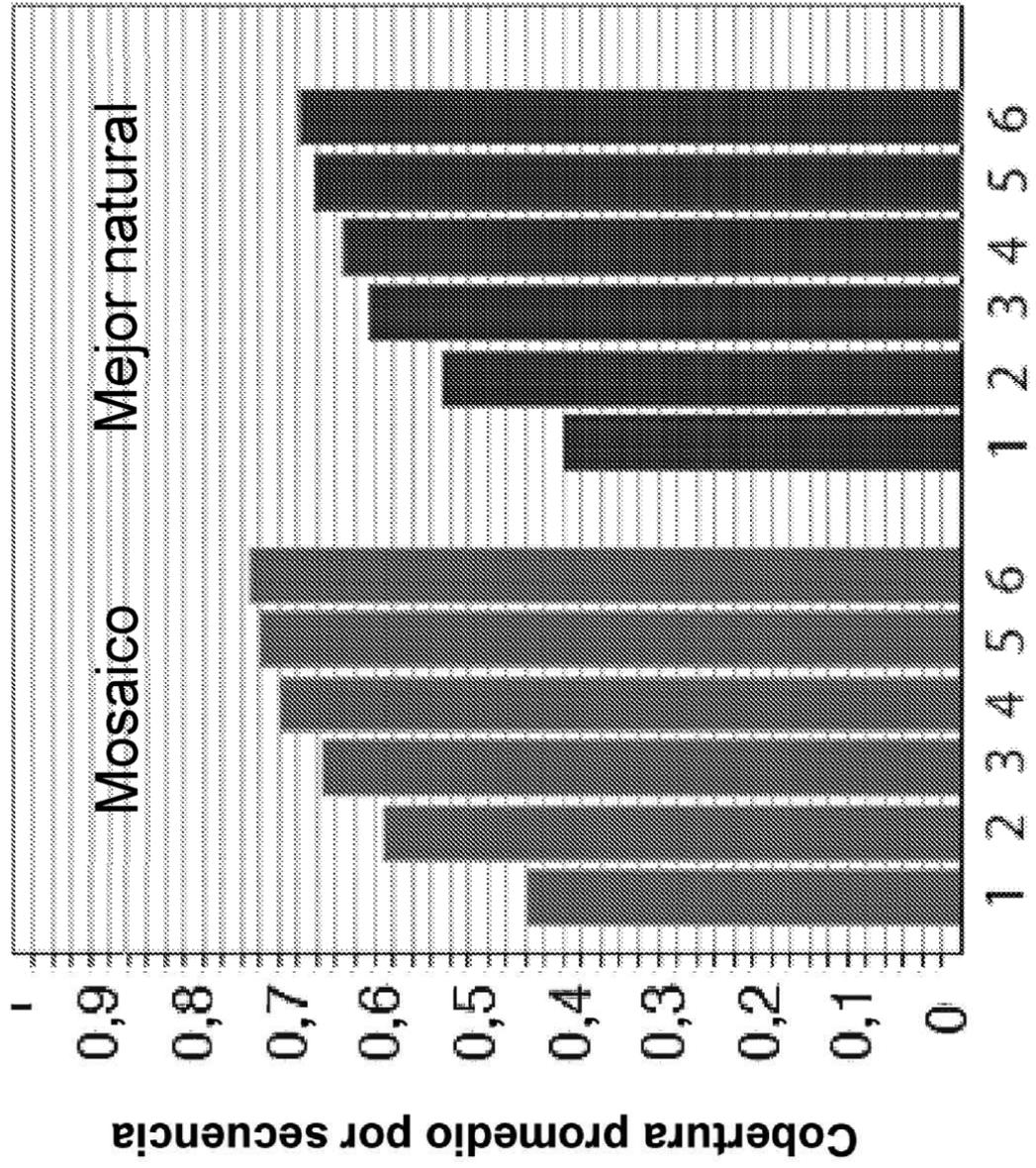


Figura 18



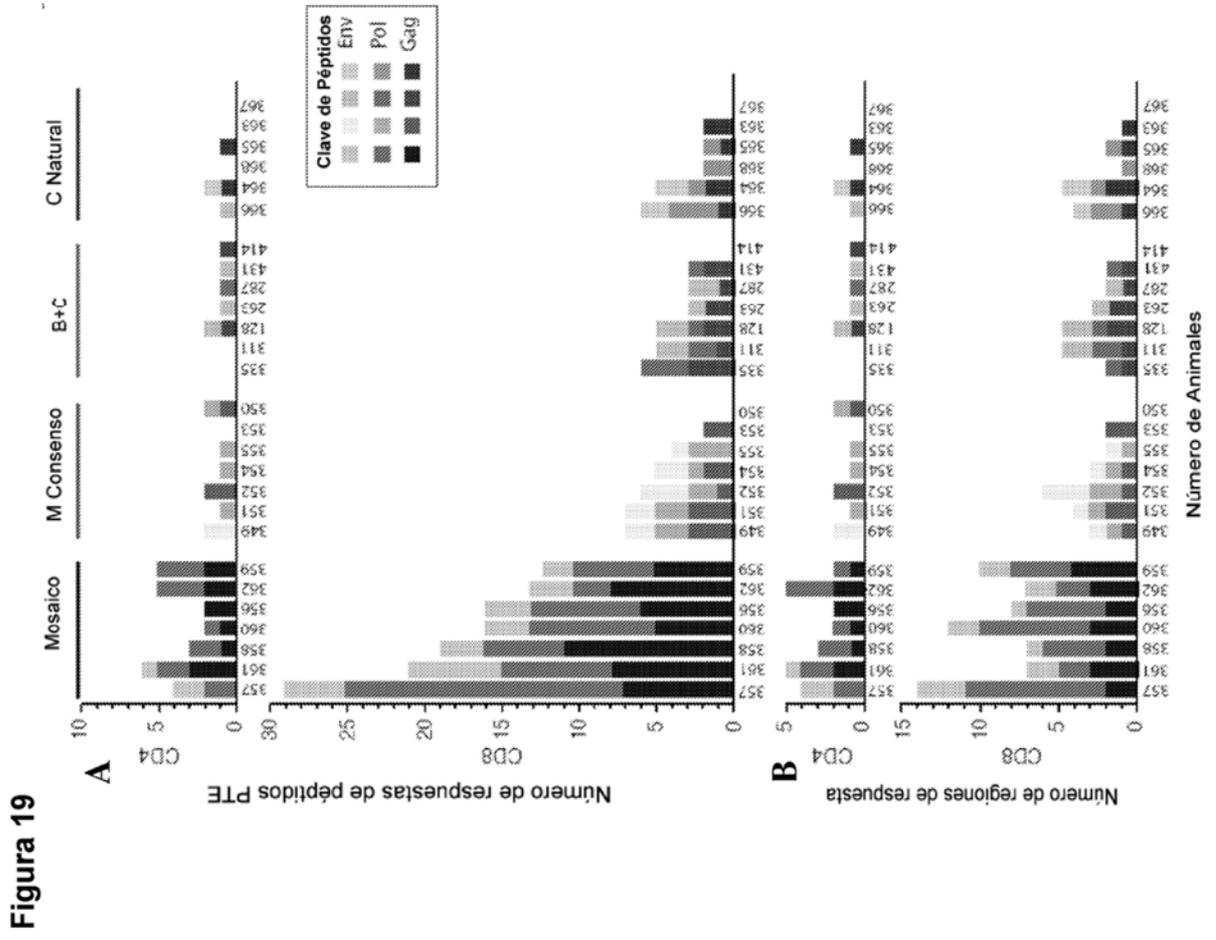


Figura 20A

CD8 Gag

C Natural
M Consenso
B+C
Mosaico

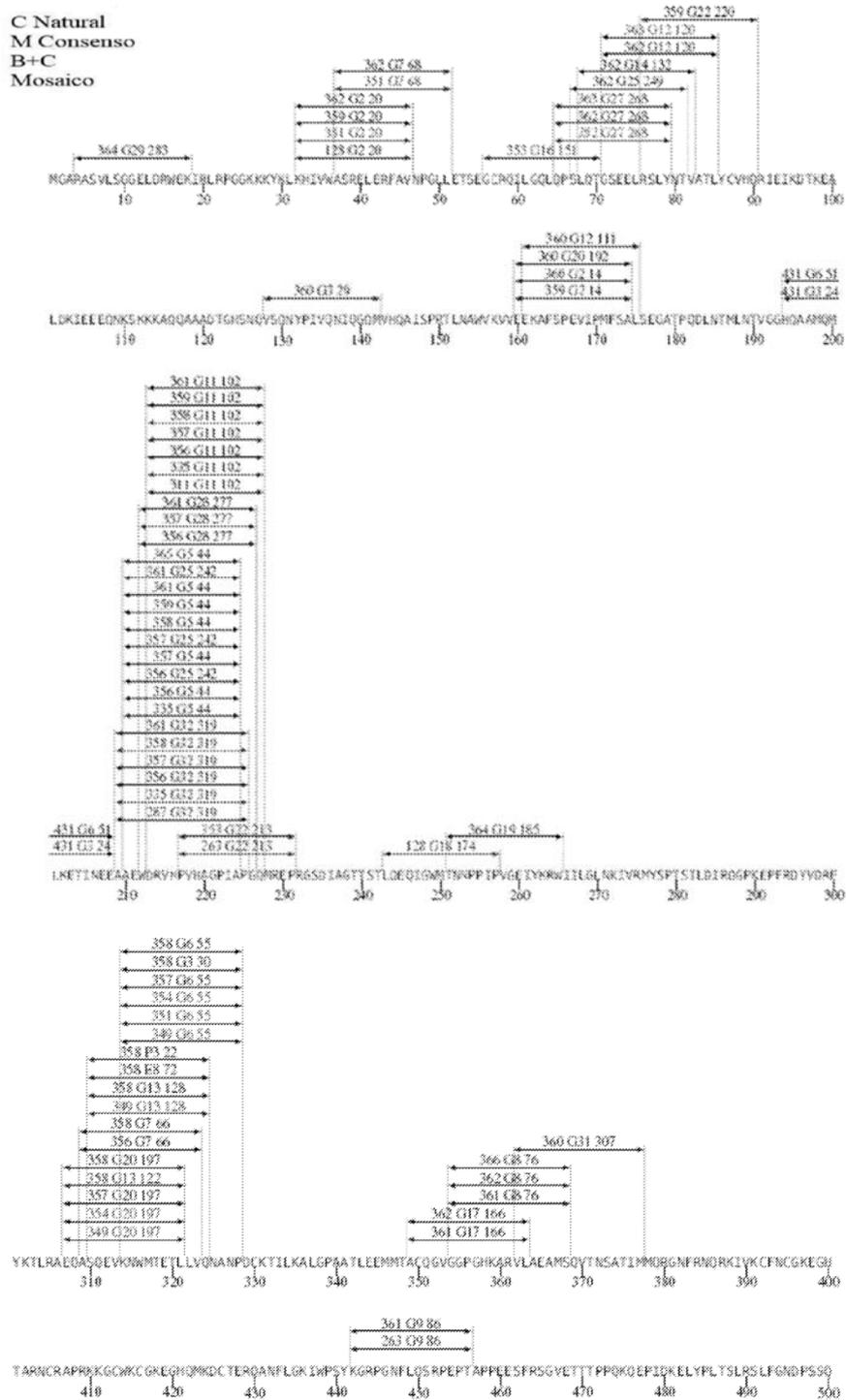


Figura 20B

CD8 Pol

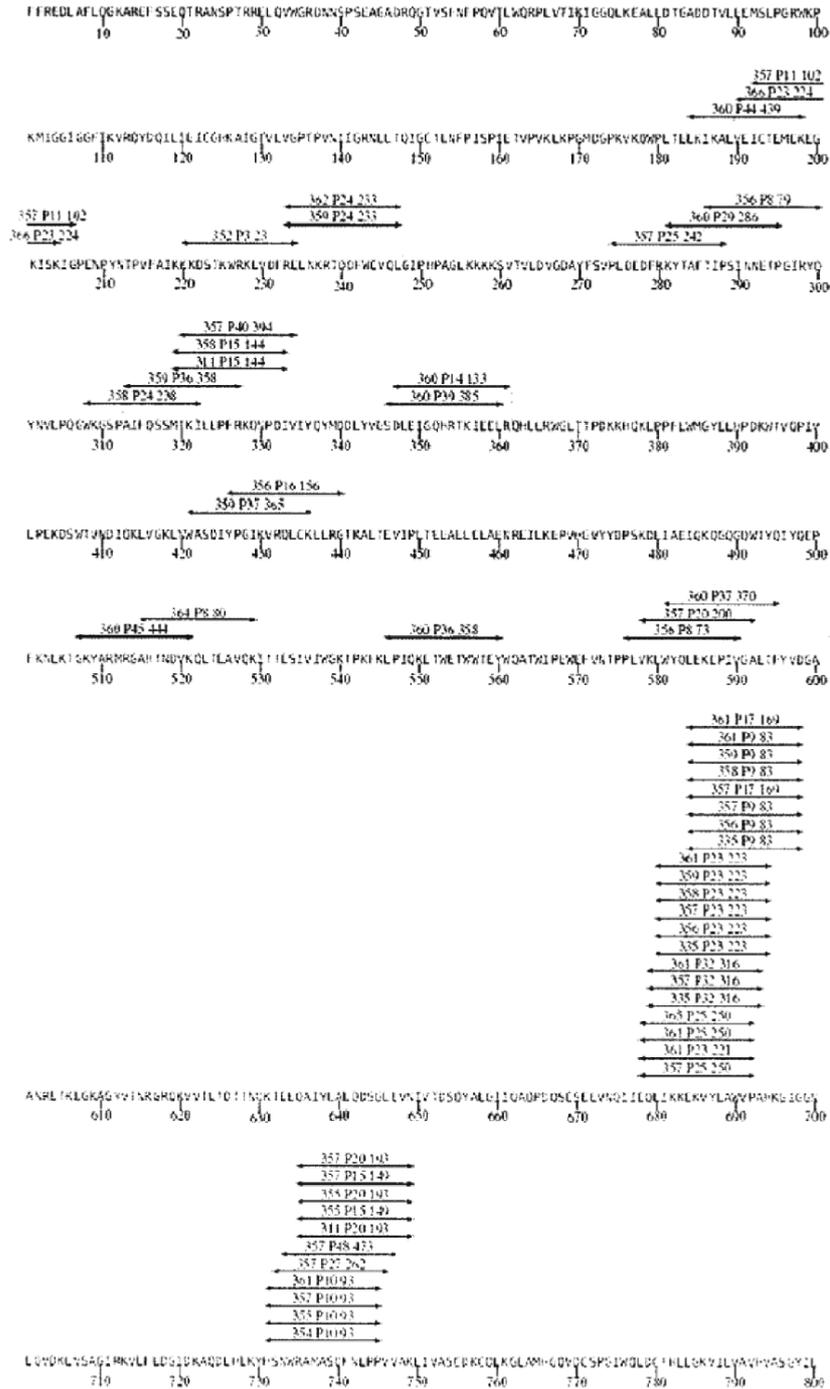


Figura 20B (continuación)

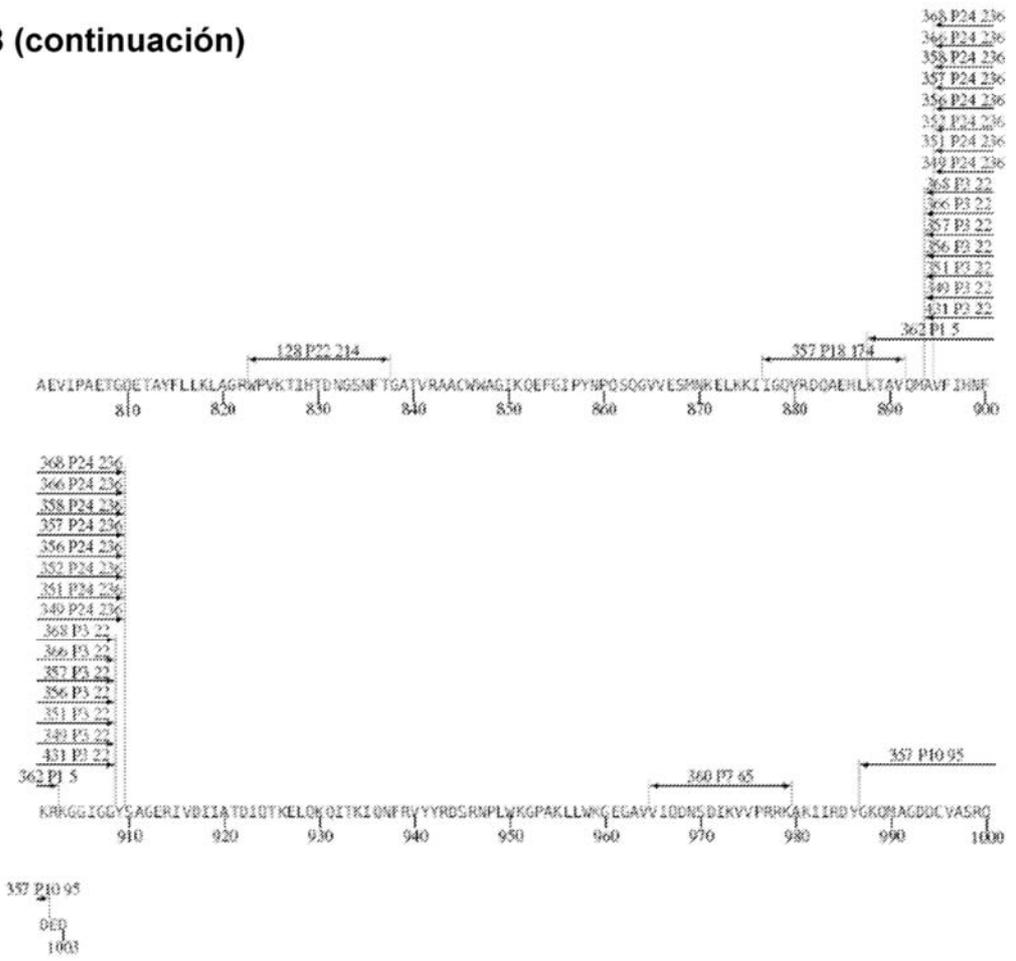


Figura 20C

CD8 Env

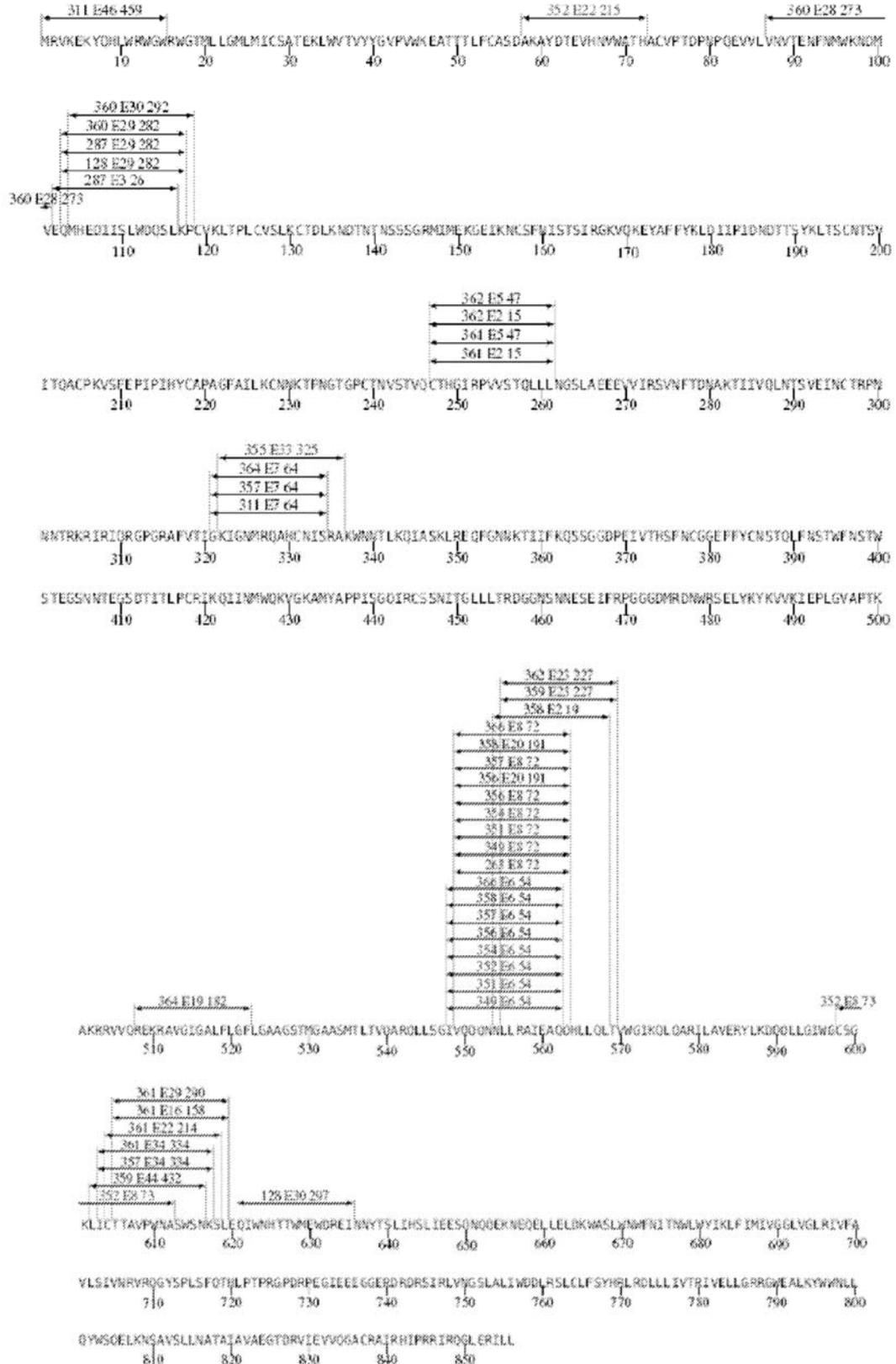


Figura 21A

CD4 Gag

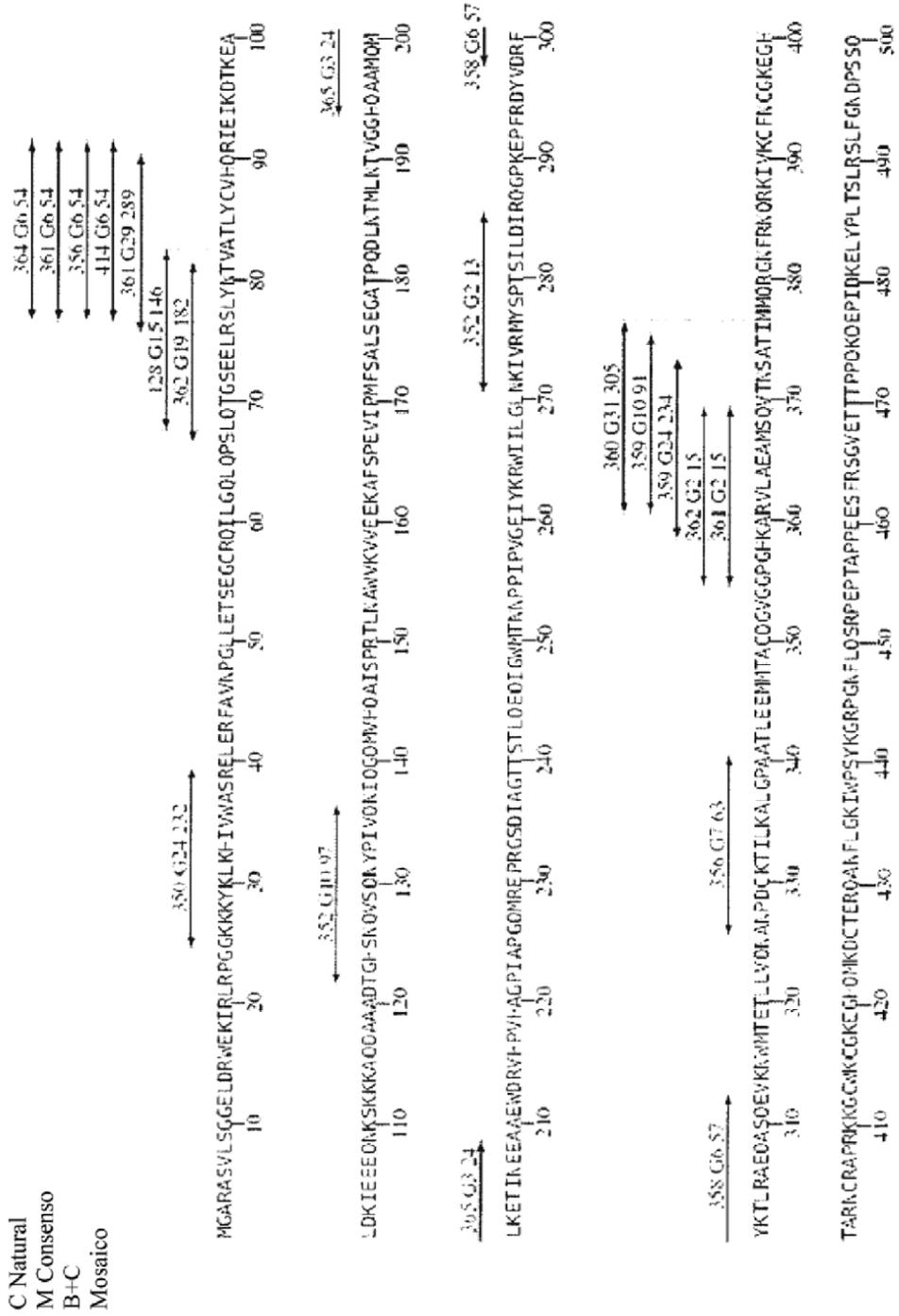


Figura 21B

CD4 Pol

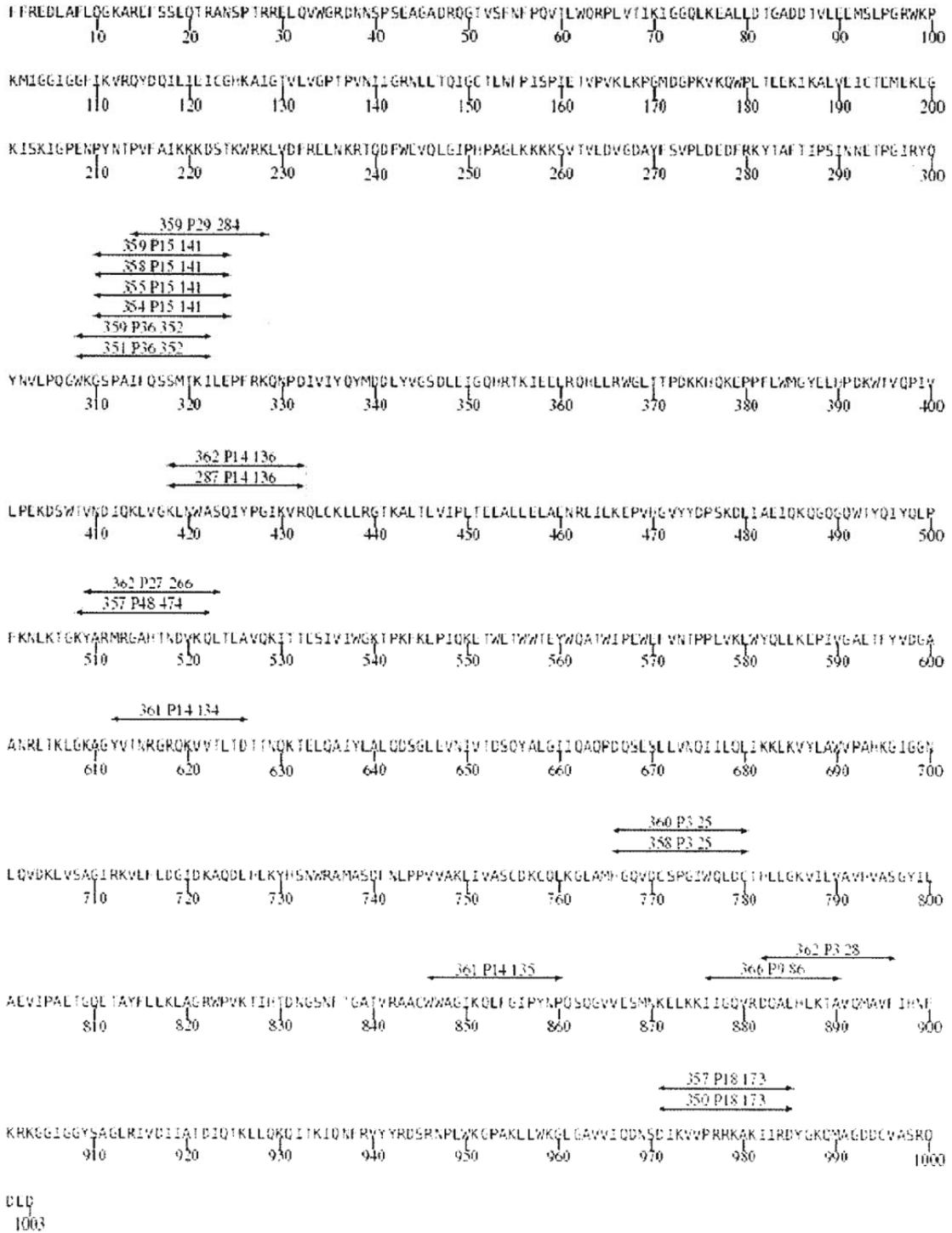


Figura 21C

CD4 Env

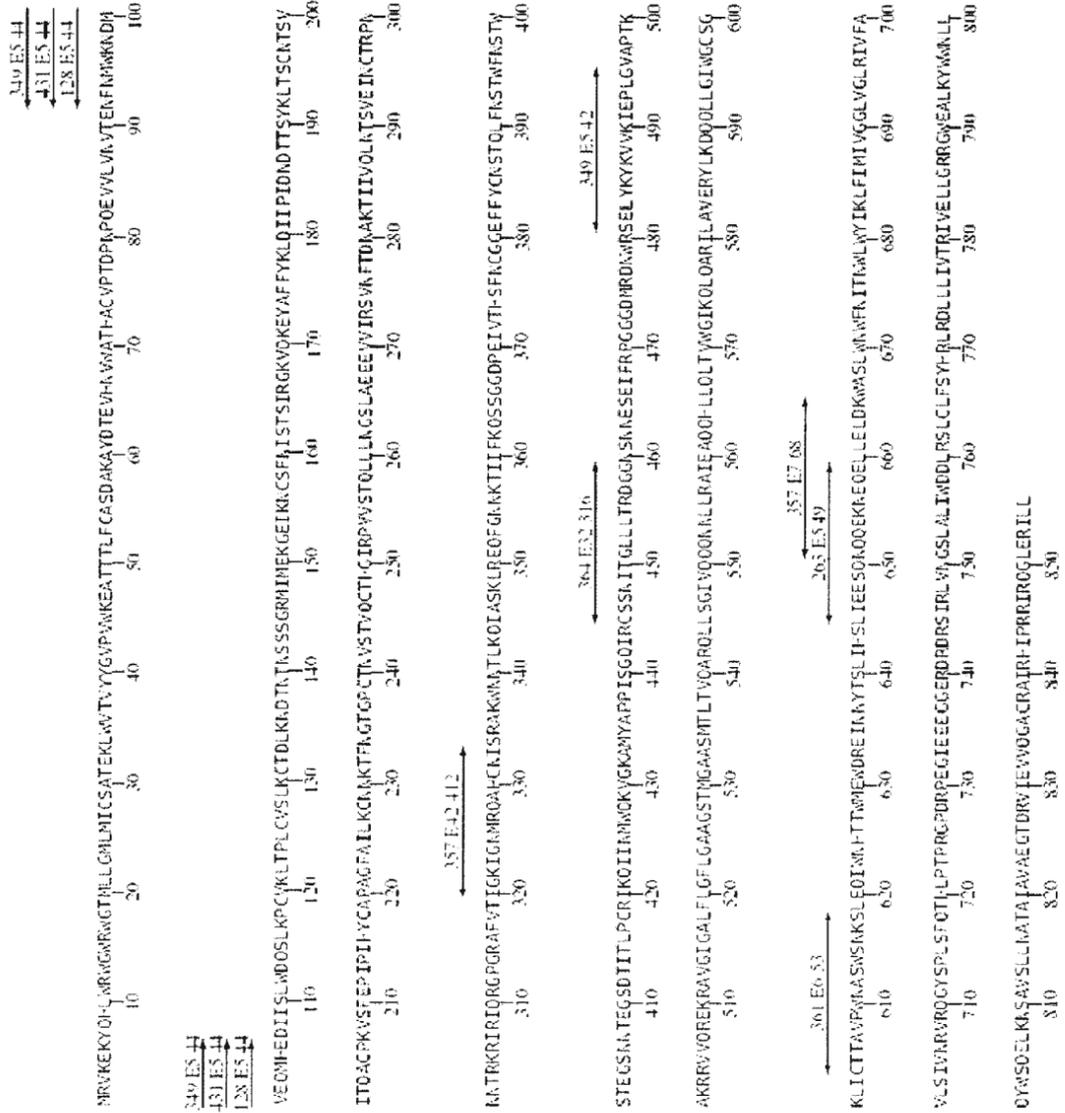


Figura 22**Péptidos reactivos y amplitud y profundidad de las respuestas: vacuna de Clado B+C****Clado B+C 287-95:**

3 CD8 PTE+, 1 CD4 PTE+
 2 regiones CD8+, 1 región CD4+
 Número de péptidos solapantes por región: CD8: 1 2 CD4: 1

3 respuestas de CD8:

C	AAEWDRHLHFPVHAGPIA			
B	AAEWDRHLHFPVHAGPIA			
G319	AAE=DRLHFPVHAGPIF Gag	209	225	
C	DQMHEDIISLWDQSLK			
B	EQMHEDIISLWDQSLK			
E26	EQMHEDIISLWDQSL Env	102	116	
E292	HMHEDIISLWDESLK Env	103	117	

1 respuesta de CD4:

C	KLNWASQIYSGIKVR			
B	KLNWASQIYFGIKVR			
P136	KLNWASQIYAGIKVK Pol	418	432	

Clado B+C 128-92:

5 CD8 PTE+, 2 CD4 PTE+
 5 regiones CD8+, 2 regiones CD4+
 Número de péptidos solapantes por región: CD8: 1 1 1 1 1 CD4: 1 1

5 respuestas de CD8:

C	KHLVWASRELERFAL			
B	KHIVWASRELERFAV			
G20	KHLVWASRELERFAL Gag	32	46	
C	LQEQIAWMTNPPFVP			
B	LQEQIGWMTNPPPIP			
G174	LAEQIAWMTSNPPIP Gag	243	257	
C	QMHEDIISLWDQSLK			
B	QMHEDIISLWDQSLK			
E292	HMHEDIISLWDESLK Env	103	117	
C	DIWDNMTWMQWDREI			
B	EIWDNMTWMEWEREI			
E297	EIWNMTWMEWEKEI Env	621	635	
C	WPKVIHTANGSNFT			
B	WPKTIHTANGSNFT			
E214	WPKVVHTDNGSNFT Pol	623	637	

2 respuestas de CD4:

C	AQTGTEELRSLYNTV			
B	LQTGSEELRSLYNTV			
G146	LQDGSEELKSLFNTV Gag	68	80	
C	YFNMWKNMVDQMHE			
B	NFNMWKNMVEQMHE			
E544	NFNMWKNMVEQMHE Env	92	106	

Figura 22 (continuación)**Clado B+C 263-00:**

3 CD8 PTE+, 1 CD4 PTE+

3 regiones CD8+, 1 región CD4+

Número de péptidos solapantes por región: CD8: 1 1 1 CD4: 1

3 respuestas de CD8:

C	GRPGNFLQSRPEPT			
B	GRPGNFLQSRPEPT			
G86	GRPGNFLQNRPEPT	Gag	442	456
C	PVHAGPIAPGQMREP			
B	PVHAGPIAPGQMREP			
G213	PVQAGPIAPGQMREP	Gag	217	231
C	VQQQSNLLRAIEAQQ			
B	VQQQNNLLRAIEAQQ			
E72	VQQQNNLLRAIEAQH	Env	549	563

1 respuesta de CD4:

C	LLEESQNQQEKNEKD			
B	LIEESQNQQEKNEQE			
E49	LIEESQNQQEKNEQD	Env	645	659

Clado B+C 311-00:

5 CD8 PTE+, 0 CD4 PTE+

5 regiones CD8+, 0 región CD4+

Número de péptidos solapantes por región: CD8: 1 1 1 1 1 CD4: 0

5 respuestas de CD8:

C	DRLHPVHAGPIAPGQ			
B	DRLHPVHAGPIAPGQ			
G102	DRVHPVHAGPIPPGQ	Gag	213	227
C	GDIIGDIRQAHCNIS			
B	GEIIGDIRQAHCNIS			
E64	GDIIGDIRQAHCNIS	Env	321	334
C	TGMLRNCQPWWIWGI			
B	KGIRKNYQHLWRWGT			
E459	KEIRKNYQHLWRWGT	Env	4	18
C	MTKILEPFRKNPEI			
B	MTKILEPFRKQNPDI			
F144	MTKILEPFRKQNPDI	Pol	319	333
C	RAMASEFNLPVVAK			
B	RAMASDFNLPVVAK			
F193	KAMASDFNLPPIVAK	Pol	735	749

Figura 22 (continuación)

Clado B+C 335-96:

6 CD8 PTE+, 0 CD4 PTE+

2 regiones CD8+, 0 región CD4+

Número de péptidos solapantes por región: CD8: 3 3 CD4: 0

6 respuestas de CD8:

C	AAEWDRLHPVHAGPIAPGQ			
B	AAEWDRLHPVHAGPIAPGQ			
G319	AAE-DRLHPVHAGPIP	Gag	209	225
G44	AEWDRLHPVHAGPIA	Gag	210	224
G102	DRVHPVHAGPIPPGQ	Gag	213	227
C	QLINKERVYLSWVPAHKGIG			
B	QLIKKEKVYLAWVPAHKGIG			
P316	KLIEKDKVYLSWVPA	Pol	679	693
P223	LIIKERVYLSWVPAH	Pol	680	694
P83	EKVYLSWVPAHKGIG	Pol	684	698

Clado B+C 414-95:

0 CD8 PTE+, 1 CD4 PTE+

0 regiones CD8+, 1 región CD4+

Número de péptidos solapantes por región: CD8: 0 CD4: 1

1 respuesta de CD4:

C	SLYNTVATLYCVHAG			
B	SLYNTVATLYCVHQK			
G54	SLYNTVATLYCVHQR	Gag	77	91

Clado B+C 431-01:

3 CD8 PTE+, 1 CD4 PTE+

2 regiones CD8+, 1 región CD4+

Número de péptidos solapantes por región: CD8: 0 1 CD4: 1

3 respuestas de CD8:

C	HQAAMQMLKDTINEE			
B	HQAAMQMLKETINEE			
G24	HQAAMQMLKDTINEE	Gag	194	208
G51	HQAAMQMLKETINEE	Gag	194	208
C	AVFIHNFKRKGGIGG			
B	AVFIHNFKRKGGIGG			
P22	AVFIHNFKRKGGIGG	Pol	894	908

1 respuesta de CD4:

C	YFNMWKNDMVDQMHE			
B	NFNMWKNNMVEQMHE			
E44	NFNMWKNNMVEQMHE	Env	92	106

Figura 22 (continuación)

Péptidos reactivos y amplitud y profundidad de las respuestas: vacuna de M Consens

M consenso 349-07:

7 CD8 PTE-, 2 CD4 PTE+

3 regiones CD8+, 2 regiones CD4+

Número de péptidos solapantes por región: CD8: 3 2 2 CD4: 1 1

3 respuestas de CD8:

Mcon:	EQATQDVKNWMTDTLLVQNANP			
G197	EQSTQEVKNWMTDTL	Gag	307	321
G128	---TQDVKNWMTDTLLIQ	Gag	310	324
G55	-----KNWMTDTLLVQNANP	Gag	314	328

Mcon:	IVQQQSNNLLRAIEAQQ			
E54	IVQQQSNNLLRAIEAQ	Env	546	562
E72	-VQQQNLLRAIEAQH	Env	549	563

Mcon:	AVFIHNFKRKGGIGGY			
P22	AVFIHNFKRKGGIGG	Pol	894	908
P236	-VLIHNFKRKGGIGGY	Pol	895	909

2 respuestas de CD4:

Mcon:	SELYKYKVVKIEPLG			
E40	SELYKYKVVKIEPLG	Env	491	495

Mcon:	NFNMWKNNMVEQMHE			
E44	NFNMWKNNMVEQMHE	Env	92	106

M consenso 350-07:

2 CD8 PTE-, 2 CD4 PTE+

0 regiones CD8+, 2 regiones CD4+

Número de péptidos solapantes por región: CD8: ninguno CD4: 1 1

2 respuestas de CD4:

Mcon:	GRKKYRLKHLVWASR			
G232	GRKKYRLKHIVWASR	Gag	25	39

Mcon:	DIKVVPRRKAKIIRD			
E173	DIKVVPRRKVKIIRD	Pol	971	985

M consenso 351-07:

7 CD8 PTE-, 1 CD4 PTE+

4 regiones CD8+, 1 región CD4+

Número de péptidos solapantes por región: CD8: 0 0 0 1 CD4: 1

Figura 22 (continuación)**4 respuestas de CD8:**

Mcon	KHLVWASRELERFALNPGLL			
G20	KHLVWASRELERFAL	Gag	32	46
G68	-----ASRELERFAVNPGLL	Gag	37	5
Mcon	IVQQQSNLLRAIEAQQ			
E6	VVQQQSNLLRAIEAQ	Env	548	562
E8	-VQQQNNLLRAIEAQH	Env	549	563
Mcon	AVFIHNFKRKGGIGGY			
P22	AVFIHNFKRKGGIGG	Pol	894	908
P236	-VLIHNFKRKGGIGGY	Pol	895	909
Mcon	KNWMTDTLLVQNANP			
G55	KNWMTDTLLVQNANP	Gag	314	328

1 respuesta de CD4:

Mcon	WKGSPAIFQSSMTKI			
P352	WKGSPAIFQASMTKI	Pol	308	322

M consenso 352-07:

6 CD8 PTE+, 2 CD4 PTE+

6 regiones CD8+, 2 regiones CD4+

Número de péptidos solapantes por región: CD8: 1 1 1 1 1 1 CD4: 1

6 respuestas de CD8:

Mcon	QPALQTGSEELRSLY			
G268	LPALKTGSEELRSLY	Gag	65	79
Mcon	IVQQQSNLLRAIEAQ			
E54	VVQQQSNLLRAIEAQ	Env	548	562
Mcon	CSGKLICTTTVPWNS			
E73	CSGKLICTTAVPWNS	Env	598	612
Mcon	AKAYDTEVHNVWATH			
E215	AKAYEKEVHNVWATH	Env	58	72
Mcon	KKDSTKWRKLVDFRE			
P23	KKDSTKWRKLVDFRE	Pol	220	234
Mcon	VFIHNFKRKGGIGGY			
P236	VLIHNFKRKGGIGGY	Pol	895	909

2 respuestas de CD4:

Mcon	NKIVRMYSFVSILDI			
G13	NKIVRMYSFVSILDI	Gag	271	285
Mcon	KGNSSKVVSQNYPIVQ			
G97	TGNSSQVVSQNYPIVQ	Gag	122	136

Figura 22 (continuación)**M consenso 353-07:**

2 CD8 PTE+, 0 CD4 PTE+

2 regiones CD8+, 0 regiones CD4+

Número de péptidos solapantes por región: CD8: 1 1 CD4: ninguno

2 respuestas de CD8:

Mcon GCKQIIGQLQFALQT

G151 GCRQILGQLQPSLQT Gag 56 70

Mcon FVHAGPIPPGQMREP

G213 PVQAGPIAPGQMREP Gag 217 231

M consenso 354-07:

5 CD8 PTE+, 1 CD4 PTE-

3 regiones CD8+, 1 región CD4+

Número de péptidos solapantes por región: CD8: 2 2 1 CD4: 1

3 respuestas de CD8:

Mcon EQATQDVKNWMTDTLLVQNANP

G197 EQSTQEVKNWMTDTL Gag 307 321

G55 -----KNWMTDTLLVQNANP Gag 314 328

Mcon IVQQQSNLLRAIEAQQ

E54 VVQQQSNLLRAIEAQ Env 548 562

E72 -VQQQNNLLRAIEAQH Env 549 563

Mcon HSNWRAMASDFNLPP

P93 HSNWRAMASDFNLPP Pol 731 745

1 respuesta de CD4:

Mcon GSPAIFQSSMTKILE

P141 GSPAIFQSSMTKILD Pol 310 324

M consenso 355-07:

3 CD8 PTE+, 1 CD4 PTE-

2 regiones CD8+, 1 región CD4+

Número de péptidos solapantes por región: CD8: 1 2 CD4: 1

2 respuestas de CD8:

Mcon IIGDIRQAHCNISGT

E325 ITGDIRQAHCNVRS Env 322 336

Mcon SNWRAMASDFNLPPIVAK

P93 SNWRAMASDFNLPP Pol 731 745

P149 ---RTMASDFNLPPVAK Pol 735 749

P193 ---KAMASDFNLPPIVAK Pol 735 749

1 respuesta de CD4:

Mcon GSPAIFQSSMTKILE

P141 GSPAIFQSSMTKILD Pol 310 324

Figura 22 (continuación)

Péptidos reactivos y amplitud y profundidad de las respuestas: vacuna de Mosaico

Mosaico 356-07:

15 CD8 PTE+, 2 CD4 PTE+

8 regiones CD8+, 2 regiones CD4+

Número de péptidos solapantes por región: CD8: 5 3 1 2 1 1 1 2 CD4: 1 1

8 respuestas de CD8:

Mos1	AAEWDRVHPVHAGPIAPGQ			
Mos2	AAEWDRHLHPVHAGPVAFGQ			
G319	AAE-DRLHPVHAGPIP	Gag	309	323
G44	-AEWDRHLHPVHAGPIA	Gag	210	224
G242	-ADWDRHLHPVHAGPVA	Gag	210	224
G277	---WDRVHPVHAGENPEG	Gag	212	226
G102	----DRVHPVHAGPIPPGQ	Gag	213	227

Mos1	IVQQQNLLRAIEAQQ			
Mos2	IVQQQSLLRAIEAQQ			
E54	VVQQQSLLRAIEAQQ	Env	543	562
E72	-VQQQNLLRAIEAQQ	Env	549	563
E191	-VQQQSLLRAIEAQQ	Env	549	563

Mos1	ASQDVKNWMTETLLV			
Mos2	ATQDVKNWMTDTLLV			
G66	ASQEVKNWMTETLLI	Gag	309	323

Mos1	AVFIHNFKRKGGIGGY			
Mos2	AVFIHNFKRKGGIGGY			
P22	AVFIHNFKRKGGIGG	Pol	894	908
P236	-VLIHNFKRKGGIGGY	Pol	895	909

Mos1	PLVKLWYQLEKDFIA			
Mos2	PLVKLWYQLEKEPIV			
P73	PLVKLWYQLEKEPIV	Pol	576	590

Mos1	TIPSTNNETPGIRYQ			
Mos2	TIPSINNETPGIRYQ			
P79	TIPSINNETPGIRYQ	Pol	266	300

Mos1	YFGIKVRQLCKLLRG			
Mos2	YAGIKVKQLCKLLRG			
P156	YFGIKVRQLCKLLRG	Pol	426	440

Mos1	LIKKERVYLSWVPAHKGIG			
Mos2	LIKKKVYLAWVPAHKGIG			
P223	LIKKERVYLSWVPAH	Pol	686	694
P13	----EKVYLSWVPAHKGIG	Pol	684	698

2 respuestas de CD4:

Mos1	SLYNTVATLYCVHQE			
Mos2	SLFNTVATLYCVHAE			
G54	SLYNTVATLYCVHQE	Gag	77	91

Mos1	ANPDCKTILKALGPA			
Mos2	ANPDCKTILRALGPG			
G63	ANPDCKTILRALGPG	Gag	326	340

Figura 22 (continuación)

Mosaico 357-07:

29 CD8 PTE+, 4 CD4 PTE+

14 regiones CD8+, 4 regiones CD4+

Número de péptidos solapantes por región: CD8: 5 2 2 5 5 2 1 1 1 1 1 1 1 1 1 CD4: 1 1 1 1

14 respuestas de CD8:

Mos1	EQLIKKERVYLSWVPAHKGIG			
Mos2	EQLIKKEKVYLSWVPAHKGIG			
P050	EPLIKKERVYLSWV	Pol	678	662
P316	-KLIKDKVYLSWVFA	Pol	679	693
P023	--LIKKEVYLSWVFAH	Pol	680	694
P169	-----EKVYLSWVPAHKGIG	Pol	684	698
P83	-----EKVYLSWVPAHKGIG	Pol	684	698
Mos1	IVQQQNNLLRAIEAQQ			
Mos2	IVQQQSNNLLRAIEAQQ			
E54	VVQQQSNNLLRAIEAQQ	Env	548	562
E70	-VQQQNNLLRAIEAQQH	Env	549	563
Mos1	EQASQDVKNWMTDTLLVQNaNP			
Mos2	EQATQDVKNWMTDTLLVQNaNP			
G197	EQSTQDVKNWMTDTL	Gag	307	321
G55	-----KNWMTDTLLVQNaNP	Gag	314	328
Mos1	HSNWRAMASDFNLPPVAK			
Mos2	HSNWRAMASEFNLPPVAK			
P93	HSNWRAMASDFNLPP	Pol	731	745
P060	-SNWRAMASEFNLPP	Pol	732	746
P473	--NWRAMASDFNLPPVI	Pol	733	747
P149	----RTMASDFNLPPVAK	Pol	735	749
P193	----KAMASDFNLPPVAK	Pol	735	749
Mos1	AEWDRVHPVHAGPIAPGQ			
Mos2	AEWDRLHPVHAGPVAPGQ			
G319	AAEDRLHPVHAGPIP	Gag	209	225
G242	ADWDRLHPVHAGPVA	Gag	210	224
G44	AEWDRLHPVHAGPIA	Gag	210	224
G277	--WDRVHPVHAGPNPPG	Gag	212	226
G100	---DRVHPVHAGPIPPGQ	Gag	213	227
Mos1	AVFIHNFRRKGGIGGY			
Mos2	AVFIHNFRRKGGIGEY			
P00	AVFIHNFRRKGGIGG	Pol	894	908
P036	-VLIHNFRRKGGIGGY	Pol	895	909
Mos1	PLDEGFRKYTAFTIP			
Mos2	PLDEDFRKYTAFTIP			
P242	PLDESFRKYTAFTIP	Pol	274	288
Mos1	TRILEPFRKNPEIV			
Mos2	TKILEPFRQNPDIV			
P394	TKILEPFRQNPELV	Pol	320	334
Mos1	ICEEMEKEGKITKIG			
Mos2	ICTEMEKEGKISKIG			
P100	ICTEMEKEGKISKIG	Pol	190	206
Mos1	VKLWYQLEKDFIAGV			
Mos2	VKLWYQLEKEPIVGA			
P030	VKLWYQLEKDFIVGA	Pol	573	590
Mos1	GDIIGDIRQAHCNIS			
Mos2	GDIIGDIRQAHCNLS			
E764	SDIIGDIRQAHCNIS	Env	301	334
Mos1	ICTTVPWNASWSNK			
Mos2	ICTTAVPWNTSWSNK			
E334	ICTTVPWNASWSNR	Env	633	617

Figura 22 (continuación)

M081	IGVFDQAEHETAV								
M082	IGVFDQAEHETAV								
P174	IGVFDQAEHETAV	P01	877						891
M081	GKQMGADCVAGQQ								
M082	GKQMGADCVAGQQ								
P185	GKQMGADCVAGQQ	P01	987						1001
4 respuestas CD4									
M081	KYAKNRGARTNDVQK								
M082	KYAKNRGARTNDVQK								
P174	KYAKNRGARTNDVQK	P01	509						522
M081	AGDIIGDIRQAHCHN								
M082	AGDIIGDIRQAHCHN								
P112	WSEIIGDIRQAHCHN	ENY	320						333
M081	NOQKNEKDLALDS								
M082	NOQKNEKDLALDS								
P088	NOQKNEKDLALDS	ENY	631						665
M081	DIRVFRERVKIKD								
M082	DIRVFRERVKIKD								
P172	DIRVFRERVKIKD	P01	971						985
Mosaico 358-07									
18 CD8 PE8+, 3 CD4 PE8+									
7 regiones CD8+, 3 regiones CD4+									
Número de péptidos solapantes por región: CD8: 2 4 2 6 3 1 1 CD4: 1 1 1									
7 respuestas CD8:									
M081	LHKERVYLSWFAHKQIS								
M082	LHKERVYLSWFAHKQIS								
P223	LHKERVYLSWFAH	P01	650						694
P033	-----EKVLSWFAHKQIS	P01	684						699
M081	IVQGNLIRAIRAQCHLQI								
M082	IVQGNLIRAIRAQCHLQI								
P04	VVQGNLIRAIRAQ	ENY	540						562
P191	-VVQGNLIRAIRAQ	ENY	549						563
P172	-VQGNLIRAIRAQH	ENY	548						563
P19	-----NLIRAIRAQCHLQI	ENY	554						569
M081	AVFTHNRKGGIGGY								
M082	AVFTHNRKGGIGGY								

Figura 22 (continuación)

P222	AVFIHNKRRKGGIGG	Pol	894	908
P236	-VLIHNKRRKGGIGY	Pol	895	909
Mcs1	EQASQDVKNWMTLLVQNAP			
Mcs2	EQATQDVKNWMTLLVQNAP			
G122	EQASQDVKNWMTLL	Gag	307	321
G197	EQSTQDVKNWMTLL	Gag	307	321
G66	--ASQTKNWTETLLI	Gag	309	323
G129	---TQDVKNWMTLLIQ	Gag	310	324
G30	-----KNWMTLLVQNAP	Gag	314	328
G55	-----KNWMTLLVQNAP	Gag	314	328
Mcs1	AEDRVLHVFHAGFAPGQ			
Mcs2	AEDRLHVFHAGFVAPGQ			
G319	AEEDRLHVFHAGFIP	Gag	209	225
G44	AEDRLHVFHAGFIA	Gag	210	224
G11	---DRVHVFHAGFIPGQ	Gag	213	227
Mcs1	WKGSPAIQCSTRI			
Mcs2	WKGSPAIQCSTTKI			
P238	WKGSPAIQCSTKI	Pol	308	322
Mcs1	MTRILEPFRANFEI			
Mcs2	MTRILEPFRQNFEI			
P144	MTRILEPFRQNFEI	Pol	319	333
3 respuestas CD4				
Mcs1	GSPAIQCSTRIILE			
Mcs2	GSPAIQCSTRIILE			
P141	GSPAIQCSTRIILD	Pol	310	324
Mcs1	HGQVDCSPGIWQLAC			
Mcs2	HGQVDCSPGIWQLAC			
P25	HGQVDCSPGIWQLDC	Pol	766	780
Mcs1	DRFYKTLRAEQASQD			
Mcs2	DRFYKTLRAEQATQD			
G57	DRFYKTLRAEQASQE	Gag	299	312
Mosaico 359-07				
1C CD8 PTE+, 5 CD4 PTE+				
10 regiones CD8+, 2 regiones CD4+				
Número de péptidos solapantes por región :				
			CD8: 1 1 1 1 1 1 1 1	CD4: 3 2
10 respuestas CD8:				
Mcs1	LIKERVYLSVFAHKGIG			

Figura 22 (continuación)

M052	LIKERVYLSWPAHKGIG	Pol	690	694
F223	LIKERVYLSWYERH	Pol	684	690
F43	---EKVYLSWPAHKGIG			
M051	AEWDRVHFVHAGPIAFQ	Gag	210	224
M050	AEWDRLHFVHAGPWAGQ	Gag	213	227
G44	AEWDRHFVHAGPIA			
G102	---EUVHFVHAGPIRFGQ			
M051	AIQCSMTRILEFFR	Pol	313	327
M052	AIQSSMTRILEFFR			
F355	AIQSSMTRILEFFR			
M051	ASGIYGIKIVQQLCK	Pol	421	436
M053	ASGIYGIKIVQQLCK			
F355	WSGIYGIKIVQQLCK			
M051	RELNRQDFWEVQL	Pol	233	247
M052	RELNRQDFWEVQL			
F233	RELNRQDFWEVQL			
M051	LEAIEAQOHLQLT	ENT	555	569
M052	LEAIEAQOHLQLT			
E217	LEAIEAQOHLQLT			
M051	LICTTVFNASWSN	ENT	602	616
M050	LICTTVFNASWSN			
E432	HICTTVFNASWSN			
M051	KHIVRSSELEEFV	Gag	32	46
M052	KHIVRSSELEEFV			
G20	KHIVRSSELEEFV			
M051	RSIYNTVALICVHQ	Gag	76	90
M052	RSIYNTVALICVHS			
G20	RSIYNTVALICVHQ			
M051	EKAFSFEVIFMFA	Gag	160	174
M050	EKAFSFEVIFMFA			
G14	EKAFSFEVIFMFA			

2 respuestas CD4:

M051	WKGSFALFQCSMTRILEFFRA	Pol	309	322
M052	WKGSFALFQCSMTRILEFFRK	Pol	310	324
F352	WKGSFALFQCSMTRILEFFR			
F144	---GSPALFQCSMTRILEFFR			

Figura 22 (continuación)

M062	WQLEKPIVQGETE				
P170	WQLENDPIQGETE	P01	581	585	
M081	ENVTFENKNNNV				
M082	ENVTFENKNNNV				
E273	ENVTFENKNNNV	ENV	87	101	
M081	VIQNSDIKVERK				
M082	VIQNSDIKVERK				
P165	VIQNSDIKVERK	P01	585	579	
M081	VSONYPIVNIQQM				
M082	VSONYPIVNIQQM				
G19	VSONYPIVNIQQM	G09	140	142	
M081	VLAEMSQVNSATM				
M082	VLAEMSQVNSATM				
G307	VLAEMSQVNSATM	G09	362	377	
2 respuestas CD4:					
M081	HQVDCSPGIMQAC				
M082	HQVDCSPGIMQAC				
P25	HQVDCSPGIMQAC	P01	788	790	
M081	RVLAMSQVNSATM				
M082	RVLAMSQVNSATM				
G305	RVLAMSQVNSATM	G09	361	376	
Mosaico 361-07:					
21 CD8 EE8+, 6 CD4 EE8+					
7 regiones CD8+, 5 regiones CD4+					
Número de péptidos solapantes por región:					
7 respuestas CD8:					
M081	EQLIKKERVYLSMVFARHGIS				
M082	EQLIKKERVYLSMVFARHGIS				
P21	EQLIKKERVYLSMVFARHGIS	P01	678	692	
P250	EQLIKKERVYLSMVFARHGIS	P01	678	692	
P316	-LILIKKERVYLSMVFARHGIS	P01	678	693	
P223	-LILIKKERVYLSMVFARHGIS	P01	680	694	
P169	-----EQLIKKERVYLSMVFARHGIS	P01	684	698	
P93	-----EQLIKKERVYLSMVFARHGIS	P01	684	698	
M081	CTHGIRPVVSTQILL				
M082	CTHGIRPVVSTQILL				
E15	CTHGIRPVVSTQILL	ENV	247	261	

CD8: 6 2 2 2 5 1 1 CD4: 2 1 1 1 1

Figura 22 (continuación)

E47	CTHCIREYVSTQLLE	ENV	247	261
M081	ICTTTFVWNASWNSKSL			
M082	ICTTAVPWNNTSWSKQ	ENV	603	617
E334	ICTTTFWASNSK	ENV	604	618
E214	-CTTVPWNSNSKT	ENV	605	619
E158	--TTAVFWASNSKSL	ENV	608	619
E290	--TTAVPWNNTSWSKSL			
M081	ACQVGGGCHKARVLAEMS			
M082	ACQVGGPSHKARVLAEMS	GAG	349	363
G166	ACQVGGGCHKARVLA	GAG	354	368
G76	-----GGPSHKARVLAEMS			
M081	AAERDRVHVHAGPIRFGQ			
M082	AAERDRLHFVHAGFVAFQ	GAG	209	225
G319	AAE-DELFVHAGPIR	GAG	210	224
G242	-AERDRLHFVHAGFVA	GAG	210	224
G44	-AERDRLHFVHAGPIA	GAG	212	226
G277	---RDFVHVHAGSNRPS	GAG	212	226
G102	---DRVHVHAGPIRFGQ	GAG	212	226
M081	HSNFWAMASDFNLPP			
M082	HSNFWAMASENLPP	POL	731	745
F93	HSNFWAMASDFNLPP			
M081	KGRGNFLQNRPEPT			
M082	KGRGNFLQSRPEPT	GAG	442	456
G86	KGRGNFLQNRPEPT			
5 respuestas CD4:				
M081	RSLENTVATLYCVHQR			
M082	RSLENTVATLYCVHNE	GAG	76	90
G289	KSENTVATLYCVHA	GAG	77	91
G54	--SLNTVATLYCVHQR			
M081	YVDRGRQKIVSLTE			
M082	YVDRGRQKWSLTD	POL	612	626
E134	YVDRGRQKWSLTD			
M081	CTTTFWNASWNSKS			
M082	CTTAVPWNNTSWSNKS	ENV	604	618
E53	CTTTFWNASWNSKS			
M081	WWSGIQQEFGIPIYP			
M082	WWSGIQQEFGIPIYP			

Figura 22 (continuación)

P135	WVAGIQCFGIPNP	Pol	846	860
Mos1	GFHKARVLAEMSQ			
Mos2	GPSHKARVLAEMSQ			
G15	GFHKARVLAEMSQ	Gag	355	369
<hr/>				
Mosaico 362-07:				
13 CD8 PTE+, 5 CD4 PTE+				
7 regiones CD8+, 5 regiones CD4+				
Número de péptidos solapantes por región: : CD8: 2 3 4 2 1 1 1 CD4: 1 1 1 1 1 1				
7 respuestas CD8:				
Mos1	CTHGIRPVVSTQLLL			
Mos2	CTHGIRPVVSTQLLL			
E15	CTHGIRPVVSTQLLL	Env	247	261
E47	CTHGIRPVVSTQLLL	Env	247	261
Mos1	KHIVWASRELEREAVNPGLL			
Mos2	KHIVWASRELEREAVNPGLL			
G20	KHIVWASRELEREAL	Gag	32	46
G68	-----ASRELEREAVNPGLL	Gag	37	51
Mos1	QPSLQGTGSEELRSLENTVAITL			
Mos2	QPALQGTGTEELRSLENTVAITL			
G268	LPALKTGSSEELRSLSL	Gag	65	79
G249	--ELQGTGSEELRSLENT	Gag	67	81
G132	---LQGTGTEELRSLENTV	Gag	68	82
G120	-----GSEELRSLENTVAITL	Gag	71	85
Mos1	ACQEVGGPGHKARVLAEMS			
Mos2	ACQEVGGPGSHKARVLAEMS			
G166	ACQEVGGPGHKARVL	Gag	349	363
G76	-----GGPSHKARVLAEMG	Gag	354	368
Mos1	RELNKRTODFWEVQL			
Mos2	RELNKRTODFWEVQL			
P233	RELNRRTODFWEVQL	Pol	233	247
Mos1	LLRAIEAQQHLLQLT			
Mos2	LLRAIEAQQHMLQLT			
E227	LLRAIEAQQHLLQLT	Env	555	569
Mos1	KTAVQMAVFIHNFKR			
Mos2	KTAVQMAVFIHNFKR			
P5	KTAVQMAVFIHNFKR	Pol	888	902

Figura 22 (continuación)

```

5 respuestas CD4:
M081 KLNWASQIYFGIYR
M082 KLNWASQIYGIKYS
F106 KLNWASQIYNSIKYK
POL 418 422

M081 YAMRHTHTNVKQL
M082 YAMRHTHTNVKQL
F060 YAMRHTHTNVKQL
POL 509 523

M081 DQAEHLKTAQVAVF
M082 DQAEHLKTAQVAVF
F10 DQAEHLKTAQVAVF
POL 642 646

M081 SLQTGSEELSLINT
M082 ALQTGSEELSLINT
G182 ALQTGSEELSLINT
GAG 67 81

M081 GPGHKARVLAEMSQ
M082 GPGHKARVLAEMSQ
G15 GPGHKARVLAEMSQ
GAG 355 369

Péptidos reactivos y amplitud y profundidad de las respuestas: vacuna de Clado C óptimo natural
Clado C natural: 363-07:
2 CD8 PE2+, 0 CD4 PE2+
1 CD8+ regions, 0 CD4+ regions
Número de péptidos solapantes por región: CD8: 0 CD4: ninguno

1 respuesta CD8:
C QPALQGTETELSLINTVATL
G208 LKALQTGSEELSLI
G120 -----GSEELSLINTVATL

Clado C natural 364-07
3 CD8 PE2+, 2 CD4 PE2+
3 CD8+ regions, 2 CD4+ regions
Number of overlapping peptides per region: CD8: 1 1 1 1 CD4: 1 1

5 respuestas CD8:
C AHNDYRQLTEAVK
F80 AHNDYRQLTEAVK POL 815 829

C GDIIQDINQAHNIS
E64 GDIIQDINQAHNIS ENV 321 334

```

Figura 22 (continuación)

C	SEKAVGICAVFLCF			
E182	SEKAVGICAVFLCF	ENV	508	522
C	RASILRGKLDNEX			
E283	RASILRGKLDNEX	Gag	4	10
C	TNNPFVYGDYRW			
E365	TNNPFVYGDYRW	Gag	251	265
2 respuestas CD4:				
C	CRSNTGILLRGG			
E316	CRSNTGILLRGG	ENV	446	459
C	SLNVTATLYVHAG			
E84	SLNVTATLYVHAG	Gag	77	91
<hr/>				
Clado C natural 365-07				
2 CD8 RT4, 1 CD4 RT4				
2 regiones CD8+, 1 región CD4+				
Número de péptidos solapantes por región: CD8: 1 1 CD4: 1				
2 respuestas CD8:				
C	EOLNKEVYLGMP			
E250	EOLNKEVYLGMP	Pol	676	692
C	RENKLEPVHAGIA			
E84	RENKLEPVHAGIA	Gag	210	224
1 respuestas CD4:				
C	QQAQMLNDTINEE			
E24	QQAQMLNDTINEE	Gag	194	208
<hr/>				
Clado C natural 366-07				
6 CD8 RT4, 1 CD4 RT4				
4 regiones CD8+, 1 región CD4+				
Número de péptidos solapantes por región: CD8: 2 2 1 1 CD4: 1				
4 respuestas CD8:				
C	IVQOSNLRIRKQ			
E24	IVQOSNLRIRKQ	ENV	548	562
E72	IVQOSNLRIRKQ	ENV	548	562
C	NVFIHFVPCIGYI			
E21	NVFIHFVPCIGYI	Pol	894	908
E236	NVFIHFVPCIGYI	Pol	894	908

Figura 22 (continuación)

```

C MAICENEKKGKIK
P224 MAICENEKKGKIK P01 190 204

C GQSHKAVLAENS
G76 GQSHKAVLAENS G49 354 368

1 respuesta CD4:

C IISVNDQAEHLKTA
P86 IISVNDQAEHLKTA P01 876 890

-----
Clado C natural 367-07:
0 CD8 PE8+, 0 CD4 PE8+
0 regiones CD8+, 0 regiones CD4
Número de péptidos solapantes por región: ninguno

0 respuesta CD8:
0 respuesta CD4:

-----
Clado C natural 368-07:
2 CD8 PE8+, 0 CD4 PE8+
1 región CD8+, 0 regiones CD4
Número de péptidos solapantes por región: CD8: 2 CD4: 0

1 respuesta CD8:
C AVFIRFKKGGIGY
P22 AVFIRFKKGGIGY P01 894 908
P236 -VLIIRFKKGGIGY P01 895 909
    
```

Figura 23

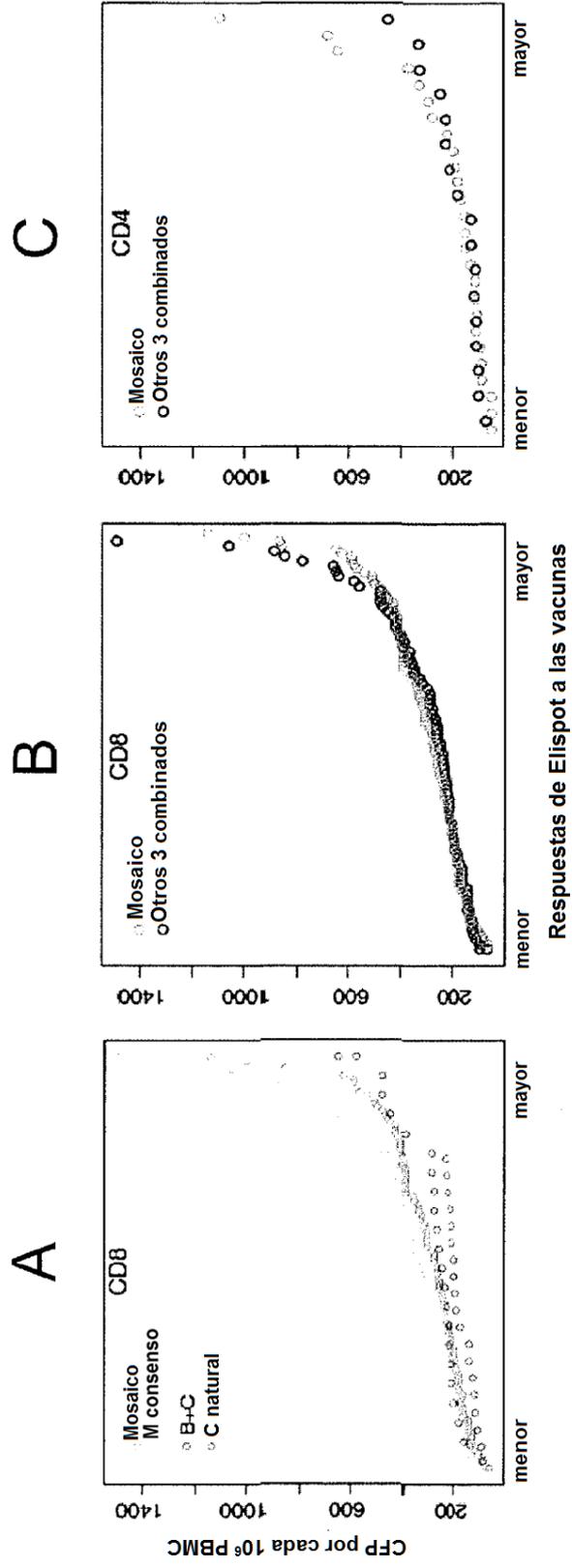


Figura 24

A		B	
Mono 366 (C natural)		Mono 367 (Mosaico)	
OptC	IVQQSNLLRAIEAQ	Mos1	EQLIKKERVYLSWVPAHKGIG
E54	VVQQSNLLRAIEAQ	Mos2	EQLIKKERVYLSWVPAHKGIG
E72	-VQQSNLLRAIEAQ	P221	EELIKKERVYLSWV
OptC	AVFIHNFRRKGGIGGY	P250	EPLIKKERVYLSWV
P22	AVFIHNFRRKGGIGG	P316	-KLIKDKRVYLSWVPA
P236	-VLIHNFRRKGGIGGY	P223	--LIKKERVYLSWVPAH
OptC	MAICEEMEKEGKITK	P169	-----EAVYLSWVPAHKGIG
P224	TAICEEMEKEGKITK	P93	-----EKVYLSWVPAHKGIG
OptC	GGPSHKARVLAEMG	Mos1	CTHGIRPVVSTQLLL
G76	GGPSHKARVLAEMG	Mos2	CTHGIRPVVSTQLLL
		E15	CTHGIRPVVSTQLLL
		E47	CTHGIRPVVSTQLLL
		Mos1	ICTTVPWNASWSNKS
		Mos2	ICTTAVPWNISWSNKSQ
		E334	ICTTVPWNASWSNR
		E214	-CTTVPWNSSWSNKT
		E158	--TTAVPWNASWSNKS
		E290	--TTAVPWNISWSNKS
		Mos1	ACQGVGGPCHKARVLAEMG
		Mos2	ACQGVGGPCHKARVLAEMG
		G166	ACQEVGGPCHKARV
		G76	-----GGPSHKARVLAEMG
		Mos1	AAEWDRVHPVHAGPIAPGQ
		Mos2	AAEWDRVHPVHAGPVAPGQ
		G319	AAE-DRVHPVHAGPI
		G242	-ADWDRVHPVHAGPVA
		G44	-AEWDRVHPVHAGPIA
		G277	---WDRVHPVHAGPVP
		G102	---DRVHPVHAGPIPPGQ
		Mos1	HSNWRAMASDFNLEPP
		Mos2	HSNWRAMASEFNLEP
		P93	HSNWRAMASDFNLEP
		Mos1	KGRPGNFIQNRPEPT
		Mos2	KGRPGNFIQSRPEPT
		G86	KGRPGNFIQNRPEPT

Figura 24 (continuación)

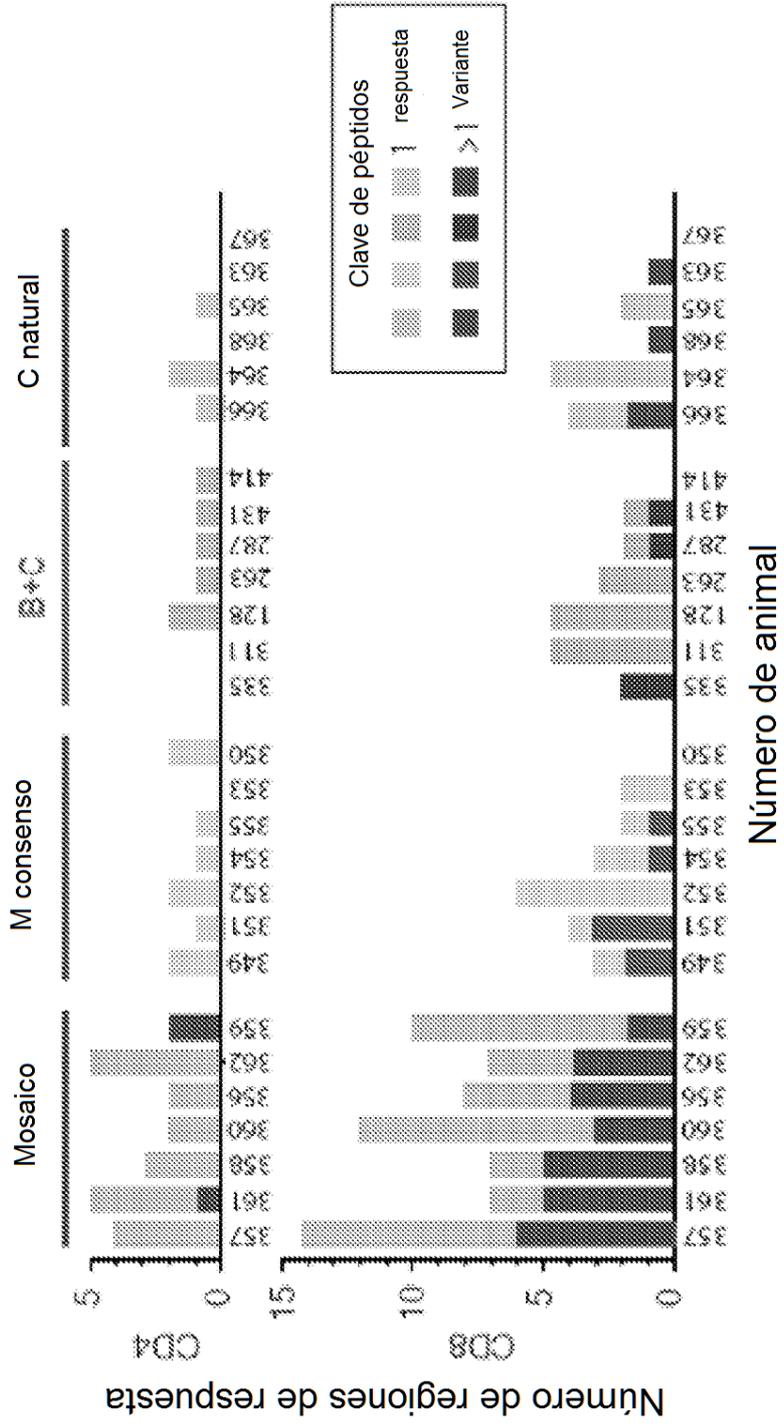


Figura 25

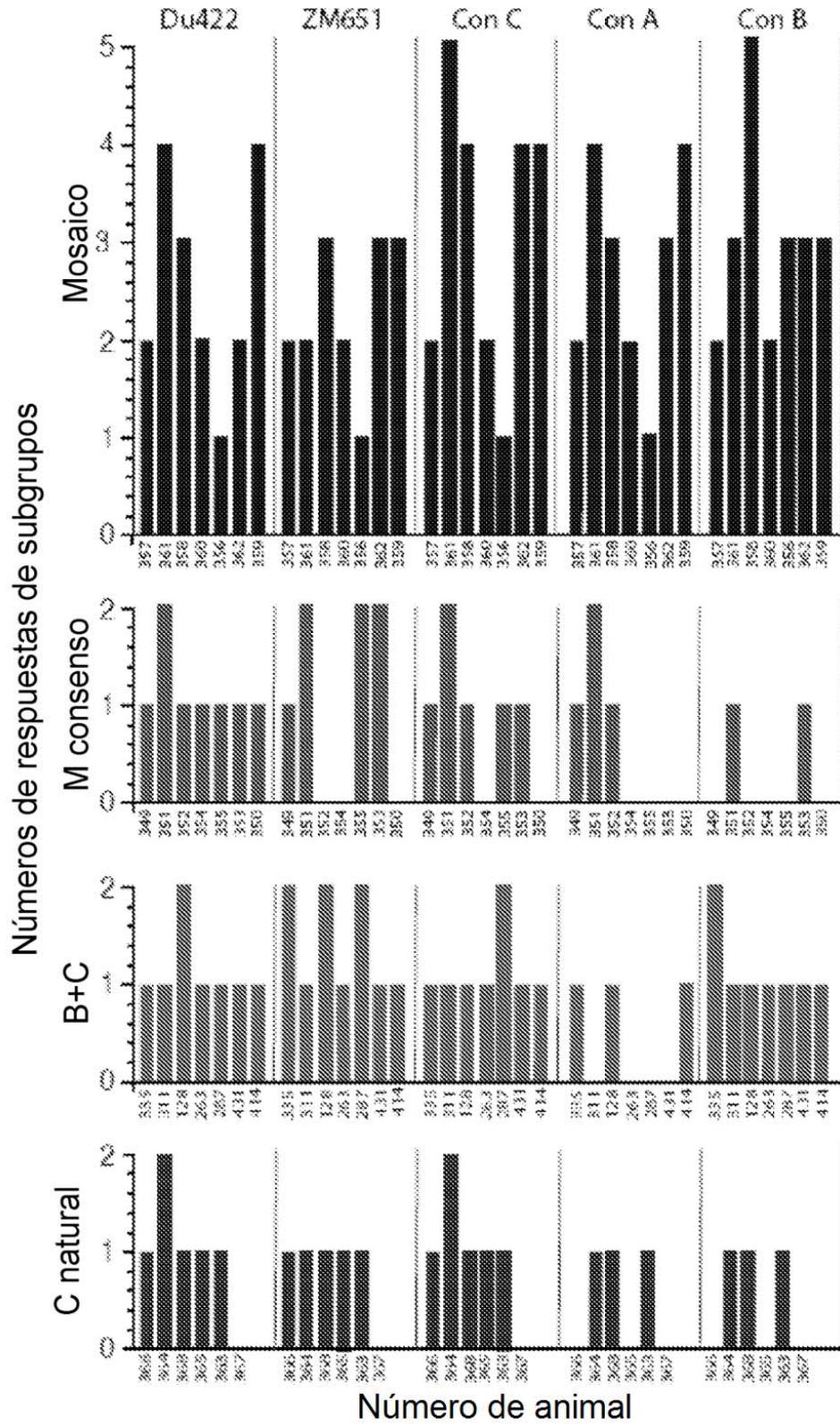
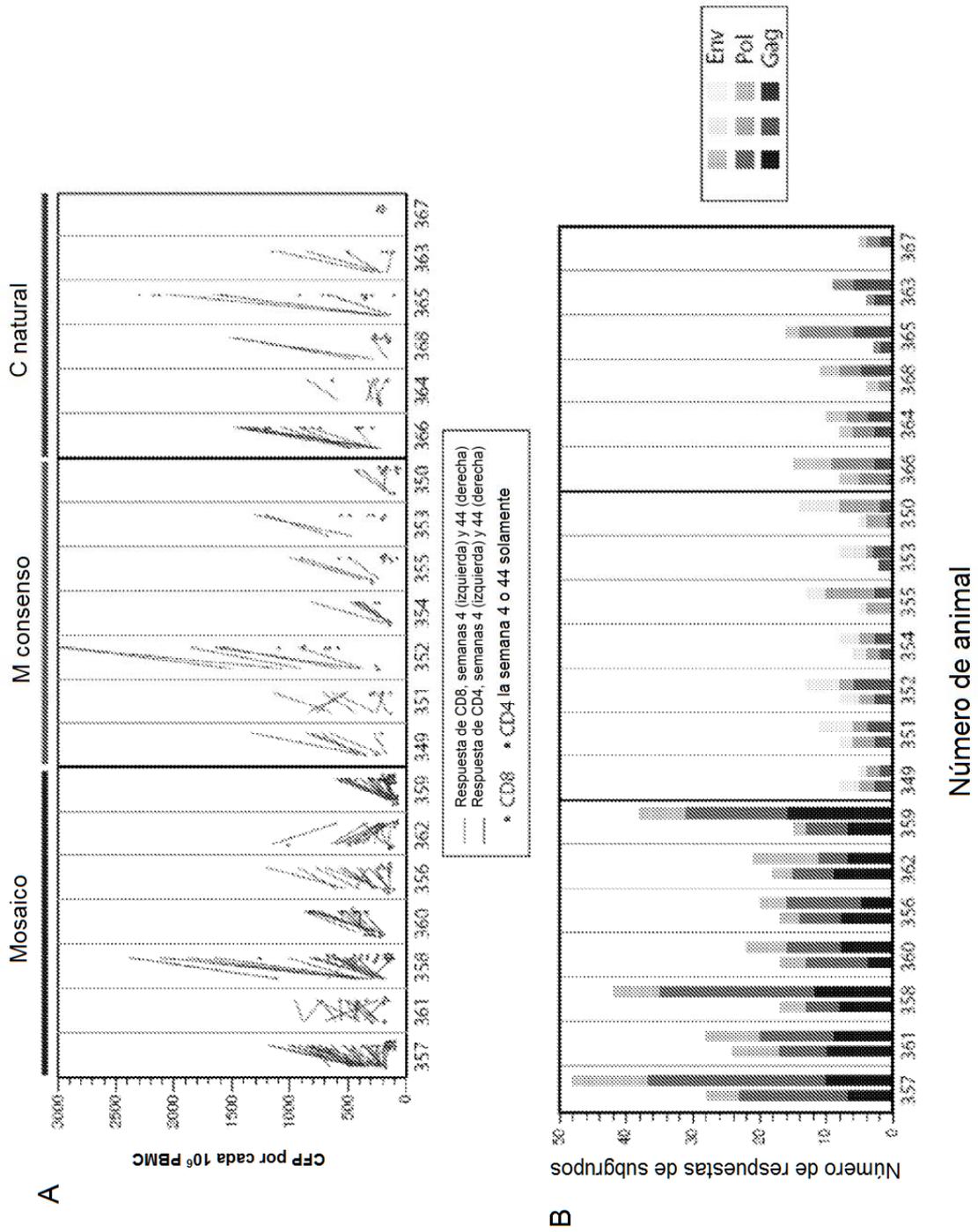


Figura 26



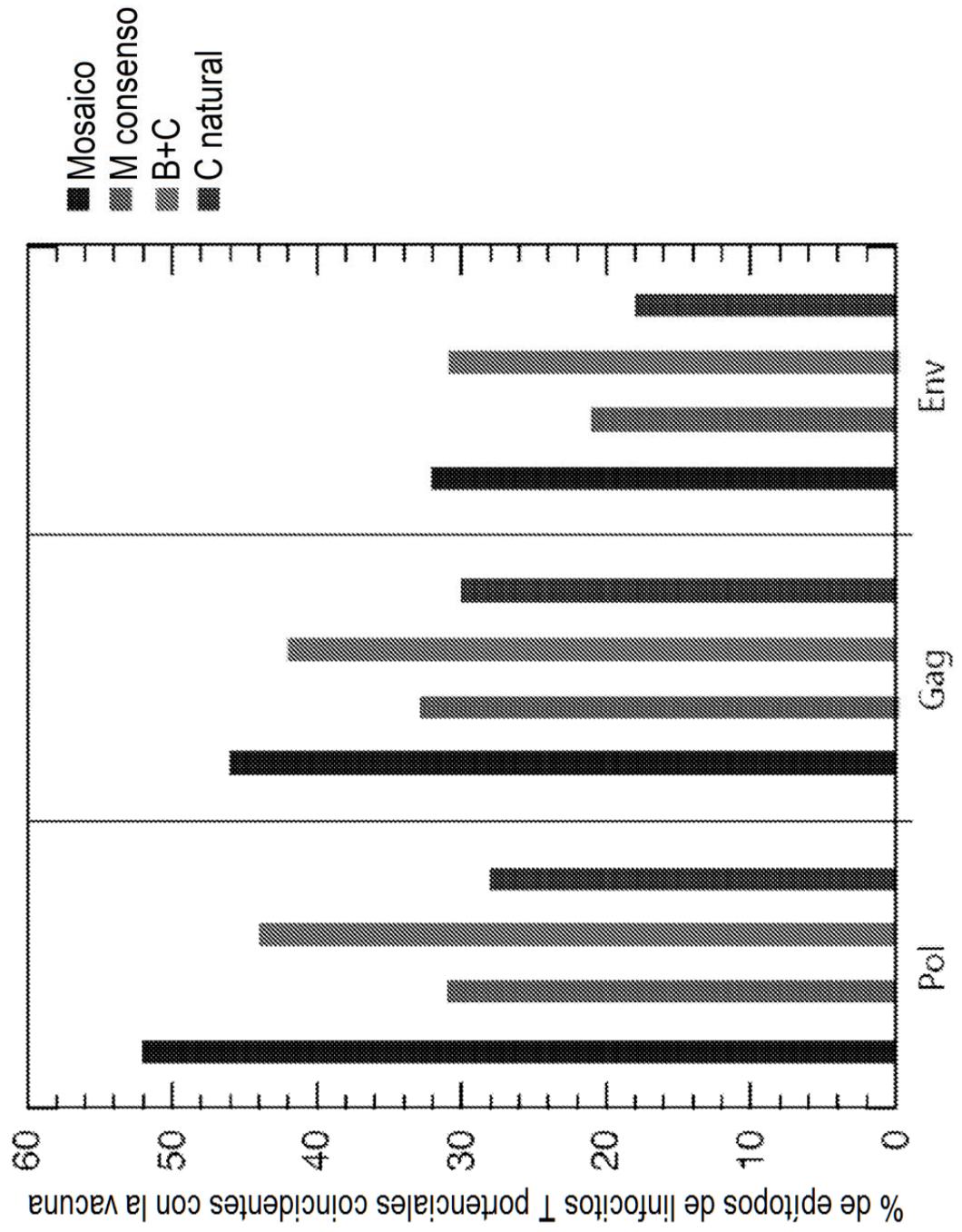


Figura 27