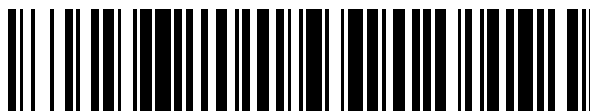


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 528 951**

51 Int. Cl.:

A61K 39/21 (2006.01)

A61K 39/12 (2006.01)

A61K 39/38 (2006.01)

C07K 14/16 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.06.2007 E 11005764 (3)**

97 Fecha y número de publicación de la concesión europea: **26.11.2014 EP 2399602**

54 Título: **Secuencias consenso, antígenos y transgenes del VIH-1 del clado A**

30 Prioridad:

02.06.2006 US 810816 P

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

13.02.2015

73 Titular/es:

**INTERNATIONAL AIDS VACCINE INITIATIVE
(100.0%)**

**110 Williams Street 27th Floor
New York, NY 10038-3901, US**

72 Inventor/es:

**GUPTA, KALPANA y
JACKSON, NICHOLAS**

74 Agente/Representante:

CARPINTERO LÓPEZ, Mario

ES 2 528 951 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

Secuencias consenso, antígenos y transgenes del VIH-1 del clado A

Campo de la invención

5 La presente invención proporciona una composición inmunogénica o de vacuna que comprende un polinucleótido que codifica un antígeno Env del VIH-1 del Clado A, en el que el antígeno Env tiene la secuencia de aminoácidos ilustrada en la Figura 17 y está codificado por la secuencia de nucleótidos ilustrada en la Figura 17. Además, la presente invención se refiere a esta composición inmunogénica o de vacuna para usar en un procedimiento para generar una respuesta inmunitaria contra el VIH-1.

10 La presente divulgación se refiere a secuencias consenso de nucleótidos y proteínas para antígenos del VIH-1 del Clado A y a secuencias de nucleótidos y proteínas para antígenos del clado A de aislados de campo circulantes del VIH-1 donde las secuencias de los antígenos están estrechamente relacionadas con esas secuencias consenso. Por ejemplo, la presente divulgación se refiere a transgenes del VIH-1 del clado A que derivan de dichas secuencias, y que codifican Env del VIH-1 del clado A. La solicitud también se refiere a vectores que contienen dichos transgenes, que incluyen en una realización preferida, vectores de adenovirus que contienen dichos transgenes. La divulgación también se refiere a composiciones inmunogénicas que contienen antígenos del VIH-1 del clado A, secuencias de nucleótidos, vectores o transgenes y a procedimientos para generar una respuesta inmunitaria contra el VIH-1 en un sujeto mediante la administración de una cantidad eficaz de dichas composiciones inmunogénicas.

Antecedentes de la invención

20 El SIDA, o Síndrome de Inmunodeficiencia Adquirida, está causado por el virus de la inmunodeficiencia humana (VIH) y se caracteriza por varias particularidades clínicas, que incluyen síndromes de consunción, degeneración del sistema nervioso central e inmunosupresión profunda que da como resultado infecciones oportunistas y neoplasias. El VIH es un integrante de la familia de los lentivirus de los retrovirus animales, que incluyen al virus visna de las ovejas y el virus de inmunodeficiencia de los bovinos, felinos y simios (VIS). Hasta el momento se han
25 identificado dos tipos de VIH, denominados VIH-1 y VIH-2 estrechamente relacionados, de los cuales el VIH-1 es por mucho la causa más común del SIDA. Sin embargo, el VIH-2, que difiere en estructura genómica y antigenicidad, produce un síndrome clínico similar.

30 Una partícula de VIH infecciosa consiste en dos cadenas idénticas de ARN, cada una de aproximadamente 9.2 kb de largo, empaquetadas dentro de un núcleo de proteínas virales. Esta estructura central está rodeada por una envoltura de bicapa de fosfolípidos derivada de la membrana de la célula huésped que también incluye proteínas de membrana codificadas viralmente (Abbas et al., Cellular and Molecular Immunology, 4^a edición, W.B. Saunders Company, 2000, p. 454). El genoma del VIH tiene la organización 5'-LTR-Gag-Pol-Env-LTR-3' característica de la familia de los retrovirus. Las repeticiones terminales largas (LTR) en cada extremo del genoma viral sirven como
35 sitios de unión para las proteínas reguladoras transcripcionales del huésped y regulan la integración viral en el genoma del huésped, la expresión de genes virales y la replicación viral.

40 El genoma del VIH codifica varias proteínas estructurales. El gen Gag codifica proteínas estructurales del núcleo de la nucleocápside y la matriz. El gen Pol codifica las enzimas transcriptasa inversa (RT), integrasa (Int) y proteasa viral necesarias para la replicación viral. El gen tat codifica una proteína que es necesaria para la elongación de los transcritos virales. El gen rev codifica una proteína que promueve la exportación nuclear de los ARN virales incompletamente empalmados o sin empalmar. El producto del gen Vif potencia la infectividad de las partículas virales. El producto del gen vpr promueve la importación nuclear de ADN viral y regula la detención del ciclo celular G2. Los genes vpu y nef codifican proteínas que regulan hacia abajo la expresión de CD4 de la célula huésped y potencian la liberación del virus desde las células infectadas. El gen Env codifica la glucoproteína de la envoltura viral que es traducida como un precursor de 160 kilodalton (kDa) (gp160) y escindida por una proteasa
45 celular para producir la glucoproteína externa de 120 kDa de la envoltura (gp120) y la glucoproteína transmembrana de 41 kDa de la envoltura (gp41), que son necesarias para la infección de las células (Abbas, págs. 454 - 456). Gp140 es una forma modificada de la glucoproteína env que contiene la porción de glucoproteína externa de 120 kDa de la envoltura y una parte de la porción gp41 de env y tiene características tanto de gp120 como de gp41. El gen Nef se conserva entre los lentivirus de primates y es uno de los primeros genes virales que
50 es transcrito después de la infección. Se han descrito varias funciones *in vitro*, que incluyen la regulación hacia abajo de la expresión superficial de CD4 y MHC de clase I, la señalización y activación de los linfocitos T alterados y una mayor infectividad viral.

55 La infección por el VIH se inicia con gp120 en la partícula viral que se une a CD4 y moléculas del receptor de quimiocinas (por ejemplo, CXCR4, CCR5) en la membrana celular de las células diana tales como los linfocitos T de CD4+, macrófagos y células dendríticas. El virus unido se fusiona con la célula diana y realiza la transcripción

inversa del genoma de ARN. El ADN viral resultante, se integra en el genoma celular, donde dirige la producción de un nuevo ARN viral y consecuentemente de proteínas virales y nuevos viriones. Esos viriones brotan de la membrana de la célula infectada y establecen infecciones productivas en otras células. Este proceso también mata a la célula infectada originalmente. El VIH también puede matar a las células indirectamente porque el receptor CD4 de los linfocitos T no infectados tiene una fuerte afinidad por gp120 expresada en la superficie de las células infectadas. En este caso, las células no infectadas se unen, a través de la interacción receptor CD4-gp120, a las células infectadas y se fusionan para formar un sincicio, que no puede sobrevivir. La destrucción de linfocitos T de CD4+, que son esenciales para la defensa inmunitaria, es una de las principales causas de la disfunción inmunitaria progresiva que es el sello característico del avance de la enfermedad del SIDA. La pérdida de linfocitos T de CD4+ perjudica seriamente la capacidad del organismo para luchar contra la mayor parte de los invasores, pero tiene un impacto especialmente grave en las defensas contra virus, hongos, parásitos y algunas bacterias, incluidas las micobacterias.

Los diferentes aislados del VIH-1 se clasificaron en tres grupos: M (principal), O (extraño) y N (no M, no O). El Grupo M del VIH-1 domina la pandemia mundial del VIH (Gaschen et al., (2002) Science 296: 2354 - 2360). Desde que el grupo M del VIH-1 inició su expansión en los humanos hace aproximadamente 70 años (Korber et al., Retroviral Immunology, Pantaleo et al., eds., Humana Press, Totowa, NJ, 2001, págs. 1 - 31), se diversificó rápidamente (Jung et al., (2002) Nature 418: 144). El Grupo M del VIH-1 está constituido por un número de diferentes clados (también conocidos como subtipos) así como variantes resultantes de la combinación de dos o más clados, conocidas como formas recombinantes circulantes (CRF). Los subtipos se definen como si tuvieran genomas que son al menos únicos al 25% (AIDS epidemic update, diciembre de 2002). Se identificaron once clados y una letra designa cada subtipo. Cuando los clados se combinan entre sí y se establecen exitosamente en el ambiente, como puede ocurrir cuando una persona está infectada con dos subtipos diferentes de VIH, el virus resultante se conoce como una CRF. Hasta ahora, se han identificado aproximadamente 13 CRF. Los clados del VIH-1 también muestran preferencia geográfica. Por ejemplo, el clado A, el segundo clado más frecuente, prevalece en África oriental, mientras que el clado B es común en Europa, las Américas y Australia. El clado C, el subtipo más común, está difundido en el sur de África, India y Etiopía (AIDS epidemic update, diciembre de 2002). Incluso entre clados existe variabilidad del virus entre diferentes cepas y aislados virales.

Esta variabilidad genética del VIH crea un reto científico para el desarrollo de vacunas. Se ha sugerido un enfoque que consiste en desarrollar secuencias consenso basadas en las secuencias de múltiples cepas diferentes del VIH y en desarrollar vacunas basadas en esas secuencias consenso. El fundamento de esos enfoques es que las secuencias consenso codificarán antígenos que están conservados entre las diferentes cepas del VIH y que dichos antígenos por lo tanto, es probable que sean útiles en la generación de respuestas inmunitarias contra múltiples cepas diferentes del VIH. Las secuencias consenso del VIH-1 del clado A fueron generadas por otros. Véase por ejemplo, Nkolola et al. (2004) Gene Ther. 2004. Jul. 11 (13): 1068 - 80, y Korber B (eds) et al. Human Retroviruses and AIDS: A Compilation and Analysis of Nucleic Acid and Amino Acid Sequences. Los Alamos National Laboratory: Los Alamos, Nuevo México, Estados Unidos, (1997) que involucran los transgenes RENTA e HIVA derivados de las secuencias consenso del clado A. Sin embargo, las secuencias consenso descritas en esos artículos parecen haber sido derivadas de la secuencia consenso del VIH-1 del clado A obtenida del laboratorio de Los Alamos, y no se generaron de la misma forma que las secuencias consenso de la presente invención. Además, esas referencias no instruyen acerca del uso de secuencias de las cepas de VIH existentes recientemente circulantes que se corresponden estrechamente con la secuencia consenso. En su lugar implican el uso de las secuencias consenso en sí mismas.

El documento WO 2005/028625 A2 desvela, por lo general, un ácido nucleico que comprende una secuencia nucleótida consenso que codifica Env del VIH-1 del Clado A y su uso en forma de vacuna en un procedimiento para generar una respuesta inmunitaria frente al VIH-1.

Sumario de la invención

La presente divulgación proporciona secuencias consenso nuevas y mejoradas para antígenos del VIH-1 del clado A y procedimientos para producir dichas secuencias consenso nuevas y mejoradas. Las secuencias consenso son especialmente ventajosas porque se basan en las secuencias de antígenos de un gran número de cepas del VIH-1 del clado A diferentes, y también porque se basan en las secuencias de antígenos de cepas recientemente aisladas del VIH-1 del clado A. En consecuencia, las secuencias consenso de la presente divulgación tienen mayor importancia biológica en comparación con las secuencias consenso del VIH-1 del clado A generadas previamente.

Otra ventaja importante de la presente divulgación es que proporciona antígenos del VIH-1 del clado A y estrategias para producir dichos antígenos, que se derivan de cepas de origen natural del VIH-1 del clado A. Esos antígenos se seleccionan de modo que estén estrechamente relacionados con, o tengan una pequeña "distancia de proteínas" respecto a, las secuencias consenso de la presente divulgación. Una ventaja de usar esas secuencias de origen natural con la máxima correspondencia con las secuencias consenso, por contraposición a las

secuencias consenso generadas artificialmente, es que se necesitan menos manipulaciones genéticas para generar esas secuencias y se asegura la importancia biológica.

5 En un primer aspecto, la presente invención está dirigida a una composición inmunogénica o de vacuna que comprende un polinucleótido que codifica un antígeno Env del VIH-1 del Clado A, en el que el antígeno Env tiene la secuencia de aminoácidos ilustrada en la Figura 17 y está codificado por la secuencia de nucleótidos ilustrada en la Figura 17. En otro aspecto, la presente invención se refiere a esta composición inmunogénica o de vacuna para usar en un procedimiento para generar una respuesta inmunitaria contra el VIH-1

10 La solicitud está dirigida a una secuencia consenso de aminoácidos para un antígeno del VIH-1 del clado A. Se refiere a secuencias consenso de aminoácidos para los antígenos para los antígenos Gag, Pol (compuesto por RT e Int), Nef y Env del VIH-1 del clado A. También se refiere a la secuencia consenso de aminoácidos de Gag de la FIG. 1, la secuencia consenso de aminoácidos de Pol de la FIG. 3, la secuencia consenso de aminoácidos de Env de la FIG. 5 y/o la secuencia consenso de aminoácidos de Nef de la FIG. 7.

15 La solicitud está dirigida a un procedimiento de identificación de una secuencia consenso de aminoácidos para un antígeno del VIH-1 del clado A de interés, que comprende determinar la secuencia de aminoácidos del antígeno de interés en varias cepas o aislados de campo circulantes del VIH-1, alinear dichas secuencias y determinar la secuencia consenso para ese antígeno.

20 Además, la solicitud se refiere a un procedimiento para identificar un antígeno del VIH-1 del clado A de una cepa circulante o un aislado de campo del VIH-1 del clado A que tenga una secuencia de aminoácidos que sea similar a la secuencia consenso de aminoácidos para ese antígeno del VIH-1 del clado A. En una realización preferida el antígeno del VIH-1 del clado A se selecciona basándose en el grado de similitud con la secuencia consenso, prefiriéndose las secuencias que tienen el mayor grado de similitud con, o la "distancia de proteínas" más pequeña respecto a, la secuencia consenso. En otra realización preferida el antígeno del VIH-1 del clado A se selecciona de una cepa o un aislado de campo recientemente circulante del VIH-1 del clado A. En otra realización la solicitud se refiere a los antígenos del VIH-1 del clado A identificados que usan dichos procedimientos.

25 Además, la solicitud se refiere a un procedimiento para identificar un antígeno del VIH-1 del clado A de una cepa o un aislado de campo circulante del VIH-1 del clado A que tiene una secuencia de aminoácidos similar a la secuencia consenso de aminoácidos para ese antígeno del VIH-1 del clado A y luego realizar mutaciones en la secuencia para anular las funciones biológicas de las secuencias. Es preferible usar un enfoque minimalista, es decir, que el número de mutaciones se mantenga en el mínimo, para que se realicen sólo las mutaciones necesarias para anular la función y facilitar la obtención de la aprobación de las autoridades de registro sanitario y se evite la alteración innecesaria de las secuencias de genes del VIH-1 originales. Por ejemplo, en una realización el componente Nef de GRIN no se altera sino más bien la fusión del extremo N-terminal de Nef con el extremo C-terminal de Int anula la función de nef conservando todas las secuencias de nucleótidos originales de Nef.

30 Además, la solicitud se refiere a un procedimiento para mejorar la estabilidad genética del transgén del VIH-1 del clado A para tecnologías de inserción en vectores virales. El componente PR (proteasa) se elimina de Gag-Pol de longitud completa-Nef (Pol de longitud completa contiene PR, e Int y RT) para que queden sólo las porciones Int y RT de Pol. Esto tiene la ventaja de mayor estabilidad genética y mejores propiedades de clonación y rescate del virus, particularmente usando Ad35 y/o Ad11. La eliminación de PR de esta forma es un enfoque minimalista puesto que se elimina sólo la subunidad funcional más pequeña de POL, conservando las subunidades funcionales más grandes IN y RT. La solicitud también se refiere a los antígenos del VIH-1 del clado A seleccionados y producidos usando estos procedimientos.

35 Por ejemplo, el antígeno es un antígeno Gag de una de las cepas indicadas en la tabla 1 y la figura 2. Preferentemente el antígeno Gag se selecciona de una cepa en la cual la "distancia de proteínas" respecto a la secuencia consenso de Gag es menor del 0,07%, o más preferentemente menor del 0,06%, o aún más preferentemente menor del 0,05%. En una realización preferida el antígeno Gag es de la cepa TZA173, la cepa 97TZ02, la cepa KNH1144 o la cepa SE7535UG del VIH-1 del clado A.

40 Otro antígeno es un antígeno Pol de una de las cepas indicadas en la tabla 2 y la figura 4. Preferentemente el antígeno Pol se selecciona de una cepa en la cual la "distancia de proteínas" respecto a la secuencia consenso de Pol es menor del 0,03%, o más preferentemente menor del 0,025%. En una realización preferida el antígeno Pol es de la cepa MSA4070, la cepa SE7245SO o la cepa SE8538 del VIH-1 del clado A.

50 Otro antígeno es un antígeno Env de una de las cepas indicadas en la tabla 3 y la figura 6. Preferentemente el antígeno Env se selecciona de una cepa en la cual la "distancia de proteínas" de la secuencia consenso de Gag es menor del 0,1, o más preferentemente menor del 0,08%, o aún más preferentemente menor del 0,07% o todavía más preferentemente menor del 0,065%. En una realización preferida el antígeno Env es de la cepa KEQ23, la cepa TZA341 o la cepa KNH1088 del VIH-1 del clado A.

55

- De forma adicional, el antígeno es un antígeno Nef de una de las cepas indicadas en la tabla 4 y la figura 8. Preferentemente el antígeno Nef se selecciona de una cepa en la cual la "distancia de proteínas" respecto a la secuencia consenso de Gag es menor del 0,1%, o más preferentemente menor del 0,08% o más preferentemente menor del 0,07%, o más preferentemente menor de 0,06, o aún más preferentemente, menor del 0,05%. En una realización preferida el antígeno Nef es de la cepa MSA4070, o la cepa KNH1211, o la cepa 97TZ03, o la cepa 99UGA070, o la cepa SE8891 UG del VIH-1 del clado A.
- Además, la presente divulgación está dirigida a las secuencias de nucleótidos que codifican los antígenos del VIH-1 del clado A de la solicitud. La solicitud también se refiere a los vectores que comprenden esas secuencias de nucleótidos. Las secuencias de nucleótidos de la solicitud y los vectores que las conforman y también los antígenos codificados por las secuencias de nucleótidos de la solicitud, son útiles para generar una respuesta inmunitaria contra antígenos del VIH del clado A *in vivo* y son útiles en la producción de vacunas contra cepas del VIH-1 del clado A. Las secuencias de nucleótidos de la solicitud también pueden ser útiles para expresar y producir los antígenos del VIH-1 del clado A que ellos codifican en las células o *in vitro*, por ejemplo, para que los antígenos se puedan producir, aislar y/o purificar.
- Los nucleótidos de la solicitud pueden estar alterados en comparación con las secuencias consenso de nucleótidos, o en comparación con las secuencias de aislados circulantes del VIH-1 que están estrechamente relacionadas con dichas secuencias consenso. Por ejemplo, en una realización las secuencias de nucleótidos se pueden mutar para anular la actividad de las proteínas codificadas *in vivo*. En otra realización a las secuencias de nucleótidos se le puede hacer optimización de codones, por ejemplo, los codones se pueden optimizar para uso humano. Preferentemente, las secuencias de nucleótidos de la solicitud se mutan para anular la función normal *in vivo* de las proteínas codificadas y se les optimizan los codones para uso humano. Por ejemplo, cada una de las secuencias de Gag, Pol, Env, Nef, RT e Int se puede modificar de esas maneras.
- Dentro de la divulgación, una sola secuencia de nucleótidos codifica una proteína de fusión que comprende los antígenos Gag, RT (parte de Pol) y Nef. Como se usa en la presente memoria descriptiva, las abreviaturas "GRN" y "GRtN" se usan indistintamente para referirse a las proteínas de fusión del VIH-1 del clado A que comprenden La secuencia de nucleótidos que codifica a GRN se inserta en un vector adecuado para permitir la expresión de la proteína de fusión GRN. El vector es un vector de adenovirus seleccionado del grupo que está constituido por Ad5, Ad35, Ad11, C6 y C7.
- Además, dentro de la divulgación se encuentra una sola secuencia de nucleótidos que codifica una proteína de fusión que comprende los antígenos Gag, Pol (que incluye RT e INT) y Nef. Como se usa en la presente memoria descriptiva, las abreviaturas "GRIN" y "GRtIN" se usan indistintamente para referirse a las proteínas de fusión del VIH-1 del clado A que comprenden los antígenos Gag, Pol y Nef y para referirse a las secuencias de nucleótidos que codifican esas proteínas de fusión. GRIN tiene la secuencia de aminoácidos que se ilustra en las figuras 16A - 16J y está codificada por la secuencia de nucleótidos que se ilustra en las figuras 16A - 16J. La secuencia de nucleótidos que codifica a GRIN se inserta en un vector adecuado para permitir la expresión de la proteína de fusión GRIN. El vector es un vector de adenovirus, por ejemplo, un vector de adenovirus seleccionado del grupo que está constituido por Ad5, Ad35, Ad11, C6 y C7.
- Dentro de la divulgación, una sola secuencia de nucleótidos de la invención codifica un antígeno Env del VIH-1 del clado A de acuerdo con la divulgación. El antígeno Env tiene la secuencia de aminoácidos que se ilustra en las Figuras 17A - 17D y es codificado por la secuencia de nucleótidos que se ilustra en las Figuras 17A - 17D. La secuencia de nucleótidos que codifica a Env se inserta en un vector adecuado para permitir la expresión de la proteína. Preferentemente el vector es un vector de adenovirus, más preferentemente un vector de adenovirus seleccionado del grupo que está constituido por Ad5, Ad35, Ad11, C6 y C7.
- La divulgación proporciona procedimientos para generar una respuesta inmunitaria contra antígenos del VIH-1 del clado A que comprenden administrar a un sujeto una secuencia de nucleótidos o un antígeno de acuerdo con la solicitud. El procedimiento para generar una respuesta inmunitaria contra el VIH-1 del clado A comprende administrar una secuencia de nucleótidos que codifica a GRIN o a GRN donde la secuencia de nucleótidos está contenida en un vector de adenovirus seleccionado del grupo que está constituido por Ad5, Ad35, Ad11, C6 y C7. Los vectores que comprenden a GRIN o GRN se administran conjuntamente con un vector que comprende una secuencia de nucleótidos que codifica un antígeno Env.
- Además, la presente divulgación proporciona composiciones inmunogénicas o composiciones de vacuna que comprenden las secuencias de nucleótidos de la divulgación.
- La presente invención proporciona una composición inmunogénica o de vacuna que comprende un polinucleótido que codifica un antígeno Env del VIH-1 del clado A como se define en la reivindicación 1, así como la composición inmunogénica o de vacuna de la reivindicación 1 para usar en un procedimiento para generar una respuesta

inmunitaria contra el VIH-1.

Cabe señalar que en esta divulgación y particularmente en las reivindicaciones o los párrafos, términos tal como "comprende", "comprendido", "que comprende" y análogos pueden tener el significado que se les atribuye en la ley de patentes de los Estados Unidos; por ejemplo, pueden significar "incluye", "incluido", "que incluye" y análogos; y que términos como "que está esencialmente constituido por" y "está constituido esencialmente por" tienen el significado que se les atribuye en la ley de patentes de los Estados Unidos, por ejemplo, dejan un margen para elementos no explícitamente enumerados, pero excluye elementos que se encuentran en la técnica anterior o que afectan una característica básica o nueva de la invención.

Estas y otras realizaciones se desvelan o son obvias a partir de, y abarcadas por, la descripción detallada siguiente.

Breve descripción de los dibujos

La descripción detallada siguiente, que se proporciona a modo de ejemplo, pero no pretende limitar la invención a las realizaciones específicas descritas, se puede comprender mejor conjuntamente con las figuras adjuntas.

La figura 1 es una secuencia consenso de aminoácidos de la proteína Gag del VIH-1 del clado A.

La figura 2 es un gráfico que ilustra la "distancia" de las secuencias de la proteína Gag de cepas circulantes del VIH del clado A respecto a la de la secuencia consenso de la proteína Gag del VIH-1 del clado A.

La figura 3 es una secuencia consenso de aminoácidos de la proteína Pol del VIH-1 del clado A.

La figura 4 es un gráfico que ilustra la "distancia" de las secuencias de la proteína Pol de cepas circulantes del VIH del clado A respecto a la de la secuencia consenso de la proteína Pol del VIH-1 del clado A.

La figura 5 es una secuencia consenso de aminoácidos de la proteína Env del VIH-1 del clado A.

La figura 6 es un gráfico que ilustra la "distancia" de las secuencias de la proteína Env de cepas circulantes del VIH del clado A respecto a la de la secuencia consenso de la proteína Env del VIH-1 del clado A.

La figura 7 es una secuencia consenso de aminoácidos de la proteína Nef del VIH-1 del clado A.

La figura 8 es un gráfico que ilustra la "distancia" de las secuencias de la proteína Nef de cepas circulantes del VIH del clado A respecto a la de la secuencia consenso de la proteína Nef del VIH-1 del clado A.

La figura 9 es una representación esquemática de los transgenes GRIN y GRN.

La figura 10 ilustra la secuencia de aminoácidos de la proteína Gag de la cepa TZA173 del VIH-1 del clado A que tiene el número de registro de Genbank AY253305.

La figura 11 ilustra la secuencia de aminoácidos de la proteína Pol de la cepa MSA4070 del VIH-1 del clado A que tiene el número de registro de Genbank AF457081.

La figura 12 ilustra la secuencia de aminoácidos de la proteína Nef de la cepa MSA4070 del VIH-1 del clado A que tiene el número de registro de Genbank AF457081.

La figura 13 ilustra la secuencia de aminoácidos de la proteína Env de la cepa TZA341 del VIH-1 del clado A que tiene el número de registro de Genbank AY253314.

Las figuras 14A - 14 C proporcionan una secuencia de GRIN como la insertada en el vector Ad35.

Las figuras 15A - 15B proporcionan una secuencia de Env como la insertada en el vector Ad35.

Las figuras 16A - 16J proporcionan secuencias de nucleótidos y aminoácidos con del transgén GRIN con codones optimizados.

Las figuras 17A - 17 D proporcionan secuencias de nucleótidos y aminoácidos del transgén Env con codones optimizados.

La figura 18 ilustra gráficamente la inmunogenia de Ad5-GRIN y Ad5-GRN en ratones, medida por el ensayo de IFN-gamma ELIspot.

La figura 19 ilustra gráficamente la inmunogenia de C7-GRIN y C7-GRN en ratones, medida por el ensayo de IFN-gamma ELIspot.

La figura 20 ilustra gráficamente la inmunogenia de C7-GRIN y C7-GRN en ratones, medida por el ensayo de IL-2 ELISpot.

La figura 21 ilustra gráficamente la inmunogenia de C6-GRIN y C6-GRN en ratones, medida por el ensayo de IFN-gamma ELISpot.

5 La figura 22 ilustra gráficamente la inmunogenia de C6-GRIN y C6-GRN en ratones, medida por el ensayo de IL-2 ELISpot.

10 La figura 23A ilustra la inmunogenia por IFN- γ ELISpot de Ad35-GRIN/ENV a la dosis de 10^{10} vp después de un programa de vacunación del mes 0 al 6 en *rhesus macaques*. Definición de respuesta positiva: para una mezcla de un solo péptido de una sola muestra: respuesta = (media del recuento de péptido - media del recuento de no péptido). Para ser positiva, una respuesta de un sólo péptido debe ser: 1. Media del recuento de péptido $>4x$ media del recuento de no péptido de la misma placa; 2. Coeficiente de variación entre recuentos repetidos $\leq 70\%$ & 3. Respuesta > 55 SFC/106. La media geométrica de las respuestas para células formadoras de manchas (SFC) por millón de PBMC para cada componente antigénico (Gag, RT, IN y ENV) aparece en el eje Y y los tiempos de sangrado en semanas en el eje X.

15 La figura 23B ilustra la inmunogenia por IFN- γ ELISpot de Ad35-GRIN/ENV a la dosis de 10^{11} vp después de un programa de vacunación del mes 0 al 6 en *rhesus macaques*. Definición de respuesta positiva: para una mezcla de un solo péptido de una sola muestra: respuesta = (media del recuento de péptido - media del recuento de no péptido). Para ser positiva, una respuesta de un sólo péptido debe ser: 1. Media del recuento de péptido $> 4x$ media del recuento de no péptido de la misma placa; 2. Coeficiente de variación entre recuentos repetidos $\leq 70\%$ 3. Respuesta > 55 SFC/106. La media geométrica de las respuestas para células formadoras de manchas (SFC) por millón de PBMC para cada componente antigénico (Gag, RT, IN y ENV) aparece en el eje Y y los tiempos de sangrado en semanas en el eje X.

Descripción detallada de la invención

25 La presente divulgación se refiere a secuencias consenso de nucleótidos y proteínas para antígenos del VIH-1 del clado A y a aislados de campo circulantes del VIH-1 que se corresponden estrechamente con esas secuencias consenso. La divulgación también se refiere a la versión alterada de esas secuencias, que puede ser alterada de modo que se anule la función de los productos génicos in vivo, a constructos y vectores que comprenden las secuencias de la invención, y a inmunógenos, composiciones inmunogénicas y vacunas preparadas usando las secuencias de la invención. La divulgación también se refiere a procedimientos para generar una respuesta inmunitaria contra antígenos del VIH-1 del clado A en un sujeto y a procedimientos para inducir inmunidad protectora contra la provocación con el VIH-1. Las diversas realizaciones de la invención se resumen precedentemente en la sección titulada "Sumario de la invención". Se proporcionan más detalles de la invención en la Descripción detallada y los Ejemplos que siguen y también en las Figuras.

30 Como se describe en el "Sumario de la invención" anterior y los "Ejemplos" siguientes, la presente divulgación proporciona antígenos consenso del VIH-1 del clado A y también antígenos de cepas circulantes del VIH-1 del clado A que están estrechamente relacionadas con esas secuencias consenso. La divulgación también proporciona transgenes y antígenos del VIH-1 codificados por esos transgenes. Esos transgenes comprenden secuencias que codifican los antígenos del VIH-1 del clado A de la divulgación, por ejemplo antígenos Gag, Pol, Env, Nef, RT, e INT de la divulgación. Por ejemplo, la presente solicitud proporciona un transgén GRIN (también denominado GRtIN) que comprende los antígenos Gag, Pol (ambos RT e Int) y Nef de la divulgación. La divulgación también proporciona un transgén GRN (también denominado GRtN) que comprende los antígenos Gag, RT y Nef de la divulgación. En otra realización, la presente divulgación proporciona un transgén Env que comprende un antígeno Env de la divulgación.

35 La presente invención proporciona una composición inmunogénica o de vacuna que comprende un polinucleótido que codifica un antígeno Env del VIH-1 del clado A, en el que el antígeno Env tiene la secuencia de aminoácidos ilustrada en la figura 17 y está codificado por la secuencia de nucleótidos ilustrada en la figura 17. Además, la presente invención se refiere a esta composición inmunogénica o de vacuna para usar en un procedimiento para generar una respuesta inmunitaria contra el VIH-1.

40 Los términos "proteína", "péptido", "polipéptido" y "secuencia de aminoácidos" se usan indistintamente en la presente memoria descriptiva para referirse a polímeros de residuos de aminoácidos de cualquier longitud. El polímero puede ser lineal o ramificado, puede comprender aminoácidos modificados o análogos de aminoácidos, y puede estar interrumpido por restos químicos funcionales distintos de los aminoácidos. Los términos también abarcan un polímero de aminoácidos que ha sido modificado de forma natural o por intervención; por ejemplo por formación de enlace disulfuro, glucosilación, lipidación, acetilación, fosforilación, o cualquier otra manipulación o modificación, como la conjugación con un componente de marcado o bioactivo.

Como se usa en la presente memoria descriptiva, los términos "antígeno" o "inmunógeno" se usan indistintamente para referirse a una sustancia, normalmente una proteína, que es capaz de inducir una respuesta inmunitaria en un sujeto. El término también se refiere a las proteínas que son inmunológicamente activas en el sentido de que una vez administradas a un sujeto (ya sea directamente o mediante la administración al sujeto de una secuencia de nucleótidos o un vector que codifiquen la proteína) es capaz de provocar una respuesta inmunitaria de tipo humoral y/o celular dirigida contra esa proteína.

Se debe entender que las proteínas y los antígenos de la divulgación pueden diferir de las secuencias exactas ilustradas y descritas en la presente memoria descriptiva. Por consiguiente, la divulgación contempla deleciones, adiciones y sustituciones a las secuencias mostradas, en tanto que las secuencias funcionen de conformidad con los procedimientos de la divulgación. En este respecto, las sustituciones particularmente preferidas serán generalmente de naturaleza conservadora, es decir, aquellas sustituciones que tienen lugar dentro de una familia de aminoácidos. Por ejemplo, los aminoácidos se dividen generalmente en cuatro familias: (1) ácidos — aspartato y glutamato; (2) básicos — lisina, arginina, histidina; (3) no polares — alanina, valina, leucina, isoleucina, prolina, fenilalanina, metionina, triptófano; y (4) sin carga polar — glicina, asparragina, glutamina, cistina, serina, treonina, tirosina. Fenilalanina, triptófano y tirosina a veces se clasifican como aminoácidos aromáticos. Es razonablemente previsible que un reemplazo aislado de leucina, con isoleucina o valina, o viceversa; un aspartato con un glutamato o viceversa; una treonina con una serina o viceversa; o un reemplazo conservador similar de un aminoácido con un aminoácido estructuralmente relacionado, no tendrá un efecto importante en la actividad biológica. Las proteínas que tienen sustancialmente la misma secuencia de aminoácidos que las secuencias ilustradas y descritas pero poseen sustituciones de aminoácidos menores que no afectan sustancialmente la inmunogenia de la proteína están comprendidas, por consiguiente, dentro del alcance de la divulgación.

La solicitud describe secuencias "consenso" de aminoácidos para antígenos del VIH-1 del clado A. Se refiere a secuencias consenso de aminoácidos para los antígenos Gag, Pol (que comprende RT e Int), Nef y Env del VIH-1 del clado A. Se refiere a una secuencia consenso de aminoácidos de Gag de la FIG. 1, la secuencia consenso de aminoácidos de Pol de la FIG. 3, a una secuencia consenso de aminoácidos de Env de la FIG. 5, y/o a una secuencia consenso de aminoácidos de Nef de la figura 7. La divulgación está dirigida a un procedimiento de identificación de una secuencia consenso de aminoácidos para un antígeno del VIH-1 del clado A de interés, que comprende obtener la secuencia de aminoácidos del antígeno de interés en varias cepas o aislados de campo circulantes del VIH-1, alinear dichas secuencias y determinar la secuencia consenso para ese antígeno. Por ejemplo, en una realización se genera una base de datos usando secuencias disponibles para cepas circulantes no recombinantes del VIH-1 del clado A, y a continuación los genes individuales del VIH-1 (por ejemplo gag, pol, nef y env) de todas las secuencias de la base de datos se alinean, con guiones insertados para mantener el alineamiento en regiones con inserciones o deleciones en la secuencia, y luego se puede derivar una secuencia consenso del 50%.

La presente divulgación también se refiere a procedimientos para identificar antígenos de cepas del VIH-1 del clado A de origen natural que tienen una secuencia de aminoácidos con una "distancia de proteínas" pequeña respecto a la secuencia consenso de aminoácidos de ese antígeno. La "distancia de proteínas" es una medida del nivel de similitud o diferencia entre dos secuencias de aminoácidos. Dos secuencias de aminoácidos que son muy similares tienen una distancia de proteínas baja. Dos secuencias de aminoácidos que son muy diferentes tienen una distancia de proteínas alta. Las distancias de proteínas se calculan preferentemente usando la matriz de sustitución de DayhoffPAM250 (M.O. Dayhoff, ed., 1978, Atlas of Protein Sequence and Structure, Vol. 5) que pondera las sustituciones según el grado de similitud bioquímica. Sin embargo, también se pueden usar otros procedimientos para determinar la distancia de proteínas.

Como se usa en la presente memoria descriptiva, los términos "secuencias de nucleótidos" y "secuencias de ácido nucleico" se refieren a ácido desoxirribonucleico (ADN) o secuencias de ácido ribonucleico (ARN), incluyendo, pero no exclusivamente, ARN mensajero (ARNm), híbridos de ADN/ARN o ácidos nucleicos sintéticos. El ácido nucleico puede ser de una sola cadena, o parcial o completamente de doble cadena (dúplex). Los ácidos nucleicos dúplex pueden ser homodúplex o heterodúplex.

Como se describe en el "Sumario de la invención" anterior y en los "Ejemplos" dados más adelante, la presente divulgación proporciona antígenos consenso del VIH-1 del clado A y las secuencias de nucleótidos que codifican esos antígenos consenso. La divulgación también se refiere a antígenos de cepas circulantes del VIH-1 del clado A que están estrechamente relacionados con esas secuencias consenso, y a las secuencias de nucleótidos que los codifican. La divulgación también proporciona transgenes del VIH-1 del clado A que comprende secuencias que codifican antígenos del VIH-1 del clado A. Según se usa en la presente memoria descriptiva, el término "transgén" se usa para referirse a secuencias de nucleótidos "recombinantes" que derivan de secuencias consenso de nucleótidos del VIH-1 del clado A de la divulgación, o de secuencias de nucleótidos que codifican los antígenos de cepas recientemente circulantes del VIH-1 del clado A que se ha identificado que se corresponden estrechamente con esas secuencias consenso. El término "recombinante" significa una secuencia de nucleótidos que ha sido

manipulada "por el hombre" y que no se encuentra en la naturaleza, o que está ligada a otra secuencia de nucleótidos o se encuentra en una distribución diferente en la naturaleza. Se entiende que manipulada "por el hombre" significa manipulada por algún medio artificial, que incluye el uso de máquinas, la optimización de codones, enzimas de restricción, etc. Por ejemplo, se proporcionan los transgenes GRIN, GRN y Env.

5 Los nucleótidos de la divulgación pueden estar alterados en comparación con las secuencias consenso de nucleótidos, o en comparación con las secuencias de aislados circulantes del VIH-1 que están estrechamente relacionadas con dichas secuencias consenso. Por ejemplo, las secuencias de nucleótidos se pueden mutar de tal forma que anula la actividad de las proteínas codificadas in vivo. Las secuencias de nucleótidos pueden tener los codones optimizados, por ejemplo, los codones se pueden optimizar para uso humano. Por ejemplo, las
10 secuencias de nucleótidos se mutan para anular la función normal in vivo de las proteínas codificadas, y se les optimizan los codones para uso humano. Por ejemplo, cada una de las secuencias de Gag, Pol, Env, Nef, RT e Int de la divulgación se puede modificar de esas maneras.

Los tipos de mutaciones que se pueden realizar para anular la función in vivo de los antígenos incluyen, pero no exclusivamente, los siguientes que también se describen en el ejemplo 7: mutación de Gly2 a Ala en Gag para eliminar un sitio de miristilación y evitar la formación de partículas semejantes a virus (VLP); mutación de Gag para evitar deslizamiento en la secuencia de desplazamiento del marco de lectura natural para dejar la secuencia de aminoácidos conservada (NFLG) intacta y permitir que se traduzca sólo el producto de la proteína GagPol de longitud completa; mutación de Asp185 de RT a Ala y mutación de Asp186 a Ala para inactivar los residuos enzimáticos activos. Mutación de Asp 64 de Int a Ala, y mutación de Asp116 a Ala y mutación de Glu 152 a Ala para inactivar los residuos enzimáticos activos.
15
20

En lo que se refiere a la optimización de codones, las moléculas de ácido nucleico de la divulgación tienen una secuencia de nucleótidos que codifica los antígenos de la divulgación y que puede ser diseñada para emplear codones que se usan en los genes del sujeto en el cual se va a producir el antígeno. Muchos virus, incluidos el VIH y otros lentivirus, usan un gran número de codones raros y, alterando esos codones para que correspondan a codones comúnmente usados en el sujeto deseado, se puede lograr una mayor expresión de los antígenos. Por ejemplo, los codones usados son codones "humanizados", es decir, los codones son los que aparecen frecuentemente en genes humanos altamente expresados (Andre et al., J. Virol. 72: 1497 - 1503, 1998) en lugar de los codones que son frecuentemente usados por el VIH. Dicho uso de codones favorece la expresión eficaz de las proteínas del VIH transgénicas en células humanas. Se puede usar cualquier procedimiento de optimización de codones adecuado. Por ejemplo los codones se pueden optimizar para uso humano según se ilustra en el ejemplo 8. No obstante, se puede usar cualquier otro procedimiento adecuado de optimización de codones. Dichos procedimientos y la selección de dichos procedimientos son bien conocidos por el experto en la técnica. Además existen varias compañías que optimizarán codones de secuencias como Geneart (geneart.com). Por lo tanto, las secuencias de nucleótidos se pueden optimizar fácilmente.
25
30

La solicitud abarca además secuencias de nucleótidos que codifican variantes y derivados funcional y/o antigénicamente equivalentes de los antígenos de la divulgación y fragmentos funcionalmente equivalentes de las mismas. Esas variantes, derivados y fragmentos funcionalmente equivalentes, tienen la capacidad de conservar la actividad antigénica. Por ejemplo, los cambios en una secuencia de ADN que no cambian la secuencia de aminoácidos codificada, así como aquellos que dan como resultado sustituciones de residuos de aminoácidos conservadoras, una o unas pocas deleciones o adiciones de aminoácidos y sustitución de residuos de aminoácidos por análogos de aminoácidos son las que no afectarán significativamente las propiedades del polipéptido codificado. Las sustituciones de aminoácidos conservadoras son glicina/alanina; valina/isoleucina/leucina; asparragina/glutamina; ácido aspártico/ácido glutámico; serina/treonina/metionina; lisina/arginina; y fenilalanina/tirosina/triptofano. Las variantes tienen al menos el 50%, al menos el 55%, al menos el 60%, al menos el 65%, al menos el 70%, al menos el 75%, al menos el 80%, al menos el 85%, al menos el 86%, al menos el 87%, al menos el 88%, al menos el 89%, al menos el 90%, al menos el 91%, al menos el 92%, al menos el 93%, al menos el 94%, al menos el 95%, al menos el 96%, al menos el 97 %, al menos el 98% o al menos el 99% de homología o identidad con el antígeno, epítipo, inmunógeno, péptido o polipéptido de interés.
35
40
45

La identidad de secuencia u homología se determina comparando las secuencias cuando se alinean de tal forma que se maximiza la superposición y la identidad minimizando al mismo tiempo los huecos de la secuencia. En particular, la identidad de secuencia se puede determinar usando cualquiera de una serie de algoritmos matemáticos. Un ejemplo no limitante de un algoritmo matemático usado para comparar dos secuencias, es el algoritmo de Karlin & Altschul, Proc. Natl. Acad. Sci. USA 1993; 90: 5873 - 5877.
50

Otro ejemplo de un algoritmo matemático usado para comparar secuencias es el algoritmo de Myers & Miller, CABIOS 1988; 4: 11 - 17. Dicho algoritmo se incorpora en el programa ALIGN (versión 2.0) que forma parte del paquete de software de alineación de secuencia GCG. Cuando se usa el programa ALIGN para comparar secuencias de aminoácidos, se pueden usar una tabla de ponderación de residuos PAM120, una penalización de
55

longitud de hueco de 12 y una penalización de hueco de 4. Aún otro algoritmo útil para identificar regiones de similitud de secuencia local y alineamiento es el algoritmo FASTA como se describe en Pearson & Lipman, Proc. Natl. Acad. Sci. USA 1988; 85: 2444 - 2448.

5 Resulta ventajoso el uso del software WU-BLAST (Washington University BLAST) versión 2.0. Los programas ejecutables WU-BLAST versión 2.0 para varias plataformas UNIX se pueden descargar de ftp://blast.wustl.edu/blast/executables. Este programa se basa en WU-BLAST versión 1.4, que a su vez se basa en el dominio público NCBI-BLAST versión 1.4 (Altschul y Gish, 1996, Local alignment statistics, Doolittle ed., Methods in Enzymology 266: 460 - 480; Altschul et al., Journal of Molecular Biology 1990; 215: 403 - 410; Gish y States, 1993; Nature Genetics 3: 266 - 272; Karlin y Altschul, 1993; Proc. Natl. Acad. Sci. EE.UU 90: 5873 - 5877).

10 Las diversas secuencias de nucleótidos y transgenes recombinantes de la divulgación se realizan usando técnicas de recombinación de ADN y de clonación convencionales. Dichas técnicas son bien conocidas por los expertos en la técnica. Véase por ejemplo, "Molecular Cloning: A Laboratory Manual", segunda edición (Sambrook et al., 1989).

15 Las secuencias de nucleótidos de la presente invención se pueden insertar en "vectores". El término "vector" es ampliamente usado y entendido por los expertos en la técnica, y, como se usa en la presente memoria descriptiva, el término "vector" se usa coherentemente con su significado para los expertos en la técnica. Por ejemplo, el término "vector" es usado comúnmente por los expertos en la técnica para referirse a un vehículo que permite o facilita la transferencia de moléculas de ácido nucleico de un ambiente a otro, o que permite o facilita la manipulación de una molécula de ácido nucleico.

20 Se puede usar cualquier vector que permita la expresión de los transgenes del VIH-1 del clado A de conformidad con la presente divulgación. En ciertas realizaciones, se pueden usar transgenes del VIH-1 del clado A de la presente divulgación in vitro (como el uso de sistemas de expresión exentos de células) y/o en cultivos celulares in vitro para producir los antígenos del VIH-1 codificados que se pueden usar para diversas aplicaciones, como en la producción de vacunas proteínicas. Para este tipo de aplicaciones, se puede usar cualquier vector que permita la expresión de los transgenes del VIH-1 del clado A in vitro o en células en cultivo.

25 Para las aplicaciones en las que se desea que los transgenes se expresen in vivo, por ejemplo cuando se usan transgenes en vacunas de ADN o que contienen ADN, se puede usar cualquier vector que permita la expresión de los transgenes del VIH-1 del clado A de la presente divulgación y sea seguro para usar in vivo. Los vectores usados son seguros para usar en seres humanos, mamíferos y animales de laboratorio.

30 Para que los transgenes de la presente solicitud se expresen, la secuencia de codificación de la proteína debe estar "operativamente unida" a las secuencias reguladoras o de control de ácido nucleico que dirigen la transcripción y traducción de la proteína. Como se usa en la presente memoria descriptiva, se dice que una secuencia de codificación y una secuencia de ácido nucleico de control o promotora están "operativamente unidas" cuando están unidas covalentemente de tal forma que colocan la expresión o la transcripción y/o la traducción de la secuencia de codificación bajo la influencia o el control de la secuencia de control de ácido nucleico. La "secuencia de control de ácido nucleico" puede ser cualquier elemento de ácido nucleico, como, pero no exclusivamente, promotores, potenciadores, IRES, intrones y otros elementos descritos que dirigen la expresión de una secuencia de ácido nucleico o secuencia de codificación que está unida operativamente a éstos. En la presente memoria descriptiva se usará el término "promotor" para referirse a un grupo de módulos de control transcripcional que se agrupan alrededor del sitio de iniciación de la ARN polimerasa II y que cuando están operativamente unidos a las secuencias de codificación de las proteínas de la divulgación dan lugar a la expresión de la proteína codificada. La expresión de los transgenes de la presente divulgación puede estar bajo el control de un promotor constitutivo o de un promotor inducible, que inicia la transcripción sólo cuando se expone a algún estímulo externo particular, como, sin limitación, antibióticos como tetraciclina, hormonas como ecdisona, o metales pesados. El promotor también puede ser específico de un determinado tipo de célula, tejido u órgano. Se conocen muchos promotores y potenciadores adecuados en el área, y se puede usar cualquier promotor o potenciador adecuado para la expresión de los transgenes de la divulgación. Por ejemplo, los promotores y/o potenciadores adecuados se pueden seleccionar de la base de datos Eukaryotic Promoter Database (EPDB).

50 Normalmente los vectores de la presente divulgación se deberían elegir de forma típica de modo que contengan una región reguladora del gen adecuada, tal como un promotor o un potenciador, para que los transgenes se puedan expresar.

55 Por ejemplo, cuando el objetivo es expresar los transgenes de la divulgación in vitro, o en cultivos de células, o en cualquier sistema procarionta o eucarionta, con el fin de producir la(s) proteína(s) codificada(s) por ese transgén, entonces se puede usar cualquier vector adecuado dependiendo de la aplicación. Por ejemplo, se pueden usar plásmidos, vectores virales, vectores bacterianos, vectores de protozoarios, vectores de insectos, vectores de expresión de baculovirus, vectores de levadura, vectores de células de mamífero y similares. Los vectores

adecuados pueden ser seleccionados por los expertos en la técnica teniendo en cuenta las características del vector y los requisitos para la expresión de los transgenes en las circunstancias identificadas.

5 Cuando el objetivo es expresar los transgenes de la divulgación in vivo en un sujeto, por ejemplo, para generar una respuesta inmunitaria contra un antígeno del VIH-1 y/o inmunidad protectora contra el VIH-1, se deben elegir los
 10 vectores de expresión que son adecuados para la expresión en el sujeto, y que son seguros para usar in vivo. Por ejemplo, en algunas realizaciones puede ser deseable expresar los transgenes de la divulgación en un animal de laboratorio, tal como para las pruebas preclínicas de las composiciones inmunogénicas del VIH-1 y las vacunas de la invención. En otras realizaciones, será deseable expresar los transgenes de la divulgación en sujetos humanos, tal como en los ensayos clínicos y para el uso clínico real de las composiciones inmunogénicas y las vacunas de la invención. Puede emplearse cualquier vector adecuado para este tipo de usos, y está dentro de las capacidades de los expertos en la técnica seleccionar un vector adecuado. En algunas realizaciones puede ser preferible que a los vectores usados para esas aplicaciones in vivo se les atenúe la amplificación en el sujeto. Por ejemplo, si se usan vectores plasmídicos, preferentemente carecerán de un origen de replicación que funcione en el sujeto con el fin de mejorar la seguridad para el uso in vivo en el sujeto. Si se usan vectores virales, preferentemente se atenúan o no se replican defectuosamente en el sujeto, una vez más, con el fin de mejorar la seguridad para el uso in vivo en el
 15 sujeto.

20 Por ejemplo, se usan vectores virales. Los vectores de expresión viral son bien conocidos por los expertos en la técnica y son, por ejemplo, virus tales como los adenovirus, virus adeno-asociados (AAV), alfavirus, herpesvirus, retrovirus y poxvirus, incluidos virus avipox, poxvirus atenuados, virus vaccinia y particularmente, el virus vaccinia Ankara modificado (MVA; N° de registro ATCC VR-1566). Dichos virus, cuando se usan como vectores de expresión son innatamente no patógenos en los sujetos seleccionados tales como los seres humanos o fueron modificados para tornarlos no patógenos en los sujetos seleccionados. Por ejemplo, son bien conocidos los alfavirus y adenovirus que tiene replicación defectuosa y se pueden usar como vectores para la administración de genes.

25 En realizaciones particularmente preferidas se usan vectores de adenovirus. Se conocen en la técnica muchos vectores de adenovirus y se puede usar cualquiera de dichos vectores adecuados. En realizaciones preferidas el vector de adenovirus usado se selecciona del grupo que está constituido por los vectores Ad5, Ad35, Ad11, C6 y C7.

30 Se ha publicado la secuencia del genoma del Adenovirus 5 ("Ad5"). (Chroboczek, J., Bieber, F., y Jacrot, B. (1992) The Sequence of the Genome of Adenovirus Type 5 and Its Comparison with the Genome of Adenovirus Type 2, Virology 186, 280 - 285). Los vectores Ad35 se describen en las patentes de los Estados Unidos N° 6.974.695, 6.913.922 y 6.869.794. Los vectores Ad11 se describen en la patente de los Estados Unidos N° 6.913.922. Los vectores del adenovirus C6 se describen en las patentes de los Estados Unidos N° 6.780.407; 6.537.594; 6.309.647; 6.265.189; 6.156.567; 6.090.393; 5.942.235 y 5.833.975. Los vectores C7 se describen en la patente de los Estados Unidos N° 6.277.558.

35 También se pueden usar vectores de adenovirus con E1 defectuoso o eliminado, E3 defectuoso o eliminado, o E4 defectuoso o eliminado. Ciertos adenovirus con mutaciones en la región E1 tienen un mayor margen de seguridad porque los mutantes de adenovirus con E1 defectuoso tienen una replicación defectuosa en células no permisivas, o, al menos, están muy atenuados. Los adenovirus con mutaciones en la región E3 pueden tener una mayor inmunogenia al interrumpir el mecanismo mediante el cual el adenovirus regula por disminución las moléculas del MHC clase I. Los adenovirus que tienen mutaciones E4 pueden tener el vector de adenovirus con menor inmunogenia debido a la supresión de la expresión génica tardía. Dichos vectores pueden ser particularmente útiles cuando se desea la revacunación repetida usando el mismo vector. Se pueden usar vectores de adenovirus con eliminación de, o mutación en, E1, E3, E4, E1 y E3 y E1 y E4.

45 Además, también se pueden emplear vectores de adenovirus "vacíos" (gutless), en los cuales se eliminan todos los genes virales. Dichos vectores requieren un virus cooperador para su replicación y requieren una línea celular 293 humana especial que exprese E1a y Cre, una condición que no existe en el ambiente natural. Dichos vectores "vacíos" son no inmunógenos y por lo tanto los vectores se pueden inocular varias veces para revacunación. Los vectores de adenovirus "vacíos" se pueden usar para la inserción de insertos/genes heterólogos como los
 50 transgenes e incluso se pueden usar para co-administración de un gran número de insertos/genes heterólogos.

La solicitud también abarca un diseño que pone a Env y GRIN en vectores separados para permitir la evaluación de si la inclusión de Env es beneficiosa o perjudicial en cuanto a la inmunidad mediada por células (CMI) y la eficacia protectora. Los beneficios o perjuicios de Env en CMI y la eficacia protectora sigue siendo una incógnita en el campo de la vacuna del VIH.

55 Las secuencias de nucleótidos y los vectores de la divulgación se pueden administrar a las células, por ejemplo si

el objetivo es expresar los antígenos del VIH-1 en las células para producir y aislar las proteínas expresadas, tal como de las células en cultivo. Para expresar los transgenes en las células se puede usar cualquier procedimiento adecuado de transfección, transformación o administración de genes. Dichos procedimientos son bien conocidos por los expertos en la técnica, y un experto en la técnica podrá seleccionar fácilmente un procedimiento adecuado según la naturaleza de las secuencias de nucleótidos, los vectores y los tipos de células usados. Por ejemplo, se podría usar transfección, transformación, microinyección, infección, electroporación, lipofección o administración mediada por liposomas. La expresión de los antígenos se puede realizar en cualquier tipo adecuado de célula huésped, como las células bacterianas, las levaduras, las células de insectos y las células de mamíferos. También se pueden expresar los antígenos del VIH-1 del clado A usando sistemas que incluyen transcripción y traducción in vitro. Todos esos procedimientos son bien conocidos por los expertos en la técnica, y un experto en la técnica podrá seleccionar fácilmente un procedimiento adecuado según la naturaleza de las secuencias de nucleótidos, los vectores y los tipos de células usados.

Después de la expresión, los antígenos de la divulgación se pueden aislar y/o purificar o concentrar mediante cualquier técnica adecuada conocida en la técnica. Por ejemplo, se puede usar cromatografía de intercambio aniónico o catiónico, cromatografía en fosfocelulosa, cromatografía de interacción hidrófoba, cromatografía por afinidad, cromatografía de inmun afinidad, cromatografía de hidroxapatita, cromatografía de lectina, cromatografía de tamiz molecular, isoelectroenfoque, electroforesis en gel, o cualquier otro procedimiento adecuado o combinación de procedimientos.

Por ejemplo, las secuencias de nucleótidos y/o los antígenos de la divulgación se administran in vivo, por ejemplo cuando el objetivo es producir una respuesta inmunogénica en un sujeto. Un "sujeto" en el contexto de la presente divulgación puede ser cualquier animal. Por ejemplo, en algunas realizaciones puede ser deseable expresar los transgenes de la divulgación en un animal de laboratorio, tal como para las pruebas preclínicas de las composiciones inmunogénicas del VIH-1 y las vacunas de la divulgación. En otras realizaciones, será deseable expresar los transgenes de la divulgación en sujetos humanos, tal como en los ensayos clínicos y para el uso clínico real de las composiciones inmunogénicas y las vacunas de la invención. De forma preferible, el sujeto es un ser humano, por ejemplo un ser humano que está infectado, o corre riesgo de infección, con el VIH-1.

Para dichas aplicaciones in vivo las secuencias de nucleótidos y/o los antígenos de la divulgación se administran preferentemente en forma de un componente de una composición inmunogénica que comprende las secuencias de nucleótidos y/o los antígenos de la divulgación en mezcla con un vehículo farmacéuticamente aceptable. Las composiciones inmunogénicas de la invención son útiles para estimular una respuesta inmunitaria contra el VIH-1 y se pueden usar como uno o más componentes de una vacuna profiláctica o terapéutica contra el VIH-1 para la prevención, la mejora o el tratamiento del SIDA. Los ácidos nucleicos y vectores de la divulgación son particularmente útiles para proporcionar vacunas genéticas, es decir, vacunas para la administración de ácidos nucleicos que codifican los antígenos del VIH-1 del clado A a un sujeto tal como un ser humano, de modo que los antígenos del VIH-1 del clado A se expresen a continuación en el sujeto para provocar una respuesta inmunitaria.

Las composiciones de la invención pueden ser suspensiones inyectables, soluciones, aerosoles, polvos liofilizados, siropes, elixires y similares. Se puede usar cualquier forma adecuada de la composición. Para preparar una composición de ese tipo, se mezcla un ácido nucleico que tenga el grado de pureza deseado, con uno o más vehículos y/o excipientes farmacéuticamente aceptables. Los vehículos y los excipientes deben ser "aceptables" en el sentido de ser compatibles con los demás ingredientes de la composición. Los vehículos, excipientes o estabilizantes aceptables son no tóxicos para los destinatarios en las dosis y las concentraciones empleadas e incluyen, pero no exclusivamente, agua, solución salina, solución salina tamponada con fosfato, dextrosa, glicerol, etanol, o combinaciones de los mismos, tampones, tales como fosfato, citrato y otros ácidos orgánicos; antioxidantes que incluyen ácido ascórbico y metionina; conservantes (tal como cloruro de octadecildimetilbencil amonio; cloruro de hexametonio; cloruro de benzalconio, cloruro de bencetonio; fenol, alcohol butílico o bencílico; alquilparabenos como metil o propilparabeno; catecol; resorcinol; ciclohexanol; 3-pentanol; y m-cresol); polipéptidos de bajo peso molecular (menos de aproximadamente 10 residuos); proteínas tales como albúmina de suero, gelatina o inmunoglobulinas; polímeros hidrófilos tales como polivinilpirrolidona; aminoácidos tales como glicina, glutamina, asparragina, histidina, arginina o lisina; monosacáridos, disacáridos y otros carbohidratos que incluyen glucosa, manosa o dextrinas; agentes quelantes como EDTA; azúcares tales como sacarosa, manitol, trehalosa o sorbitol; contraiones formadores de sales tales como sodio; complejos metálicos (por ejemplo, complejos de Zn-proteína); y/o tensioactivos no iónicos tales como TWEEN™, PLURONICS™ o polietilenglicol (PEG).

Una composición inmunogénica o inmunológica también se puede formular en forma de una emulsión de aceite en agua. La emulsión de aceite en agua se puede basar, por ejemplo, en aceite de parafina líquida ligera (tipo Farmacopea Europea); aceite de isoprenoides como escualano, escualeno, EICOSANE™ o tetratetracontano; aceite resultante de la oligomerización de alqueno(s), p. ej., isobuteno o deceno; ésteres de ácidos o de alcoholes que contienen un grupo alquilo lineal, tales como aceites vegetales, oleato de etilo, di(caprilato/caprato)de

propilenglicol, tri(caprilato/caprato) de glicerilo o dioleato de propilenglicol; ésteres de ácidos grasos o alcoholes ramificados, por ejemplo, ésteres de ácido isosteárico. El aceite se usa ventajosamente en combinación con emulsionantes para formar la emulsión. Los emulsionantes pueden ser tensioactivos no iónicos, como ésteres de sorbitán, mannide (p. ej., oleato de anhidromanitol), glicerol, poliglicerol, propilenglicol y ácido oleico, isosteárico, ricinoleico o hidroxisteárico, opcionalmente etoxilados, y copolímeros de bloque de polioxipropileno-polioxi-etileno, como los productos Pluronic®, por ejemplo, L121. El adyuvante puede ser una mezcla de emulsionante(s), agente de formación de micela y aceite como el que se comercializa con el nombre de Provac® (IDEC Pharmaceuticals, San Diego, CA).

Las composiciones inmunogénicas de la invención pueden contener sustancias adicionales, como humectantes o emulsionantes, soluciones tamponadores o adyuvantes para mejorar la eficacia de las vacunas (Remington's Pharmaceutical Sciences, 18ª edición, Mack Publishing Company, (ed.) 1980).

También se pueden incluir adyuvantes. Los adyuvantes incluyen, pero no exclusivamente, sales minerales (por ej., $\text{Al}(\text{SO}_4)_2$, $\text{AlNa}(\text{SO}_4)_2$, $\text{AlNH}(\text{SO}_4)_2$, sílice, alum, $\text{Al}(\text{OH})_3$, $\text{Ca}_3(\text{PO}_4)_2$, caolín, o carbón), polinucleótidos con o sin complejos inmunoestimulantes (ISCOMs) (por ej., oligonucleótidos CpG, tal como los descritos en Chuang, T.H. et al, (2002) J. Leuk. Biol. 71(3): 538 - 44; Ahmad-Nejad, P. et al (2002) Eur. J. Immunol. 32(7): 1958 - 68; ácidos poli IC o poli AU, poliarginina con o sin CpG (también conocida en la técnica como IC31; véase Schellack, C. et al (2003) Proceedings of the 34th Annual Meeting of the German Society of Immunology; Lingnau, K. et al (2002) Vaccine 20 (29 - 30): 3498 - 508), JuvaVax™ (patente de los EE.UU. N° 6.693.086), ciertas sustancias naturales (por ej., cera D de *Mycobacterium tuberculosis*, sustancias encontradas en *Corynebacterium parvum*, *Bordetella pertussis*, o componentes del género *Brucella*), flagelina (Toll-like receptor 5 ligand; véase McSorley, S.J. et al (2002) J. Immunol. 169(7): 3914 - 9), saponinas tales como QS21, QS17 y QS7 (patentes de los EE.UU. N° 5.057.540; 5.650.398; 6.524.584; 6.645.495), monofosforil lípido A, en particular, 3-de-O-acilado monofosforil lípido A (3D-MPL), imiquimod (también conocido en la técnica como IQM disponible en el mercado como Aldara®; patentes de los EE.UU. N° 4.689.338; 5.238.944; Zuber, A.K. et al (2004) 22 (13 - 14): 1791 - 8), y el inhibidor CCR5 CMPD167 (véase Veazey, R.S. et al (2003) J. Exp. Med. 198: 1551 - 1562).

Comúnmente se usa hidróxido o fosfato de aluminio (alumbre) en una solución del 0,05% al 0,1% en solución salina tamponada con fosfato. Otros adyuvantes que se pueden usar, especialmente con las vacunas de ADN, son la toxina del cólera, especialmente CTA1-DD/ISCOM (véase Mowat, A.M. et al (2001) J. Immunol. 167(6): 3398 - 405), polifosfazenos (Allcock, H.R. (1998) App. Organometallic Chem. 12(10-11): 659 - 666; Payne, L.G. et al (1995) Pharm. Biotechnol. 6: 473 - 93), citoquinas tales como, pero no exclusivamente, IL-2, IL-4, GM-CSF, IL-12, IL-15 IGF-1, IFN- α , IFN- β e IFN- γ (Boyer et al., (2002) J. Liposome Res. 121: 137 - 142; documento WO01/095919), proteínas inmunorreguladoras tales como CD40L (ADX40; véase, por ejemplo, el documento WO03/063899) y el ligando CD1a de linfocitos citotóxicos naturales (también conocido como CRONY o α -galactosilceramida; véase Green, T.D. et al, (2003) J. Virol. 77(3): 2046 - 2055), proteínas de fusión inmunoestimulantes tales como IL-2 fusionada con el fragmento Fc de las inmunoglobulinas (Barouch et al., Science 290: 486 - 492, 2000) y moléculas coestimulantes B7.1 y B7.2 (Boyer), todos los cuales se pueden administrar como proteínas o en forma de ADN, en los mismos vectores de expresión que los que codifican los antígenos de la divulgación o en vectores de expresión independientes.

Las composiciones inmunogénicas se pueden diseñar para introducir los antígenos, ácidos nucleicos o vectores de expresión del VIH-1 del clado A en un sitio de acción deseado y liberarlo a una velocidad apropiada y controlable. En la técnica se conocen procedimientos de preparación de formulaciones de liberación controlada. Por ejemplo, se pueden producir preparaciones de liberación controlada mediante el uso de polímeros para formar complejos o absorber el inmunógeno y/o la composición inmunogénica. Se pueden preparar formulaciones de liberación controlada usando macromoléculas apropiadas (por ejemplo, poliésteres, ácidos poliamino, polivinilo, pirrolidona, etilenoivinilacetato, metilcelulosa, carboximetilcelulosa o sulfato de protamina) que se sabe que proporcionan las características de liberación controlada o el perfil de liberación deseados. Otro procedimiento posible para controlar la duración de la acción por una preparación de liberación controlada es incorporar los principios activos en las partículas de un material polimérico tal como, por ejemplo, poliésteres, ácidos poliamino, hidrogeles, ácido poliláctico, ácido poliglicólico, copolímeros de esos ácidos o copolímeros de vinilacetato de etileno. Alternativamente, en lugar de incorporar esos ingredientes activos en partículas poliméricas, es posible atrapar esos materiales en microcápsulas preparadas, por ejemplo, mediante técnicas de coacervación o por polimerización interfacial, por ejemplo, microcápsulas de hidroximetilcelulosa o gelatina y poli-(metilmetacrilato), respectivamente, en sistemas de administración de fármacos coloidales (por ejemplo, liposomas, microesferas de albúmina, microemulsiones, nanopartículas y nanocápsulas) o en macroemulsiones. Dichas técnicas se dan a conocer en New Trends and Developments in Vaccines, Voller et al. (eds.), University Park Press, Baltimore, Md., 1978 y Remington's Pharmaceutical Sciences, 16ª edición.

Las dosis adecuadas de los antígenos, ácidos nucleicos y vectores de expresión (colectivamente, los inmunógenos) del VIH-1 del clado A de la divulgación en la composición inmunogénica de la invención, pueden ser

determinadas fácilmente por los expertos en la técnica. Por ejemplo, la dosificación de los inmunógenos puede variar dependiendo de la vía de administración y el tamaño del sujeto. Las dosis adecuadas pueden ser determinadas por los expertos en la técnica, por ejemplo, midiendo la respuesta inmunitaria de un sujeto, tal como un animal de laboratorio, usando técnicas inmunológicas convencionales, y ajustando las dosis según corresponda. Dichas técnicas para medir la respuesta inmunitaria del sujeto incluyen, pero no exclusivamente, ensayos de liberación de cromo, ensayos de unión al tetrámero, ensayos de IFN- γ ELISPOT, ensayos de IL-2 ELISPOT, ensayos de citoquinas intracelulares y otros ensayos de detección inmunológicos, por ejemplo, como los detallados en el texto "Antibodies: A Laboratory Manual" por Ed Harlow y David Lane.

Cuando se proporcionan profilácticamente las composiciones inmunogénicas de la invención se administran idealmente a un sujeto antes de la infección por el VIH, o de evidencia de infección por el VIH, o antes de cualquier síntoma debido al SIDA, especialmente en sujetos de alto riesgo. La administración profiláctica de las composiciones inmunogénicas puede servir para proporcionar inmunidad protectora a un sujeto contra la infección por el VIH-1 o para evitar o atenuar el avance del SIDA en un sujeto ya infectado con el VIH-1. Cuando se proporcionan terapéuticamente, las composiciones inmunogénicas pueden servir para mejorar y tratar los síntomas del SIDA y se usan ventajosamente después de la infección tan pronto como sea posible, preferentemente antes de la aparición de cualquier síntoma del SIDA pero también se pueden usar en (o después de) la aparición de los síntomas de la enfermedad.

Las composiciones inmunogénicas se pueden administrar usando cualquier procedimiento de administración adecuado, incluyendo, pero no exclusivamente, la administración intramuscular, intravenosa, intradérmica, por la mucosa y tópica. Dichas técnicas son bien conocidas por los expertos en la técnica. Ejemplos más específicos de procedimientos de administración son la inyección intramuscular, la inyección intradérmica y la inyección subcutánea. Sin embargo, la administración no necesita estar limitada a los procedimientos de inyección. Además, la administración de ADN al tejido animal se logró mediante liposomas catiónicos (Watanabe et al., (1994) Mol. Reprod. Dev. 38: 268 - 274; y documento WO 96/20013), inyección directa de ADN desnudo en el tejido muscular del animal (Robinson et al., (1993) Vaccine 11: 957 - 960; Hoffman et al., (1994) Vaccine 12: 1529 - 1533; Xiang et al., (1994) Virology 199: 132 - 140; Webster et al., (1994) Vaccine 12: 1495 - 1498; Davis et al., (1994) Vaccine 12: 1503 - 1509; y Davis et al., (1993) Hum. Mol. Gen. 2: 1847 - 1851), o inyección intradérmica de ADN usando tecnología de "pistola de genes" (Johnston et al., (1994) Meth. Cell Biol. 43: 353 - 365). Alternativamente, las vías de administración pueden ser oral, intranasal o cualquier otra vía adecuada. La administración también se puede llevar a cabo a través de una superficie de mucosa como la mucosa oral, vaginal o anal.

Los programas de vacunación (o pautas de administración) para los animales (incluidos los humanos) son bien conocidos y se pueden determinar fácilmente para el sujeto y la composición inmunogénica particulares. Por lo tanto, se pueden administrar los inmunógenos una o más veces al sujeto. Preferentemente, hay un intervalo de tiempo establecido entre administraciones separadas de la composición inmunogénica. Si bien este intervalo varía para cada sujeto, normalmente varía desde 10 días hasta varias semanas y a menudo es de 2, 4, 6 u 8 semanas. Para los seres humanos, el intervalo es normalmente de 2 a 6 semanas. Las pautas de administración de la vacunación habitualmente implican de 1 a 6 administraciones de la composición inmunogénica, pero pueden ser tan sólo uno o dos o cuatro. Los procedimientos para inducir una respuesta inmunitaria también pueden incluir la administración de un adyuvante con los inmunógenos. En algunos casos, las vacunas de refuerzo anuales, bianuales o a otros intervalos largos (5 - 10 años) pueden complementar el protocolo de vacunación inicial.

Los presentes procedimientos también incluyen una diversidad de pautas de administración de sensibilización/refuerzo, especialmente pautas de administración de sensibilización con ADN / refuerzo con adenovirus. En esos procedimientos, una o más vacunas de sensibilización son seguidas por una o más vacunas de refuerzo. La composición inmunogénica real puede ser la misma o diferente en cada vacunación y el tipo de composición inmunogénica (por ejemplo, que contiene proteína o vector de expresión), la vía y la formulación de los inmunógenos también pueden variar. Por ejemplo, si se usa un vector de expresión para las etapas de sensibilización y refuerzo, puede ser del mismo tipo o de otro tipo (por ejemplo, ADN o vector de expresión bacteriano o viral). Una pauta de administración de sensibilización/refuerzo útil prevé dos vacunas de sensibilización, separadas cuatro semanas, seguidas de dos vacunas de refuerzo 4 y 8 semanas después de la última vacunación de sensibilización. También debe ser fácilmente evidente para un experto en la técnica que existen varias permutaciones y combinaciones que se engloban usando los vectores de expresión de ADN bacterianos y virales de la divulgación, para proporcionar las pautas de administración de sensibilización y refuerzo.

Una realización específica de la invención proporciona procedimientos para inducir una respuesta inmunitaria contra el VIH en un sujeto mediante la administración de una composición inmunogénica o de vacuna coo se define en la reivindicación 1, una o más veces a un sujeto en el que se expresan los antígenos Env del VIH-1 del clado A a un nivel suficiente para inducir una respuesta inmunitaria específica en el sujeto. Dichas vacunas se pueden repetir varias veces a intervalos de tiempo de al menos 2, 4 o 6 semanas (o más) de acuerdo con una

pauta de administración de vacunación deseada.

Las composiciones inmunogénicas o de vacuna de la invención se pueden administrar solas, o se pueden coadministrar o administrar secuencialmente con otros inmunógenos del VIH o composiciones inmunogénicas del VIH, por ejemplo, con "otras" composiciones inmunológicas, antigénicas o para vacuna o terapéuticas, mediante las cuales se proporcionan composiciones multivalentes o "cocktail" o una combinación de composiciones, y procedimientos para emplearlas. Una vez más, los ingredientes y la forma (secuencial o coadministración) de administración, así como las dosis, se pueden determinar teniendo en cuenta factores tales como la edad, el género, el peso, la especie y la afección particular del sujeto, y la vía de administración.

Cuando se usan en combinación, los otros inmunógenos del VIH se pueden administrar al mismo tiempo o en momentos diferentes como parte de una pauta de administración de la vacunación global, por ejemplo, como parte de una pauta de administración de sensibilización-refuerzo u otro protocolo de vacunación. Se conocen muchos otros inmunógenos del VIH en la técnica, uno de dichos inmunógenos preferidos es HIVA (descrito en el documento WO 01/47955), que se puede administrar como una proteína, en un plásmido (por ejemplo, pThr.HIVA) o en un vector viral (por ejemplo, MVA.HIVA). Otro de dichos inmunógenos del VIH es RENTA (descrito en el documento PCT/US2004/037699), que también se puede administrar como una proteína, en un plásmido (por ejemplo, pThr.RENTA) o en un vector viral (por ejemplo, MVA.RENTA).

Por ejemplo, un procedimiento para inducir una respuesta inmunitaria contra el VIH en un sujeto humano comprende administrar al menos una dosis sensibilizante de un inmunógeno del VIH y al menos una dosis de refuerzo de un inmunógeno del VIH, en el que el inmunógeno de cada dosis puede ser el mismo o diferente, siempre que al menos uno de los inmunógenos sea un antígeno del VIH-1 del clado A de la divulgación, un ácido nucleico que codifica un antígeno del VIH-1 del clado A de la divulgación o un vector de expresión, preferentemente un vector de adenovirus, que codifica un antígeno del VIH-1 del clado A de la divulgación, y en el que los inmunógenos se administran en una cantidad o se expresan a un nivel suficiente para inducir una respuesta inmunitaria específica del VIH en el sujeto. La respuesta inmunitaria específica del VIH puede incluir una respuesta inmunitaria de linfocitos T específicos del VIH o una respuesta inmunitaria de linfocitos B específicos del VIH. Dichas vacunaciones se pueden realizar a intervalos, preferentemente de al menos 2 a 6 o más semanas.

Los ejemplos no limitantes siguientes se dan con el fin de ilustrar diversas realizaciones de la divulgación.

Ejemplos

Ejemplo 1: secuencia consenso para Gag del VIH del clado A

Tabla 1

	Distancia con respecto al consenso	País	Año
A Consensu	0		
A_97TZ02_1	0,04081	TZ	1997
A_TZA173_1	0,0425	TZ	2001
A_KNH1144_	0,04259	KE	2000
A_SE7535UG	0,04303	UG	1994
A_KNH1211_	0,04463	KE	2000
A_KSM4024_	0,04684	KE	2000
A_KNH1207_	0,04701	KE	2000
A_SE6594UG	0,04709	UG	1993
A_92UG037_	0,05079	UG	1992
A_TZA195_1	0,05127	TZ	2001
A_MSA4079_	0,05279	KE	2000

ES 2 528 951 T3

A_TZA341_1	6,05523	TZ	2001
A_MSA4072_	0,05583	KE	2000
A_MSA4076_	0,056	KE	2000
A_KNH1199	0,05687	KE	2000
A_MSA407_	0,05947	KE	2000
A_98UG5713	0,06038	UG	1998
A_KEQ23-17	0,06072	KE	1994
A_KNH1209_	0,06101	KE	2000
A_NKU3005_	0,06108	KE	2000
A_SE7253SO	0,06113	SO	1994
A_98UG5713	0,06119	UG	1998
A_SE8538_	0,06137	TZ	1995
A_KNH1088_	0,06262	KE	1999
A_KER2608_	0,065	KE	2000
A_99UGA070	0,06531	UG	1999
A_KER2012-	0,06654	KE	2000
A_KER2009_	0,0674	KE	2000
A_99UGG033	0,06871	UG	1999
A_KSM4030-	0,07026	KE	2000
A_KSM4021-	0,07145	KE	1999
A_98UG5713	0,07189	UG	1998
A_SE8891UG	0,07197	UG	1995
A_SE8131UG	0,07462	UG	1995
A_97TZ03_1	0,07653	TZ	1997
A_KNH1135_	0,07687	KE	1999
A_98UG5714	0,0781	UG	1998
A_UGU455_1	0,08349	UG	1985
A_MSA4069_	0,08867	KE	2000

5 Se analizaron las secuencias de aminoácidos de las proteínas Gag de 39 cepas no recombinantes del VIH del clado A. La tabla 1 enumera las 39 cepas usadas y se refiere a cada una por su número de registro de Genbank. La tabla 1 también identifica el país y el año del aislamiento de cada una de estas 39 cepas. 20 de las cepas eran de Kenia, 12 de Uganda, 6 de Tanzania y 1 de Somalia. 20 de las cepas se aislaron entre 2000 y 2002, 10 se aislaron entre 1997 y 1999, 6 se aislaron entre 1994 y 1996 y 3 se aislaron antes de 1993.

Las secuencias de la proteína Gag se alinearon agregando espacios para preservar el alineamiento en regiones con inserciones o deleciones. Se derivó una secuencia consenso del 50%. La secuencia consenso de aminoácidos

se muestra en la figura 1. En la figura 1 los espacios que se agregaron para conservar el alineamiento en regiones con inserciones o deleciones están representados por guiones, y las posiciones para las cuales no se logró un consenso del 50% están representadas por una "X".

5 Para cada una de las 39 secuencias usadas para generar la secuencia consenso, la "distancia" de esa secuencia respecto a la secuencia consenso se calculó usando la matriz de sustitución de Dayhoff PAM250, que pondera las sustituciones según el grado de similitud bioquímica. Como se muestra en la tabla 1, la distancia de la secuencia de cada cepa respecto a la secuencia consenso varió entre el 4% y el 9%.

10 La figura 2 ilustra en forma gráfica la distancia de la secuencia de aminoácidos de cada cepa respecto a la secuencia consenso de aminoácidos e identifica las cuatro cepas que tienen las secuencias más similares a las secuencias consenso. Esas cuatro cepas son la cepa 97TZ02 de un individuo de bajo riesgo en la región de Mbeya del suroeste de Tanzania en 1997 que tiene el número de registro de Genbank AF361872, la cepa TZA173 obtenida de un donante de sangre anónimo en la región de Mbeya al suroeste de Tanzania en 2001 que tiene el número de registro de Genbank AY253305, la cepa KNH1144 obtenida de un donante de sangre anónimo en Kenia meridional en el año 2000 que tiene el número de registro de Genbank AF4587006, y la cepa SE7535 obtenida en 15 1994 en Suecia de un individuo que se pensaba que había sido infectado en Uganda que tiene el número de registro de Genbank AF069671.

Ejemplo 2: secuencia consenso para Pol del VIH del clado A

20 Se analizaron las secuencias de aminoácidos de las proteínas Pol de 36 cepas no recombinantes del VIH del clado A. La tabla 2 enumera las 36 cepas usadas y se refiere a cada una por su número de registro de Genbank. La tabla 2 también identifica el país y el año del aislamiento de cada una de esas 36 cepas. 20 de las cepas fueron de Kenia, 9 de Uganda, 6 de Tanzania y 1 de Somalia. 19 de las cepas se aislaron entre 2000 y 2002, 10 se aislaron entre 1997 y 1999, 4 se aislaron entre 1994 y 1996 y 3 se aislaron antes de 1993.

25 Se alinearon las secuencias de las proteínas Pol. No hubo inserciones ni deleciones. Se derivó una secuencia consenso del 50%. La secuencia consenso de aminoácidos se muestra en la figura 3. En la figura 3 las posiciones para las cuales no se logró un consenso del 50% están representadas por una "X". Había 4 de dichas posiciones en los 947 residuos de aminoácidos. Para cada una de las 36 secuencias usadas para generar la secuencia consenso, la "distancia" de esa secuencia respecto a la secuencia consenso se calculó usando la matriz de sustitución de Dayhoff PAM250, que pondera las sustituciones según el grado de similitud bioquímica. Como se muestra en la tabla 2, la distancia de la secuencia de cada cepa con respecto a la secuencia consenso varió entre 30 el 1,5% y el 4,8%.

Tabla 2

	Distancia con respecto al consenso	País	Año
A_pol,cons	0		
A_MSA4070_	0,01479	KE	2000
A_SE7253SO	0,01582	SO	1994
A_SE8538TZ	0,01898	TZ	1995
A_KER2012-	0,02329	KE	2000
A_97TZ02_3	0,0235	TZ	1997
A_KEQ23-17	0,02445	KE	1994
A_KNH1211_	0,02449	KE	2000
A_TZA341_3	0,0246	TZ	2001
A_KSM4024_	0,02528	KE	2000
A_97TZ03_3	6,02544	TZ	1997
A_KNH1088_	0,02544	KE	1999

ES 2 528 951 T3

A_MSA4076_	6,02564	KE	2000
A_KNH1207_	6,0265	KE	2000
A_NKU3005_	0,02661	KE	2000
A_TZA173_3	0,02756	TZ	2001
A_MSA4079_	0,02762	KE	2000
A_KER2009_	0,02765	KE	2000
A_TZA195_3	0,02881	TZ	2001
A_KSM4021-	0,02881	KE	1999
A_SE7535UG	0,02883	UG	1994
A_MSA4069_	0,02886	KE	2000
A_SE6594UG	0,02889	UG	1993
A_98UG5713	0,02975	UG	1998
A_KNH1135_	0,0299	KE	1999
A_92UG037_	0,02993	UG	1992
A_KNH1209_	0,03202	KE	2000
A_99UGG033	0,03291	UG	1999
A_KER2008_	0,03294	KE	2000
A_KSM4030-	0,0343	KE	2000
A_KNH1199_	0,03439	KE	2000
A_99UGA070	6,03537	UG	1999
A_MSA4072	0,03625	KE	2000
A_KNH1144_	0,03863	KE	2000
A_98UG5713	0,04178	UG	1998
A_UGU455_3	0,04294	UG	1985
A_98UG5713	0,04808	UG	1998

La figura 4 ilustra en forma gráfica la distancia de la secuencia de aminoácidos de cada cepa con respecto a la secuencia consenso de aminoácidos e identifica las tres cepas que tienen las secuencias más similares a las secuencias consenso. Esas tres cepas son la cepa MSA4070 de un donante de sangre anónimo en Kenia meridional en el año 2000, la cepa SE7235SO que se obtuvo en 1994 de un individuo en Suecia que se pensaba que había sido infectado en Somalia, y la cepa SE8538 que se obtuvo en 1995 de un individuo en Suecia que se pensaba que había sido infectado en Tanzania.

5

Ejemplo 3: Secuencia consenso para Env del VIH del clado A

Tabla 3

	Distancia con respecto al consenso	País	Año
--	------------------------------------	------	-----

ES 2 528 951 T3

A,cons	0		
A_KEQ23-17	0,06307	KE	1994
A_TZA341_1	0,06413	TZ	2001
A_KNH1088_	0,06524	KE	1999
A_KNH1209_	0,0699	KE	2000
A_KHNH1144_	0,07088	KE	2000
A_99UGA070	0,07365	UG	1999
A_MSA4072_	0,07516	KE	2000
A_KSM4021-	0,0778	KE	1999
A_97TZ021_1	0,07825	TZ	1997
A_KNH1199_	0,07883	KE	2000
A_MSA4079_	0,07944	KE	2000
A_SE7535UG	0,08375	UG	1994
A_SE8538TZ	0,08432	TZ	1995
A_98UG5713	0,08462	UG	1998
A_97TZ03_1	0,08541	TZ	1997
A_MSA4070_	0,0874	KE	2000
A_NKU3005_	0,0884	KE	2000
A_TZA173_12	0,09046	TZ	2001
A_KNH1207_	0,09106	KE	2000
A_TZA195_1	0,09389	TZ	2001
A_MSA4076_	0,09517	KE	2000
A_92UG037_	0,098	UG	1992
A_98UG5714	0,09816	UG	1998
A_SE7253SO	0,09886	SO	1994
A_KER2012-	0,09984	KE	2000
A_98UG5713	0,10139	UG	1998
A_SE6594UG	0,10195	UG	1993
A_SE8891UG	0,10225	UG	1995
A_UGU455_1	0,10314	UG	1985
A_KER2009_	0,10338	KE	2000
A_KNH1211_	0,11319	KE	2000
A_SE8131UG	0,11321	UG	1995

A_MSA4069_	0,11507	KE	2000
A_99UGG033	0,11653	UG	1999
A_KNH1135_	0,11713	KE	1999
A_KER2008_	0,12689	KE	2000

5 Se analizaron las secuencias de aminoácidos de las proteínas Env de 36 cepas no recombinantes del VIH del clado A. La tabla 3 enumera las 36 cepas usadas y se refiere a cada una por su número de registro de Genbank. La tabla 3 también identifica el país y el año del aislamiento de cada una de esas 36 cepas. 18 de las cepas eran de Kenia, 11 de Uganda, 6 de Tanzania y 1 de Somalia. 17 de las cepas se aislaron entre 2000 y 2002, 10 se aislaron entre 1997 y 1999, 6 se aislaron entre 1994 y 1996 y 3 se aislaron antes de 1993.

10 Las secuencias de las proteínas Env se alinearon agregando espacios para preservar el alineamiento en regiones con inserciones o deleciones. Había muchas regiones con una amplia heterogeneidad en la longitud de las inserciones/deleciones. Se derivó una secuencia consenso del 50%. La secuencia consenso de aminoácidos se muestra en la figura 5. En la figura 5 los espacios que se agregaron para conservar el alineamiento en regiones con inserciones o deleciones están representados por guiones, y las posiciones para las cuales no se logró un consenso del 50% están representadas por una "X". Hubo muchas posiciones de aminoácidos para las cuales no se alcanzó un consenso del 50%.

15 Para cada una de las 36 secuencias usadas para generar la secuencia consenso, la "distancia" de esa secuencia con respecto a la secuencia consenso se calculó usando la matriz de sustitución de Dayhoff PAM250, que pondera las sustituciones según el grado de similitud bioquímica. Como se muestra en la tabla 3, la distancia de la secuencia de cada cepa respecto a la secuencia consenso varió entre el 6,3% y el 12,7%.

20 La figura 6 ilustra en forma gráfica la distancia de la secuencia de aminoácidos de cada cepa con respecto a la secuencia consenso de aminoácidos e identifica las tres cepas que tienen las secuencias más similares a las secuencias consenso. Esas tres cepas eran KEQ23 de una trabajadora social diplomada en Kenia en 1994, TZA341 que era de un donante de sangre anónimo en Tanzania en 2002 y KNH1088 que era de un donante de sangre anónimo en Kenia en 1999.

Ejemplo 4: Secuencia consenso para Nef del VIH del clado A

25 Se analizaron las secuencias de aminoácidos de las proteínas Nef de 38 cepas no recombinantes del VIH del clado A. La tabla 4 enumera las 38 cepas usadas y se refiere a cada una por su número de registro de Genbank. El país y el año de aislamiento de cada una de esas 38 cepas se describen en las tablas 1 a 3 en los ejemplos anteriores. Más de la mitad de las cepas eran de Kenia, con una parte sustancial proveniente de Uganda y unas pocas cepas de Tanzania. Aproximadamente la mitad de la cepas se aislaron entre 2000 y 2002.

Tabla 4

	A. cons
A MSA4070_	0,0318
A KNH1211_	0,04807
A 97TZ03_1	0,0535
A 99UGA070	0,05354
A SE8891UG	0,05383
A KEQ23-17	0,06476
A 98UG5713	0,07043

ES 2 528 951 T3

A NKU3005	0,0709
A SE7535UG	0,07117
A 98UG5714	0,07613
A SE6594UG	0,07634
A TZA341 1	0,0805
A MSA4069	0,08097
A KNH1199	0,08213
A 97TZ02 1	0,08276
A KSM4030-	0,08704
A KSM4021-	0,08795
A MSA4076	0,08873
A KNH1209	0,0899
A KER2012-	0,09224
A KNH1144	0,09577
A KER2008	0,09703
A MSA4072	0,09892
A 98UG5713	0,09892
A 99UGG033	0,09967
A KNH1088	0,10303
A 92UG037	0,10654
A SE8538TZ	0,10996
A KER2009	0,1102
A MSA4079	0,11083
A KSM4024	0,11126
A SE8131UG	0,11326
A SE7253SO	0,11453
A KNH1207	0,11549
A TZA173 1	0,13766

A 98UG5713	0,1399
A UGU455 1	0,15688
A KNH1135	0,16076
A.cons	0

5 Las secuencias de las proteínas Nef se alinearon con espacios para preservar el alineamiento en regiones con inserciones o deleciones. Se derivó una secuencia consenso de 50%. La secuencia consenso de aminoácidos se muestra en la figura 7. En la figura 7 los espacios que se agregaron para conservar el alineamiento en regiones con inserciones o deleciones están representados por guiones, y las posiciones para las cuales no se logró un consenso del 50% están representadas por una "X". Había 6 posiciones de aminoácidos para las cuales no se alcanzó un consenso del 50%.

10 Para cada una de las 38 secuencias usadas para generar la secuencia consenso la "distancia" de esa secuencia respecto a la secuencia consenso se calculó usando la matriz de sustitución de Dayhoff PAM250, que pondera las sustituciones según el grado de similitud bioquímica. Como se muestra en la tabla 4, la distancia de la secuencia de cada cepa respecto a la secuencia consenso varió entre el 3,2 y el 16,1% con una distancia media del 9,3%.

15 La figura 8 ilustra en forma gráfica la distancia de la secuencia de aminoácidos de cada cepa respecto a la secuencia consenso de aminoácidos e identifica las cinco cepas que tienen las secuencias más similares a las secuencias consenso. Esas cinco cepas fueron MSA4070 y KNH1211, las cuales fueron de donantes anónimos en Kenia meridional y se obtuvieron en el año 2000, 97TZ03 de un individuo de bajo riesgo en la región de Mbeya del suroeste de Tanzania se obtuvo en 1997, y UGA070 y SE8891, ambas eran de individuos en Uganda y se obtuvieron en 1999 y 1995, respectivamente.

Ejemplo 5. Cepas del VIH del clado A que son las más similares a las secuencias consenso del VIH del clado A.

20 Como se describe en los ejemplos 1 a 4 antes y como se resume en la tabla 5, se identificaron las cepas del VIH el clado A que tienen las secuencias Gag, Pol, Env y Nef más similares a las secuencias consenso de cada una de esas proteínas. Además, se identificaron las cepas que eran más similares globalmente a la secuencia consenso puntuando a cada una de las cepas según su similitud con la secuencia consenso de una proteína particular donde la cepa puntuada con el número 1 fue aquella cuya secuencia para esa proteína era la más similar a la de la secuencia consenso y después sumando las puntuaciones para cada cepa para las 4 proteínas (es decir Gag, Pol, Env y Nef). Las seis cepas que fueron más similares globalmente a la secuencia consenso en las cuatro proteínas estudiadas se indican a continuación en la tabla 6. Se puede observar que la cepa 97TZ02 tiene una secuencia que es globalmente más similar a las secuencias consenso de cada una de los genes de Gag, Pol, Env y Nef.

Tabla 5

Gag	Pol	Env	Nef
97TZ02	MSA4070	KEQ23	MSA4070
TZA173	SE7245SO	TZA341	KNH1211
KNH1144	SE8538	KNH1088	97TZ03
SE7535UG			99UGA070
			SE8891UG

30

Tabla 6

	gag	pol	env	nef	sum
A_97TZ02_1	1	5	9	15	30
A_KEQ23-17	18	6	1	6	31
A_MSA4070_	16	1	16	1	34
A_TZA341_1	12	8	2	12	34
A_SE7535UG	4	20	12	9	45
A_KNH1211_	5	7	31	2	45

Ejemplo 6. Construcción de los transgenes GRIN, GRN y Env.

5 Se hicieron constructos de transgenes con secuencias de proteínas del VIH del clado A derivadas de los aislados de campo circulantes del VIH-1 identificados más recientemente que se correspondían mejor con la secuencia consenso del VIH del clado A para cada una de esas proteínas. Esta estrategia se desarrolló para maximizar la importancia biológica de las secuencias del VIH del clado A usadas. Se debe comprender que también se pueden usar otras secuencias, es decir secuencias distintas de las secuencias específicas descritas en este ejemplo, de conformidad con esta divulgación. Si se usan otras secuencias es preferible que las secuencias se seleccionen de modo que deriven de aislados de campo recientes y que tengan secuencias que sean similares a las secuencias consenso del VIH del clado A descritas en la presente memoria descriptiva, o a las secuencias consenso del VIH del clado A que se puedan generar en el futuro.

15 Se hicieron constructos denominados GRIN y GRN. El constructo GRIN contenía secuencias del VIH del clado A que codifican las proteínas Gag, Pol (RT e integrasa) y Nef. El transgén GRN contenía secuencias que codifican las proteínas Gag, RT y Nef. Los constructos GRIN y GRN se representan esquemáticamente en la figura 9. Los transgenes GRIN y GRN se prepararon usando la secuencia de la proteína Gag de la cepa TZA173 que tiene el número de registro de Genbank AY253305, la secuencia de Pol (que comprende las secuencias RT e Int) de la cepa MSA4070 que tiene el número de registro de Genbank AF457081 y la secuencia de Nef de la cepa MSA4070 que tiene el número de registro de Genbank AF457081. Esas secuencias se seleccionaron porque eran de los aislados de campos circulantes del VIH-1 más recientemente identificados que tenían la mayor correspondencia con la secuencia consenso para cada una de Gag, Pol y Nef, respectivamente. Esas secuencias se ilustran en las figuras 10, 11 y 12, respectivamente.

25 También se preparó un constructo de Env con la secuencia de codificación de Env del aislado de campo circulante del VIH-1 más recientemente identificado que tenía la mayor correspondencia con la secuencia consenso de Env, es decir, la secuencia de la proteína Env de la cepa TZA341 que tiene el número de registro de Genbank AY253314. Esta secuencia se ilustra en la figura 13.

30 A todas las secuencias se les hizo optimización de codones para expresión en humanos en GeneArt (Alemania). Las secuencias con los genes optimizados permiten un alto nivel de expresión y la expresión estable de las proteínas en humanos u otras células de mamíferos. Otros detalles del proceso de optimización de codones se proporcionan en el ejemplo 8. Los transgenes también se modificaron por ingeniería genética para incorporar mutaciones específicas o dispuestas en un orden específico para anular la función normal de los productos génicos in vivo. Los detalles de esas mutaciones, y los efectos biológicos de cada una, se describen en el ejemplo 7 dado más adelante.

35 Para los transgenes GRIN y GRN, las secuencias de codificación para cada una de las proteínas Gag, Pol (RT e Int) o RT, y Nef se unieron en el marco de tal forma que cada uno de los constructos del transgén (es decir de GRIN o de GRN) codifica una sola proteína de fusión. Se realizaron búsquedas Blast para asegurarse de que no se formaran neoepítomos en las uniones. Aunque no se usa en este ejemplo, se debe señalar que también es posible insertar secuencias espaciadoras entre las secuencias de codificación para los componentes individuales de la proteína de fusión final para permitir el plegado óptimo del dominio de la proteína, por ejemplo, se puede agregar una región espaciadora entre Gag y Pol para permitir que los dominios de la proteína se plieguen en una conformación más natural. Asimismo, se agregaron sitios de restricción únicos en los extremos 5' y 3' de cada secuencia con el fin de facilitar la unión de cada secuencia (por ejemplo, la unión del extremo 5' de Nef con el extremo 3' de Pol, etc.).

5 Para usar in vivo se insertaron los transgenes GRIN, GRN y Env en los vectores de adenovirus Ad5, Ad35, Ad11, C6 o C7. Para facilitar la clonación se agregaron sitios de restricción únicos en esos vectores en los extremos 5" y 3' de los constructos GRIN, GRN o Env. Las figuras 14A - 14 C proporcionan la secuencia de GRIN insertada en el vector Ad35 y muestran los sitios de restricción usados para clonar la secuencia de GRIN en el vector Ad35 (subrayados y en negrita). La secuencia también incluye la secuencia promotora CMV cadena arriba de la secuencia de GRIN. Las figuras 15A - 15B proporcionan la secuencia de Env insertada en el vector Ad35 y muestran los sitios de restricción usados para clonar la secuencia de Env en el vector Ad35 (subrayados y en negrita). La secuencia también incluye la secuencia promotora CMV cadena arriba de la secuencia de Env.

10 Se usaron técnicas convencionales de clonación y recombinación del ADN para generar todos los constructos anteriores. Dichas técnicas son bien conocidas por los expertos en la técnica. Véase por ejemplo, "Molecular Cloning: A Laboratory Manual", segunda edición (Sambrook et al., 1989).

Ejemplo 7: Mutaciones para anular la función normal in vivo de las proteínas del VIH del clado A

15 La tabla 7 resume las mutaciones modificadas por ingeniería genética en las secuencias de GRIN y GRN para anular la función in vivo de sus productos génicos. Esas mutaciones se realizaron usando técnicas convencionales de recombinación del ADN. Dichas técnicas son bien conocidas por los expertos en la técnica.

Diseño	Gen	Mutación / justificación
Mutación	gag	Gly2 → Ala : elimina el sitio de miristilación evitando la formación de VLP.
Mutación	gag	<i>Para evitar el deslizamiento en la secuencia de cambio de marco natural, la secuencia de ADN se mutó de una manera que deja intacta la secuencia de aminoácidos conservada (NFLG) y permite que sólo se traduzca el producto de la proteína GagPol de longitud completa.</i>
Mutación	RT	Asp185 → Ala & Asp186 → Ala: inactiva los residuos enzimáticos activos.
Mutaciones	Integrasa (IN)	Asp 64 → Ala, Asp116 → Ala & Glu 152 → Ala: inactiva los residuos enzimáticos activos.
Sin cambio	Nef	La fusión del extremo N-terminal de nef con el extremo C-terminal de IN evita la miristilación y la acción selectiva sobre la membrana anulando la función de Nef.

20 La proteína Gag se expresa como una poliproteína precursora de 55 kDa (Pr55^{gag}) y es escindida por la proteasa viral del VIH-1. Cuatro proteínas virales principales resultan de la escisión; Matriz (MA), cápside (CA), nucleocápside (NC) y p6; así como dos polipéptidos espaciadores p2 y p1, que representan secuencias entre CA y NC y entre NC y p6, respectivamente.

25 MA desempeña una función clave en varias etapas de la replicación del virus, que incluyen la mediación crítica del ensamblado de la partícula viral y el brote de la membrana celular a través de la formación de partículas similares a virus (VLP) (véase Gheysen, D., E. Jacobs, F. de Foresta, D. Thiriart, M. Francotte, D. Thines, y M. De Wilde. (1989). Assembly and release of HIV-1 precursor pr55gag virus-like particles from recombinant baculovirus-infected cells. Cell 59: 103 - 112).

30 Tanto Pr55^{gag} como MA (p17) se miristilan, es decir, forman enlace amida con el ácido mirístico. Véase Veronese di Marzo, F., Copeland, T. D., Oroszlan, S., Gallo, R. C. y Sarngadharan, M. G. (1988). J. Virol. 62, 795 - 801. Véase también la sección sobre Nef en el ejemplo 7 para una descripción completa del proceso de miristilación. Diferentes aislados del VIH-1 demuestran que el aceptor de miristilo es el residuo de glicina N-terminal (Gly2). Véase Bryant y Ratner. (1990). Myristoylation-dependent replication and assembly of human immunodeficiency virus 1. Proc. Natl. Acad. Sci. EE.UU; 87: 523 - 527.

35 Bryant y Ratner (1990) demostraron que la sustitución de Gly2 con Ala eliminó la replicación del virus de un clon del VIH-1. El Pr55^{gag}, que carece de la glicina aceptor de miristilo, se acumuló en células Hela infectadas y no se procesó en la cápside del virión maduro. Se concluyó que la miristilación de la Gly2 es necesaria para la asociación estable de la membrana plasmática y el ensamblado posterior de los viriones. Otros grupos demostraron de manera semejante la importancia de la miristilación de Gly2 en la MA. Véase Göttlinger H G, Sodroski J G, Haseltine W A. (1989). Role of capsid precursor processing and myristoylation in morphogenesis and infectivity of human immunodeficiency virus type 1. Proc Natl Acad Sci USA; 86:5781-5785, y Paul Spearman,

Jaang-Jiun Wang, Nancy Vander Heyden y Lee Ratner. (1994). Identification of Human Immunodeficiency Virus Type 1 Gag Protein Domains Essential to Membrane Binding and Particle Assembly. *J. Virol*; 68 (5): 3232 - 3242.

5 Si se muta la glicina N-terminal aceptora de miristilo (Gly2) en MA, se anula la unión a la membrana y se evita el ensamblado de partículas. Por consiguiente, Gag del clado A se modifica por ingeniería genética para cambiar Gly2 → Ala. Esto provoca la pérdida de la función biológica de Gag.

10 La transcriptasa inversa (RT) es una enzima viral esencial para la replicación. RT convierte el ARN+ viral entrante en ADNbc, catalizada por las actividades de polimerasa dependiente de ARN y ADN, y de ARNasa H de la enzima. RT es un heterodímero compuesto por las proteínas subunidades p66 y p51. Véase Alfredo Jacobo-Molina et al. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. USA*; 90: 6320 - 6324. p66 tiene 2 dominios, la polimerasa y la ARNasa H. p51 tiene el mismo dominio polimerasa.

15 Los residuos Asp-110, Asp-185 y Asp-186 catalíticamente esenciales se encuentran ubicados en el sitio activo altamente conservado de la ADN polimerasa. Se cree que esos tres residuos denominados "la tríada catalítica" se unen a cationes divalentes necesarios para la función de catálisis. Véase Alfredo Jacobo-Molina et al. (1993). Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc. Natl. Acad. Sci. USA*; 90: 6320 - 6324.

20 Se demostró que la mutación de los ácidos aspárticos en los residuos 185 y 186 en la asparragina o el glutamato da lugar a proteínas mutantes que fueron catalíticamente inactivas. Véase Lowe DM, Parmar V, Kemp SD, Larder BA. (1991). Mutational analysis of two conserved sequence motifs in HIV-1 reverse transcriptase. *FEBS Lett.*; 6;282(2): 231 - 4.

La mutación de Asp 185 → Ala y Asp 186 → Ala en la RT del clado A inactivará a la enzima polimerasa de RT interrumpiendo la "tríada catalítica". Esto eliminará la función biológica de la RT del clado A.

25 El ADNc proviral generado por RT se integra en el genoma de la célula huésped a través de la acción de la enzima viral integrasa (Int). Int contiene un dominio de recombinasa de ADN que cataliza dos reacciones endonucleóticas distintas. La primera reacción, el procesamiento de 3', elimina dinucleótidos de cada extremo de los ADNc produciendo dos extensiones 5' de dos nucleótidos en ambos extremos. En la segunda reacción, Int escinde no específicamente el ADN de la célula huésped y une los grupos libres del extremo 3' del ADNc a los grupos del 5' del ADN escindido de la célula huésped. Las enzimas celulares reparan los huecos dando lugar a un genoma viral totalmente integrado en el ADN de la célula huésped. Véase Coffin JM. *Retroviridae and their Replication*. Capítulo 27. págs. 645 - 708 & Wong-Staal F. *Human Immunodeficiency Viruses and Their Replication*. Capítulo 28. págs. 709 - 723. In Fields, BN. & Knipe DM. 2ª edición *Fundamental Virology*. Raven Press. Véase también Engelman A, Mizuuchi K, Craigie R. (1991). HIV-1 DNA integration: mechanism of viral DNA cleavage and DNA strand transfer. *Cell*; 67(6): 1211 - 1221. Los residuos 50 a 212 del dominio catalítico, contienen una tríada de residuos Asp-64, Asp-116 y Glu-152 (denominada motivo D,D-35-E) que compromete el sitio activo de la enzima. Véase Esposito, D., y R. Craigie. (1999). HIV integrase structure and function. *Adv. Virus Res.* 52: 319 - 333. Véase también Khan, E., J. P. G. Mack, R. A. Katz, J. Kulkosky, y A. M. Skalka. (1991). Retroviral integrase domains: DNA binding and the recognition of LTR sequences. *Nucleic Acids Res.* 19: 851 - 860.

35 A través de diversas técnicas, algunos grupos han demostrado la anulación de la función de la endonucleasa y/o integración de IN a través de la mutación dirigida al sitio de los residuos Asp-64, Asp-116 y Glu-152 en el motivo D, D-35-E. Véase Drelich M, Wilhelm R, Mous J. (1992). Identification of amino acid residues critical for endonuclease and integration activities of HIV-1 IN protein in vitro. *Virology*; 188(2): 459 - 468. Véase también LaFemina RL, Schneider CL, Robbins HL, Callahan PL, LeGrow K, Roth E, Schleif WA, Emini EA. (1992). Requirement of active human immunodeficiency virus type 1 integrase enzyme for productive infection of human T-lymphoid cells. *J Virol*; 66(12): 7414 - 7419. Véase también Leavitt AD, Shiue L, Varmus HE. (1993). Site-directed mutagenesis of HIV-1 integrase demonstrates differential effects on integrase functions in vitro. *J Biol Chem*; 268(3): 2113 - 2119.

40 La mutación de Asp-64→ Ala, Asp-116→ Ala y Glu-152→ Ala en la Int del clado A inactivará a la enzima activa Int interrumpiendo el motivo crítico D,D-35-E. Esto eliminará la función biológica de Int del clado A .

50 La proteína de factor negativo (Nef) (27-kDa) es la primera proteína viral que se acumula en la célula recién infectada. Véase Haseltine, W. (1991). *Molecular biology of the human immunodeficiency virus type 1*. FASEB. Vol. 5. 2349 - 2360. A través de la miristilación, Nef es capaz de ubicarse del lado del citosol de la membrana celular. Véase Yu G, Felsted RL. (1992). Effect of myristoylation on p27 nef subcellular distribution and suppression of HIV-LTR transcription. *Virology*. 187(1): 46 - 55. Véase también Kaminchik, J., N. Bashan, A. Itach, N. Sarver, M. Gorecki, y A. Panet. (1991). Genetic characterization of human immunodeficiency virus type 1 nef gene products translated in vitro and expressed in mammalian cells. *J. Virol.* 65: 583 - 588. La miristilación de las

55

proteínas es un fenómeno cotraduccional e implica la transferencia de miristato desde la miristil-coenzima A al motivo amino-terminal MGXXX de las proteínas por la enzima N-miristil transferasa (NMT). Véase Towler, D. A., S. P. Adams, S. R. Eubanks, D. S. Towery, E. Jackson-Machelski, L. Glaser y J. I. Gordon (1987). Purification and characterization of yeast myristoyl CoA:protein N- myristoyltransferase. *Proc Natl Acad Sci USA* 84: 2708 - 2712.

5 La metionina líder del polipéptido es escindida por la metionina amino peptidasa durante la traducción y reconoce a NMT el grupo amino terminal de la glicina recién generado del péptido emergente después de que aproximadamente veinte residuos se liberan del ribosoma. NMT transfiere miristato al residuo de glicina (el aceptor de miristilo) y se completa la miristilación. El reemplazo de la penúltima glicina aceptora de miristilo por otro residuo de aminoácido inhibe la miristilación. Véase Towler, D. A., S. R. Eubanks, D. S. Towery, S. P. Adams & L. Glaser (1987). Amino-terminal processing of proteins by N-myristoylation. Substrate specificity of N-myristoyl transferase. *J Biol Chem* 262: 1030 - 1036.

Nef es una proteína multifuncional capaz de modular una serie de moléculas superficiales de la célula infectada, como CD4 (véase Garcia, J. V. y A. D. Miller. (1991). Serine phosphorylation-independent downregulation of cell surface CD4 by nef. *Nature* 350: 508 - 511; y Mariani R y Skowronski J. (1993). CD4 down-regulation by nef alleles isolated from human immunodeficiency virus type 1-infected individuals *Proc. Natl. Acad. Sci. USA*. Vol. 90, págs. 5549 - 5553; y Aiken C, Konner J, Landau NR, Lenburg ME, Trono D (1994). Nef induces CD4 endocytosis: requirement for a critical dileucine motif in the membrane-proximal CD4 cytoplasmic domain. *Cell*. 11; 76(5): 853 - 64), CD28 (véase Swigut, T., N. Shohdy y J. Skowronski. (2001). Mechanism for down-regulation of CD28 by Nef. *EMBO J*. 20: 1593 - 1604), MHC-I (véase Schwartz, O., V. Marechal, S. Le Gall, F. Lemonnier y J. M. Heard. (1996). Endocytosis of major histocompatibility complex class I molecules is induced by the HIV-1 Nef protein. *Nat. Med.* 2: 338 - 342), the macrophage-e 5 xpressed MHC 1b protein HFE (véase Drakesmith H, Chen N, Ledermann H, Sreaton G, Townsend A, Xu XN. (2005). HIV-1 Nef down-regulates the hemochromatosis protein HFE, manipulating cellular iron homeostasis. *Proc Natl Acad Sci USA*. 102(31): 11017 - 22), MHC-II (véase Stumpfner-Cuvelette, P., S. Morchoisne, M. Dugast, S. Le Gall, G. Raposo, O. Schwartz y P. Benaroch. (2001). Nef del HIV-1 daña la presentación del antígeno de MHC clase II y la expresión superficial. *Proc. Natl. Acad. Sci. USA* 98: 12144 - 12149), así como interrumpe las vías de transducción de la señal (véase Tolstrup, M., L. Ostergaard, A. L. Laursen, S. F. Pedersen y M. Duch. (2004). VIH/VIS escapan de la vigilancia inmunitaria enfocada en Nef. *Curr. HIV Res.* 2: 141 - 151) a través de la asociación con múltiples quinasas y otras proteínas de la superficie celular de la membrana celular. Los mecanismos de esas acciones y los motivos de nef involucrados aún permanecen sin elucidar completamente.

De forma específica, un mutante de Nef con delección de los 19 aminoácidos N-terminales, que incluye la señal de miristilación N-terminal eliminó CD4 y la regulación por disminución de MHC-1, manteniendo simultáneamente la mayor parte de los epitopes de CTL, linfocitos T auxiliares y linfocitos B (véase Peng B, Robert Guroff M (2001). Deletion of N-terminal myristoylation site of HIV Nef abrogates both MHC-1 and CD4 down-regulation. *Immunol Lett.* 20 78(3): 195 - 200). Otros grupos demostraron que la mutación de la glicina amino terminal de Nef (Gly2) en alanina evita la miristilación (véase Liang, X. et al (2002). Development of HIV-1 Nef vaccine components: immunogenicity study of Nef mutants lacking myristylation and dileucine motif in mice. *Vaccine* 20: 3413 - 3421, y Kaminchik, J. et al. (1991). Genetic Characterization of Human Immunodeficiency Virus Type 1 nef Gene Products Translated in vitro and Expressed in Mammalian Cells. *J. of Virol.* 65(2): 583 - 588).

40 Puesto que el motivo amino terminal MGXXX de Nef del clado A está incluido en la proteína de fusión GRIN, no hay ninguna metionina incipiente para ser escindida por la metionina amino peptidasa durante la traducción. Por consiguiente, no se produce ningún grupo amino-terminal recién generado de glicina y NMT es incapaz de ejecutar la miristilación. En conclusión, la incapacidad de Nef en GRIN para experimentar miristilación anula la función biológica de Nef.

45 **Ejemplo 8: Optimización de codones para GRIN (GagPolNef) y Env**

El uso de codones para cada uno de GRIN y Env se adaptó al sesgo de codones de los genes humanos. La secuencia de nucleótidos y aminoácidos de la secuencia GRIN con codones optimizados se proporciona en las figuras 16A - 16J. La secuencia de nucleótidos y aminoácidos de la secuencia de Env con codones optimizados se proporciona en las Figuras 17A - 17 D.

50 Se evitaron las regiones con contenido muy alto (más del 80%) o muy bajo (menos del 30%) de GC cuando fue posible. Durante el proceso de optimización se evitaron los motivos que actúan en cis siguientes: cajas internas TATA, sitios chi, sitios de entrada ribosómicos, tramos de secuencia ricos en AT o en GC, elementos de secuencia ARE, INS o CRS, secuencias repetidas, estructuras secundarias de ARN, de sitios dadores y aceptores de empalme críptico, puntos de ramificación y sitios de restricción HindIII, NcoI, BglII y BclI excepto como se indica en las secuencias provistas en las figuras 16 y 17. También, se introdujo una secuencia Kozak en dirección al extremo del ATG de inicio para cada una de GRIN y Env para aumentar la iniciación de la traducción y se agregaron dos codones de parada a cada uno de GRIN y Env para asegurar una terminación eficaz. También se

agregaron sitios de restricción para facilitar la subclonación, como se indica en las figuras 16 y 17.

Ejemplo 9: Estudio de primates no humanos

5 Se realizó un estudio en primates no humanos (macacos rhesus chinos) con el objetivo principal de evaluar la inmunogenia de GRIN y ENV en un sistema de administración de un vector de adenovirus humano tipo 35 (Ad35). Los animales recibieron una dosis cada vez mayor de Ad35-GRIN/ENV (10^9 , 10^{10} y 10^{11} partículas de virus [vp]; vía intramuscular) y recibieron dos vacunas en el mes 0 y el mes 6 (con 8 animales por grupo para la primera vacuna y 4 animales por grupo para la segunda vacuna). En diversos momentos (desde la semana 0 hasta la semana 50), se extrajo sangre de los animales y se midió la inmunogenia por IFN-gamma ELISpot (véanse las figuras 23A y 23B para las dosis de vp de 10^{10} y 10^{11} , respectivamente).

10 Se observó una respuesta a la dosis (no se muestran los datos para 10^9 vp), por ELISPot, tanto en intensidad como en frecuencia de los que responden tras la primera vacunación (no se muestran los datos). Se observaron respuestas a todos los componentes antigénicos de la vacuna de GRIN/ENV y las respuestas de IFN γ ELISPOT se reforzaron después de la segunda vacunación en el mes 6.

15 La divulgación se describe más detalladamente mediante los párrafos numerados siguientes:

1. Una secuencia consenso de nucleótidos para antígenos del VIH-1 del clado A, donde la secuencia comprende secuencias de nucleótidos que codifican a Gag, Pol (RT e Int) y Nef ("GRIN) del VIH-1 del clado A, Gag, RT y Nef ("GRN") del VIH-1 del clado A, o Env del VIH-1 del clado A.

20 2. Una secuencia consenso de nucleótidos de acuerdo con el párrafo 1, en la que la proteína codificada Gag tiene la secuencia de aminoácidos de la figura 1.

3. Una secuencia consenso de nucleótidos de acuerdo con el párrafo 1, en la que la proteína codificada Pol tiene la secuencia de aminoácidos de la figura 3.

4. Una secuencia consenso de nucleótidos de acuerdo con el párrafo 1, en la que la proteína codificada Env tiene la secuencia de aminoácidos de la figura 5.

25 5. Una secuencia consenso de nucleótidos de acuerdo con el párrafo 1, en la que la proteína codificada Nef tiene la secuencia de aminoácidos de la figura 7.

30 6. Un procedimiento para identificar un antígeno del VIH-1 del clado A de una cepa o un aislado de campo circulante del VIH-1 que tiene una secuencia de aminoácidos que es similar a la secuencia consenso de aminoácidos para ese antígeno del VIH-1 del clado A, que comprende comparar las secuencias de aminoácidos de los antígenos de cepas o aislados de campo circulantes del VIH-1 con la secuencia consenso de aminoácidos para esa proteína y seleccionar un antígeno de cepas o aislados de campo circulantes del VIH-1 que tengan una distancia de proteínas pequeña respecto a la secuencia consenso.

7. Un antígeno del VIH-1 del clado A identificado mediante el procedimiento del párrafo 6.

35 8. Un procedimiento para producir un antígeno transgénico del VIH-1 del clado A que comprende seleccionar un antígeno del VIH-1 del clado A usando el procedimiento del párrafo 6 y mutando la secuencia de nucleótidos que codifica el antígeno donde la mutación anula la función de ese antígeno.

9. Un procedimiento para generar una respuesta inmunitaria contra el VIH-1 que comprende administrar a un sujeto una composición que comprende una secuencia de nucleótidos o antígeno de acuerdo con cualquiera de los párrafos anteriores.

40

REIVINDICACIONES

1. Una composición inmunogénica o de vacuna que comprende un polinucleótido que codifica un antígeno Env del VIH-1 del Clado A, en el que el antígeno Env tiene la secuencia de aminoácidos ilustrada en la Fig. 17 y está codificada por la secuencia de nucleótidos ilustrada en la Fig. 17.
- 5 2. Una composición inmunogénica o de vacuna de acuerdo con la reivindicación 1, que comprende además agentes humectantes o emulgentes, agentes de tamponamiento o adyuvantes.
3. Una composición inmunogénica o de vacuna de acuerdo con la reivindicación 1 ó 2, que comprende además uno o más vehículos y / o excipientes farmacéuticamente aceptables.
- 10 4. Una composición inmunogénica o de vacuna de acuerdo con cualquiera de las reivindicaciones 1 a 3, en la que dicha composición es una suspensión inyectable, una solución, un aerosol, un polvo liofilizado, un sirope, un elixir, o una emulsión de aceite-en-agua.
- 15 5. Una composición inmunogénica o de vacuna de acuerdo con cualquiera de las reivindicaciones 1 a 4, en la que dicha composición está diseñada para introducir el polinucleótido que codifica un antígeno Env del VIH-1 del Clado A a un sitio de acción deseado y liberarlo a una velocidad adecuada y controlable.
- 20 6. Una composición inmunogénica o de vacuna de acuerdo con la reivindicación 1 a 5, en la que dicha composición es adecuada para ser administrada a un sujeto antes de la infección por VIH, o evidencia de infección por VIH, o antes de cualquier síntoma debido al SIDA.
7. Una composición inmunogénica o de vacuna de acuerdo con la reivindicación 1 a 6, en la que dicha composición es adecuada para ser administrada por vía intramuscular, intravenosa, intradérmica, mucosal, o tópica.
8. La composición inmunogénica o de vacuna de la reivindicación 1 para su uso en un procedimiento para generar una respuesta inmunitaria contra el VIH-1.
9. La composición inmunogénica o de vacuna de acuerdo con la reivindicación 8, en la que dicha composición es para ser administrada una o más veces a un sujeto.

Consenso Gag

MGARASVLSGGKLDWEKIRLRPGGKKYRLKHLVWASRELERFALNPSLLETAEGCQQIM
EQLQPALKTGTEELRSLFNTVATLYCVHQRIDVKDTKEALDKIEEIQNKSKQK---TQQ--
AAADTGXSSKVS---
QNYPIVQNAQQGMIHQXLSRPTLNAWVKVIEEKAFSPEVIPMFSALSEGATPODLNMMMLNIVG
GHQAAMQMLKDTINEEAEDRLHPVHAGPIPPGQMPREPRGSDIAGTTSTPQEQGAWMTG
NPIPIVGDIIYKRWILGLNKIVRMYSPVILDIKQGPKEPFRDYVDRFFKTLRAEQATQEVKGW
MTETLLVQANAPDCKSILRALGXGATLEEMMTACQGVGGPGGHKARVLAEMSQQQTN--
IMM-QRGNFRGQKR-
IKCFNCGKEGHLARNCRAPRKKGCWKCQKEGHQMKDCTERQANFLGKIWPSSKGRPGNFP
QSRPEPTAPPAEI.FGMGEEIASPPKQEQK--DREQXXPPLYSLKSLFGNDPLSQ

FIG. 1

Consenso Pol

PQJTLWQRPLVTVKIGGQLKEALLDTGADDTVLEDINLPGKWPKMIGGIGGFIKVKQYD
 QILIEICGKKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPIPIETVPVKKLPGMDGPKV
 KQWPLTEEKIKALTEICTEMEKEGKISKIGPENPNYNTPIFAIKKKDSTKWRKLVDFRELN
 KRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLDEFRKYTAFTIPSTNNEIPG
 IRYQYVLPQGWKGSFAIFQSSMTKILEPFRSKNPEIHYQYMDLTVVGSLEIGQHRTK
 IEELRAHLLSWGFTTPDKKHQKEPPFLWMGYELHPDKWTVQPIXLPEKESWTVNDIQKLV
 GKLNWASQYAGIKVKQLCKLLRGAKALTDIVTLTEEALELEAENREILKDPVHGVYYDP
 SKDLIAEIQKQGQDQWTYQYQEPFKNLKTGKYARKRSAHINDVKQLAEVVQKVVVMESIV
 IWGKTPKFKLPQKETWEIWMMDYWQATWPEWEFVNTIPLVKLWYQLEKDPXGAETFY
 VDGAANRETKLGKAGYVTDGRQKVVSLTETTQKTELHAXLALQDSGSEVNIIVTDSQY
 ALGIIQAQDRSESELVNOIEKLLIGKDKVYLSWVPAHKGIGGNEQVDKL VSSGIRKVLV
 LDGIDKAQEEHERYHSNWRXMASDFNLPPIVAKEIVASCDKCKQLKGEAMHGQVDCSPGIW
 QLDCTHLEGKVLVAHVASGYIEAEVIPAETGQETAYFLKLAGRWPVKVVHTDNGSNF
 TSAAFKACWWANIQQEFGIPYNPQSQGVVESHMKNELKKGQVREQAEHLKTAVQMAVF
 IHNFKKGGIGGY'SAGERIIDIIATDIQTKELQKQITKIQNFVYYRDSRDPWKGPAKL
 LWKGEAVVIQDNSDIKVVPRRKAKIIRDYGKQMGAGDDDCVAGRQDED

FIG. 3

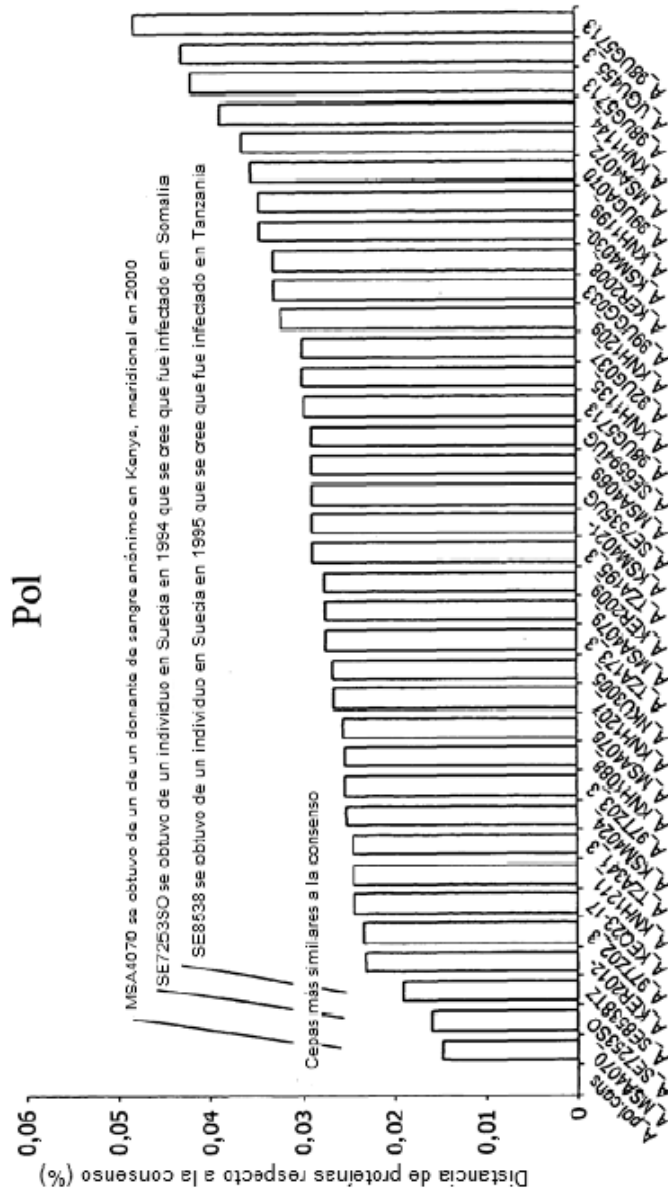


FIG. 4

Consenso Env
 MRV MG IQRNCQHLLRWG-TMILGMIIICS--XAENLWTVTVYGVVWVKDAETTLCASDA
 KAYTEXHNVWATHACVPTDPNPQEI XLNVTEEFNMWKNDMVEQMHTDIISLWDQSLKP
 CVKLTPLCVTLXCXXXXXXXXXXXXXXXXXXXXXXXXXXXX--IKNC-SFNMTTEL RDKKQ
 KVSLSFYRLDVVQI-----XXXXXXXXXXSYRLINCNTSAITQACPKVSFEPIPIHY
 CAPAGFALIKCXDXEFNGTGPKCNVSTVQCTHGKIPVSTQLLLNGSLAE.XXVX-IRSEN
 ITNNAKXIVQLXXPVXINCFRPNNTRK---SIRIGPGQAKYATGDIIGDIRQAHCNVS
 RxxWNXTLQXVAXQLRXXXFXNKTIHFXSSGGDLEITTHSFNCGGEFFYCNTSGLFNST
 W-----XXXXXXXXXXSXXSNDTITLXCRIKQIVNMWQRXGQAMYAPPIQGVIRCES
 NITGLILTRDGGXXXXXXXXNETFRPGGDMRDNRSELKYKVVV KIEPLGVAPTRAKRR
 VVEREKRAV-GIGAVFLGFLGAAGSTMGAASITLTVQARQLLSGIVQQQSNLLRAIEAQQ
 HLLKLTVWGKQLQARVLAVERYLRDQQLGIWGCSGKLICTINVPWNSSWSNKSXXEIW
 DNMTWLQWDKEISNYTQIIYXLIIEESQNQKEKNEQDLLALDKWANLWNWFDISNWL YWIK
 IFIMVGGGLIGLRIVFAVLSIINRVRQGYSPFSFQTHTPNPRGLDRPGRIIEEGEGEQGRD
 RSIRLVSGFLALAWDDDLRSLCLFSYHRLRDFILIAARTVELLGHSSLKGLRLGWEGLYL
 WNLLXYWGRELKISAINLXDTIAIAVAGWTDTRVIEIGQRIGRAILHIPRRIRQGLERALL

FIG. 5

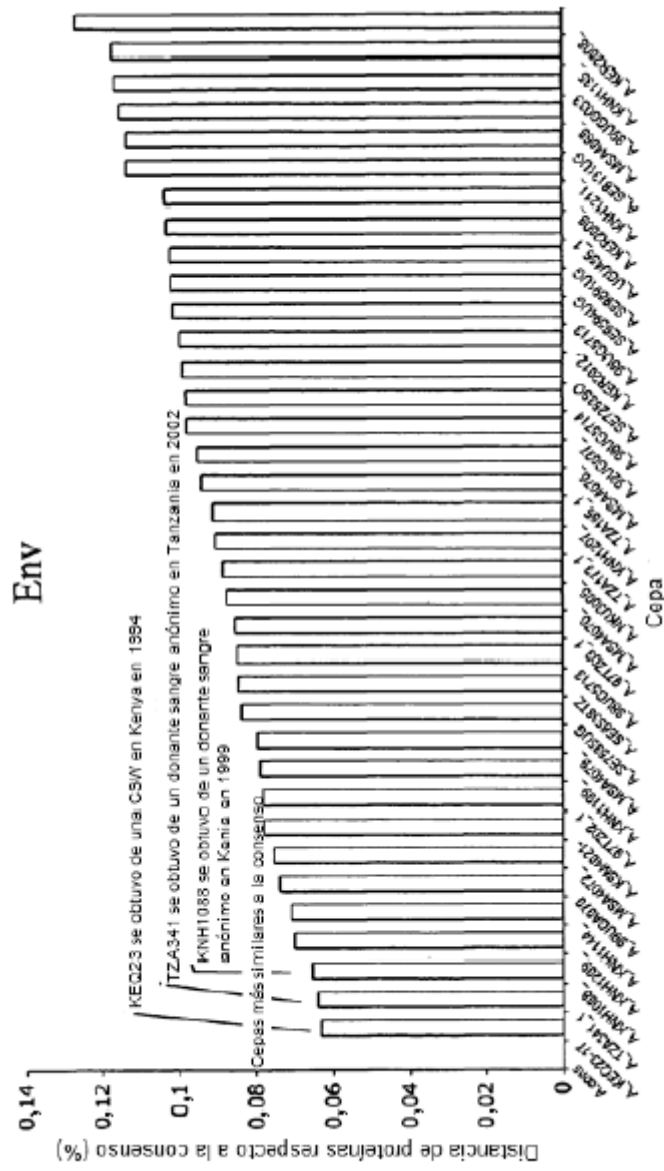


FIG. 6

Consenso .Nef
 MGGKWSKSSIVGWPEVRERMRRTPXAAAX-----
 GVGAVSQDLKKGHGAITSSNIN
 H--PSCVWLEAQEEEE--
 VGFFVRPQVFLRPMITYKGAXDLSHFLKEKGGLDGLIYSRK
 RQ
 EILDLVVYHTIQGYFPDWQNYTPGPGXRYPLTFGWCFKLV
 PVDPEVEKATEGENNSLLHP
 ICQHGMDDDEEREVLXWKFDSSLALAKHRAXELHPEFYKD

Se muestra la secuencia consenso del 50%. Hay 5 posiciones en las que no se pudo alcanzar el 50% de consenso. La media de la distancia de proteínas en nef fue del 9,3%, intervalo del 3,2% al 16,1%.

FIG. 7

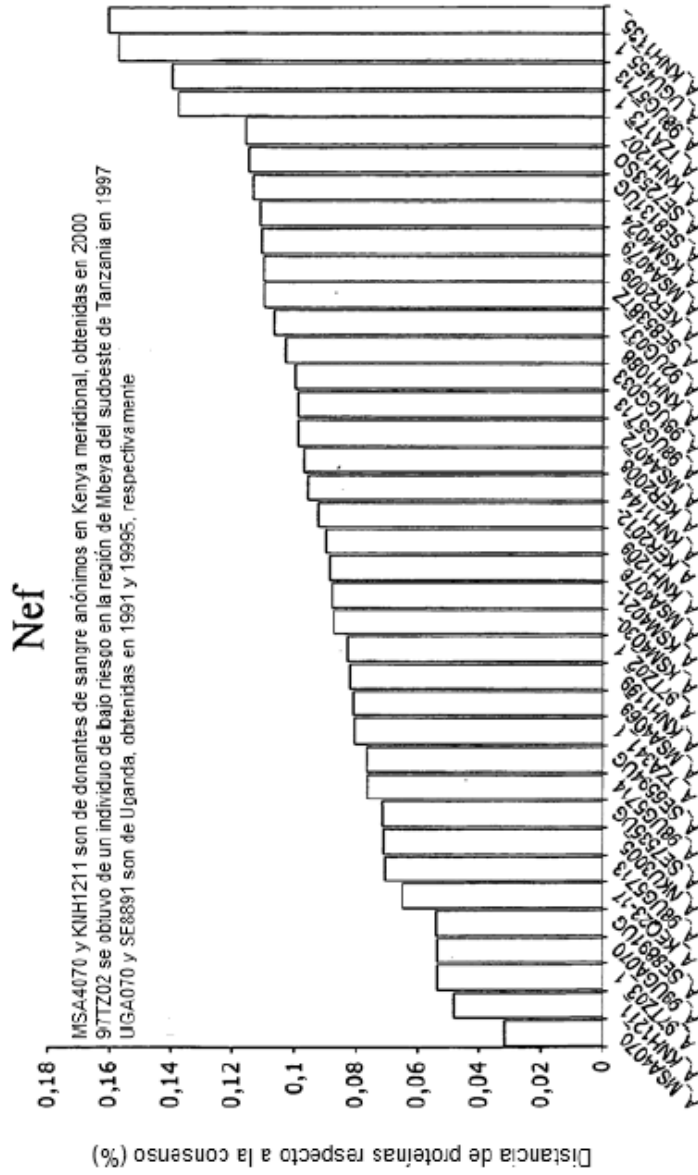


FIG. 8

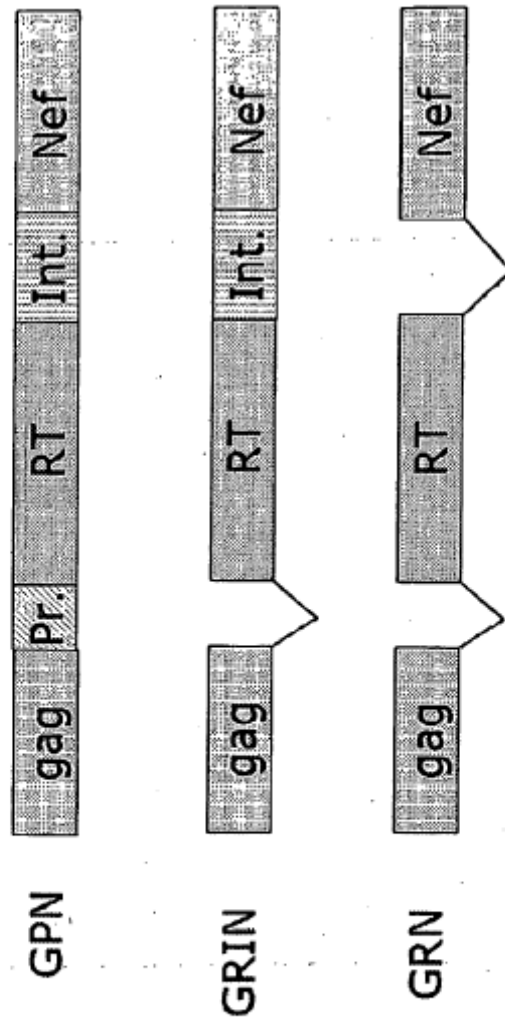


FIG. 9

GAG

AY253305 8766 bp ADN lineal VRL 26-AGOSTO-2004

Cepa 01TZA173 del VIH-1 de Tanzania: genes de la proteína gag (gag) y la proteína pol (pol), cds parciales; y genes de la proteína vif (vif), la proteína vpr (vpr), la proteína tat (tat), la proteína rev (rev), la proteína vpu (vpu), la glucoproteína de la envoltura (env) y la proteína (nef), cds completos

*MG ARASILSGGKLLDAWEKIRLRPGGKKKYRLKHLVWASRELDRLFAL
 NPSLLETTEGCQQIMNQLQPAVKTGTTEEIKSLFNTVATLYCVHQRIDVKDTKEALDKI
 EEIQNKSKQKTQQAADTGDSSKVSQNYPIQNAQQQMIHQNLSPRTLNAWVKVIEEK
 AFSPEVIPMFSALSEGATPQDLNVMNLNIVGGHQAAMQMLKDTNEEAAEWDRLHPVQA
 GPIPPGQIREPRGSDIAGTTSTPQEQLQWMTGNPPIPVGNLYKRWILGLNKIVRMYS
 PVSILDIKQGPKEFRDYVDRFFKALRAEQATQDVKGWMTETLLVQNAHPDCKSILKA
 LGSGATLEEMMTACQGVGGPGHKARVLAEMSQAQQTNIMMQRGNFRGQKRKCFNCG
 KEGHLARNCRAPRKKGCWCKEGHQMKDCITERQANFLGKIWPSSKGRPGNFPQSRPE
 PTAPPAELFGMGEGIASLPKQEQLDREQVPPPLVSLKSLFGNDPLSQ

*MG falta en la entrada de Genbank, artefacto del iniciador amplicon

FIG. 10

POL

AF457081 8827 bp ADN lineal VRL 11-octubre-2002
 Cepa 00KE_MSA4070 del VIH-1 de Kenya, genoma parcial

PQIILWQRPLVTVKIGGQLKEALLDTGADDTVLEDINLPGKWKPRM
 IGGIGGFIKVKQYDQILIEICGKKAIGTVLVGPTPVNIIGRNMLTQIGCTLNFPISPI
 ETVPVTLKPGMDGPKVKQWPLTEEKIKALTEICTEMEKESKISKIGPENPYNTPIFAI
 KKKDSTKWRKLVDFRELNKRITQDFWEVQLGIPHPAGLKKKKSVTVLDVGDAYFSVPLD
 ENFRKYTAFTIPSTNNETPGVRYQYNVLPQGWKGSFAIFQSSMTKILEPFRSKNPEII
 IYQYMDLDLYVGSDEIGQHRTKIEELRAHLLSWGFTTPDKKHQKEPPFLWMGYELHPD
 KWTVPIMLPDKESWTVNDIQKLVGKLNWASQIYAGIKVVKQLCRLLRGAKALTDIVTL
 TEEAELELAENREILKDPVHGYYDPSKDLVAEIQKQGQDQWYQIYQEPFKNLKTGK
 YARKRSAHTNDVRQLAEVVQKVAMESIVWVKTPKFKLPIQKETWETWWMQDYWQATWI
 PEWEFVNTPLVKLWYQLEKDPILGAETFYVDGAANRETKLKGAGYVTDGRQKVVSL
 TETTQKTELHAILLALQDSGSEVNIIVDSQYALGIHQAPDRSESELVNQIEKLLIG
 KDKIYLSWVPAHKGIGGNEQVDKLVSSGIRKVLFLDGIDKAQEDHERYHSNWRTMASD
 FNLPPIVAKEIVASCDCQKLGAEAMHGQVDCSPGIWQLDCTHLEKGVILVAVHVASGY
 IEAEVIPAETGQETAFFLLKLAGRWPVKVHTDNGSNFTSAAVKACWVANIQQEFGI
 PYNPQSQGVVESHMKNELKKIIGQVRDQAEHLKTAVQMAVFIHFNFRKKGIGGY'SAGER
 IIDIIATDIQTKELQKQITKIQNFRVY'YRDSRDPWKGPAKLLWKGE'GAVVIQD'NSDI
 KVVPRRKAKILRDYGYKQMGAGDDCVAGRQDED

FIG. 11

NEF

AF457081 8827 bp ADN lineal VRL 11-octubre-2002
Cepa 00KE_MSA4070 del VIH-1 de Kenya, genoma parcial

MGGKWSKGSIVGWPEIRERMRRAPAAAPGVGA VSQDLDKHGAI
SSNINNPSCVWLEAQEEEEVGFVVRPQVPLRPMTYKGAFDLSHFLKEKGGLDGLIYSR
KRQEILDWVYHTQGYFPDWQNYTPGPGVRYPLTFGWCFKLVPMEPDEVEKATEGENN
SLLHPICQHGMDDEEREVLIWKFDSRLALKHRAQELHPEFYKDC

FIG. 12

ENV

AY253314_8758_BP_ADN_LINEAL_VRL_26-AGOSTO-2004

Cepa 01TZA341 del VIH-1 de Tanzania: genes de la proteína gag (gag) y la proteína pol (pol), cds parciales; y genes de la proteína vif (vif), la proteína vpr (vpr), la proteína tat (tat), la proteína rev (rev), la proteína vpu (vpu), la glucoproteína de la envoltura (env) y la proteína (nef), cds completos

La secuencia Env gp140 siguiente NO incluye la región transmembrana

```

MRVMEIQRNCQHLLRWGIMLGMIIICSTADNLWVTVYYGVPVW
RDAETTLFCASDAKAYSTEKHNVWATHACVPIDPNPQEIPLDNTVEFNMWKNNMVDQ
MHEDIISLWDQSLKPCVQLTPLCVTLNCSNARVNATFNSTEDREGMKNC SFNMTTELK
DKKQQVYSLFYRLDIEKINSSNNSEYRLVNCNTSAITQACPVTPEPIPHYCAPAG
FAILKCNDFEFGTGPCKNVSTVQCTHGKIPVSTQLLNGSLAEREVRIRSENIANN
AKNIIVQFASPVKINCRPNNTKSYRIGPGQTFYATDIVGDIRQAHCNVSRIDWNN
TLRLVANQLRKYFSNKTIIFINSSGGDEIITHSFCGGEFFYCNLSGLFNSTWTINN
MQESNDTSNGHTILPCRKQIIRMWQRVQGQAMYAPPIEGVIRCESNTGLILTRDGGN
NNSANETFRPGGGDIRDNWRSELYKYKVVVIEPLGVAPTRAKRRVVEREKRAVGIGAV
FLGFLGAAAGSTMGAASITLVQARQLLSGIVQQSNLRAIEAQQQLKLTVVWGIKQL
QARVLAVERYLRDQQLGIWGCCKLICITNVPWSSWSNKSYYDDIWNQNMVTLQWDKE
ISNYTDIHSYLIIEESQNQQEKNEQDLLALDKWANLWNWFDISKWLWYI
    
```

FIG. 13

Secuencia ensamblada del inserto de GRIN (DZU36984)

AGTCTTCTGTTTTACGTAGGTGTCAGCCTAGGTGGTCAATATTGGCCATTAGCC
 ATATTATTCATTGGTTATATAGCATAAATCAATATTGGCTATTGGCCATTGCATAC
 GTTGTATCCATATCATAATATGTACATTTATATTGGCTCATGTCCAACATTACCGC
 CATGTTGACATTGATTATTGACTAGTTATTAATAGTAATCAATTACGGGGTCATTA
 GTTCATAGCCCATATATGGAGTTCGCGTTACATAACTTACGGTAAATGGCCCGC
 CTGGCTGACCGCCCAACGACCCCGCCATTGACGTCAATAATGACGTATGTTCC
 CCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTAC
 GGTAACCTGCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCC
 CTATTGACGTCAATGACGGTAAATGGCCCGCTGGCATTATGCCCAGTACATGA
 CCTTATGGGACTTTCCACTTGGCAGTACATCTACGTATTAGTCATCGCTATTAC
 CATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTTGACTC
 ACGGGGATTTCCAAGTCTCCACCCATTGACGTCAATGGGAGTTTGTGGCA
 CCAAAATCAACGGGACTTTCCAAAATGTCGTAACAACCTCCGCCCATGACGCA
 AATGGGCGGTAGGCGGTACGGTGGGAGGTCTATATAAGCAGAGCTCGTTTAGT
 GAACCGTCAGATCGCCTGGAGACGCCATCCACGCTGTTTTGACCTCCATAGAAG
 ACACCGGGACCGATCCAGCCTCCGCGGCGGGAACGGTGCATTGGAAGCTTG
 CCGCCACCATGGCCGCCAGAGCCAGCATCCTGAGCGGGGGCAAGCTGGACGC
 CTGGGAGAAGATCAGACTGAGGCCTGGCGGCAAGAAGAAGTACCGGCTGAAGC
 ACCTGGTGTGGGCCAGCAGAGAGCTGGATCGCTTCGCCCTGAATCCTAGCCTG
 CTGGAGACCACCGAGGGCTGCCAGCAGATCATGAACCAGCTGCAGCCCGCCGT
 GAAAACCGGCACCGAGGAGATCAAGAGCCTGTTCAACACCGTGGCCACCCTGT
 ACTCGGTGACCCAGCGGATCGACGTGAAGGATACCAAGGAGGCCCTGGACAAG
 ATCGAGGAGATCCAGAACAAGAGCAAGCAGAGAAAACCCAGCAGGCCCGCTGCCGA
 CACCGGCGACAGCAGCAAAGTGAGCCAGAACTACCCCATCATCCAGAATGCC
 AGGGCCAGATGATCCACCAGAACCTGAGCCCCAGAACCCTGAATGCCTGGGTG
 AAAGTGATCGAGGAAAAGGCCTTCAGCCCCGAAGTGATCCCTATGTTTCAGCGCC
 CTGAGCGAGGGCGCCACCCCCAGGACCTGAACGTGATGCTGAACATTGTGGG
 CGGACACCAGGCCGCCATGCAGATGCTGAAGGACACCATCAATGAGGAGGCCG
 CCGAGTGGGACAGACTGCACCCCGTGCAGGCCGGACCCATCCCCCTGGCCA
 GATCAGAGAGCCCAGAGGCAGCGACATCGCCGGCACCACCTCCACCCCTCAAG
 AACAGCTGCAGTGGATGACCGGCAACCCTCCCATCCCTGTGGGCAACATCTACA
 AGCGGTGGATCATCCTGGGCCTGAACAAGATTGTGCGGATGTACAGCCCCGTG
 TCCATCCTGGATATCAAGCAGGGCCCCAAGGAGCCCTTCAGAGACTACGTGGA
 CCGGTTCTTCAAGGCCCTGAGAGCCGAGCAGGCCACCCAGGACGTGAAGGGCT
 GGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCTGACTGCAAGAGCATC
 CTGAAGGCCCTGGGCAGCGGCCACACTGGAGGAGATGATGACCGCCTGCC
 AGGGAGTGGGCGGACCCGGCCACAAGGCCAGAGTGCTGGCCGAGGCCATGAG
 CCAGGCCAGCAGACCAACATCATGATGCAGCGGGGCAACTTCAGAGGCCAGA
 AGCGGATCAAGTGCTTCAACTGCGGCAAGGAGGGCCACCTGGCCAGAACTGC
 AGAGCCCCAGGAAGAAGGGCTGCTGGAAGTGTGGCAAGGAAGGGCACCAGA
 TGAAGGACTGCACCGAGAGGCAGGCCAATTTCTGGGCAAGATTTGGCCTAGC

FIG. 14A

AGCAAGGGCAGACCCGGCAATTTCCCCCAGAGCAGACCCGAGCCCACCGCCCC
 TCCCCCGGAGCTGTTTCGGCATGGGCGAGGGCATCGCCAGCCTGCCAAGCAG
 GAGCAGAAGGACAGAGAGCAGGTGCCCCCCCTGGTGTCCCTGAAGTCCCTGTT
 CGGCAACGATCCTCTGAGCCAGGGATCCCCCATCAGCCCCATCGAGACCGTGC
 CCGTGACCCTGAAGCCCGGCATGGATGGCCCCAAAGTGAAACAGTGGCCCCTG
 ACCGAGGAGAAGATTAAGGCCCTGACCGAAATCTGTACCGAGATGGAGAAGGA
 GGGCAAGATCAGCAAGATCGGCCCGGAGAACCCCTACAACACCCCCATCTTCG
 CCATCAAGAAGAAGGACAGCACCAAGTGGCGGAAACTGGTGGACTTCCGGGAG
 CTGAACAAGAGGACCCAGGACTTCTGGGAAGTGCAGCTGGGCATCCCCACCC
 TGCCGGCCTGAAGAAGAAGAAGTCCGTGACAGTGTGGATGTGGGCGACGCCT
 ACTTCAGCGTGCCCCCTGGACGAGAACTTCAGGAAGTACACCGCCTTACCATCC
 CCAGCACCAACAACGAGACCCCCGGAGTGAGATAACAGTACAACGTGTGCCT
 CAGGGCTGGAAGGGCAGCCCCGCCATCTTCCAGAGCAGCATGACCAAGATCCT
 GGAGCCCTTCCGGAGCAAGAACCCCGAGATCATCATCTACCAGTACATGGCCG
 CCCTGTATGTGGGCAGCGATCTGGAGATCGGCCAGCACAGGACCAAGATCGAA
 GAGCTGAGGGCCACCTGCTGAGCTGGGGCTTACCACCCCCGATAAGAAGCA
 CCAGAAGGAGCCCCCTTCTGTGGATGGGCTACGAGCTGCACCCCGATAAGT
 GGACCGTGCAGCCCATCATGCTGCCCGATAAGGAGAGCTGGACCGTGAACGAC
 ATCCAGAAACTGGTGGGCAAGCTGAATTGGGCCAGCCAAATCTACGCCGGCATT
 AAAGTGAAGCAGCTGTGCAGGCTGCTGAGAGGCCCAAAGCCCTGACAGACAT
 CGTGACACTGACAGAGGAGGCCGAGCTGGAGCTGGCCGAGAACAGGGAGATC
 CTGAAGGACCCCGTGCACGGCGTGTACTACGACCCAGCAAGGACCTGGTGGC
 CGAGATTCAGAAGCAGGGCCAGGACCAGTGGACCTACCAAATCTACCAGGAGC
 CTTTCAAGAACCTGAAAACCGGGAAGTACGCCAGGAAGAGAAGCGCCACACC
 AACGATGTGAGGCAGCTGGCCGAAGTGGTGCAGAAAGTGGCTATGGAGAGCAT
 CGTGATCTGGGGCAAGACCCCAAGTTCAAGCTGCCCATCCAGAAGGAGACCT
 GGGAAACCTGGTGGATGGACTACTGGCAGGCCACCTGGATTCTGAGTGGGAG
 TTCGTGAACACCCCCCTCTGGTGAAGCTGTGGTATCAGCTGGAGAAGGACCC
 CATCCTGGGCGCCGAGACCTTCTACGTGGACGGAGCCGCCAATAGAGAGACCA
 AGCTGGGCAAGGCCGGCTACGTGACCGACAGAGGCAGACAGAAAGTGGTGTCT
 CTGACCGAGACAACCAACCAGAAAACCGAGCTGCACGCCATCCTGCTGGCCCT
 GCAGGACAGCGGCAGCGAAGTGAACATCGTGACCGACTCCCAGTACGCCCTGG
 GCATCATTGAGGCCAGCCCGATAGAAGCGAGAGCGAGCTGGTGAACCAGATC
 ATCGAGAAGCTGATCGGCAAGGACAAAATCTACCTGAGCTGGGTGCCCGCCCA
 CAAGGGCATCGGCGGCAACGAGCAGGTGGACAAGCTGGTGTCCAGCGGCATC
 CGGAAAGTGCTGTTTCTGGACGGCATCGACAAGGCCAGGAGGACCACGAGAG
 ATACCACAGCAACTGGCGGACAATGGCCAGCGACTTCAACCTGCCTCCCATCGT
 GGCCAAGGAGATCGTGGCCAGCTGCGATAAGTGTGAGCTGAAGGGCGAGGCCA
 TGCACGGCCAGGTGGACTGCAGCCCTGGCATCTGGCAGCTGGCCTGCACCCAC
 CTGGAGGGCAAAGTGATTCTGGTGGCCGTGCACGTGGCCAGCGGCTACATCGA
 GGCCGAAGTGATTCCC GCCGAGACCGGCCAGGAGACCGCCTACTTCTGCTGA
 AGCTGGCCGGCAGATGGCCCGTGAAGTGGTGCACACCGCCAACGGCAGCAA
 CTTACCTCTGCCGCCGTGAAGGCCGCTGTTGGTGGGCCAATATCCAGCAGG

FIG. 14B

AGTTCGGCATCCCCTACAACCCTCAGAGCCAGGGCGTGGTGGCCAGCATGAAC
AAGGAGCTGAAGAAGATCATCGGCCAGGTGAGGGACCAGGCCGAGCACCTGAA
AACAGCCGTGCAGATGGCCGTGTTTCATCCACAACCTCAAGCGGAAGGGCGGCA
TTGGCGGCTACAGCGCCGGAGAGCGGATCATCGACATCATCGCCACCGATATC
CAGACCAAGGAACTGCAGAAGCAGATCACCAAGATTCAGAACTTCAGAGTGTAC
TACCGGGACAGCAGGGACCCCATCTGGAAGGGCCCTGCCAAGCTGCTGTGGAA
GGGCGAAGGGCGCCGTGGTGATCCAGGACAACAGCGACATCAAAGTGGTGCCCC
GGAGGAAGGCCAAGATTCTGCGGGACTACGGCAAACAGATGGCCGGCGATGAC
TGCCTGGCCGGCAGGCAGGATGAGGACAGATCTATGGGCGGCAAGTGGTCCAA
GGGCAGCATTGTGGGCTGGCCCAGATCCGGGAGAGAATGAGAAGAGCCCT
GCCGCCGCTCCTGGAGTGGGCGCCGTGTCTCAGGATCTGGATAAGCACGGCG
CCATCACCAGCAGCAACATCAACAACCCAGCTGTGTGTGGCTGGAGGCCAG
GAAGAGGAGGAAGTGGGCTTCCCTGTGAGACCCAGGTGCCCTGAGACCCAT
GACCTACAAGGGCGCCTTCGACCTGAGCCACTTCTGAAGGAGAAGGGCGGCC
TGGACGGCCTGATCTACAGCCGGAAGCGGCAGGAGATCCTGGATCTGTGGGTG
TACCACACCCAGGGCTACTTCCCCGACTGGCAGAATTACACCCCTGGCCCTGGA
GTGCGGTATCCCCTGACCTTCGGCTGGTGCTTCAAGCTGGTGCCTATGGAGCC
CGACGAAGTGGAGAAGGCCACAGAGGGCGAGAACAACAGCCTGCTGCACCCTA
TCTGCCAGCACGGCATGGACGATGAGGAGCGGGAAGTGTGATCTGGAAGTTC
GACAGCAGGCTGGCCCTGAAGCACAGAGCCCAGGAACTGCACCCAGAGTTCTA
CAAGGACTGCTGATGATCATAATA**TCTAGAC**GAGATCCGAACTTGTTTATTGCA
GCTTATAATGGTTACAAATAAAGCAATAGCATCACAAATTCACAAATAAAGCATT
TTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTATCATG
TCTAGATCTGAGGTATGATGATACGAGATCGAGGGTGCGCATGCGAATGCG
GAGGCAAGCATGCCAGGTTCCAGC

Nota: los sitios de clonación están subrayados y en negrita

Apéndice 2: Secuencia ensamblada del inserto de Env (DSP33447_01)

ATTATTGAAGCATTATCAGGGTTATTGTCTCATGAGCGGATACATATTTGAATG
TATTTAgAAAAATAACAAATAGGGTTCCGCGCACATTTCCCGAAAAGTGCCA
CCTGACGTCTAAGAAACCATIATTATCATgACATTAACCTATAAAAAATAGGCGIAT
CACGAGGCCCTTCGTCTTCAAGAATTGGTCGATGGCAAACAGCTATtATGGGTA
TTATGGGTTTCGAATTAATTAATCGACATCATCAATAATATACCTTATAGATGGAAT
GGTGCCAATATGTAAATGAGGTGATTTTAAAAAGTGTGGGCCGTGTGGTGATTG
GCTGTGGGGTTAACGGTTAAAAGGGGCGGCGCGGCCGTGGGAAAATGACGTT
TTATGGGGGTGGAGTTTTTTTGAAGTTGTCGCGGGAAATGTTACGCATAAAAA
GGCTTCTTTTCTCACGGAACACTTAGTTTTCCACGGTATTTAACAGGAAATGA
GGTAGTTTTGACCGGATGCAAGTGAAAATTGCTGATTTTCGCGCGAAAACCTGAA
TGAGGAAGTGTTTTCTGAATAATGTGGTATTTATGGCAGGGTGGAGTATTTGTT
CAGGGCCAGGTAGACTTTGACCCATTACGTGGAGGTTTCGATTACCGTGTTTTT
TACCTGAATTTCCGCGTACCGTGTCAAAGTCTTCTGTTTTTACGTAGGTGTCAGC
CTAGGTGGTCAATATTGGCCATTAGCCATATTATTCATTGGTTATATAGCATAAA
TCAATATTGGCTATTGGCCATTGCATACGTTGTATCCATATCATAATATGTACATT
TATATTGGCTCATGTCCAACATTACCGCCATGTTGACATTGATTATTGACTAGTT
ATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCG
CGTTACATAACTTACGGTAAATGGCCCCGCTGGCTGACCGCCCAACGACCCCC
GCCATTGACGTCAATAATGACGTATGTTCCCATAGTAACGCCAATAGGGACTT
TCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACTGCCCACTGGCAGTAC
ATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAAAT
GGCCCCGCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCTACTTGGC
AGTACATCTACGTATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTA
CATCAATGGGCGTGGATAGCGGTTTACTCACGGGGATTTCCAAGTCTCCACC
CCATTGACGTCAATGGGAGTTTGTGTTTTGGCACAAAATCAACGGGACTTTCCAA
AATGTCGTAACAACCTCCGCCCATGACGCAAATGGGCGGTAGGCGTGTACGG
TGGGAGGTCTATATAAGCAGAGCTCGTTTTAGTGAACCGTCAGATCGCCTGGAG
ACGCCATCCACGCTGTTTTGACCTCCATAGAAGACACCGGGACCGATCCAGCC
TCCGCGGCCGGGAACGGTGCATTGGAAGCTTGCCGCCACCATGAGGGTGATG
GAGATCCAGCGGAACCTGCCAGCACCTGCTGAGATGGGGCATCATGATCCTGGG
CATGATTATCATCTGCAGCACCGCCGACAACCTGTGGGTGACCGTGTACTACG
GCGTGCCTGTGTGGAGAGATGCCGAGACCACCCTGTTCTGCGCCAGCGACGC
CAAGGCCTACAGCACCGAGAAGCACAATGTGTGGGCCACCCACGCCTGCGTG
CCTACCGATCCCAACCCTCAGGAGATCCCCCTGGACAACGTGACCGAGGAGTT
CAACATGTGGAAGAACAACATGGTGGACCAGATGCACGAGGACATCATCAGCC
TGTGGGACCAGAGCCTGAAGCCCTGCGTGCAGCTGACCCCCCTGTGCGTGAC
CCTGAACTGCAGCAACGCCAGAGTGAACGCCACCTTCAACTCCACCGAGGACA
GGGAGGGCATGAAGAACTGCAGCTTCAACATGACCACCGAGCTGCGGGATAAG
AAGCAGCAGGTGTACAGCCTGTTCTACCGGCTGGACATCGAGAAGATCAACAG
CAGCAACAACAACAGCGAGTACCGGCTGGTGAACCTGCAATACCAGCGCCATCA
CCCAGGCCTGCCCTAAGGTGACCTTCGAGCCCATCCCCATCCACTACTGCGCC

FIG. 15A

CCTGCCGGCTTCGCCATCCTGAAGTGCAACGACACCGAGTTCAATGGCACCGG
 CCCCTGCAAGAATGTGAGCACCGTGCACTGCACCCACGGCATCAAGCCCGTG
 GTGTCCACCCAGCTGCTGCTGAACGGCAGCCTGGCCGAGAGAGAAGTGC
 TCAGGAGCGAGAACATCGCCAACAACGCCAAGAACATCATCGTGAGTTCGCC
 AGCCCCGTGAAGATCAACTGCATCCGGCCCAACAACAATACCCGGAAGAGCTA
 CAGAATCGGCCCTGGCCAGACCTTCTACGCCACCGACATTGTGGGCGACATCA
 GACAGGCCCACTGCAACGTGTCCAGGACCGACTGGAACAACACCCTGAGACTG
 GTGGCCAACCCAGCTGCGGAAGTACTTCAGCAACAAGACCATCATCTTCACCAAC
 AGCAGCGGCGGAGACCTGGAGATCACCACCCACAGCTTCAATTGTGGCGGCG
 AGTTCTTCTACTGCAACACCTCCGGCCCTGTTCAATAGCACCTGGACCACCAACA
 ACATGCAGGAGTCCAACGACACCAGCAACGGCACCATCACCTGCCCTGCCGG
 ATCAAGCAGATCATCCGGATGTGGCAGCGCGTGGCCAGGCCATGACGCC
 CTCCCATCGAGGGCGTGATTCGCTGCGAGAGCAACATCACCGGCCTGATCTG
 ACCAGAGATGGCGGCAACAACAATTCCGCCAACGAGACCTTCAGACCTGGCGG
 CGGAGATATCCGGGACAACCTGGCGGAGCGAGCTGTACAAGTACAAGGTGGT
 AAGATCGAGCCCCTGGGCGTGGCCCCACCAGAGCCAAGAGAAGAGTGGTGG
 AGCGGGAGAAGAGAGCCGTGGGCATCGGCGCCGTGTTTCTGGGCTTCTGGG
 AGCCGCCGGATCTACAATGGGAGCCGCCAGCATCACCTGACCGTGCAGGCC
 AGACAGCTGCTGAGCGGCATCGTGACGAGCAGAGCAATCTGCTGAGAGCCAT
 CGAGGCCCAGCAGCAGCTGCTGAAGCTGACAGTGTGGGGCATCAAGCAGCTG
 CAGGCCAGGGTGTGGCCGTGGAGAGATACCTGAGGGACCAGCAGCTCCTGG
 GCATCTGGGGCTGCAGCGCAAGCTGATCTGCACCACCAACGTGCCCTGGAAT
 AGCAGCTGGAGCAACAAGAGCTACGACGACATCTGGCAGAACATGACCTGGCT
 GCAGTGGGACAAGGAGATCAGCAACTACACCGACATCATCTACAGCCTGATCG
 AGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGATCTGCTGGCCCTGGA
 CAAGTGGGCCAACCTGTGGAAGTGGTTCGACATCAGCAAGTGGCTGTGGTACA
 TCAGATCTTGATAATCTAGACGAGATCCGAACTTGTTTATTGCAGCTTATAATGG
 TTACAAATAAAGCAATAGCATCACAAATTTACAAATAAAGCATTTTTTTCACTGC
 ATTCTAGTTGTGGTTTGTCCAACTCATCAATGTATCTTATCATGTCTAGATCTGA
 GGTATGATGATACGAGATCGAGGGTGCAGCATGCGAATGCGGAGGCAAGCA
 TGCCAGGTTCCAGCCGGTGTGTGTAGATGTGACCGAAGATCTCAGACCGGATC
 ATTTGGTTATTGCCCGCACTGGAGCAGAGTTCGGATCCAGTGGAGAAGAACT
 GACTAAGGTGAGTATTGGGAAAACCTTTGGGGTGGGATTTTCAGATGGACAGATT
 GAGTAAAAATTTGTTTTTCTGTCTTGCAGCTGACATGACTGGAAATGCTTCTTT
 TAAGGGGGGAGTCTTCAGCCCTTATCTGACAGGGCGTCTCCATCCTGGGCA
 GGAGTTCGT

Nota: los sitios de clonación están subrayados y en negrita

FIG. 15B

```

HindIII      NcoI                                  BstNI
1  AAGCTTGCCGCCACCATGGCCGCCAGAGCCAGCATCCTGAGCGGGGCAAGCTGGACGCC
   +-----+-----+-----+-----+-----+-----+
1  TTCGAACGGCGGTGGTACCGGCGGTCTCGGTCTGTTAGGACTCGCCCCGTTCGACCTGCGG
   M  A  A  R  A  S  I  L  S  G  G  K  L  D  A

                               StuI
                               BstNI
61  TGGGAGAAGATCAGACTGAGGCCCTGGCGGCAAGAAGAAGTACCGGCTGAAGCACCTGGTG
   +-----+-----+-----+-----+-----+-----+
61  ACCCTCTTCTAGTCTGACTCCGGACCGCCGTTCTTCTTCATGGCCGACTTCGTGGACCAC
   W  E  K  I  R  L  R  P  G  G  K  K  K  Y  R  L  K  H  L  V

                                               HinfI      BsaI
121  TGGGCCAGCAGAGAGCTGGATCGCTTCGCCCTGAATCCTAGCCTGCTGGAGACCACCGAG
   +-----+-----+-----+-----+-----+-----+
121  ACCCGGTCGTCTCTCGACCTAGCGAAGCGGGACTTAGGATCGGACGACCTCTGGTGGCTC
   W  A  S  R  E  L  D  R  F  A  L  N  P  S  L  L  E  T  T  E

                               PvuII
                               PstI
181  GGCTGCCAGCAGATCATGAACCAGCTGCAGCCCGCCGTGAAAACCGGCACCGAGGAGATC
   +-----+-----+-----+-----+-----+-----+
181  CCGACGGTCGTCTAGTACTTGGTCGACGTCGGGCGGCACFTTTGGCCGTGGCTCCTCTAG
   G  C  Q  Q  I  M  N  Q  L  Q  P  A  V  K  T  G  T  E  E  I

241  AAGAGCCTGTTCAACACCGTGGCCACCCGTACTGCGTGCACCAGCGGATCGACGTGAAG
   +-----+-----+-----+-----+-----+-----+
241  TTCTCGGACAAGTTGTGGCACCGGTGGGACATGACGCACGTGGTCGCCTAGCTGCACTTC
   K  S  L  F  N  T  V  A  T  L  Y  C  V  H  Q  R  I  D  V  K

                               BstNI
301  GATACCAAGGAGGCCCTGGACAAGATCGAGGAGATCCAGAACAAGAGCAAGCAGAAAACC
   +-----+-----+-----+-----+-----+-----+
301  CTATGGTTCCTCCGGGACCTGTTCTAGCTCCTCTAGGTCTTGTTCGTTTCGTTCTTTTGG
   D  T  K  E  A  L  D  K  I  E  E  I  O  N  K  S  K  O  K  T

361  CAGCAGGCCGCTGCCGACACCGGCGACAGCAGCAAAGTGAGCCAGAACTACCCCATCATC
   +-----+-----+-----+-----+-----+-----+
361  GTCGTCCGGCGACGGCTGTGGCCGCTGTCGTGTTTCACTCGGTCTTGATGGGGTAGTAG
   Q  Q  A  A  A  D  T  G  D  S  S  K  V  S  Q  N  Y  P  I  I

                               BstNI                                  BstNI
421  CAGAATGCCAGGGCCAGATGATCCACCAGAACCTGAGCCCCAGAACCTGAATGCCTGG
   +-----+-----+-----+-----+-----+-----+
421  GTCTTACGGTCCCGGTCTACTAGGTGGTCTTGGACTCGGGGCTTGGGACTTACGGACC
   Q  N  A  Q  G  Q  M  I  H  Q  N  L  S  P  R  T  L  N  A  W

```

FIG. 16A

```

                StuI                               HaeII
481 GTGAAAGTGATCGAGGAAAAGGCC TTCAGCCCCGAAGTGATCCCTATG TTCAGCGCCCTG
-----+-----+-----+-----+-----+-----+-----+
CACTTTCAGTACTCCTTTTCCGGAAGTCGGGGCTTCACTAGGGATACAAGTCGCGGGAC
V_K_V_I_E_E_K_A_F_S_P_E_V_I_P_M_F_S_A_L_

                NarI
                KasI
                HaeII          BstNI                               BstNI
541 AGCGAGGGCGCCACCCCCAGGACCTGAACGTGATGCTGAACATTGTGGGCGGACACCAG
-----+-----+-----+-----+-----+-----+-----+
TCGCTCCCGCGGTGGGGGTCTGGACTTGCACTACGACTTGTAACACCCGCCTGTGGTC
S_E_G_A_T_P_Q_D_L_N_V_M_L_N_I_V_G_G_H_Q_

601 GCCGCCATGCAGATGCTGAAGGACACCATCAATGAGGAGGCCGCCGAGTGGGACAGACTG
-----+-----+-----+-----+-----+-----+-----+
CGGCGGTACGCTACGACTTCCTGTGGTAGTTACTCCTCCGGGGGCTCACCTGTCTGAC
A_A_M_Q_M_L_K_D_T_I_N_E_E_A_A_E_W_D_R_L_

                BstNI
661 CACCCCGTGCAGGCCGACCCATCCCCCTGGCCAGATCAGAGAGCCCAGAGGCAGCGAC
-----+-----+-----+-----+-----+-----+-----+
GTGGGGCACGTCGGCCTGGGTAGGGGGACCGGTCTAGTCTCTCGGGTCTCCGTGCTG
H_P_V_Q_A_G_P_I_P_P_G_Q_I_R_E_P_R_G_S_D_

                PvuII
                PstI                               BstXI
721 ATCGCCGGCACCACCTCCACCCCTCAAGAACAGCTGCAGTGGATGACCGGCAACCCCTCC
-----+-----+-----+-----+-----+-----+-----+
TAGCGGCCGTGGTGGAGGTGGGGAGTTCTTGTGCGACGTCACCTACTGGCCGTTGGGAGGG
I_A_G_T_T_S_T_P_Q_E_Q_L_Q_W_M_T_G_N_P_P_

                BstNI
781 ATCCCTGTGGGCAACATCTACAAGCGGTGGATCATCCTGGGCCTGAACAAGATTGTGCGG
-----+-----+-----+-----+-----+-----+-----+
TAGGGACACCCGTTGTAGATGTTGCCACCTAGTAGGACCCGGACTTGTCTAACACGCC
I_P_V_G_N_Y_Y_K_R_W_I_I_L_G_L_N_K_I_V_R_

                EcoRV
                BstNI          ApaI
841 ATGTACAGCCCCGTGTCATCCTGGATATCAAGCAGGGCCCCAAGGAGCCCTTCAGAGAC
-----+-----+-----+-----+-----+-----+-----+
TACATGTCGGGGCACAGGTAGGACCTATAGTTCCGTCGGGGTTCCTCGGGAAGTCTCTG
M_Y_S_P_V_S_I_L_D_I_K_Q_G_P_K_E_P_F_R_D_

                AgeI                               BstNI
901 TACGTGGACCGGTTCTTCAAGGCCCTGAGAGCCGAGCAGGCCACCCAGGACGTGAAGGGC
-----+-----+-----+-----+-----+-----+-----+
ATGCACCTGGCCAAGAAGTTCGGGACTCTCGGCTCGTCCGGTGGGTCTGCACTTCCCG
Y_V_D_R_F_F_K_A_L_R_A_E_Q_A_T_Q_D_V_K_G_

```

FIG. 16B


```

          BsaI
961  TGGATGACCGAGACCCTGCTGGTGCAGAACGCCAACCCCGACTGCAAGAGCATCTGAAG
-----+-----+-----+-----+-----+-----+-----+-----+
ACCTACTGGCTCTGGGACGACCACGTCTTGGCGTTGGGGCTGACGTTCTCGTAGGACTTC
W_M_T_E_T_L_L_V_Q_N_A_N_P_D_C_K_S_I_L_K_

          NarI
          KasI
          PflMI
          BstNI
          HaeII
          BstNI
1021  GCCCTGGGCAGCGGCCACACTGGAGGAGATGATGACCGCCTGCCAGGGAGTGGGCGGA
-----+-----+-----+-----+-----+-----+-----+
CGGGACCCGTCGCCGCGGTGTGACCTCCTTACTACTGGCGGACGGTCCCTCACCCGCCT
A_L_G_S_G_A_T_L_E_E_M_M_T_A_C_Q_G_V_G_G_

          BstXI
          BstNI
1081  CCCGGCCACAAGGCCAGAGTGTGGCCGAGGCCATGAGCCAGGCCAGCAGACCAACATC
-----+-----+-----+-----+-----+-----+-----+
GGGCCGGTGTCCGGTCTCACGACCGGCTCCGGTACTCGGTCCGGGTCGTCTGGTTGTAG
P_G_H_K_A_R_V_L_A_E_A_M_S_Q_A_Q_Q_T_N_I_

1141  ATGATGCAGCGGGGCAACTTCAGAGGCCAGAAGCGGATCAAGTGCTTCAACTGCGGCAAG
-----+-----+-----+-----+-----+-----+-----+
TACTACGTCGCCCCGTTGAAGTCTCCGGTCTTCGCCTAGTTCACGAAGTTGACGCCGTTTC
M_M_Q_R_G_N_F_R_G_Q_K_R_I_K_C_F_N_C_G_K_

          BstNI
          PstI
          BstNI
1201  GAGGGCCACCTGGCCAGAACTGCAGAGCCCCCAGGAAGAAGGGCTGCTGGAAGTGTGGC
-----+-----+-----+-----+-----+-----+-----+
CTCCCGGTGGACCGTCTTTGACGTCCTCGGGGTCTTCTTCCCGACGACCTTCACACCG
E_G_H_L_A_R_N_C_R_A_P_R_K_K_G_C_W_K_C_G_

          BstXI
          BstNI
1261  AAGGAAGGGCACCAGATGAAGGACTGCACCGAGAGGCAGGCCAATTTCTGGGCAAGATT
-----+-----+-----+-----+-----+-----+-----+
TTCCTTCCCGTGGTCTACTTCTGACGTGGCTCTCCGTCGGTTAAAGGACCCGTTCTAA
K_E_G_H_Q_M_K_D_C_T_E_R_Q_A_N_F_L_G_K_I_

1321  TGGCCTAGCAGCAAGGGCAGACCCGCAATTTCCCCCAGAGCAGACCCGAGCCCACCGCC
-----+-----+-----+-----+-----+-----+-----+
ACCGGATCGTCGTTCCCGTCTGGGCCGTTAAAGGGGGTCTCGTCTGGGCTCGGGTGGCGG
W_P_S_S_K_G_R_P_G_N_F_P_Q_S_R_P_E_P_T_A_

1381  CCTCCCGCCGAGCTGTTCCGCATGGGCGAGGGCATCGCCAGCCTGCCAAGCAGGAGCAG
-----+-----+-----+-----+-----+-----+-----+
GGAGGGCGGCTCGACAAGCCGTACCCGCTCCCGTAGCGGTGGACGGGTTGTCCTCGTC
P_P_A_E_L_F_G_M_G_E_G_I_A_S_L_P_K_Q_E_Q_

          BspMI
          BstNI
1441  AAGGACAGAGAGAGGTGCCCCCTGGTGTCCCTGAAGTCCCTGTTCCGCAACGATCCT
-----+-----+-----+-----+-----+-----+-----+
TTCCTGTCTCTCGTCCACGGGGGGACCACAGGGACTTCAGGGACAAGCCGTTGCTAGGA
K_D_R_E_Q_V_P_P_L_V_S_L_K_S_L_F_G_N_D_P_

```

FIG. 16C

```

          BstNI   NcoI
          BamHI
          BsaI   BstEII
1501 CTGAGCCAGGGATCCATGGCCCCCAGATCACCCCTGTGGCAGAGACCCCTGGTGACCGTG
-----+-----+-----+-----+-----+-----+-----+
GACTCGGTCCCTAGGTACCGGGGGTCTAGTGGGACACCGTCTCTGGGGACCACTGGGCAC
L_S_Q_G_S_M_A_P_Q_I_T_L_W_Q_R_P_L_V_T_V_

          NarI
          KasI
          PvuII   HaeII
1561 AAGATCGGCGCCAGCTGAAGGAAGCCCTGCTGGATACAGGCCTGATGATACCGTGCTG
-----+-----+-----+-----+-----+-----+-----+
TTCTAGCCCGGTCGACTTCCTTCGGGACGACCTATGTCCCGGCTACTATGGCAGCAGAC
K_I_G_G_Q_L_K_E_A_L_L_D_T_G_A_D_D_T_V_L_

          BspMI
1621 GAGGACATCAACCTGCCCGCAAGTGAAGCCTAGAATGATCGGCGGCATCGGGGGCTTC
-----+-----+-----+-----+-----+-----+-----+
CTCCTGTAGTTGGACGGGCGTTCACCTTCGGATCTTACTAGCCCGCTAGCCCCGAAG
E_D_I_N_L_P_G_K_W_K_P_R_M_I_G_G_I_G_G_F_

1681 ATCAAAGTGAAGCAGTACGACCAGATCCTGATCGAGATTTGCGGGAAGAAGGCCATCGGC
-----+-----+-----+-----+-----+-----+-----+
TAGTTTCACTTCGTTCATGCTGGTCTAGGACTAGCTCTAAACGCCCTTCTCCGGTAGCCG
I_K_V_K_Q_Y_D_Q_I_L_I_E_I_C_G_K_K_A_I_G_

          ApaI   EagI
1741 ACCGTGCTGGTGGGCCCCACCCCTGTGAATATCATCGGCCGGAACATGCTGACCCAGATC
-----+-----+-----+-----+-----+-----+-----+
TGGCACGACCACCCGGGGTGGGGACACTTATAGTAGCCGGCCTTGTACGACTGGGTCTAG
T_V_L_V_G_P_T_P_V_N_I_I_G_R_N_M_L_T_Q_I_

          BsaI
1801 GGCTGCACCCCTGAACCTCCCCATCAGCCCCATCGAGACCGTGCCCGTGACCCTGAAGCCC
-----+-----+-----+-----+-----+-----+-----+
CCGACGTGGGACTTGAAGGGGTAGTCGGGGTAGCTCTGGCACGGGCACTGGGACTTCGGG
G_C_T_L_N_F_P_I_S_P_I_E_T_V_P_V_T_L_K_P_

1861 GGCATGGATGGCCCCAAGTGAACAGTGGCCCCCTGACCGAGGAGAAGATTAAGGCCCTG
-----+-----+-----+-----+-----+-----+-----+
CCGTACCTACCGGGTTTCACTTTGTCACCGGGGACTGGCTCCTCTTCTAATTCGGGGAC
G_M_D_G_P_K_V_K_Q_W_P_L_T_E_E_K_I_K_A_L_

1921 ACCGAAATCTGTACCGAGATGGAGAAGGAGGGCAAGATCAGCAAGATCGGCCCGGAGAAC
-----+-----+-----+-----+-----+-----+-----+
TGGCTTTAGACATGGCTCTACCTCTTCTCCCGTTCTAGTCGTTCTAGCCGGGCTCTTG
T_E_I_C_T_E_M_E_K_E_G_K_I_S_K_I_G_P_E_N_

1981 CCCTACAACACCCCATCTTCGCCATCAAGAAGAAGGACAGCACCAAGTGGCGGAAACTG
-----+-----+-----+-----+-----+-----+-----+
GGGATGTTGTGGGGGTAGAAGCGGTAGTTCTTCTTCTGTCGTTTACCGCCTTTGAC
P_Y_N_T_P_I_F_A_I_K_K_K_D_S_T_K_W_R_K_L_

```

FIG. 16D

```

                                BstNI          PvuII
2041 GTGGACTTCCGGGAGCTGAACAAGAGGACCCAGGACTTCTGGGAAGTGCAGCTGGGCATC
-----+-----+-----+-----+-----+-----+-----+
CACCTGAAGGCCCTCGACTTGTTCCTCTGGGTCTTGAAGACCCCTCACGTCGACCCGTAG
V_D_F_R_E_L_N_K_R_T_Q_D_F_W_E_V_Q_L_G_I_

                                BstNI
2101 CCCCACCTGCCGGCCTGAAGAAGAAGAAGTCCGTGACAGTGC TGGATGTGGGCGACGCC
-----+-----+-----+-----+-----+-----+-----+
GGGGTGGGACGGCCGGACTTCTTCTTCTTCAGGCACTGTCACGACCTACACCCGCTGCGG
P_H_P_A_G_L_K_K_K_K_S_V_T_V_L_D_V_G_D_A_

                                BstNI
2161 TACTTCAGCGTGCCCTGGACGAGAACTTCAGGAAGTACACCGCCTTCACCATCCCCAGC
-----+-----+-----+-----+-----+-----+-----+
ATGAAGTCGCACGGGGACCTGCTCTTGAAGTCCTTCATGTGGCGGAAGTGGTAGGGGTCG
Y_F_S_V_P_L_D_E_N_F_R_K_Y_T_A_F_T_I_P_S_

                                BsaI
2221 ACCAACACGAGACCCCCGGAGTGAGATAACCAGTACAACGTGCTGCCTCAGGGCTGGAAG
-----+-----+-----+-----+-----+-----+-----+
TGGTTGTTGCTCTGGGGCCTCACTCTATGGTCATGTTGCACGACGGAGTCCCGACCTTC
T_N_N_E_T_P_G_V_R_Y_Q_Y_N_V_L_P_Q_G_W_K_

                                BstXI BstNI
2281 GGCAGCCCCGCCATCTTCCAGAGCAGCATGACCAAGATCCTGGAGCCCTCCGGAGCAAG
-----+-----+-----+-----+-----+-----+-----+
CCGTCGGGGCGGTAGAAGGTCTCGTCGTACTGGTTCTAGGACCTCGGGAAGGCCTCGTTC
G_S_P_A_I_F_Q_S_S_M_T_K_I_L_E_P_F_R_S_K_

                                PflMI
2341 AACCCCGAGATCATCATCTACCAGTACATGGCCGCCCTGTATGTGGGCAGCGATCTGGAG
-----+-----+-----+-----+-----+-----+-----+
TTGGGGCTCTAGTAGTAGATGGTCATGTACCGCGGGACATACACCCGTCGCTAGACCTC
N_P_E_I_I_I_Y_Q_Y_M_A_A_L_Y_V_G_S_D_L_E_

                                ApaI BspMI
2401 ATCGGCCAGCACAGGACCAAGATCGAAGACCTGAGGGCCACCTGCTGAGCTGGGGCTTC
-----+-----+-----+-----+-----+-----+-----+
TAGCCGGTCGTGCTCTGGTTCTAGCTTCTCGACTCCCGGGTGGACGACTCGACCCCGAAG
I_G_Q_H_R_T_K_I_E_E_L_R_A_H_L_L_S_W_G_F_

2461 ACCACCCCGATAAGAAGCACCAGAAGGAGCCCCCTTTCCTGTGGATGGGCTACGAGCTG
-----+-----+-----+-----+-----+-----+-----+
TGTTGGGGCTATTCTTCGTGGTCTTCTCGGGGAAAGGACACCTACCCGATGCTCGAC
T_T_P_D_K_K_H_Q_K_E_P_P_F_L_W_M_G_Y_E_L_

2521 CACCCCGATAAGTGGACCGTGCAGCCCATCATGCTGCCCGATAAGGAGAGCTGGACCGTG
-----+-----+-----+-----+-----+-----+-----+
GTGGGGCTATTCACCTGGCACGTCGGGTAGTACGACGGGCTATTCCTCTCGACCTGGCAC
H_P_D_K_W_T_V_Q_P_I_M_L_P_D_K_E_S_W_T_V_

```

FIG. 16E

```

                PflMI
2581 AACGACATCCAGAACTGGTGGGCAAGCTGAAT1TGGGCCAGCCAAATCTACGCCGGCATT
-----+-----+-----+-----+-----+-----+-----+-----+
TTGCTGTAGGTCTTTGACCACCGTTCGACTTAACCCGGTCGGTTTAGATGCGGCCGTAA
N_D_I_Q_K_L_V_G_K_L_N_W_A_S_Q_I_Y_A_G_I_

                NarI
                KasI
                PvuII
                HaeII
2641 AAAGTGAAGCAGCTGTGCAGGCTGCTGAGAGGCCGCAAAGCCCTGACAGACATCGTGACA
-----+-----+-----+-----+-----+-----+-----+-----+
TTTCACTTCGTCGACACGTCCGACGACTCTCCGCGGTTTCGGGACTGTCTGTAGCACTGT
K_V_K_Q_L_C_R_L_L_R_G_A_K_A_L_T_D_I_V_T_

2701 CTGACAGAGGAGGCCGAGCTGGAGCTGGCCGAGAACAGGGAGATCCTGAAGGACCCCGTG
-----+-----+-----+-----+-----+-----+-----+-----+
GACTGTCTCCTCCGGCTCGACCTCGACCGGCTCTTGTCCTCTAGGACTTCTGGGGCAC
L_T_E_E_A_E_L_E_L_A_E_N_R_E_I_L_K_D_P_V_

                BstNI
                HinfI
                BstNI
2761 CACGGCGTGTACTACGACCCAGCAAGGACCTGGTGGCCGAGATTCAGAAGCAGGGCCAG
-----+-----+-----+-----+-----+-----+-----+-----+
GTGCCGCACATGATGCTGGGGTCGTTCTCTGGACCACCGGCTCTAAGTCTTCGTCGCCGGTC
H_G_V_Y_Y_D_P_S_K_D_L_V_A_E_I_Q_K_Q_G_Q_

                BstNI
2821 GACCAGTGGACCTACCAAATCTACCAGGAGCCTTCAAGAACCTGAAAACCGGGAAGTAC
-----+-----+-----+-----+-----+-----+-----+-----+
CTGGTCACCTGGATGGTTTAGATGGTCTCGGAAAGTTCPTGGACTTTTGGCCCTTCATG
D_Q_W_T_Y_Q_I_Y_Q_E_P_F_K_N_L_K_T_G_K_Y_

                BstNI
                HaeII
                PvuII
2881 GCCAGGAAGAGAAGCGCCACACCAACGATGTGAGGCAGCTGGCCGAAGTGGTGCAGAAA
-----+-----+-----+-----+-----+-----+-----+-----+
CGGTCTTCTCTTCGCGGGTGTGGTTGCTACACTCCGTCGACCGGCTTCACCACGTCTTT
A_R_K_R_S_A_H_T_N_D_V_R_Q_L_A_E_V_V_Q_K_

2941 GTGGCTATGGAGAGCATCGTGATCTGGGGCAAGACCCCAAGTTCAGCTGCCATCCAG
-----+-----+-----+-----+-----+-----+-----+-----+
CACCGATACCTCTCGTAGCACTAGACCCCGTTCGGGGTTCAAGTTCGACGGGTAGGTC
V_A_M_E_S_I_V_I_W_G_K_T_P_K_F_K_L_P_I_Q_

                BstNI
                BsaI
                BstNI
                HinfI
                BstNI
3001 AAGGAGACCTGGGAAACCTGGTGGATGGACTACTGGCAGGCCACCTGGATTCTGAGTGG
-----+-----+-----+-----+-----+-----+-----+-----+
TTCTCTGGACCCCTTGGACCACCTACCTGATGACCGTCCGGTGGACCTAAGGACTCACC
K_E_T_W_E_T_W_W_M_D_Y_W_O_A_T_W_I_P_E_W_

```

FIG. 16F

```

                                     PvuII                               BstNI
3061  GAGTTCGTGAACACCCCTCTGGTGAAGCTGTGGTATCAGCTGGAGAAGGACCCATC
-----+-----+-----+-----+-----+-----+-----+-----+
CTCAAGCACTTGTGGGGGGAGACCACTTCGACACCATAAGTCGACCTCTCCTGGGGTAG
E_F_V_N_T_P_P_L_V_K_L_W_Y_Q_L_E_K_D_P_I_

      NarI
      KasI
      HaeII BsaI                               BsaI
3121  CTGGGCGCCGAGACCTTCTACGTGGACGGAGCCGCAATAGAGAGACCAAGCTGGGCAAG
-----+-----+-----+-----+-----+-----+-----+
GACCCGCGGCTCTGGAAGATGCACCTGCCTCGGCGGTATCTCTCTGGTTCGACCCGTTTC
L_G_A_E_T_F_Y_V_D_G_A_A_N_R_E_T_K_L_G_K_

3181  GCCGGCTACGTGACCGACAGAGGCAGACAGAAAGTGGTGTCTCTGACCGAGACAACCAAC
-----+-----+-----+-----+-----+-----+-----+
CGGCCGATGCACTGGCTGTCTCCGTCTGTCTTTCACCACAGAGACTGGCTCTGTGGTTG
A_G_Y_V_T_D_R_G_R_Q_K_V_V_S_L_T_E_T_T_N_

                                     BstXI                               PstI
3241  CAGAAAACCGAGCTGCACGCCATCCTGCTGGCCCTGCAGGACAGCGGCAGCGAAGTGAAC
-----+-----+-----+-----+-----+-----+-----+
GTCITTTGGCTCGACGTGCGGTAGGACGACCGGGACGTCCCTGTCGCGCTCGCTTCACTTG
Q_K_T_E_L_H_A_I_L_L_A_L_Q_D_S_G_S_E_V_N_

      HinfI                               BstNI
3301  ATCGTGACCGACTCCCAGTACGCCCTGGGCATCATTACAGGCCAGCCCGATAGAAGCGAG
-----+-----+-----+-----+-----+-----+-----+
TAGCACTGGCTGAGGGTCAATGCGGGACCCGTAGTAAGTCCGGGTCCGGCTATCTTCGCTC
I_V_T_D_S_O_Y_A_L_G_I_I_O_A_O_P_D_R_S_E_

3361  AGCGAGCTGGTGAACCAGATCATCGAGAAGCTGATCGGCAAGGACAAAATCTACCTGAGC
-----+-----+-----+-----+-----+-----+-----+
TCGCTCGACCACTTGGTCTAGTAGCTCTTCGACTAGCCGTTCCCTGTTTATAGTGGACTCG
S_E_L_V_N_Q_I_I_E_K_L_I_G_K_D_K_I_Y_L_S_

                                     BspMI
3421  TGGGTGCCCGCCCAAGGGCATCGGGCGCAACGAGCAGGTGGACAAGCTGGTGTCCAGC
-----+-----+-----+-----+-----+-----+-----+
ACCCACGGGCGGGTGTTCCTCGTAGCCGCGTTCGCTCGTCCACCTGTTTCGACCACAGGTGC
W_V_P_A_H_K_G_I_G_G_N_E_Q_V_D_K_L_V_S_S_

                                     BstNI
3481  GGCATCCGAAAGTGCTGTTTCTGGACGGCATCGACAAGGCCAGGAGGACCACGAGAGA
-----+-----+-----+-----+-----+-----+-----+
CCGTAGGCCCTTCACGACAAAGACCTGCCGTAGCTGTTCCGGGTCCCTGGTGTCTCTCT
G_I_R_K_V_L_F_L_D_G_I_D_K_A_Q_E_D_H_E_R_

```

FIG. 16G

```

                                     BspMI
3541 TACCACGCAACTGGCGGACAATGGCCAGCGACTTCAACCTGCCTCCCATCGTGGCCAAG
-----+-----+-----+-----+-----+-----+-----+
ATGGTGTGCGTTGACCGCCTGTTACCGGTGCGCTGAAGTTGGACGGAGGGTAGCACCGGTTG
Y_H_S_N_W_R_T_M_A_S_D_F_N_L_P_P_I_V_A_K

          PvuII          PvuII          BstNI
3601 GAGATCGTGGCCAGCTGCGATAAGTGTGAGCTGAAGGGCGAGGCCATGCACGGCCAGGTG
-----+-----+-----+-----+-----+-----+
CTCTAGCACCGGTGCGACGCTATTCACAGTCGACTTCCCGCTCCGGTACGTGCCGGTCCAC
E_I_V_A_S_C_D_K_C_Q_L_K_G_E_A_M_H_G_Q_V

          PstI   BstNI          PvuII          BstNI          HinfI
3661 GACTGCAGCCCTGGCATCTGGCAGCTGGCCCTGCACCCACCTGGAGGGCAAAGTGATTCTG
-----+-----+-----+-----+-----+-----+
CTGACGTCGGGACCGTAGACCGTCGACCGGACGTTGGGTGGACCTCCCGTTTACTAAGAC
D_C_S_P_G_I_W_Q_L_A_C_T_H_L_E_G_K_V_I_L

                                     HinfI   BsaI   BstNI
3721 GTGGCCGTGCACGTGGCCAGCGGCTACATCGAGGCCGAAGTGATCCCGCCGAGACCGGC
-----+-----+-----+-----+-----+-----+
CACCGGCACGTGCACCGGTGCGCGATGTAGCTCCGGCTTACTAAGGGCGGCTCTGGCCG
V_A_V_H_V_A_S_G_Y_I_E_A_E_V_I_P_A_E_T_G

          BsaI
3781 CAGGAGACCGCCTACTTCTGCTGAAGCTGGCCGGCAGATGGCCCGTGAAGTGGTGCAC
-----+-----+-----+-----+-----+-----+
GTCCTCTGGCGGATGAAGGACGACTTCGACCGGCCGCTACCGGGCACITTCACCACGTG
Q_E_T_A_Y_F_L_L_K_L_A_G_R_W_P_V_K_V_V_H

3841 ACCGCCAACGGCAGCAACTTCACCTCTGCCCGCGTGAAGGCCGCTGTTGGTGGCCAAT
-----+-----+-----+-----+-----+-----+
TGGCGGTTGCCGTCGTTGAAGTGGAGACGGCGGCACTTCCGGCGGACAACCACCGGTTA
T_A_N_G_S_N_F_T_S_A_A_V_K_A_A_C_W_W_A_N

                                     PflMI
                                     BstNI
3901 ATCCAGCAGGAGTTCGGCATCCCTACAACCTCAGAGCCAGGGCGTGGTGGCCAGCATG
-----+-----+-----+-----+-----+-----+
TAGGTGCGTCTCAAGCCGATAGGGGATGTTGGGAGTCTCGGTCCCGCACCACCGGTCGTAC
I_Q_Q_E_F_G_I_P_Y_N_P_O_S_O_G_V_V_A_S_M

          BstNI          BstNI
3961 AACAAAGGAGCTGAAGAAGATCATCGGCCAGGTGAGGGACCAGGCCGAGCACCTGAAAACA
-----+-----+-----+-----+-----+-----+
TTGTTCTCGACTTCTTCTAGTAGCCGGTCCACTCCCTGGTCCGGCTCGTGGACTTTTGT
N_K_E_L_K_K_I_I_G_Q_V_R_D_Q_A_E_H_L_K_T

4021 GCCGTGCAGATGGCCGTGTTTCATCCACAACCTCAAGCGGAAGGGCGGCATTGGCCGGCTAC
-----+-----+-----+-----+-----+-----+
CGGCACGTCACCGCACAAAGTAGGTGTTGAAGTTCGCCTTCCCGCGTAACCGCCGATG
A_V_O_M_A_V_F_I_H_N_F_K_R_K_G_G_I_G_G_Y

```

FIG. 16H

```

HaeII                      EcoRV                      PstI
4081  AGCGCCGGAGCGGATCATCGACATCATCGCCACCGATATCCAGACCAAGGAAC TGCAG
-----+-----+-----+-----+-----+-----+-----+
TCGCGGCCTCTCGCCTAGTAGCTGTAGTAGCGGTGGCTATAGGTCTGGTTCCTTGACGTC
S_A_G_E_R_I_I_D_I_I_A_T_D_I_O_T_K_E_L_O_

HinfI
4141  AAGCAGATCACCAAGATTCAGAACTTCAGAGTGTACTACCGGGACAGCAGGGACCCCATC
-----+-----+-----+-----+-----+-----+-----+
TTCGTCTAGTGGTTCTAAGTCTTGAAGTCTCACATGATGGCCCTGTCGTCCCTGGGGTAG
K_Q_I_T_K_I_Q_N_F_R_V_Y_Y_R_D_S_R_D_P_I_

NarI
KasI
HaeII                      BstNI
4201  TGGAAGGCCCTGCCAAGCTGCTGTGGAAGGGCGAAGGCGCCGTGGTGTCCAGGACAAC
-----+-----+-----+-----+-----+-----+-----+
ACCTTCCCGGGACGGTTCGACGACACCTTCCCGCTCCCGCGGCACCAC TAGGTCCTGTTG
W_K_G_P_A_K_L_L_W_K_G_E_G_A_V_V_I_O_D_N_

HinfI
4261  AGCGACATCAAAGTGGTGCCCCGGAGGAAGGCCAAGATCTGCGGGACTACGGCAAACAG
-----+-----+-----+-----+-----+-----+-----+
TCGCTGTAGTTTACCACGGGGCCTCCTTCCGGTCTAAGACGCCCTGATGCCGTTGTGTC
S_D_I_K_V_V_P_R_R_K_A_K_I_L_R_D_Y_G_K_Q_

BglII
4321  ATGGCCGGCGATGACTGCGTGGCCGGCAGGCAGGATGAGGACAGATCTATGGCGGCAAG
-----+-----+-----+-----+-----+-----+-----+
TACCGCCGCTACTGACGCACCGCCGTCCTACTCCTGTCTAGATACCCGCGGTTCC
M_A_G_D_D_C_V_A_G_R_Q_D_E_D_R_S_M_G_G_K_

4381  TGGTCCAAGGGCAGCATGTGGGCTGGCCCGAGATCCGGGAGAGAATGAGAAGAGCCCT
-----+-----+-----+-----+-----+-----+-----+
ACCAGGTTCCCGTCGTAACACCCGACCGGGCTCTAGGCCCTCTCTTACTCTTCTCGGGGA
W_S_K_G_S_I_V_G_W_P_E_I_R_E_R_M_R_R_A_P_

NarI                      NarI
KasI                      KasI
BstNI                      HaeII                      HaeII
4441  GCCGCCGCTCCTGGAGTGGGCGCGTGTCTCAGGATCTGGATAAGCACGGCGCCATCACC
-----+-----+-----+-----+-----+-----+-----+
CGGCGGCGAGGACCTCACCCGGGCACAGAGTCC TAGACCTATTCGTGCCGCGGTAGTGG
A_A_A_P_G_V_G_A_V_S_Q_D_L_D_K_H_G_A_I_T_

PvuII                      BstNI
4501  AGCAGCAACATCAACAACCCAGCTGTGTGGCTGGAGGCCAGGAAGAGGAGGAAGTG
-----+-----+-----+-----+-----+-----+-----+
TCGTGTTGTAGTTGTTGGGGTCGACACACCCGACCTCCGGGTCCTTCTCCTCCTCAC
S_S_N_I_N_N_P_S_C_V_W_L_E_A_O_E_E_E_V_

```

FIG. 16I

```

                                     BsaI  BstNI          BsaI          NarI
                                     KasI
                                     HaeII
4561  GGCTTCCCTGTGAGACCCAGGTGCCCTGAGACCCATGACCTACAAGGGCGCCTTCGAC
-----+-----+-----+-----+-----+-----+-----+
CCGAAGGGGACACTCTGGGGTCCACGGGGACTCTGGGTACTGGATGTTCCCGCGGAAGCTG
G_F_P_V_R_P_Q_V_P_L_R_P_M_T_Y_K_G_A_F_D

                                     BstNI
4621  CTGAGCCACTTCTGAAGGAGAAGGGCGGCCTGGACGGCCTGATCTACAGCCGGAAGCGG
-----+-----+-----+-----+-----+-----+-----+
GACTCGGTGAAGGACTTCTCTTCCCGCCGGACCTGCCGGACTAGATGTCGGCCTTCGCC
L_S_H_F_L_K_E_K_G_L_D_G_L_I_Y_S_R_K_R

                                     BstNI          BstNI
4681  CAGGAGATCCTGGATCTGTGGGTGTACCACACCCAGGGCTACTTCCCCGACTGGCAGAAT
-----+-----+-----+-----+-----+-----+-----+
GTCCTCTAGGACCTAGACACCCACATGGTGTGGGTCCCGATGAAGGGGCTGACCGTCTTA
Q_E_I_L_D_L_W_V_Y_H_T_Q_G_Y_F_P_D_W_Q_N

                                     BstNI  BstNI
4741  TACACCCCTGGCCCTGGAGTGGGTATCCCCTGACCTTCGGCTGGTCTCAAGCTGGTG
-----+-----+-----+-----+-----+-----+-----+
ATGTGGGGACCGGGACCTCACGCCATAGGGGACTGGAAGCCGACCACGAAGTTCGACCAC
Y_T_P_G_P_G_V_R_Y_P_L_T_F_G_W_C_F_K_L_V

4801  CCTATGGAGCCCGACGAAGTGGAGAAGGCCACAGAGGGCGAGAACAACAGCCTGCTGCAC
-----+-----+-----+-----+-----+-----+-----+
GGATACCTCGGGCTGCTTCACCTCTTCCGGTGTCTCCCGCTCTTGTGTGGGACGACGTG
P_M_E_P_D_E_V_E_K_A_T_E_G_E_N_N_S_L_L_H

4861  CCTATCTGCCAGCACGGCATGGACGATGAGGAGCGGGAAGTCTGATCTGGAAGTTCGAC
-----+-----+-----+-----+-----+-----+-----+
GGATAGACGGTCGTGCCGTACTGCTACTCCTCGCCCTTACGACTAGACCTTCAAGCTG
P_I_C_Q_H_G_M_D_D_E_E_R_E_V_L_I_W_K_F_D

                                     BstNI
4921  AGCAGGCTGGCCCTGAAGCACAGAGCCCAGGAACATGCACCCAGAGTCTACAAGGACTGC
-----+-----+-----+-----+-----+-----+-----+
TCGTCCGACCGGGACTTCGTGTCTCGGGTCTTGACGTGGGTCTCAAGATGTTCTGACG
S_R_L_A_L_K_H_R_A_Q_E_L_H_P_E_F_Y_K_D_C

                                     BclI          XbaI
4981  TGATGATCATAATAATCTAGAA
-----+-----+-----+
ACTACTAGTATTATTAGATCTT
*

```

FIG. 16J

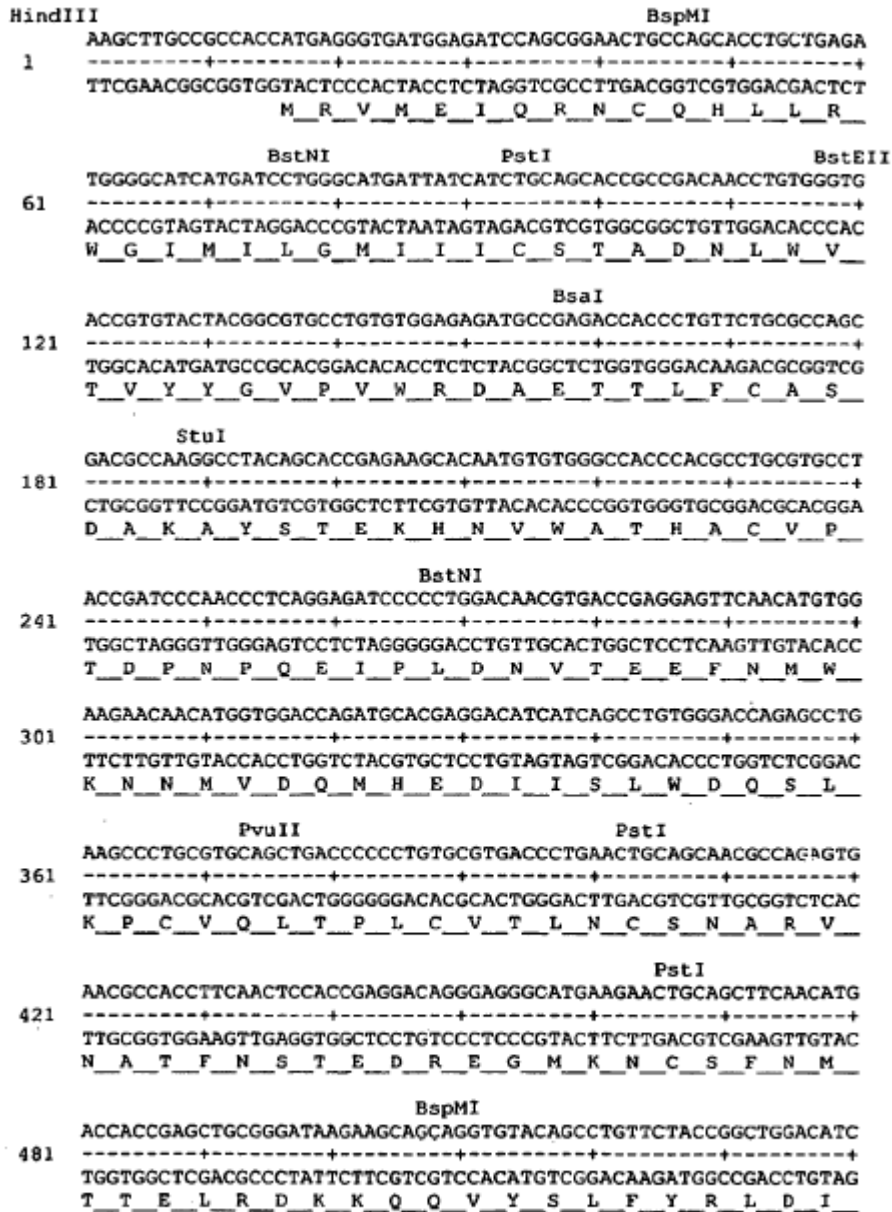


FIG. 17A

```

                                                    HaeII
541 GAGAAGATCAACAGCAGCAACAACAACAGCGAGTACCGGCTGGTGAACGCAATACCAGC
-----+-----+-----+-----+-----+-----+
CTCTTCTAGTTGTCGTCGTTGTTGTTGTCGCTCATGGCCGACCACTTGACGTTATGGTCG
E_K_I_N_S_S_N_N_N_S_E_Y_R_L_V_N_C_N_T_S_

          StuI
          BstNI          BstEII
601 GCCATCACCCAGGCTGCCCTAAGGTGACCTTCGAGCCCATCCCCATCCACTACTGCGCC
-----+-----+-----+-----+-----+
CGGTAGTGGGTCGGACGGGATTCCACTGGAAGCTCGGGTAGGGGTAGGTGATGACGCGG
A_I_T_Q_A_C_P_K_V_T_F_E_P_I_P_I_H_Y_C_A_

661 CCTGCCGGCTTCGCCATCCTGAAGTGCAACGACACCGAGTTC AATGGCACCCGGCCCTGC
-----+-----+-----+-----+-----+
GGACGGCCGAAGCGGTAGGACTTCACGTTGCTGTGGCTCAAGTTACCGTGGCCGGGGACG
P_A_G_F_A_I_L_K_C_N_D_T_E_F_N_G_T_G_P_C_

                                                    PvuII
721 AAGAATGTGAGCACCGTGCAGTGCACCCACGGCATCAAGCCCGTGGTGTCCACCCAGCTG
-----+-----+-----+-----+-----+
TTCTTACACTCGTGGCACGTCACGTGGGTGCCGTAGTTCGGGCACCCACAGGTGGGTCGAC
K_N_V_S_T_V_Q_C_T_H_G_I_K_P_V_V_S_T_Q_L_

          BstNI
781 CTGCTGAACGGCAGCCTGGCCGAGAGAGAAGTGCGGATCAGGAGCGAGAACATCGCCAAC
-----+-----+-----+-----+-----+
GACGACTTGCCGTCGGACCGGCTCTCTCTTCACGCCTAGTCCCTCGCTCTGTAGCGGTTG
L_L_N_G_S_L_A_E_R_E_V_R_I_R_S_E_N_I_A_N_

841 AACGCCAAGAACATCATCGTGCAGTTCGCCAGCCCCGTAAGATCAACTGCATCCGGCCC
-----+-----+-----+-----+-----+
TTGCGGTTCTGTAGTAGCACGTC AAGCGGTCGGGGCACTTCTAGTTGACGTAGGCCGGG
N_A_K_N_I_I_V_Q_F_A_S_P_V_K_I_N_C_I_R_P_

          HinfI  BstNI
901 AACACAATACCCGGAAGAGCTACAGAATCGGCCCTGGCCAGACCTTCTACGCCACCGAC
-----+-----+-----+-----+-----+
TTGTTGTTATGGGCCTTCTCGATGTCTTAGCCGGGACCGGTCTGGAAGATGCGGTGGCTG
N_N_N_T_R_K_S_Y_R_I_G_P_G_Q_T_F_Y_A_T_D_

          BstNI
961 ATTGTGGGCGACATCAGACAGGCCCACTGCAACGTGTCCAGGACCGACTGGAACAACACC
-----+-----+-----+-----+-----+
TAACACCCGCTGTAGTCTGTCCGGGTGACGTTGCACAGGTCCTGGCTGACCTTGTGTGG
I_V_G_D_I_R_Q_A_H_C_N_V_S_R_T_D_W_N_N_T_

          PvuII  ScaI
1021 CTGAGACTGGTGGCCAACCAGCTGCGGAAGTACTTCAGCAACAAGACCATCATCTTCACC
-----+-----+-----+-----+-----+
GACTCTGACCACCGGTTGGTCGACGCCTTCATGAAGTCGTTGTTCTGGTAGTAGAAGTGG
L_R_L_V_A_N_O_L_R_K_Y_F_S_N_K_T_I_I_F_T_

```

FIG. 17B

```

                BstNI
                BsaI
1081 AACAGCAGCGGGCGGAGACCTGGAGATCACCACCCACAGCTTCAATTGTGGCGGGCAGTTC
-----+-----+-----+-----+-----+-----+-----+
TTGTCGTCGCCGCTCTGGACCTCTAGTGGTGGGTGTCGAAGTTAACACCGCCGCTCAAG
N_S_S_G_G_D_L_E_I_T_T_H_S_F_N_C_G_G_E_F_

                BstNI
                HinfI
1141 TTCTACTGCAACACCTCCGGCCTGTTCATAGCACCTGGACCACCAACAACATGCAGGAG
-----+-----+-----+-----+-----+-----+-----+
AAGATGACGTTGTGGAGGCCGGACAAGTTATCGTGGACCTGGTGGTTGTTGTACGTCTC
F_Y_C_N_T_S_G_L_F_N_S_T_W_T_T_N_N_M_Q_E_

1201 TCCAACGACACCAGCAACGGCACCATCACCTGCCCTGCCGGATCAAGCAGATCATCCGG
-----+-----+-----+-----+-----+-----+-----+
AGGTTGCTGTGGTCGTTGCCGTTAGTGGGACGGGACGGCCTAGTTCGTCTAGTAGGCC
S_N_D_T_S_N_G_T_I_T_L_P_C_R_I_K_Q_I_I_R_

                BstNI
                HinfI
1261 ATGTGGCAGCGCGTGGGCCAGGCCATGTACGCCCCCTCCCATCGAGGGCGTGATTCGCTGC
-----+-----+-----+-----+-----+-----+-----+
TACACCGTCGCGCACCCGGTCCGGTACATGCGGGGAGGGTAGCTCCCGCACTAAGCGAGC
M_W_Q_R_V_G_Q_A_M_Y_A_P_P_I_E_G_V_I_R_C_

1321 GAGAGCAACATCACCGCCTGATCCTGACCAGAGATGGCGGCAACAACAATCCGCCAAC
-----+-----+-----+-----+-----+-----+-----+
CTCTCGTTGTAGTGGCCGGACTAGGACTGGTCTCTACCGCGTTGTTGTTAAGCGGGTTG
E_S_N_I_T_G_L_I_L_T_R_D_G_G_N_N_N_S_A_N_

                BsaI
                BstNI
                EcoRV
1381 GAGACCTTCAGACCTGGCGGCGGAGATATCCGGGACAACCTGGCGGAGCGAGCTGTACAAG
-----+-----+-----+-----+-----+-----+-----+
CTCTGGAAGTCTGGACCGCCGCTCTATAGGCCCTGTTGACCGCCTCGCTCGACATGTTT
E_T_F_R_P_G_G_G_D_I_R_D_N_W_R_S_E_L_Y_K_

                BstNI
1441 TACAAGGTGGTGAAGATCGAGCCCCTGGGCGTGGCCCCCACCAGAGCCAAGAGAAGAGTG
-----+-----+-----+-----+-----+-----+-----+
ATGTTCCACCACTTCTAGCTCGGGGACCCGCACCGGGGGTGGTCTCGGTTCTTCTCAC
Y_K_V_V_K_I_E_P_L_G_V_A_P_T_R_A_K_R_R_V_

                NarI
                KspI
                HaeII
                BstNI
1501 GTGGAGCGGGAGAAGAGAGCCGTGGGCATCGGGCGCCGTGTTTCTGGGCTTCCTGGGAGCC
-----+-----+-----+-----+-----+-----+-----+
CACCTCGCCCTCTTCTCTCGGCACCCGTAGCCGCGGCACAAAGACCCGAAGGACCCCTCGG
V_E_R_E_K_R_A_V_G_I_G_A_V_F_L_G_F_L_G_A_

```

FIG. 17C

```

                                PvuII
1561 GCCGGATCTACAATGGGAGCCGCCAGCATCACCTGACCGTGCAGGCCAGACAGCTGCTG
-----+-----+-----+-----+-----+-----+-----+
CGGCCTAGATGTTACCCTCGGCCGGTCGTAGTGGGACTGGCACGTCCGGTCTGTGACGAC
A_G_S_T_M_G_A_A_S_I_T_L_T_V_Q_A_R_Q_L_L

                                PvuII
1621 AGCGGCATCGTGCAGCAGCAGAGCAATCTGCTGAGAGCCATCGAGGCCAGCAGCAGCTG
-----+-----+-----+-----+-----+-----+-----+
TCGCCGTAGCACGTGCTGCTCGTTAGACGACTCTCGGTAGCTCCGGGTCGTGCTGAC
S_G_I_V_Q_Q_Q_S_N_L_L_R_A_I_E_A_Q_Q_Q_L

                                PvuII      BstXI
                                PstI      BstNI
1681 CTGAAGCTGACAGTGTGGGGCATCAAGCAGCTGCAGGCCAGGGTGCTGGCCGTGGAGAGA
-----+-----+-----+-----+-----+-----+-----+
GACTTCGACTGTCACACCCCGTAGTTCGTGACGTCCGGTCCCACGACCCGGCACCTCTCT
L_K_L_T_V_W_G_I_K_Q_L_Q_A_R_V_L_A_V_E_R

                                BstNI      PstI
1741 TACCTGAGGGACCAGCAGCTCCTGGGCATCTGGGGCTGCAGCGCAAGCTGATCTGCACC
-----+-----+-----+-----+-----+-----+-----+
ATGGACTCCCCTGGTCGTGAGGACCCGTAGACCCCGACGTCCCGTTCGACTAGACGTGG
Y_L_R_D_Q_Q_L_L_G_I_W_G_C_S_G_K_L_I_C_T

                                BstNI      PvuII
1801 ACCAACGTGCCCTGGAATAGCAGCTGGAGCAACAAGAGCTACGACGACATCTGGCAGAAC
-----+-----+-----+-----+-----+-----+-----+
TGGTTGCACGGGACCTTATCGTCGACCTCGTTGTTCTCGATGCTGCTGTAGACCGTCTTG
T_N_V_P_W_N_S_S_W_S_N_K_S_Y_D_D_I_W_Q_N

                                PstI
                                BstNI
1861 ATGACCTGGCTGCAGTGGGACAAGGAGATCAGCAACTACACCGACATCATCTACAGCCTG
-----+-----+-----+-----+-----+-----+-----+
TACTGGACCGACGTACCCCTGTTCCCTCTAGTCGTTGATGTGGCTGTAGTAGATGTCGGAC
M_T_W_L_Q_W_D_K_E_I_S_N_Y_T_D_I_I_Y_S_L

                                BstNI
1921 ATCGAGGAGAGCCAGAACCAGCAGGAGAAGAACGAGCAGGATCTGCTGGCCCTGGACAAG
-----+-----+-----+-----+-----+-----+-----+
TAGCTCCTCTCGGTCTTGGTCGTCCTCTTCTTGCTCGTCCTAGACGACCGGGACCTGTT
I_E_E_S_Q_N_Q_Q_E_K_N_E_Q_D_L_L_A_L_D_K

                                PflMI      BglII
1981 TGGGCCAACCTGTGGAACCTGGTTCGACATCAGCAAGTGGCTGTGGTACATCAGATCTTGA
-----+-----+-----+-----+-----+-----+-----+
ACCCGGTTGGACACCTTGACCAAGCTGTAGTCGTTACCCGACACCATGTAGTCTAGAACT
W_A_N_L_W_N_W_F_D_I_S_K_W_L_W_Y_I_R_S_*

                                XbaI
2041 TAATCTAGAA
-----+
ATTAGATCTT

```

FIG. 17D

**IAVA – Evaluación de los constructos AdH5
clonados en CB6F1; DÍA 10 – Elispot**

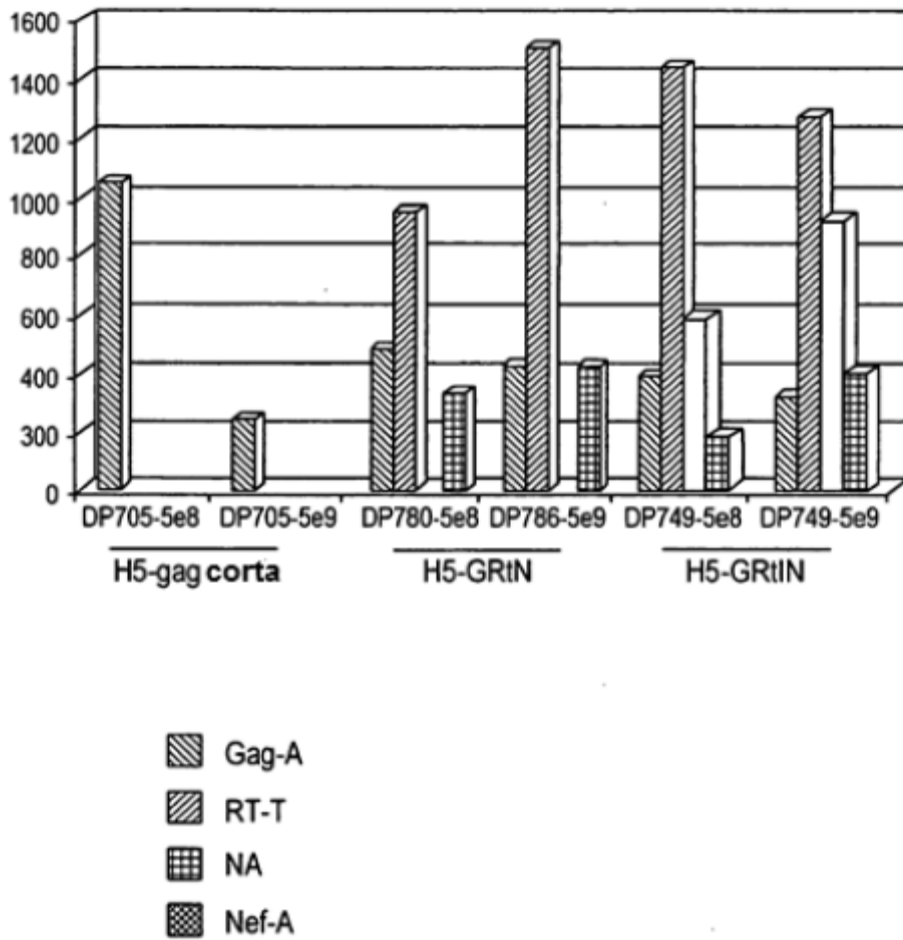


FIG. 18

FCC/05.002: Media de las respuestas a las mezclas de la biblioteca de péptidos del VIH del clado A por IFN-gamma Elispot 10 días post-primovacunación

Cada barra representa la media de 4 u 8 muestras (cada una analizada por triplicado) +/- 1 Error Estándar

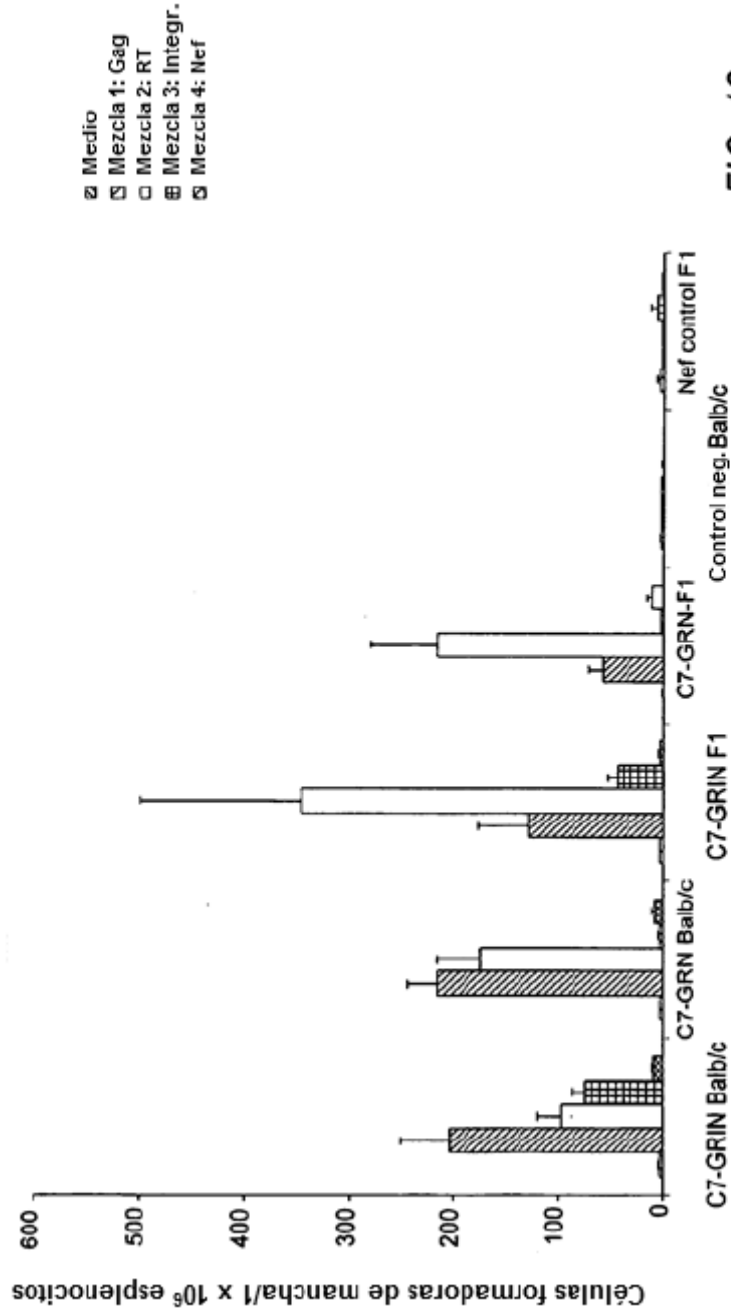


FIG. 19

FCC/02.005: Media de las respuestas a las mezclas de la biblioteca de péptidos del VIH del clado A por IL-2 Elispot 10 días post-primovacunación

Cada barra representa la media de 4 u 8 muestras (cada una analizada por triplicado) +/- 1 Error Estándar

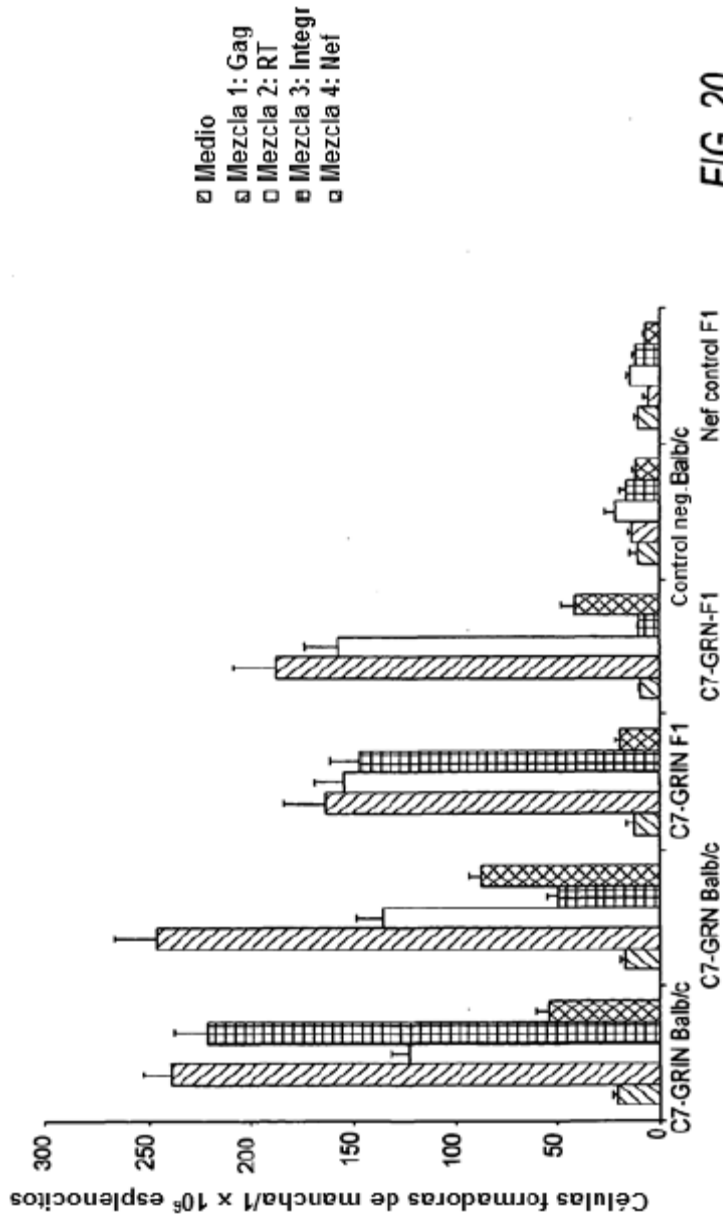


FIG. 20

FCC/05.003: Media de las respuestas a las mezclas de la biblioteca de péptidos de VIH del clado A por IFN-gamma Elispot 10 días post-primovacunación

Cada barra representa la media de 8 muestras (cada una analizada por triplicado) +/- 1 Error Estándar

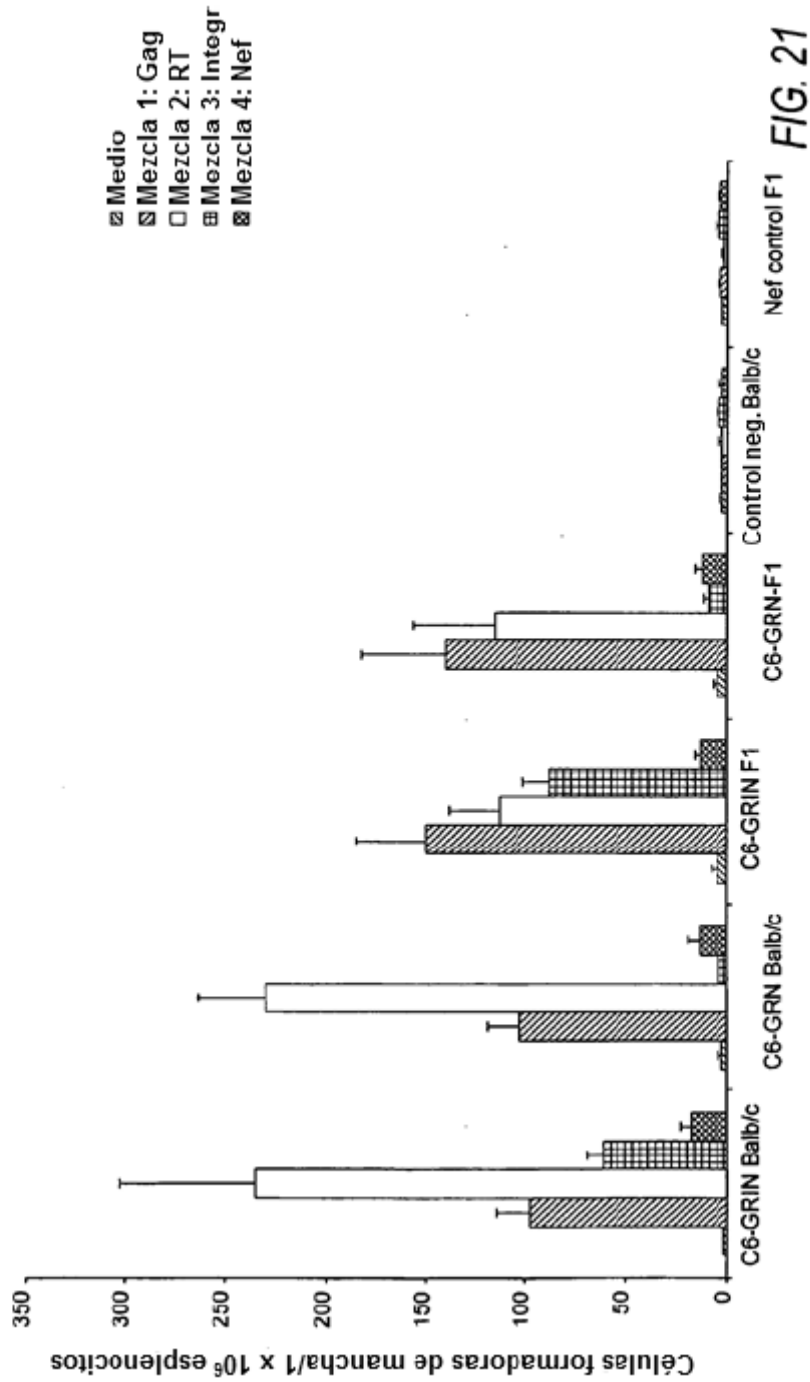


FIG. 21

FCC/05.003: Media de las respuestas a las mezclas de la biblioteca de péptidos del VIH del clado A por IL-2 Elispot 10 días post-primovacunación

Cada barra representa la media de 4 muestras (cada una analizada por triplicado) +/- Error Estándar

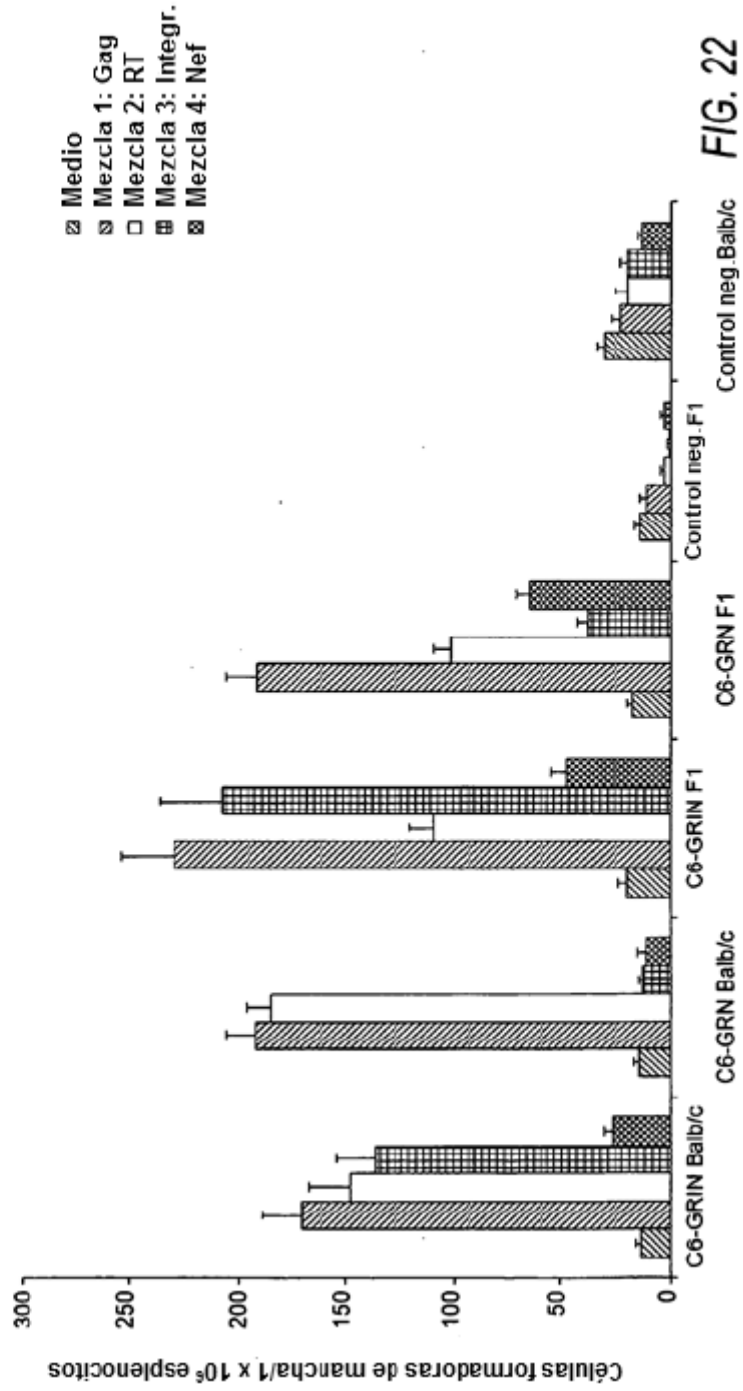


FIG. 22

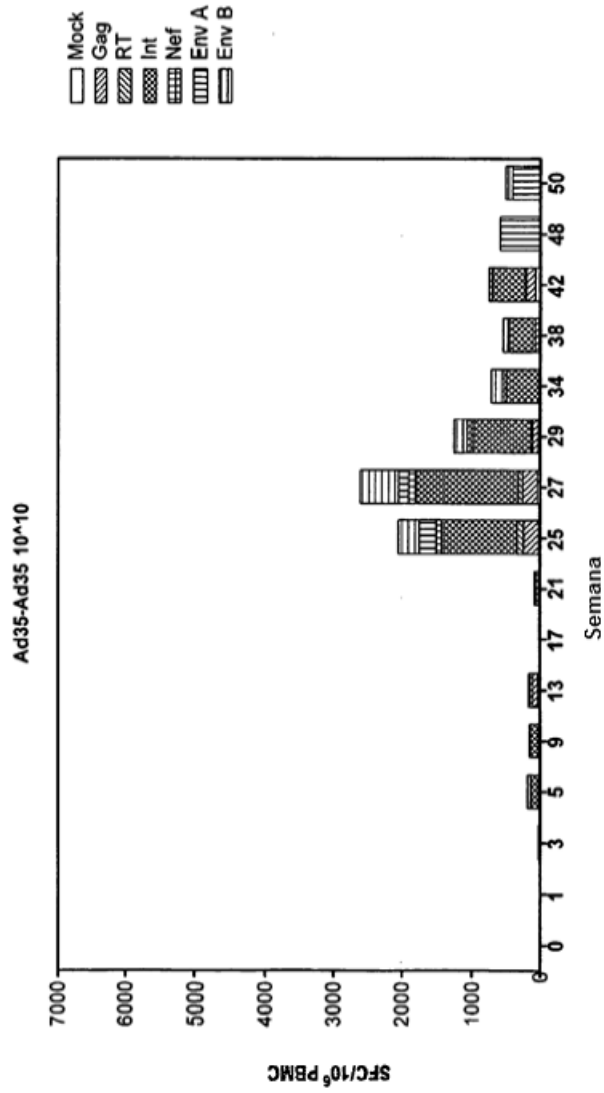


FIG. 23A

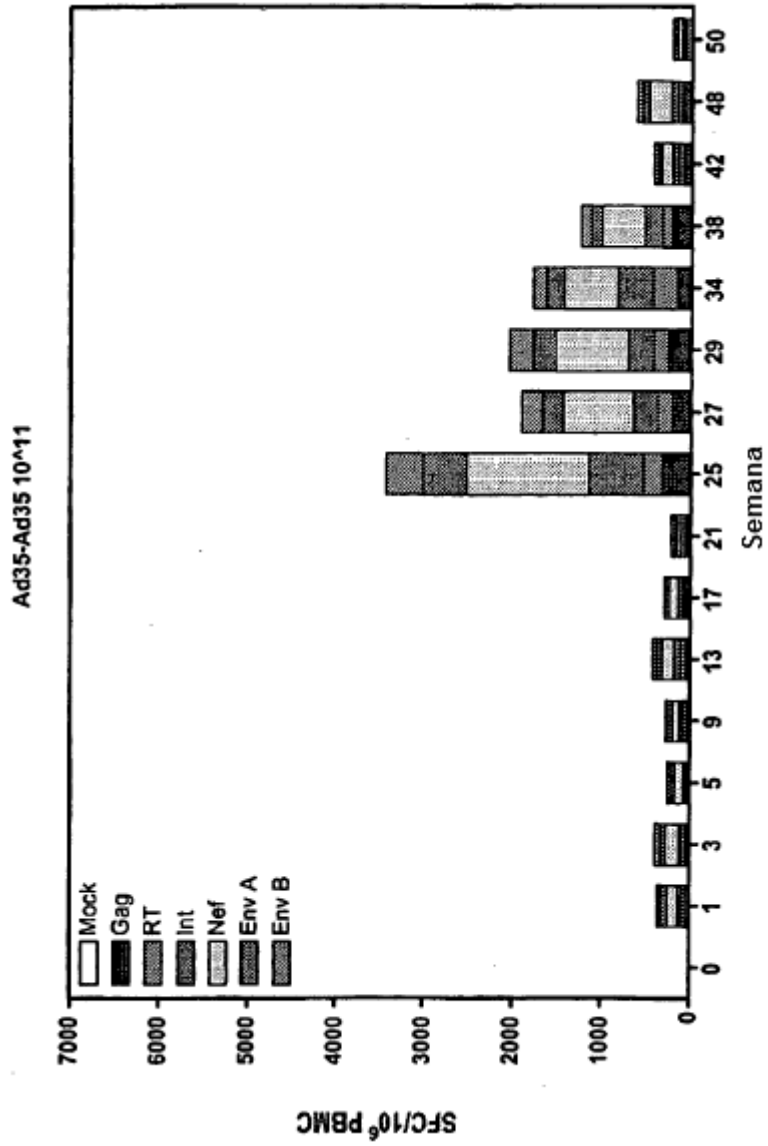


FIG. 23B