

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 577 929**

51 Int. Cl.:

C12N 15/867 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **27.05.2011** **E 11722021 (0)**

97 Fecha y número de publicación de la concesión europea: **27.04.2016** **EP 2576798**

54 Título: **Recombinasa elaborada a medida para recombinar unos sitios diana asimétricos en una pluralidad de cepas de retrovirus**

30 Prioridad:

27.05.2010 EP 10005499

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
19.07.2016

73 Titular/es:

**HEINRICH-PETTE-INSTITUT LEIBNIZ-INSTITUT
FÜR EXPERIMENTELLE VIROLOGIE-STIFTUNG
BÜRGERLICHEN RECHTS - (50.0%)
Martinistrasse 52
20251 Hamburg, DE y
MAX-PLANCK-GESELLSCHAFT ZUR
FÖRDERUNG DER WISSENSCHAFTEN E.V.
(50.0%)**

72 Inventor/es:

**HAUBER, JOACHIM;
CHEMNITZ, JAN;
BUCHHOLZ, FRANK y
CHUSAINOW, JANET**

74 Agente/Representante:

CURELL AGUILÁ, Mireia

ES 2 577 929 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

Recombinasa elaborada a medida para recombinar unos sitios diana asimétricos en una pluralidad de cepas de retrovirus.

La presente invención se refiere a un procedimiento para la preparación de un vector de expresión codificante de una recombinasa elaborada a medida ("tailored"), que puede recombinarse con secuencias diana asimétricas dentro de la repetición terminal larga (LTR) de ADN provírico de una pluralidad de cepas de VIH-1 que pueden insertarse en el genoma de una célula hospedadora, así como al vector de expresión obtenido, células transfectadas por el mismo, recombinasa expresada y composiciones farmacéuticas que comprenden el vector de expresión, células y/o recombinasa. Las composiciones farmacéuticas resultan útiles, por ejemplo, en el tratamiento y/o la prevención de la infección por VIH-1. En particular, se dan a conocer secuencias diana asimétricas presentes en una pluralidad de cepas de VIH-1, así como recombinasas elaboradas a medida que pueden combinarse con dichas secuencias (Tre 3.0 y 4.0) y vectores de expresión codificantes de las mismas.

Antecedentes de la técnica

Las infecciones retrovíricas, tales como, por ejemplo, las infecciones por el virus de la inmunodeficiencia humana (VIH) todavía son de las enfermedades humanas más importantes y más extendidas.

Un enfoque al tratamiento de los retrovirus, por ejemplo el VIH, es focalizarlo en el provirus insertado en el genoma de la célula hospedadora. La extracción del ADN provírico del genoma del huésped, por ejemplo, evitaría una replicación posterior del VIH y difiere de los procedimientos actuales en que presenta el potencial de erradicar incluso los virus latentes presentes en el genoma del huésped.

Una clase de proteínas que se ha considerado para la utilización en dicho enfoque alternativo son las recombinasas específicas de sitio (Flowers et al., 1997). Las recombinasas específicas de sitio median en una multitud de funciones en la naturaleza, desde reorganizaciones génicas a la segregación genómica, tal como, por ejemplo, la extracción, la inversión o la integración de unidades de ADN definidas (revisión en Stark et al., 1992).

Una de las recombinasas más simples y mejor entendidas es la recombinasa Cre del bacteriófago P1 que resuelve los dímeros genómicos en monómeros mediante recombinación entre dos sitios de ADN de doble cadena idénticos de una secuencia particular (Hoess y Abremski, 1985). La recombinasa Cre ha encontrado un amplio uso en la genética del ratón (Nagy, 2000). Cre es una proteína de 38 kDa denominada de esta manera por su función, ya que causa la recombinación (Sternberg y Hamilton, 1981). Un requisito previo para esta recombinación es la alineación de dos sitios de recombinación reconocidos por Cre en orientación antiparalela que seguidamente son ligadas por cuatro subunidades idénticas de Cre que se unen formando un anillo en el que cada subunidad entra en contacto con dos subunidades contiguas y un hemisito ("half-site") de un sitio de recombinación (Hoess y Abremski, 1985). El sitio de recombinación reconocido por Cre es una secuencia de ADN de doble cadena de 34 pb conocido como loxP (por sus siglas en inglés, "[locus of crossing over (x), P1]" [locus de entrecruzamiento (x), P1], Sternberg y Hamilton, 1981), que es palindrómico con la excepción de sus ocho pares de bases más internos (a los que se denomina "el espaciador"), que proporcionan direccionalidad al sitio.

Algunos sistemas de recombinación específica de sitio, incluyendo el sistema Cre/loxP, funcionan sin proteínas accesorias o cofactores y funcionan bajo una amplia diversidad de condiciones celulares. Sin embargo, debido a que las recombinasas específicas de sitio funcionan mediante interacciones específicas de las subunidades del enzima recombinasa con sus secuencias de ADN diana afines, la utilización de estos enzimas se encuentra restringida por el requisito de que las regiones de ADN diana deben contener sitios diana adecuadamente situados (Lewandoski, 2001). Todavía no se ha identificado ninguna recombinasa de tipo salvaje que reconozca secuencias retrovíricas naturales como sus secuencias diana de ADN.

En los últimos años se han llevado amplios análisis mutacionales y estructurales de las recombinasas específicas de sitio con el fin de alterar sus propiedades y alcanzar una mejor comprensión de los complejos mecanismos de estos enzimas (para una revisión ver Van Duyne, 2001, y Coates et al., 2005). Muchos estudios se han centrado en la recombinasa Cre para explorar su capacidad de evolución. Varios estudios han demostrado que la especificidad de la diana Cre podía alterarse al modificar unos cuantos nucleótidos en su sitio de reconocimiento loxP (Buchholz y Stewart, 2001; Santoro y Schultz, 2002; Rufer y Sauer, 2002). Algunos estudios adicionales se refieren a la manipulación de sitios diana de loxP mutado que contienen secuencias del LTR del VIH-1 para desarrollar posibles sitios diana para la utilización de Cre como estrategia antivírica (Lee y Park, 1998; Lee et al., 2000).

El procedimiento de evolución dirigida es un potente procedimiento para seleccionar enzimas con especificidades alteradas (revisión en Yuan et al., 2005, y Johannes y Zhao, 2006). Inicialmente este procedimiento se utilizó para aislar enzimas mejoradas basándose en el ARN mediante la selección de las moléculas de ARN con sitio de sustrato alterados. La utilización de procedimientos basados en la PCR permite el cribado de colecciones muy grandes y la recuperación de las regiones codificantes deseadas a partir de un grupo de candidatos. En contraste, en la evolución dirigida de proteínas, el cribado y la recuperación de mutantes mejorados, los cuales se identifican en alteraciones

de las propiedades de la proteína, requiere un procedimiento para recuperar la secuencia de ácidos nucleicos codificante de la proteína. El enlace entre la proteína y su secuencia codificante con frecuencia se ha mantenido mediante compartimentalización. En consecuencia, el cribado de colección ("library") en la evolución dirigida de proteínas se ha limitado a enfoques de "uno a uno" que mantienen los compartimientos, por lo que no se ha dispuesto de las ventajas asociadas al cribado de grupos de candidatos.

Dicha limitación se ha superado mediante el desarrollo de procedimientos que permiten el entrecruzamiento de proteínas con sus ARN mensajeros (ARNm) respectivos utilizando fusiones de ARNm-proteína y la expresión ribosómica. De esta manera, los cribados funcionales para propiedades mejoradas de las proteínas se han acoplado a la recuperación directa de las moléculas codificantes correspondientes y se han cribado grandes grupos in vitro (ver, por ejemplo, Buchholz et al., 1998). Una mejora adicional de la evolución dirigida de proteínas se ha conseguido mediante la denominada evolución de proteínas ligada a sustrato (SLiPE, Buchholz y Stewart, 2001), en la que el sustrato de la recombinasa se introduce en la misma molécula de ADN que la región codificante de proteína. De esta manera, al expresarse la recombinasa dentro de un compartimiento, su acción altera el sustrato de ADN contiguo a su propia región codificante. En consecuencia, puede cribarse una colección como grupo mediante PCR, amplificando únicamente regiones codificantes candidatas que se encuentran en posición contigua a un sustrato alterado. Lo anterior permite el cribado de grandes colecciones convenientemente para la rápida recuperación de regiones codificantes deseadas. Este procedimiento se ha aplicado a la alteración de la especificidad de ADN de la recombinasa Cre y su adaptación a un nuevo sitio diana de reconocimiento (Buchholz y Stewart, 2001).

A partir del potencial de las recombinasas específicas de sitio y la necesidad de encontrar una terapia para el SIDA que erradica el provirus de VIH-1 a partir del genoma de la célula hospedadora, el documento WO 2008/083931 da a conocer la generación de una recombinasa elaborada a medida (TRE) que es capaz de recombinar sitios diana asimétricos dentro de la LTR del ADN provirico de un retrovirus insertado en el genoma de una célula hospedadora, extrayendo de esta manera el provirus del genoma de la célula hospedadora. La recombinasa manipulada que se da a conocer en los ejemplos, TRE, reconoce un sitio específico de tipo loxP presente en una cepa de VIH-1 particular. El documento nº WO 2008/083931 aprecia que, debido a la elevada variabilidad de secuencias de los retrovirus, en particular del VIH, para el tratamiento de un paciente con una cepa de VIH diferente, podría resultar necesario adaptar una recombinasa elaborada a medida diferente, o preparar una colección de recombinasas que contenga recombinasas específicas para una diversidad de secuencias diana.

A la luz de lo anteriormente expuesto, los inventores ahora han dado pasos para resolver el problema de proporcionar una recombinasa elaborada a medida capaz de extraer una pluralidad de cepas de VIH. De esta manera, la recombinasa generada podría utilizarse para una pluralidad de infecciones por VIH sin generar una nueva recombinada para cada cepa. Este problema ha sido resuelto por la presente invención.

Descripción de la invención

Los presentes inventores ha conseguido proporcionar por primera vez un procedimiento para generar un vector de expresión codificante de una recombinasa elaborada a medida capaz de recombinar secuencias diana asimétricas dentro de la LTR del ADN provirico de una pluralidad de cepas de VIH-1 de una especie insertadas en el genoma de una célula hospedadora. Las recombinasas han sido adaptadas para reconocer sitios diana asimétricos diferentes de sus sitios diana simétricos naturales, mediante la división del sustrato en varios nuevos subconjuntos con diferencias más pequeñas respecto a la diana original y adaptando en etapas las recombinasas para reconocer dichos subconjuntos (documento nº WO 2008/083931). Un enfoque combinatorial permite la selección de moléculas funcionales que reconocen el sitio diana asimétrico dentro de una secuencia dada. De esta manera, utilizando el enfoque de recorrer varios intermediarios de sustrato durante la evolución molecular dirigida, resulta posible producir enzimas con nuevas especificidades remotas de diana asimétrica. La presente invención utiliza este enfoque al problema de proporcionar recombinasas elaboradas a medida capaces de recombinar una pluralidad de cepas de VIH-1. Han encontrado que, a pesar de la elevada variabilidad de las secuencias de los retrovirus, resulta posible identificar secuencias diana asimétricas presentes en una elevada proporción de los virus de un subtipo particular.

La invención proporciona un procedimiento para preparar un vector de expresión codificante de una recombinasa elaborada a medida, la cual es capaz de recombinar secuencias diana asimétricas dentro de la LTR del ADN provirico de una pluralidad de cepas de VIH-1 que pueden encontrarse insertadas en el genoma de una célula hospedadora, que comprende las etapas de identificar en la secuencia de la LTR del ADN provirico, una pluralidad de secuencias de cepas retrovíricas con una homología de por lo menos 30% respecto a la secuencia de hemisito izquierda y la secuencia de hemisito derecha de por lo menos un sitio diana de recombinasa conocido, en el que las secuencias homólogas se encuentran separadas por un espaciador de 5 a 12 nucleótidos, y en el que la secuencia diana asimétrica se encuentra en una pluralidad de cepas de VIH-1, en la que la secuencia diana asimétrica presenta la secuencia SEC ID nº 1, y de generar, mediante etapas repetidas de:

- i) evolución molecular dirigida en por lo menos una recombinasa que reconoce el sitio diana homólogo conocido utilizando como sustrato, secuencias diana modificadas basadas en la secuencia de la secuencia diana asimétrica pero modificadas para que contengan únicamente un número limitado de variaciones

respecto a la secuencia diana conocida, en la que, en cada ronda, la secuencia diana puede variar respecto a la secuencia diana sobre la que es conocido que actúa la recombinasa, en uno, dos o tres nucleótidos, y

- ii) mezclar ("shuffling") las colecciones de recombinasa para obtener colecciones de recombinasa capaces de recombinar secuencias diana más homólogas respecto a la secuencia diana asimétrica,

hasta obtener por lo menos una recombinasa que sea activa sobre la secuencia diana asimétrica dentro de la LTR del ADN retrovírico; aislar el ácido nucleico de por lo menos una recombinasa obtenida, y clonar el ácido nucleico codificante de la recombinasa en un vector de expresión adecuado, tal como se reivindica.

El procedimiento de generación de una recombinasa elaborada a medida dado a conocer en el documento n° WO 2008/083931 puede utilizarse para generar una recombinasa elaborada a medida capaz de recombinar una secuencia diana asimétrica.

La invención proporciona además un procedimiento para preparar un vector de expresión codificante de una recombinasa elaborada a medida, la cual es capaz de recombinar secuencias diana asimétricas dentro de la LTR de ADN provírico de una pluralidad de cepas de VIH-1 que pueden encontrarse insertadas en el genoma de una célula hospedadora, comprendiendo las etapas de:

- (a) identificación en la secuencia de la LTR del ADN provírico, una pluralidad de secuencias de cepa de VIH-1 con una homología de por lo menos 30% respecto a la secuencia de hemisio izquierdo y la secuencia de hemisio derecho de por lo menos un sitio diana de recombinasa conocido, en el que las secuencias homólogas se encuentran separadas por un espaciador de 5 a 12 nucleótidos, y en el que la secuencia diana asimétrica se encuentra en una pluralidad de cepas retrovíricas y presenta la secuencia SEC ID n° 1,

- (b) identificación de dos secuencias, en la que la primera secuencia corresponde a la secuencia de la secuencia diana asimétrica de la etapa (a) homóloga respecto al hemisio izquierdo de dicho sitio diana conocido y que se denomina "secuencia 1 de hemisio" y en el que la segunda secuencia corresponde a la secuencia de la secuencia diana asimétrica de la etapa (a) homóloga respecto al hemisio derecho y se denomina "secuencia 2 de hemisio",

- (c) determinación de los nucleótidos dentro de las secuencias de la etapa (b) que se apartan de las secuencias homólogas correspondientes del hemisio izquierdo y el hemisio derecho del sitio diana homólogo conocido de la etapa (a),

- (d) generación de un primer subconjunto de dos ácidos nucleicos diana que comprenden secuencias diana, en la que la primera secuencia diana se denomina subconjunto 1 y comprende, contiguas entre sí y en orden 5' a 3', la secuencia 1 de hemisio de la etapa (b), la secuencia espaciadora de la secuencia diana asimétrica y una repetición invertida de la secuencia 1 de hemisio, y en la que la segunda secuencia diana se denomina subconjunto 2 y comprende, contiguas y en orden 5' a 3', una repetición invertida de la secuencia 2 de hemisio, la secuencia espaciadora de la secuencia diana asimétrica y la secuencia 2 de hemisio de la etapa (b),

- (e) generación de un segundo subconjunto de ácidos nucleicos diana que comprenden secuencias diana modificadas, basada en las secuencias diana en el primer subconjunto de la etapa (d),

en la que, en secuencias basadas en el subconjunto 1, en la secuencia de hemisio izquierdo, una parte de los nucleótidos diferentes respecto de la secuencia de hemisio homóloga correspondiente del sitio o sitios diana conocidos de la etapa (a), se sustituye por los nucleótidos naturales presentes en dicho sitio diana conocido hasta que dicha secuencia de hemisio contenga uno, dos o tres nucleótidos diferentes de los presentes en dicho sitio diana conocido, en el que el hemisio derecho de dicha secuencia diana modificada está formado por una repetición invertida de dicha secuencia modificada del hemisio izquierdo, que se encuentra separada de dicha secuencia modificada de hemisio izquierdo por la secuencia espaciadora de la secuencia diana asimétrica, y

en la que, en las secuencias basadas en el subconjunto 2, en la secuencia de hemisio derecho, una parte de los nucleótidos diferentes respecto de la secuencia homóloga correspondiente de hemisio del sitio o sitios diana conocidos de la etapa (a) se sustituye por los nucleótidos naturales presentes en dicho sitio diana conocido hasta que dicha secuencia de hemisio contenga uno, dos o tres nucleótidos diferentes de los presentes en dicho sitio diana conocido, en el que el hemisio izquierdo de dicha secuencia diana modificada está formado por una repetición invertida de dicha secuencia modificada del hemisio derecho, que se encuentra separada de dicha secuencia modificada del hemisio derecho por la secuencia espaciadora de la secuencia diana asimétrica,

de manera que en la totalidad de las secuencias de hemisio modificadas que se originan a partir de una secuencia diana del primer subconjunto de la etapa (d) consideradas conjuntamente pueden encontrarse todos los nucleótidos diferentes, mientras que ninguna de dichas secuencias modificadas de hemisio por sí

sola comprende todos los nucleótidos diferentes,

(f) aplicar separadamente la evolución molecular dirigida a por lo menos una recombinasa que reconoce un sitio diana homólogo conocido según la etapa (a) utilizando como sustrato cada ácido nucleico del segundo subconjunto obtenido en la etapa (e),

(g) mezclar las colecciones de recombinasa evolucionadas en la etapa (f), en la que todas las colecciones de recombinasa evolucionadas en las secuencias basadas en el subsitio 1 se combinan y se mezclan, y en la que todas las colecciones de recombinasa evolucionadas en las secuencias basadas en el subsitio 2 se combinan y se mezclan,

(h) aplicar la evolución molecular dirigida a las colecciones mezcladas que se han obtenido en la etapa (g) utilizando como sustrato cada ácido nucleico del subconjunto según la etapa (d),

(i) mezclar las colecciones de recombinasa evolucionadas en la etapa (h),

(j) aplicar la evolución de proteínas ligada a sustrato a la colección mezclada que se ha obtenido en la etapa (g), utilizando un ácido nucleico que comprende la secuencia diana asimétrica de la etapa (a) como sustrato, hasta la obtención de por lo menos una recombinasa que es activa sobre la secuencia diana asimétrica dentro de la LTR del ADN retroviro de la etapa (a),

(k) aislar el ácido nucleico de la recombinasa o recombinasas obtenidas en la etapa (j) a partir de la colección, y

(l) clonar el ácido nucleico obtenido en la etapa (k) en un vector de expresión adecuado.

En la etapa (a) del procedimiento de la presente invención, puede determinarse la secuencia de la LTR del ADN provírico, tal como mediante, por ejemplo, secuenciación del ADN utilizando inhibidores terminadores de cadena (Sanger et al., 1977). Sin embargo, en el caso de que la secuencia de la ARTL del ADN retroviro insertado en el genoma del huésped ya haya sido determinado, la secuencia puede determinarse mediante referencia a una base de datos. Basándose en la información de secuencia, se llevó a cabo un análisis computerizado de la información de secuencia con el fin de identificar en la misma las secuencias con una homología de por lo menos 30% respecto a las secuencias del hemisitio izquierdo y del hemisitio derecho de los sitios diana conocidos, respectivamente, de recombinasas conocidas que están separadas por un espaciador adecuado de 5 a 12 nucleótidos, en el que la secuencia de diana asimétrica está presente en una pluralidad de cepas retrovíticas. Preferentemente, la homología respecto a las secuencias del hemisitio izquierdo y del hemisitio derecho de los sitios diana conocidos es de por lo menos 40% o de por lo menos 50%. Preferentemente, dichas cepas de VIH-1 son de un subtipo. Preferentemente, una pluralidad de cepas comprende más de 10 cepas, más preferentemente, más de 100 cepas, más de 130 cepas, más de 200 cepas o más de 300 cepas de VIH-1. Las cepas son de un subtipo del virus VIH-1, por ejemplo VIH-1 subtipo A y B o VIH-1 subtipo B. De esta manera, la recombinasa obtenida o vector de expresión codificante de la misma puede utilizarse para el tratamiento de la infección por una pluralidad de cepas, por ejemplo más de 50%, más de 70%, más de 80%, más de 90% o la totalidad de las cepas conocidas de un retrovirus o subtipo del mismo.

El término "recombinasa" tal como se utiliza en la presente memoria se refiere a una proteína que participa en la recombinación. De esta manera, dichas recombinasas reconocen y se unen a dos secuencias de ADN específicas denominadas "sitios de recombinación" o "sitios diana" y median en la recombinación entre estos dos sitios diana. De acuerdo con lo anterior, el término "recombinasa" pretende referirse a cualquier componente proteína de cualquier sistema recombinante que media en las reorganizaciones de ADN en un locus específico del ADN. Las recombinasas naturales reconocen sitios diana simétricos que consisten de dos secuencias idénticas denominadas "medio sitio" de aproximadamente 9 a 20 pb que forman una repetición invertida, en la que las secuencias de hemisitio están separadas por una secuencia espaciadora de 5 a 12 pb. Las recombinasas de la familia de la tirosina integrasa se caracterizan porque presentan una tirosina como el nucleófilo de sitio activo que es utilizado para el corte del ADN, mientras que las recombinasas de la familia de la serina integrasa utilizan una serina en lugar de una tirosina.

En una forma de realización de la presente invención, la recombinasa o recombinasas conocidas con la secuencia diana utilizada en la etapa (a) y en las que se aplica la evolución molecular dirigida en las etapas (h) y (j) pertenecen a la familia de las serina integrasas. Las recombinasas preferentes pertenecientes a la familia de las serina integrasas se seleccionan de entre el grupo que consiste en phiC31 integrasa (Combes et al., 2002), cualquier componente de los sistemas de recombinación Gin o Hin, la Tn3 resolvasa (Krasnow y Cozzarelli, 1983) o cualquier otro elemento de las serina recombinasas grandes, Rag1, Rag2 o cualquier otro componente del sistema de recombinación VDJ o variantes de los mismos.

En otra forma de realización, dicha recombinasa pertenece a la familia de las tirosina integrasas. Las recombinasas preferentes pertenecientes a la familia de las tirosina integrasas se seleccionan de entre el grupo que consiste en Cre del fago P1 (Abremski et al., 1983, 1984), la recombinasa FLP de la levadura (Volkert y Broach, 1986), Dre del fago D6 (Sauer y McDermott, 2004), la recombinasa R del plásmido pSR1 de *Zygosaccharomyces rouxii*, la

recombinasa A del plásmido pKD1 de *Kluveromyces drosophilarius*, una recombinasa del plásmido pKW1 de *Kuveromyces waltii*, Tnpl del transposón Tn4430 de *Bacillus*, cualquier componente del sistema de recombinación Int de λ o variantes de los mismos. Preferentemente, dicha recombinasa es la recombinasa Cre o una variante de la misma. Por ejemplo, puede utilizarse una recombinasa elaborada a medida que se da a conocer en el documento WO 2008/083931 (Tre).

El término variante en el presente contexto se refiere a proteínas que se derivan de las proteínas anteriormente indicadas mediante delección, sustitución y/o adición de aminoácidos y que conservan algunas o todas las funciones inherentes a la proteína de la que se derivan.

En una forma de realización preferida, la recombinasa conocida es una recombinasa quimérica obtenida mediante, por ejemplo, "mezcla de familias", tal como se indica en Cramer et al. (1998). El prerequisite para la utilización de la mezcla de familias es una homología significativa entre las recombinasas utilizadas para generar las recombinasas quiméricas. Un ejemplo de una recombinasa quimérica que puede utilizarse en la presente invención es una recombinasa quimérica que consiste en secuencias de la recombinasa Cre y de la recombinasa Dre, respectivamente.

En una forma de realización más preferida, la recombinasa es la recombinasa Cre que reconoce un sitio diana simétrico de 34 pb conocido como loxP (SEC ID nº 3). El sitio loxP (y también otros sitios de recombinación de recombinasas de tipo salvaje) es palindrómico con dos repeticiones de 13 pb separadas por los ocho pares de bases más internos, que representan el denominado espaciador, que proporciona direccionalidad al sitio. La recombinación tiene lugar mediante corte dentro de la secuencia espaciadora. Dependiendo de la localización y orientación relativas de los dos sitios loxP participantes, Cre cataliza la integración, extracción o reorganización del ADN (Hoess y Abremski, 1985).

Una recombinasa útil es Zre (SEC ID nº 26) aislada a partir de *Salmonella enterica*, o variantes, fragmentos u homólogos de la misma, por ejemplo con una homología de por lo menos aproximadamente 70%, de por lo menos aproximadamente 80%, de por lo menos aproximadamente 90% o de por lo menos aproximadamente 95% respecto a la secuencia de tipo salvaje, y que presenta función de recombinasa. Las recombinasas Zre recombinan el ADN en los sitios zox (figura 1). Pueden utilizarse para iniciar el procedimiento de la invención solos o en el contexto de una colección.

En una forma de realización, se utiliza una colección de recombinasa como punto inicial para la evolución molecular, por ejemplo una colección de recombinasa que comprende diferentes recombinasas de tipo salvaje y/o adaptadas/mezcladas, por ejemplo tal como se indica posteriormente o en el Ejemplo 2. Dicha colección se utiliza preferentemente como punto de partida para la generación de recombinasas elaboradas a medida capaces de reconocer la SEC ID nº 1.

La recombinasa elaborada a medida es capaz de recombinar secuencias diana asimétricas dentro de la LTR del ADN provírico de una pluralidad de cepas de VIH-1. El ADN provírico que es la diana de la recombinasa puede insertarse en el genoma de una célula hospedadora. Alternativamente, la recombinasa elaborada a medida de la invención puede recombinar secuencias diana asimétricas dentro de la LTR del ADN provírico de una pluralidad de cepas de VIH-1 que no se encuentran (todavía) integradas en el genoma de una célula hospedadora, es decir, que se encuentran presentes como complejo pre-integración (CPI) no integrado. De esta manera, el VIH que todavía no se ha integrado en el genoma de la célula hospedadora, así como el VIH que ya se ha integrado, puede ser inactivado por la recombinasa elaborada a medida de la invención.

Debe indicarse que en la presente invención y también en la técnica, las expresiones "secuencia diana", "sitio diana" y "sitio de recombinación" se utilizan intercambiamente.

Al contrario que las recombinasas naturales que reconocen un sitio diana simétrico, el procedimiento de la presente invención proporciona recombinasas elaboradas a medida que reconocen sitios diana, que no consisten de secuencias palindrómicas separadas por un espaciador. Por el contrario, en los sitios diana asimétricos, las secuencias no forman una repetición invertida simétrica. De acuerdo con lo anterior, una recombinasa elaborada a medida capaz de reconocer un sitio diana asimétrico debería reconocer y recombinar sitios diana que consisten de medios sitios de secuencia variable.

Dentro de un sitio diana simétrico, las secuencias denominadas "medio sitio izquierdo" y "medio sitio derecho" se definen a partir de su homología con el hemisito izquierdo y derecho, respectivamente, de un sitio diana conocido. La secuencia situada entre las secuencias homólogas de los medios sitios izquierdo y derecho de un sitio diana conocido se denomina espaciador.

Sin embargo, en el caso de que se encuentren secuencias en la LTR que presenten homología únicamente con la secuencia de hemisito izquierda o la de hemisito derecha de un sitio diana conocido, estas secuencias podrían utilizarse todavía en la práctica de la presente invención. El tamaño del sitio diana perteneciente a la recombinasa, la secuencia diana natural de la cual muestra homología respecto a secuencias dentro de la LTR, es conocido por el

experto en la materia. Por ejemplo, en el caso de que se encuentre homología dentro de la secuencia de la LTR respecto a una secuencia diana reconocida por la recombinasa Cre, un sitio diana asimétrico que debe ser reconocido por la recombinasa Cre debería consistir de 34 nucleótidos con dos secuencias de hemisio de 13 nucleótidos cada una separadas por un espaciador de 8 nucleótidos. De acuerdo con lo anterior, la secuencia homóloga dentro de la LTR se define como el hemisio izquierdo o derecho o el espaciador del sitio diana asimétrico dependiendo de la homología respecto de la secuencia del sitio diana conocido. De esta manera, las secuencias con homología respecto del hemisio izquierdo de una secuencia diana conocida se definen como hemisio izquierdo y las secuencias con homología respecto al hemisio derecha de una secuencia diana conocida se definen como hemisio derecha. Partiendo de dicha definición, las otras partes de los sitios diana asimétricos se definen considerando la estructura del sitio diana conocido. De esta manera, tras definir, por ejemplo, una secuencia de hemisio derecha dentro de la LTR con homología respecto a un sitio loxP (reconocido por la recombinasa Cre), pueden definirse con facilidad las otras secuencias correspondientes al espaciador y al hemisio izquierdo de la secuencia diana asimétrica. La secuencia del espaciado se define, por ejemplo, contando 8 nucleótidos cadena arriba del extremo 5' de la secuencia definida como secuencia de hemisio derecha, mientras que la secuencia de hemisio izquierda se define de manera similar contando 13 nucleótidos cadena arriba del extremo 5' de la secuencia del espaciador definida anteriormente.

La homología en el presente contexto, así como en toda la solicitud, se refiere a identidad de secuencia. Una comparación preferente para los fines de homología es comparar por lo menos dos secuencias utilizando técnicas estándares conocidas de la técnica, incluyendo, aunque sin limitación, el algoritmo de homología local de Smith y Waterman (1981), el algoritmo de alineación de homología de Needleman y Wunsch (1970) o el procedimiento de búsqueda de similitud de Pearson y Lipman (1988). Para los fines de la presente solicitud, la homología de secuencias se determina preferentemente utilizando el programa informático ClustalW, disponible del European Bioinformatics Institute (EBI), a menos que se indique lo contrario.

En vista del requisito de que deben encontrarse presentes dos sitios diana idénticos en el genoma del provirus para permitir que la recombinasa extraiga la secuencia entre estos dos sitios diana, se escanean las secuencias del ADN provírico en la etapa (a) del procedimiento de la presente invención que se encuentran presentes por lo menos dos veces en el genoma. Estas secuencias son, por ejemplo, las secuencias de la LTR del ADN provírico. De acuerdo con lo anterior, se escanea preferentemente la secuencia de la LTR, ya que la 5'-LTR y la 3'-LTR del ADN provírico son idénticas. Un sitio diana asimétrico presente en la 5'-LTR también se encuentra presente en la 3'-LTR y, de esta manera, permite la extracción del ADN provírico situado entre las LTR.

Aparte de las secuencias identificadas dentro de la secuencia de la LTR que presentan suficiente homología respecto a los sitios diana conocidos, preferentemente se seleccionan secuencias que presentan la homología máxima respecto a la secuencia del sitio diana de las recombinasas conocidas. Sin embargo, también resulta posible seleccionar secuencias diferentes de las que presentan la homología máxima, por ejemplo las que se encuentran presentes en el máximo número de cepas retrovíricas, o en las cepas retrovíricas de interés, por ejemplo en el caso de que el paciente se encuentre infectado por una cepa particular.

Debe indicarse que el potencial del procedimiento de la presente invención incluso permite adaptar recombinasas que reconocen sitios diana asimétricos con una homología inferior al 30% respecto a los sitios diana conocidos, por ejemplo con una homología de por lo menos 11% o de por lo menos 20%. Sin embargo, con el fin de garantizar la presencia de una actividad de recombinación residual con el sitio diana asimétrico respectivo, preferentemente se escanea para secuencias que presentan una homología de por lo menos 30% respecto a las secuencias del hemisio izquierdo y del hemisio derecho de los sitios diana conocidos de las recombinasas conocidas. En formas de realización preferentes adicionales, se seleccionan secuencias diana asimétricas con una homología de por lo menos 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79 o 80%, más preferentemente 85%, particularmente preferentemente 90% y todavía más preferentemente 95% respecto a las secuencias del hemisio izquierdo y del hemisio derecho de sitios diana conocidos de las recombinasas conocidas.

La secuencia seleccionada en la LTR presenta homología respecto a los sitios diana loxP simétricos reconocidos por la recombinasa específica de sitio Cre.

En una forma de realización, se utiliza una colección de recombinasas como punto de partida para la evolución molecular, por ejemplo una colección de recombinasa que comprende diferentes recombinasas de tipo salvaje y/o adaptadas/mezcladas, tales como la colección indicada en el Ejemplo 2. Una colección ejemplar comprende Cre y recombinasas derivadas de la misma. Puede comprender además Tre, Dre, recombinasas de *Salmonella* y de *Shewanella* y/o recombinasas derivadas de las mismas. La colección puede comprender, por ejemplo, Cre, Dre, Dre "tipo Cre" (SEC ID nº 24), recombinasa de *Shewanella* (Shew), Shew "tipo Cre" (SEC ID nº 25) y/o Zre (SEC ID nº 26). En la figura 3A se muestra una colección. Tre es una recombinasa elaborada a medida tal como se da a conocer en el documento nº WO 2008/083931, a la que también se hace referencia como Tre 1.0.

En una forma de realización, todas las recombinasas en la colección reconocen una secuencia diana con la misma longitud de espaciador. La longitud total de las secuencias de hemisio 1 y 2, incluyendo el espaciador,

preferentemente es de 34 nucleótidos.

En el caso de que la recombinasa o recombinasas sean una colección de recombinasa, la homología será homología respecto al grupo de sitios diana de recombinasa conocidos (es decir, la homología en una posición dada respecto a por lo menos una de las secuencias diana se define como homología). En consecuencia, en la etapa (c), únicamente los nucleótidos que no correspondan a un nucleótido en por lo menos una de las secuencias diana conocidas se definen como nucleótidos divergentes. En el caso de una colección de recombinasas, un "nucleótido nativo" en la etapa (c) puede ser un nucleótido presente en dicha posición en cualquiera de las secuencias diana conocidas, preferentemente es un nucleótido presente en dicha posición en varias o en la mayoría de las secuencias diana conocidas.

Con el fin de identificar las secuencias diana presentes en una pluralidad de cepas retrovíricas, pueden utilizarse los sitios de reconocimiento de recombinasas conocidos, los cuales han sido descritos en la literatura, como criterio para una búsqueda de secuencias diana asimétricas conservadas frente a un tramo genómico. Sin embargo, dada la naturaleza repetitiva de las regiones, está excluida la utilización de herramientas estándares de búsqueda de similitud de secuencias. Sarkar et al., 2007, han utilizado BLAST (Altschul et al., 1997) para buscar un sitio de unión de tipo lox en diferentes cepas de VIH. La búsqueda BLAST para el sitio de tipo lox realizada en secuencias de LTR de VIH-1 ha resultado en que se ha encontrado únicamente un sitio presente en una única cepa. Sin embargo, en el caso de que deban utilizarse recombinasas como agentes terapéuticos contra genomas retrovíricos, resulta crítico manipular las recombinasas para que presenten como diana sitios de reconocimiento presentes en el máximo número posible de cepas del retrovirus.

Debido a que BLAST no funciona bien con dichas secuencias redundantes cortas, se consideró la utilización de HMMER (Eddy et al., 1998), RepeatMasker o el programa de palíndromos del grupo Emboss de paquetes. Se desarrolló HMMER para encontrar homólogos remotos basándose en un modelo probabilístico de un patrón de secuencias que se busca, lo que no es la naturaleza de la búsqueda que se pretende llevar a cabo. Podría forzarse que HMMER realizase esta búsqueda, pero la cantidad de pre- y post-procesamiento de los datos y parámetros incrementaría la tendencia a errores y resultaría ineficiente. RepeatMasker sólo realiza búsquedas de repeticiones y regiones de baja complejidad que ya han sido bien caracterizadas, lo que nuevamente no es la naturaleza de esta búsqueda. El programa de palíndromos Emboss es el que más se aproxima a resolver el problema de las búsquedas, pero no permite definir la búsqueda sino que, por el contrario, resulta en una lista de posibilidades de sitios de tipo lox. Estos deberían emparejarse a continuación con la signatura de tipo lox de interés. Claramente esta estrategia sólo complicaría y haría más engorrosa la búsqueda. Para resolver la cuestión de la falta de un programa y un procedimiento para encontrar sitios de tipo lox, resultó necesario desarrollar un programa capaz de realizar búsquedas de sitios de tipo lox degenerados en secuencias genómicas.

Puede identificarse una secuencia diana asimétrica en una pluralidad de cepas retrovíricas con dicho programa, utilizando un algoritmo basado en una matriz de pesos de las posiciones para las regiones flanqueantes basada en un sitio de reconocimiento conocido de una recombinasa. Preferentemente, para que la búsqueda resultase computacionalmente eficiente, se utilizan operaciones binarias en las secuencias tras transformarlas en cadenas de bits.

En la invención, el retrovirus es el VIH-1. Para el VIH-1, se determinaron secuencias diana asimétricas adecuadas que presentaban la secuencia indicada como SEC ID nº 1 o SEC ID nº 2, posteriormente. La invención reivindicada se basa en la SEC ID nº 1.

Las secuencias de hemisito izquierdo y de hemisito derecho se encuentran subrayadas y el espaciador se muestra en negrita:

SEC ID nº 1 AACCCACTGCTTA**AGCCTCAATA**AAGCTTGCCTT

SEC ID nº 2 CTGGGCGGGACTG**GGGAGTGGC**GAGCCCTCAGAT

La SEC ID nº 1 es idéntica en 92% de las cepas de VIH-1 subtipo B (348/379) y en 80% de las cepas de VIH-1 subtipo A (32/40) buscadas. La SEC ID nº 2 es idéntica en 76% de las cepas de subtipo B buscadas (288/379). El 82% de las cepas de subtipo C comprendía también dicha secuencia. La SEC ID nº 2 no se encontraba presente en ninguna de las cepas de subtipo A buscadas.

Tal como se muestra en la figura 1, la SEC ID nº 1 presentaba una homología de 54% respecto a un grupo de sitios diana de recombinasa conocidos y la SEC ID nº 2 presentaba una homología de 42% respecto al grupo de dichas secuencias (con respecto a los hemisitos izquierdo y derecho, respectivamente). La homología respecto a sitios diana conocidos individuales era más baja, por ejemplo de por lo menos 30% para la SEC ID nº 1 y de por lo menos 11% para la SEC ID nº 2. En particular en el caso de la homología individual baja respecto a los sitios diana conocidos, puede resultar ventajoso utilizar una colección de recombinasas como material de partida, por ejemplo para generar una recombinasa elaborada a medida capaz de recombinar SEC ID nº 1 o SEC ID nº 2, una colección que comprende Cre, Fre, Dre, Zre y Tre.

En la etapa (b) del procedimiento de la invención, la secuencia del sitio diana asimétrico dentro de la LTR de los provirus que es homóloga respecto al hemisito izquierdo del sitio diana conocido se define como "secuencia 1 de hemisito". La secuencia del sitio diana asimétrico dentro de la LTR de los provirus que es homóloga respecto al hemisito derecho del sitio diana conocido se define como secuencia 2 de hemisito. La secuencia entre las secuencias que representan los medios sitios izquierdo y derecho se denomina espaciador.

En la etapa (c), los nucleótidos en la "secuencia 1 de hemisito" y en la "secuencia 2 de hemisito" de las secuencias de la etapa (b) que divergen de las secuencias de las secuencias homólogas correspondientes de los medios sitios izquierdo y derecho, respectivamente, de la diana conocida se determinan mediante alineación y comparación de las secuencias. En este contexto, la secuencia de la "secuencia 1 de hemisito" se compara con el hemisito natural correspondiente, que preferentemente es la secuencia de hemisito izquierda, mientras que la secuencia de la "secuencia 2 de hemisito" se compara con el otro hemisito que forma el sitio diana natural palindrómico, que preferentemente es la secuencia de hemisito derecha.

La figura 1 muestra el resultado de dicha comparación para las SEC ID nº 1 y nº 2, en comparación con una colección de recombinasas. Los nucleótidos divergentes se muestran en negrita.

Dicha comparación no debe llevarse a cabo necesariamente después de la etapa (b) y antes de la etapa (d) del procedimiento de la invención, sino que también puede llevarse a cabo en una etapa diferente del procedimiento después de la etapa (a) y antes de la etapa (e).

En la etapa (d), se genera un primer subconjunto de dos ácidos nucleicos diana que comprenden secuencias diana, en el que la primera secuencia diana se denomina subsitio 1 y comprende, contiguos entre sí y en orden 5' a 3', la secuencia 1 de hemisito de la etapa (b), la secuencia espaciadora de la secuencia diana asimétrica y una repetición invertida de la secuencia 1 de hemisito, y en la que la secuencia diana se denomina subsitio 2 y comprende, contiguos entre sí y en orden 5' a 3', una repetición invertida de la secuencia 2 de hemisito, la secuencia espaciadora de la secuencia diana asimétrica y la secuencia 2 de hemisito de la etapa (b). Las secuencias diana del primer subconjunto son secuencias oligonucleótidas palindrómicas que presentan la estructura de un sitio diana simétrico. Estos sitios diana simétricos artificiales se sintetizan basándose en las secuencias de hemisito de la etapa (b) mediante la complementación de la secuencia de hemisito faltante en cada secuencia oligonucleótida como repetición invertida, en la que la "secuencia 1 de hemisito" y la "secuencia 2 de hemisito", respectivamente, se utilizan para complementar la secuencia del segundo hemisito en el extremo opuesto de la secuencia espaciadora. De acuerdo con lo anterior, la primera secuencia diana en el primer subconjunto (denominada "subsitio 1") comprende una repetición invertida que consiste en la "secuencia 1 de hemisito" y la "secuencia 1 de hemisito" repetida inversamente, separadas por la secuencia espaciadora, mientras que la segunda secuencia diana en el primer subconjunto (denominada "subsitio 2") comprende una repetición invertida que consiste en la "secuencia 2 de hemisito" repetida inversamente y la "secuencia 2 de hemisito" separadas por la secuencia espaciadora. En el "subsitio 1", la secuencia se organiza de la manera siguiente: 5'-"secuencia 1 de hemisito"-espaciador-"repetición invertida de la secuencia 1 de hemisito"-3', en el "subsitio 2", la secuencia se organiza de la manera siguiente: 5'-"repetición invertida de la secuencia 2 de hemisito"-espaciador-"secuencia 2 de hemisito"-3'.

Las secuencias espaciadoras dentro de cada dos secuencias diana sintéticas del primer subconjunto preferentemente son idénticas y corresponden a la secuencia de la LTR que representa o define la secuencia espaciadora del sitio diana asimétrico. Sin embargo, en una forma de realización adicional, las secuencias espaciadoras pueden comprender una o dos divergencias de secuencia originadas en sustituciones de nucleótidos.

Generalmente, dicha etapa representa una primera división de las secuencias del sitio diana asimétrico seleccionadas para adaptar una recombinasa específica (ver la figura 1 del documento nº WO 2008/083931, que se incorpora por completo en la presente memoria como referencia, y la figura 2 de la presente solicitud). Se generan secuencias en esta etapa que incluyen sitios diana simétricos derivados de los medios sitios del sitio diana asimétrico seleccionado para adaptar una recombinasa específica. En consecuencia, cada mutación (es decir, cada diferencia con el sitio diana reconocido por la recombinasa de tipo salvaje) presente en un hemisito de dicho sitio diana asimétrico ahora se ha extendido entre las secuencias diana simétricas en el primer subconjunto.

En la etapa (e) del procedimiento de la invención, se genera un segundo subconjunto de ácidos nucleicos diana que comprende secuencias diana modificadas, basándose en las secuencias diana en el primer subconjunto de la etapa (d). En las secuencias basadas en el subsitio 1, en la secuencia de hemisito izquierda, una parte de los nucleótidos divergentes respecto a la secuencia de hemisito homóloga correspondiente del sitio o sitios diana conocidos de la etapa (a) se sustituye por los nucleótidos naturales que se encuentran en dicho sitio diana conocido, hasta que dicha secuencia de hemisito contenga uno, dos o tres (preferentemente dos) nucleótidos divergentes respecto a dicho sitio diana conocido, en el que el hemisito derecho de dicha secuencia diana modificada está formado o una repetición invertida de dicha secuencia de hemisito izquierdo modificada, que se encuentra separada de dicha secuencia modificada de hemisito izquierdo por la secuencia espaciadora de la secuencia diana asimétrica.

En secuencias basadas en el subsitio 2, en la secuencia de hemisito derecha, una parte de los nucleótidos

divergentes respecto a la secuencia de hemisio homóloga correspondiente de por lo menos un sitio diana conocido de la etapa (a) se sustituye por los nucleótidos naturales que se encuentran en dicho sitio diana conocido, hasta que dicha secuencia de hemisio contenga uno, dos o tres (preferentemente dos) nucleótidos divergentes respecto a dicho sitio diana conocido, en el que el hemisio izquierdo de dicha secuencia diana modificada está formada de una repetición invertida de dicha secuencia de hemisio derecho modificada, que se encuentra separada de dicha secuencia de hemisio derecho modificada por la secuencia espaciadora de la secuencia diana asimétrica.

Por ejemplo, en el caso de que un subsitio comprenda seis nucleótidos divergentes, tal como ambos subsitos basados en la SEC ID nº 1 o el subsitio 2 de SEC ID nº 2 con respecto a la colección de recombinasas mostrada en la figura 1, pueden generarse tres secuencias diana modificadas basándose en el subsitio, conteniendo cada una dos nucleótidos divergentes (diferentes) en el hemisio izquierdo (en caso de basarse en el subsitio 1) o el hemisio derecho (en caso de basarse en el subsitio 2). En consecuencia, en cada secuencia diana modificada, la secuencia del subsitio respectivo se modifica para que corresponda con la secuencia de la secuencia diana conocida (o por lo menos una secuencia diana conocida) en cuatro nucleótidos (figura 2). Evidentemente también resulta posible generar seis secuencias diana modificadas, conteniendo cada una, uno de los nucleótidos divergentes, o dos secuencias diana, conteniendo cada una, tres de los nucleótidos divergentes.

En otro ejemplo, en el caso de que un subsitio comprenda nueve nucleótidos divergentes, tal como el subsitio 1 de SEC ID nº 2, con respecto a la colección de recombinasas mostrada en la figura 1, pueden generarse tres secuencias diana modificadas basándose en el subsitio, conteniendo cada una tres nucleótidos divergentes (diferentes) en el hemisio.

En consecuencia, en todas las secuencias de hemisio modificadas que originan a partir de una secuencia diana del primer subconjunto de la etapa (d) conjuntamente, pueden encontrarse todos los nucleótidos divergentes, mientras que ninguna de dichas secuencias de hemisio modificadas por sí sola comprende la totalidad de los nucleótidos divergentes.

Nuevamente, se genera una repetición invertida basándose en la secuencia de hemisio modificada, de manera que la secuencia espaciadora separa ambas secuencias que forman la repetición invertida (ver la figura 2). Las secuencias espaciadoras dentro de cada secuencia diana modificada de un nuevo subconjunto que se derivan de una secuencia diana de un subconjunto más elevado preferentemente son idénticas y corresponden a la secuencia de la LTR que representa o que define la secuencia espaciadora del sitio diana asimétrico. Sin embargo, en una forma de realización adicional, las secuencias espaciadoras pueden comprender una o más divergencias de secuencia originadas en sustituciones de nucleótidos. Utilizando este enfoque, el número de mutaciones (es decir, diferencias respecto al sitio diana reconocido por la recombinasa de tipo salvaje) en las secuencias diana que representan cada subconjunto es menor que en la secuencia diana asimétrica inicial, aunque todas las mutaciones todavía se encuentran representadas en una de las secuencias diana (ver la figura 1 del documento WO 2008/083931, figura 2 de la presente solicitud).

La expresión "nucleótido que se diferencia" tal como se utiliza en la presente memoria se refiere a un nucleótido en la secuencia diana asimétrica identificado o definido en la LTR o en una secuencia diana de un subconjunto generado según la presente invención que se diferencia (es decir, es diferente) del nucleótido presente en la misma posición en la secuencia homóloga correspondiente de la secuencia diana simétrica homóloga conocida de una recombinasa conocida seleccionada en la etapa (a) del procedimiento de la presente invención. En el presente contexto, las expresiones "nucleótidos que se diferencian" y "mutaciones" se utilizan intercambiamente.

El documento WO 2008/08393 enseña que las recombinasas pueden adaptarse utilizando evolución molecular dirigida utilizando secuencias diana como sustratos, en el caso de que la secuencia diana utilizada como sustrato difiera en no más de 3 nucleótidos respecto a la secuencia diana natural. De esta manera, la generación de subconjuntos de diferentes órdenes indicados anteriormente sirve para reducir el número de nucleótidos divergentes por cada secuencia diana a 3 o menos (ver la figura 1 del documento nº WO 2008/083931). La reducción en etapas del número de nucleótidos divergentes finalmente rinde un número de subconjuntos de secuencias diana de diferentes órdenes con un número decreciente de nucleótidos divergentes hasta crear un subconjunto final que puede utilizarse como sustrato para la evolución molecular dirigida. Durante la creación de los diferentes subconjuntos, reduciendo de esta manera el número de nucleótidos divergentes, las diferencias respecto al sitio diana reconocidas por la recombinasa de tipo salvaje se distribuyen entre varias secuencias diana que no comprenden más de 3 de dichos nucleótidos divergentes cada una, mientras que las secuencias diana del orden final globalmente siguen representando la totalidad de los nucleótidos divergentes.

Opcionalmente, en el procedimiento de la invención, pueden generarse subconjuntos adicionales de secuencias diana partiendo de las secuencias diana del segundo subconjunto mediante la repetición por etapas del procedimiento de la etapa (e), es decir, dividiendo las secuencias diana en las secuencias de hemisio respectivas y generando nuevas estructuras palindrómicas basadas en dichas secuencias de hemisio tras alterar la secuencia de hemisio derivado de una secuencia diana del segundo subconjunto, generando cada vez un nuevo subconjunto de secuencias diana, en las que las secuencias de hemisio utilizadas para generar las repeticiones invertidas contienen menos nucleótidos divergentes respecto a la secuencia de hemisio homóloga correspondiente del sitio o

sitios diana conocidos. Estas secuencias diana adicionales pueden utilizarse para etapas adicionales de evolución molecular dirigida y la mezcla de colecciones de recombinasas. Evidentemente, dicha etapa adicional únicamente puede llevarse a cabo para algunas de las secuencias, por ejemplo para secuencias para las que se obtienen recombinasas de baja eficiencia de recombinación. En el caso de que se generen subconjuntos adicionales y se evolucionen recombinasas a partir de ellas, las colección evolucionada de recombinasas se utiliza en la etapa (f) del procedimiento de la invención.

Partiendo del segundo subconjunto de secuencias diana obtenido en la etapa (e), puede generarse un tercer subconjunto, seguido de un cuarto, quinto, sexto, etc. subconjuntos, en caso necesario. Sin embargo, la generación del tercer subconjunto generalmente sólo resulta necesaria en el caso de que las secuencias diana del segundo subconjunto sigan conteniendo más de tres nucleótidos divergentes. Lo mismo se aplica a la generación de los siguientes subconjuntos, los cuales únicamente resultan necesarios en el caso de que las secuencias diana del subconjunto anterior todavía contengan más de tres nucleótidos divergentes. De acuerdo con lo anterior, según el número de nucleótidos divergentes en cada secuencia de hemisitio, puede diferir el número de subconjuntos generados para cada secuencia de hemisitio del sitio diana asimétrico. Puede resultar necesario, por ejemplo, generar únicamente dos subconjuntos para la secuencia de hemisitio izquierda, mientras que deben generarse tres o cuatro subconjuntos para el hemisitio derecho con el fin de extender los nucleótidos divergentes entre varias secuencias diana de manera que una única secuencia diana no comprenda más de 3 de dichos nucleótidos divergentes.

El principio de generar subconjuntos adicionales de las secuencias diana para reducir el número de nucleótidos divergentes a un número inferior a tres se ilustra en la figura 1 del documento WO 2008/083931, y la figura 2 de la presente solicitud proporciona ejemplos específicos de secuencias diana modificadas.

En la etapa (f), se aplica un procedimiento de evolución molecular dirigida a la recombinasa o recombinasas que reconocen un sitio diana homólogo conocido de la etapa (a), utilizando una secuencia diana del subconjunto final o del segundo subconjunto en la etapa (e) que contiene uno, dos o tres nucleótidos que divergen de la secuencia de hemisitio homóloga correspondiente de dicho sitio diana homólogo conocido como sustrato.

La expresión "subgrupo final" tal como se utiliza en la presente memoria se refiere al último subconjunto generado en la etapa (e), es decir, en el caso de que no se generen subconjuntos adicionales en el segundo subconjunto. Dependiendo del número de nucleótidos divergentes en el sitio diana asimétrico y del número de subconjuntos que debían generarse para reducir el número de nucleótidos divergentes por cada secuencia diana a menos de 3, el "subgrupo final" puede corresponder a cualquier subconjunto, por ejemplo el segundo, tercer, cuarto o posterior subconjunto, y puede ser diferente para las secuencias de hemisitio de la secuencia diana asimétrica en el LTR. En el caso de que las recombinasas hayan sido previamente evolucionadas en subconjuntos adicionales de secuencias diana modificadas con menos nucleótidos divergentes respecto a la secuencia de hemisitio homóloga correspondiente de dicho sitio diana homólogo conocido, se utiliza la recombinasa obtenida en dicha etapa.

Evidentemente resulta posible iniciar el procedimiento de la invención con una recombinasa específica para una secuencia diana modificada específica y con otra recombinasa (o una colección) para otra secuencia diana modificada específica.

Los procedimientos de evolución molecular dirigida, también denominados de evolución de laboratorio o evolución in vitro, son conocidos de la técnica (para una revisión, ver Yuan et al., 2005, y referencias en el mismo; Johannes y Zhao, 2006).

En una primera etapa de evolución molecular dirigida, se generan mediante procedimientos conocidos de la técnica colecciones de secuencias de recombinasa mutadas aleatoriamente, por ejemplo mediante la utilización de PCR con tendencia a errores y reorganización de ADN (revisión en, por ejemplo, Yuan et al., 2005) o los procedimientos dados a conocer en la solicitud de patente internacional nº WO 2002/44409. Los plásmidos de cada colección que comprende la recombinasa mutada contienen además una de las secuencias diana del subconjunto final obtenido en la etapa (f). Tras la transfección de la colección de plásmidos generada en células apropiadas, se activa la expresión de la recombinasa y se lleva a cabo la evolución molecular dirigida tal como es conocido por el experto en la materia.

En una forma de realización preferida, la evolución molecular dirigida que se utiliza en la etapa (f) del procedimiento de la presente invención es evolución de proteínas ligada a sustrato (EPLIS, Buchholz y Stewart, 2001; solicitud de patente internacional nº WO 02/44409). La evolución de proteínas ligada a sustrato puede llevarse a cabo tal como se indica en detalle en los ejemplos del documento nº WO 2008/083931. Brevemente, las secuencias diana obtenidas en la etapa (e) se clonan en un plásmido (el denominado vector de evolución) conjuntamente con una secuencia codificante de la recombinasa mutada aleatoriamente. La mutación aleatoria se lleva a cabo mediante PCR con tendencia a errores (ver Buchholz y Stewart, 2001). A continuación, la colección de plásmidos generada se transfecta en células de *E. coli* para que resulte posible la expresión de la recombinasa. Mediante la utilización de un promotor inducible para controlar la expresión de la recombinasa, resulta posible ajustar los niveles de expresión. Tras la incubación durante la noche, se aísla el ADN plasmídico a partir de las células y se digiere con NdeI para

5 cortar los plásmidos que no se habían recombinado y únicamente los plásmidos recombinados se amplificaron posteriormente con cebadores. El producto de PCR de la forma recombinada del plásmido produjo una banda de 1,7 Kb. Se digirió el producto de PCR con BsrGI y XbaI y se subclonó nuevamente en un vector de evolución digerido de manera similar, para el siguiente ciclo de evolución.

10 En la etapa (g), las colecciones de recombinasas evolucionadas en la etapa (f) se agrupan y mezclan. La tecnología de mezcla de ADN es conocida de la técnica (para una revisión ver Minshull y Stemmer, 1999; Stemmer, 1994). Las colecciones de recombinasas evolucionadas en secuencias diana modificadas basadas en el subsitio 1 se agrupan y se mezclan y, separadamente, las colecciones de recombinasas evolucionadas en secuencias diana modificadas basadas en el subsitio 2 se agrupan y se mezclan.

15 A continuación, las colecciones agrupadas y mezcladas se clonan en una nueva generación de vectores que comprenden las secuencias diana del subconjunto superior siguiente, es decir, en el caso de que se generen dos subconjuntos, el subconjunto generado en la etapa (d). Por ejemplo, el vector para la colección evolucionada en las secuencias basadas en el subsitio 1 comprende la secuencia del subsitio 1 como secuencia diana y el vector para la colección evolucionada en las secuencias basada en el subsitio 2 comprende la secuencia del subsitio 2 como secuencia diana.

20 En la etapa (h), el procedimiento de evolución de proteínas ligada a sustrato se aplica en las colecciones mezcladas obtenidas en la etapa (g) utilizando la secuencia diana en el subconjunto superior siguiente, que, tal como se ha comentado, puede ser el subconjunto según la etapa (d).

25 Dicha etapa rinde recombinasas que reconocen y se recombinan con secuencias diana que incluyen la combinación (y de esta manera un número creciente) de mutaciones de las diferentes secuencias diana del subconjunto inferior. La combinación de mutaciones de diferentes colecciones de un subconjunto inferior de secuencias diana resulta en efectos sinérgicos y conduce a la generación de recombinasas, las cuales ahora recombinan secuencias diana de un subconjunto superior, demostrando que puede utilizarse una estrategia de evolución a través de intermediarios para alcanzar la actividad deseada.

30 En las etapa (i), (g), es decir la combinación y mezcla de colecciones de recombinasa, y (j), es decir la aplicación de evolución molecular dirigida a las colecciones agrupadas y mezcladas, se repiten hasta conseguir por lo menos una recombinasa que sea activa sobre la secuencia diana asimétrica presente en la LTR del ADN provírico.

35 Por ejemplo, en un procedimiento en el que resultó necesaria la generación de dos subconjuntos de secuencias diana para generar secuencias diana con únicamente una, dos o tres divergencias de nucleótido, las colecciones de recombinasas evolucionadas, por ejemplo, para el segundo subconjunto de secuencias diana, se agruparon y se mezclaron y se aplicó la evolución molecular dirigida a dicha colección mezclada utilizando las secuencias diana del primer subconjunto. En la etapa siguiente (y final), la secuencia diana asimétrica de la etapa (a) en la LTR del ADN provírico se utiliza para evolucionar la colección de recombinasas que comprende recombinasas que reconocen las secuencias diana del primer subconjunto mediante evolución molecular dirigida con el fin de obtener por lo menos una recombinasa que es activa sobre la secuencia diana asimétrica en la LTR del ADN provírico. En esta etapa, el procedimiento de evolución molecular dirigida preferentemente es el procedimiento de evolución de proteínas ligada a sustrato.

45 En la etapa (k), se aísla a partir de la colección el ácido nucleico de una recombinasa que presenta actividad sobre la secuencia diana asimétrica de la etapa (a) en la LTR del ADN retrovírico. El ácido nucleico se aísla del plásmido respectivo en la colección utilizando enzimas de restricción apropiados. El procedimiento de digestión con endonucleasa de restricción es conocido por el experto en la materia. El ácido nucleico codificante de la recombinasa seguidamente puede recuperarse mediante procedimientos conocidos, tales como, por ejemplo, la electroforesis en gel.

50 El ácido nucleico puede almacenarse (preferentemente a temperaturas inferiores a -80°C) o puede clonarse opcionalmente en la etapa (l) en un vector de expresión para la utilización en el análisis posterior en procedimientos de expresión de proteínas, o para la administración en un sujeto para el tratamiento y/o la prevención de la infección por retrovirus, en particular, de la infección por VIH y/o SIDA. Se conocen del estado de la técnica vectores de expresión adecuados o se dan a conocer posteriormente.

60 Preferentemente la secuencia diana asimétrica identificada en la etapa (a) se localiza tanto en la 5'-LTR como en la 3'-LTR del provirus para permitir la extracción del ADN provírico del genoma de la célula hospedadora.

65 Utilizando la evolución dirigida ligada a sustrato y las secuencias identificadas en la SEC ID nº 1 como sustrato, los presentes inventores producen una recombinasa elaborada a medida que se recombina con dicha secuencia diana de ADN asimétrica presente en las repeticiones terminales largas del VIH-1 de una pluralidad de cepas del VIH-1. El desarrollo de dichas recombinasas elaboradas a medida con diana específicamente en secuencias asimétricas en una pluralidad de LTR del VIH-1 permite la extracción del provirus respectivo de su integración cromosómica para la mayoría de sujetos infectados por el VIH-1. De esta manera, en una forma de realización, la recombinasa elaborada

a medida se deriva de una colección de recombinasas que comprende la recombinasa Cre, por ejemplo una colección tal como se indica en la figura 3, y con diana en una pluralidad de cepas de VIH-1, reconociendo las secuencias diana que presentan la secuencia indicada como SEC ID nº 1. La evolución molecular dirigida es evolución de proteínas ligada a sustrato (EPLiS). El vector de expresión codificante de la recombinasa elaborada a medida preferentemente se deriva de un vector lentivírico. Este vector de expresión, las células transfectadas con el mismo y/o la proteína recombinasa derivada de las mismas, presentan usos médicos, por ejemplo en el tratamiento y/o la prevención de una infección por VIH-1.

La invención proporciona un procedimiento para preparar un vector de expresión codificante de una recombinasa elaborada a medida, la cual es capaz de recombinar secuencias diana asimétricas en la LTR del ADN provírico de una pluralidad de cepas de VIH-1 que pueden insertarse en el genoma de una célula hospedadora, comprendiendo las etapas de:

- (a) identificar la SEC ID nº 1 que presenta homología respecto a la secuencia de hemisio izquierdo y la secuencia de hemisio derecho de los sitios diana de recombinasa de una colección de recombinasas que comprende las recombinasas especificadas en la figura 1,
- (b) identificar dos secuencias, en las que la primera secuencia corresponde a la secuencia de la secuencia diana asimétrica de la etapa (a) homóloga respecto al hemisio izquierdo de dicho sitio diana conocido y que se denomina "secuencia 1 de hemisio" y en el que la segunda secuencia corresponde a la secuencia de la secuencia diana asimétrica de la etapa (a) homóloga respecto al hemisio derecho y que se denomina "secuencia 2 de hemisio",
- (c) determinar los nucleótidos en las secuencias de la etapa (b) que divergen respecto de las secuencias homólogas correspondientes del hemisio izquierdo y del hemisio derecho del sitio diana homólogo conocido de la etapa (a), tal como se muestra en la figura 1,
- (d) generar un primer subconjunto de dos ácidos nucleicos diana que comprenden secuencias diana, en el que la primera secuencia diana se denomina subsitio 1 y comprende la SEC ID nº 8 y en el que la segunda secuencia diana se denomina el subsitio 2 y comprende la SEC ID nº 12,
- (e) generar un segundo subconjunto de ácidos nucleicos diana que comprenden secuencias diana modificadas basándose en las secuencias diana en el primer subconjunto de la etapa (d), respectivamente,
- (f) aplicar separadamente la evolución molecular dirigida a dicha colección de recombinasas utilizando cada ácido nucleico del segundo subconjunto obtenido en la etapa (e) como sustrato,
- (g) mezclar las colecciones de recombinasas evolucionadas en la etapa (f), en donde todas las colecciones de recombinasa evolucionadas en secuencias basadas en el subsitio 1 se agrupan y se mezclan, y en el que todas las colecciones de recombinasas evolucionadas en secuencias basadas en el subsitio 2 se agrupan y se mezclan,
- (h) aplicar la evolución molecular dirigida a las colecciones mezcladas obtenidas en la etapa (g) utilizando cada ácido nucleico del subconjunto según la etapa (d) como sustrato,
- (i) mezclar las colecciones de recombinasa evolucionadas en la etapa (h),
- (j) aplicar la evolución molecular dirigida a la colección mezclada que se ha obtenido en la etapa (g) utilizando un ácido nucleico que comprende la secuencia diana asimétrica de la etapa (a) como sustrato, hasta obtener por lo menos una recombinasa que es activa sobre la secuencia diana asimétrica en la LTR del ADN retrovírico de la etapa (a),
- (k) aislar el ácido nucleico de la recombinasa o recombinasas obtenidas en la etapa (j) de la colección, y
- (l) clonar el ácido nucleico obtenido en la etapa (k) en un vector de expresión adecuado,

en el que el procedimiento de evolución molecular dirigida es evolución de proteínas ligada a sustrato.

La recombinasa TRE adaptada para reconocer la SEC ID nº 1 se denomina TRE 3.0.

Sin embargo, resulta evidente para el experto en la materia que pueden generarse otras recombinasas específicas de sitio adaptadas que recombinan sitios diana divergentes presentes en el genoma de una pluralidad de provirus retrovíricos insertadas en el genoma de la célula hospedadora. Pueden determinarse otras secuencias diana candidatas que se encuentran presentes en el genoma de una pluralidad de cepas retrovíricas.

El ADN provírico que puede insertarse en el genoma de una célula hospedadora o que todavía no puede insertarse,

es el ADN de un retrovirus, a saber, VIH-1. Los retrovirus comprenden una familia grande y diversa de virus ARN con cubierta. La característica distintiva de la familia es su estrategia de replicación, que incluye como etapas esenciales la transcripción inversa del ARN vírico en ADN de doble cadena lineal y la posterior integración de este ADN (ADN provírico) en el genoma de la célula hospedadora. Los retrovirus se subdividen en siete grupos, definidos por su grado de relación evolutiva. Cinco de estos grupos (retrovirus alfa, beta, delta, epsilon y gamma) representan retrovirus con potencial oncogénico y los otros dos grupos son los lentivirus y los espumavirus. Los virus de la leucemia de células T humanas patogénicas humanas de tipo I y de tipo II (VLTH-I y VLTH-II) pertenecen al grupo de los retrovirus delta, mientras que los virus del SIDA, de la inmunodeficiencia humana de tipo 1 y de tipo 2 (VIH-1 y VIH-2) pertenecen al grupo de los lentivirus (para una revisión ver el libro de texto estándar "Retroviruses", de Coffin J.M., Hughes S.H., Varmus H.E. (editores), 1997, Cold Spring Harbor Laboratory Press, New York).

En una forma de realización más preferida, la secuencia diana asimétrica identificada en la etapa (a) del procedimiento de la presente invención se localiza tanto en la 5'-LTR como en la 3'-LTR de un provirus VIH. Preferentemente, dicha secuencia diana asimétrica localizada tanto en la 5'-LTR como en la 3'-LTR de un provirus de VIH-1 presenta la secuencia indicada como SEC ID nº 1.

El procedimiento de evolución molecular dirigida aplicada en el procedimiento de la presente invención es el procedimiento de evolución de proteínas ligada a sustrato (EPLiS; Buchholz y Stewart, 2001; ver también el documento WO 02/44409).

Llevando a cabo el procedimiento de la invención tal como se indica en la presente memoria, los inventores generaron varios ácidos nucleicos codificantes de una recombinasa elaborada a medida, y las recombinasas elaboradas a medida mismas. De esta manera, la invención proporciona una recombinasa elaborada a medida que comprende una secuencia según cualquiera de las SEC ID nº 40 a nº 64 o que consiste en la misma, o ácidos nucleicos codificantes de la misma.

Inesperadamente se ha descubierto que las recombinasas elaboradas a medida preferentemente comprenden determinadas secuencias de consenso, tal como se muestra en la figura 5. En particular, todas las recombinasas elaboradas a medida que se han analizado capaces de recombinarse con una secuencia diana asimétrica comprende una secuencia según SEC ID nº 37 (secuencia de consenso "Tre común"). Se ha encontrado además que todas las recombinasas elaboradas a medida que se han analizado capaces de recombinarse con la secuencia diana según la SEC ID nº 1 comprende una secuencia según la SEC ID nº 38. El 38,95% de todas las recombinasas elaboradas a medida que se han analizado capaces de recombinar la secuencia diana según la SEC ID nº 1 comprende una secuencia según SEC ID nº 39. En la SEC ID nº 37, nº 38 y nº 39, los aminoácidos variables están representados por una X (ver la figura 5).

De esta manera, la invención proporciona una recombinasa elaborada a medida capaz de recombinar secuencias diana asimétricas en la LTR del ADN provírico de una cepa de VIH-1 que puede insertarse en el genoma de una célula hospedadora (es decir, una recombinasa elaborada a medida funcional) según las reivindicaciones 7 a 11, que preferentemente comprende la SEC ID nº 37. Dicha recombinasa elaborada a medida puede obtenerse, por ejemplo, siguiendo el procedimiento de la invención. Dicha recombinasa elaborada a medida presenta una identidad de secuencia de 99% respecto a la SEC ID nº 38 y comprende los intercambios de aminoácidos definidos siguientes respecto a la secuencia de Cre (SEC ID nº 36): P12S, P15L, M44V, K86N, G93A, A175S y P307A. Estos intercambios provocan que el enzima resulte particularmente adecuado para la recombinación en una secuencia diana de SEC ID nº 1 o una secuencia diana que presenta una identidad de secuencia elevada respecto a la SEC ID nº 1 (por ejemplo una identidad de secuencia respecto a la SEC ID nº 1 de por lo menos 80%, de por lo menos 90% o de por lo menos 95%).

En una forma de realización, la secuencia de la recombinasa elaborada a medida de la invención, en particular la recombinasa elaborada a medida que comprende la SEC ID nº 37 no se da a conocer en el documento WO 2008/083931. Preferentemente, la secuencia no comprende la SEC ID nº 65 de la presente solicitud (que es idéntica a la SEC ID nº 3 del documento WO 2008/083931) o las posiciones 11 a 351 de la SEC ID nº 65, y la secuencia del ácido nucleico codificante de una recombinasa elaborada a medida de la invención no codifica una proteína que comprende las posiciones 11 a 351 de la SEC ID nº 65. La secuencia de la recombinasa elaborada a medida de la invención que comprende la SEC ID nº 37 varía también respecto a las recombinasas naturales, tales como Cre, Dre, Fre o Zre, lo que resulta evidente a partir de la característica de que es capaz de recombinar secuencias diana asimétricas en la LTR del ADN provírico de una cepa retrovítica que puede insertarse en el genoma de una célula hospedadora y, preferentemente, de que es capaz de recombinar secuencias diana asimétricas en la LTR del ADN provírico de una pluralidad de cepas retrovíticas que puede insertarse en el genoma de una célula hospedadora.

La recombinasa elaborada a medida preferentemente comprende la SEC ID nº 38 o la SEC ID nº 39.

Las recombinasas elaboradas a medida funcionales capaces de recombinar secuencias diana asimétricas en la LTR del ADN provírico de una pluralidad de cepas retrovíticas que puede insertarse en el genoma de una célula hospedadora, las cuales pueden obtenerse, por ejemplo, mediante el procedimiento de la invención, pueden variar respecto a dichas secuencias, aunque las secuencias proporcionan una guía valiosa para que el experto en la

materia produzca una recombinasa elaborada a medida capaz de recombinar sitios diana asimétricos, tales como la SEC ID nº 1, incluso sin llevar a cabo el procedimiento de la invención. Un determinado porcentaje de identidad de secuencia respecto a la secuencia SEC ID nº 37, nº 38 o nº 39 en el contexto de la presente invención se refiere a que en la identidad de secuencia respecto a las posiciones definidas en la secuencia respectiva, por ejemplo con una identidad de secuencia de 97%, puede variar el 3% de las posiciones aminoácidas definidas de dichas secuencias. Preferentemente, los intercambios de aminoácidos con respecto a las secuencias de referencia son sustituciones conservadoras, las cuales son bien conocidas por el experto en la materia (por ejemplo Creighton, Proteins, W.H. Freeman and Company (ed.), 1984). Por ejemplo, las sustituciones conservadoras sustituyen un aminoácido del grupo de aminoácidos cargados negativamente por otro.

Las recombinasas elaboradas a medida capaces de recombinar secuencias diana asimétricas en la LTR del ADN provírico de una pluralidad de cepas retrovíricas que pueden insertarse en el genoma de una célula hospedadora también pueden comprender una combinación de 2, 3, 4 o más secuencias seleccionadas de entre el grupo que consiste en las SEC ID nº 40 a nº 64, por ejemplo una parte C-terminal de cualquiera de dichas secuencias, por ejemplo la SEC ID nº 40, y la parte N-terminal de cualquier otra de dichas secuencias, por ejemplo la SEC ID nº 64. La parte C-terminal puede presentar una longitud de entre 1 y 342 aminoácidos. En una combinación de dos secuencias, la parte N-terminal puede presentar una longitud de entre 1 y 342 aminoácidos. La combinación en cualquier caso comprende un motivo de consenso de TRE, por ejemplo la SEC ID nº 37, o preferentemente, la SEC ID nº 38 o la SEC ID nº 39.

La invención proporciona además una composición que comprende dos o más recombinasas elaboradas a medida que comprenden diferentes secuencias según la invención o ácidos nucleicos codificantes de las mismas. En una forma de realización, la composición comprende dos o más, tres o más, cuatro o más, cinco o más, diez o más, 20 o más o 25 recombinasas que comprenden secuencias según cualquiera de las SEC ID nº 40 a nº 64 o ácidos nucleicos codificantes de las mismas. Dichas composiciones pueden ser composiciones farmacéuticas tal como se indica posteriormente.

La invención proporciona además un ácido nucleico codificante de una recombinasa elaborada a medida capaz de recombinar secuencias diana asimétricas en la LTR del ADN provírico de una pluralidad de cepas de VIH-1 que puede insertarse en el genoma de una célula hospedadora, comprendiendo la recombinasa elaborada a medida una secuencia de aminoácidos tal como se ha definido anteriormente. En formas de realización preferente, dicha secuencia de aminoácidos comprende la SEC ID nº 38 o la SEC ID nº 39. La invención proporciona además un ácido nucleico codificante de una recombinasa elaborada a medida que comprende una secuencia de aminoácidos seleccionada de entre el grupo que consiste en las SEC ID nº 40 a nº 64 o una combinación de dichas secuencias.

En el contexto de la invención, un ácido nucleico o una proteína que comprende una secuencia también puede consistir de dicha secuencia.

Puede comprender además secuencias adicionales, por ejemplo una secuencia de señal que proporciona la expresión/localización en un compartimiento celular específico, tal como una señal de localización nuclear (por ejemplo la señal en las posiciones 2 a 9 de la SEC ID nº 65). En el caso de que deba utilizarse una proteína en una composición farmacéutica, resulta especialmente preferente expresarla en forma de proteína de fusión con un dominio de transducción de proteínas, tal como el dominio de transducción de la proteína tat, que permite la transducción de proteínas de células diana. Preferentemente, una recombinasa elaborada a medida de la invención destinada a la utilización en una composición farmacéutica se prepara en forma de proteína de fusión con una secuencia de localización nuclear y con un dominio de transducción de proteínas, por ejemplo a partir de tat, y un ácido nucleico codificante de una recombinasa elaborada a medida de la invención puede codificar dicha proteína de fusión. Por ejemplo, los dominios de transducción de proteínas siguientes pueden utilizarse en una proteína de fusión con una recombinasa elaborada a medida de la invención, que preferentemente incluye además una señal de localización nuclear:

- Dominio básico del transactivador Tat del VIH-1 (Fawell S., Seery J., Daikh Y., Moore C., Chen L.L., Pepinsky B., Barsoum J., Tat-mediated delivery of heterologous proteins into cells, Proc. Natl. Acad. Sci. U.S.A. 91(2):664-8, 18 de junio de 1994).
- Homeodominio de *Drosophila antennapedia* (Antp) (Derossi D., Joliot A.H., Chassaing G., Prochiantz A., The third helix of the Antennapedia homeodomain translocates through biological membranes, J. Biol. Chem. 269(14):10444-50, 8 de abril de 1994).
- Factor de transcripción VP22 del VHS (Elliott G., O'Hare P., Intercellular trafficking and protein delivery by a herpesvirus structural protein, Cell 88(2):223-33, 24 de enero de 1997).
- Motivo de translocación permeable celular (MTL) del antígeno de superficie PreS2 del virus de la hepatitis B (VHB) (Oess S., Hildt E., Novel cell permeable motif derived from the PreS2-domain of hepatitis-B virus surface antigens, Gene Ther. 7(9):750-8, mayo de 2000).

En el caso de que deba purificarse la proteína, también puede añadirse una etiqueta facilitadora de la purificación de proteína, tal como una etiqueta de His.

El uso de los codones del ácido nucleico de la invención codificante de una recombinasa Tre tal como se ha definido anteriormente puede ser seleccionado por el experto en la materia. Por ejemplo, puede seleccionarse un uso de codones adecuados para la expresión en una célula humana, en particular en el caso de que se pretenda la expresión en una célula humana, por ejemplo con fines terapéuticos. El uso de los codones también puede basarse en el uso de codones de, por ejemplo, la recombinasa Cre.

La recombinasa elaborada a medida o ácido nucleico codificante de dicha recombinasa elaborada a medida puede obtenerse mediante el procedimiento de la invención tal como se indica en la presente memoria, o puede ser obtenible mediante dicho procedimiento. También puede obtenerse basándose en la secuencias dadas a conocer en la presente memoria, opcionalmente mediante la combinación y/o la variación adicional de dichas secuencias, opcionalmente sometiendo a ensayo la actividad en recombinación de los sitios diana asimétricos, SEC ID nº 1.

La invención proporciona además una composición, por ejemplo una colección, que comprende dos o más de los ácidos nucleicos codificantes de una recombinasa elaborada a medida tal como se ha definido anteriormente, por ejemplo codificante de dos o más recombinasas elaboradas a medida que comprenden diferentes secuencias según SEC ID nº 38 o SEC ID nº 39. En una forma de realización, la composición comprende ácidos nucleicos codificantes de recombinasas elaboradas a medida que comprenden dos o más, tres o más, cuatro o más, cinco o más, diez o más, 20 o más o 25 o más recombinasas que comprenden secuencias según cualquiera de las SEC ID nº 40 a nº 64 o combinaciones de dichas secuencias. Dichas composiciones, en particular composiciones en las que el ácido nucleico es un vector de expresión, pueden resultar particularmente adecuadas como composiciones farmacéuticas tal como se indica posteriormente.

En el procedimiento de la presente invención, el ácido nucleico codificante de por lo menos una recombinasa elaborada a medida que es activa en la secuencia diana asimétrica en la LTR del ADN retroviral se clona en un vector de expresión. Los vectores de expresión son constructos genéticos para expresar las proteínas codificadas por los ácidos nucleicos en el vector. Dichos vectores de expresión pueden ser vectores extracromosómicos autorreplicantes o vectores que se integran en un genoma del huésped. Generalmente, entre estos vectores de expresión se incluyen ácidos nucleicos reguladores transcripcionales y traduccionales ligados operablemente al ácido nucleico codificante de la recombinasa elaborada a medida de la presente invención.

La expresión "secuencias de control" se refiere a secuencias de ADN necesarias para la expresión de una secuencia codificante operablemente ligada en un organismo huésped particular. Entre las secuencias de control que resultan adecuadas para los procariontes se incluyen, por ejemplo, un promotor, opcionalmente una secuencia de operador, y un sitio de unión ribosómica. Es conocido que las células eucarióticas utilizan promotores, señales de poliadenilación e intensificadores.

Un ácido nucleico se encuentra "ligado funcionalmente" en el caso de que se sitúe en una relación funcional con otra secuencia de ácidos nucleicos. Por ejemplo, un promotor o intensificador se liga de manera funcional a una secuencia codificante en el caso de que afecte a la transcripción de la secuencia, o un sitio de unión ribosómica se encuentra ligada operablemente a una secuencia codificante en el caso de que se sitúe de manera que facilite la traducción. El ligamiento se consigue mediante ligación en sitios de restricción convenientes. En el caso de que no existan sitios de este tipo, se utilizan adaptadores o conectores oligonucleótidos sintéticos de acuerdo con la práctica convencional. El ácido nucleico regulador de la transcripción y de la traducción generalmente resulta apropiado a la célula hospedadora utilizada para expresar la recombinasa elaborada a medida. Se conocen de la técnica numerosos tipos de vectores de expresión apropiados y secuencias reguladoras adecuadas para una diversidad de células hospedadoras.

El vector de expresión utilizado en la presente invención puede ser un vector retroviral, un vector lentiviral, un vector espumavirus o un vector adenoviral. Sin embargo, en una forma de realización preferente, el vector de expresión es un vector lentiviral seleccionado de entre el grupo que consiste en vectores lentivirales derivados de VIH-1, VIS, VIF o VAIE. Los vectores lentivirales se encuentran descritos en, por ejemplo, Schambach et al. (2006) o en la solicitud de patente europea nº 11000751.5.

En formas de realización preferidas de la presente invención, el vector de expresión comprende un vector celular, bacteriano, vírico o híbrido.

En general, para el propósito de la presente invención, el promotor puede ser un promotor constitutivo o inducible. Además, los promotores pueden ser promotores naturales, tales como un promotor bacteriano, celular o vírico, o un promotor híbrido. Los promotores híbridos, los cuales combinan elementos de más de un promotor, son conocidos de la técnica y resultan útiles en la presente invención. Además, el promotor utilizado en la presente invención también puede ser un derivado de un promotor natural. Un "derivado" de un promotor natural tal como se utiliza en la presente memoria puede ser una combinación de elementos activos en cis obtenidos a partir de promotores o secuencias de origen diferente o, alternativamente, puede obtenerse mediante delección o mutación de elementos

activos en cis dentro de un promotor natural específico (Edelman et al., 2000; Alper et al., 2006; Hartenbach y Fussenegger, 2006).

En una forma de realización de la presente invención, el promotor constitutivo o derivado del mismo se selecciona o se deriva del grupo que consiste en promotores de citomegalovirus, virus del sarcoma de Rous, retrovirus relacionados con el virus de la leucemia murina, gen fosfoglicerocinasa, virus formador de foco en el bazo murino o factor-1 alfa de elongación humano.

En una forma de realización adicional de la presente invención, el promotor inducible o derivado del mismo se selecciona o se deriva del grupo que consiste en las LTR o derivados de las mismas derivados de lentivirus, espumavirus y delta-retrovirus.

En el contexto de la presente memoria, el término "LTR" se refiere a las repeticiones terminales largas tanto 5' como 3' de provirus que presentan función de promotor (para una revisión ver el libro de texto estándar "Retrovirus" (Coffin J.M., Hughes S.H., Varmus H.E. (editores), 1997, Cold Spring Harbor Laboratory Press, New York)).

Preferentemente, el promotor inducible o derivado del mismo se selecciona o se deriva de la LTR o derivados de la misma derivados de VIH-1, VIH-2, VMV, VAIE, VAEC, VIS, VIF, VIB, VLTH-I y VLTH-II.

La presente invención proporciona además un procedimiento para preparar una recombinasa elaborada a medida, en el que dicho procedimiento comprende el procedimiento anteriormente indicado para preparar un vector de expresión codificante de una recombinasa elaborada a medida, y la etapa adicional de expresar la recombinasa elaborada a medida o un polipéptido de fusión que comprende la secuencia de aminoácidos de dicha recombinasa elaborada a medida a partir del ácido nucleico codificante de la recombinasa insertada en el vector de expresión obtenido en el procedimiento anteriormente indicado para preparar un vector de expresión codificante de una recombinasa elaborada a medida en una célula hospedadora adecuada.

Preferentemente, las recombinasas finalmente obtenidas se someten a ensayo en células de mamífero para garantizar que funcionan en un medio de células de mamífero. Además, con el fin de obtener una buena expresión en células de mamífero, las recombinasas pueden optimizarse para la expresión en estas células (por ejemplo la optimización del uso de los codones utilizando procedimientos bien conocidos de la técnica. Ver, por ejemplo, Shimshek et al., 2002) o pueden añadirse secuencias de señal necesarias para dirigir la proteína hasta el núcleo de la célula de mamífero, tal como la secuencia NLS (Macara, 2001), al ácido nucleico de la recombinasa elaborada a medida.

La expresión del ácido nucleico codificante de la recombinasa elaborada a medida que se ha clonado en un vector de expresión, por ejemplo según la etapa (1) del procedimiento de preparación de un vector de expresión codificante de una recombinasa elaborada a medida, puede llevarse a cabo utilizando, por ejemplo, sistemas de expresión bacterianos, de insecto o de mamífero. Sin embargo, también pueden utilizarse otros sistemas de expresión conocidos en la técnica. Los procedimientos de introducción de ácidos nucleicos exógenos en huéspedes mamíferos, insectos o bacterias, así como otros huéspedes, también son bien conocidos de la técnica, y variarán con la célula hospedadora utilizada. Entre las técnicas se incluyen la transfección mediada por dextrano, la precipitación con fosfato de calcio, la transfección mediada por polibreno, la fusión de protoplasto, la electroporación, la infección vírica, el encapsulado del polinucleótido o polinucleótidos en liposomas y la microinyección directa del ADN en los núcleos.

Las proteínas de fusión se preparan mediante procedimientos bien conocidos de la técnica. Por ejemplo, el vector de expresión en el que se ha clonado el ácido nucleico codificante de la recombinasa elaborada a medida ya comprende una secuencia de ácidos nucleicos codificante de un segundo polipéptido o proteína. Mediante la clonación del ácido nucleico codificante de la recombinasa elaborada a medida en el mismo marco que la secuencia del segundo polipéptido o proteína, se expresarán ambas secuencias en forma de proteína de fusión.

Las células hospedadoras utilizadas para expresar la recombinasa elaborada a medida a partir del vector de expresión preferentemente son células hospedadoras que incluyen células procariotas, tales como, por ejemplo, células bacterianas o células de levadura, o células eucariotas, tales como, por ejemplo, células de insecto o células de mamífero. La célula hospedadora puede ser una célula hematopoyética, por ejemplo una célula troncal hematopoyética adulta o una célula T, por ejemplo una célula CD4+. La célula puede derivarse de un sujeto infectado por el retrovirus, y la célula puede administrarse nuevamente en el sujeto tras la transformación y, opcionalmente, el cultivo y/o la propagación.

La presente invención proporciona además un procedimiento para preparar una célula troncal adulta transformada, en el que dicho procedimiento comprende el procedimiento anteriormente indicado para preparar un vector de expresión codificante de una recombinasa elaborada a medida y la etapa adicional de introducir el vector de expresión obtenido en el procedimiento anteriormente indicado para preparar un vector de expresión codificante de una recombinasa elaborada a medida in vitro en una célula troncal adulta adecuada.

En un aspecto adicional, la presente invención se refiere al ácido nucleico tal como puede obtenerse con el procedimiento anteriormente indicado de la presente invención. Los ácidos nucleicos codificantes de una recombinasa elaborada a medida definida por una secuencia también se proporcionan en la presente memoria.

Un "ácido nucleico" tal como se utiliza en la presente memoria es un compuesto polimérico que comprende subunidades unidas covalentemente denominadas nucleótidos. El ácido nucleico incluye ácido polirribonucleico (ARN) y ácido polidesoxirribonucleico (ADN), los cuales pueden ser de cadena sencilla o de doble cadena. El ADN incluye el ADNc, el ADN genómico, el ADN sintético y el ADN semisintético.

El término "proteína" tal como se utiliza en la presente memoria incluye proteínas, polipéptidos y péptidos. Tal como apreciará el experto en la materia, las secuencias de ácidos nucleicos de la invención pueden utilizarse para generar secuencias de proteína. Un aspecto adicional de la invención es la proteína recombinasa elaborada a medida tal como puede obtenerse mediante, por ejemplo, el procedimiento anteriormente indicado de la presente invención, la cual opcionalmente puede ser una proteína de fusión que comprende una recombinasa funcional. En una forma de realización, la proteína recombinasa elaborada a medida puede prepararse en forma de polipéptido de fusión, utilizando técnicas bien conocidas. En una forma de realización preferente, la proteína recombinasa elaborada a medida se une a un segundo polipéptido. Preferentemente el polipéptido de fusión se obtiene con el procedimiento anteriormente indicado de la presente invención, en el que la recombinasa elaborada a medida se une a un segundo polipéptido.

En una forma de realización, la proteína recombinasa elaborada a medida se prepara en forma de un polipéptido de fusión para incrementar la expresión. En una forma de realización adicional, la proteína recombinasa elaborada a medida se prepara en forma de un polipéptido de fusión para permitir la introducción del polipéptido en células vivas. Típicamente, las proteínas purificadas no pueden entrar en las células debido a que no pueden pasar a través de la membrana celular debido a su tamaño. Sin embargo, la fusión de secuencias peptídicas específicas con proteínas puede resultar en la incorporación de estas proteínas de fusión en las células. En la célula, la proteína seguidamente puede llevar a cabo su función. Se han administrado con éxito recombinasas específicas de sitio, incluyendo la recombinasa Cre, en células utilizando este procedimiento (Peitz et al., 2002). La recombinasa permeable a la célula ha sido descrita adicionalmente por Nolden et al. (2006) y Lin et al. (2004). Por lo tanto, esta estrategia puede utilizarse para administrar las recombinasas elaboradas a medida en el interior de las células para eliminar los provirus de las células infectadas. De esta manera, el segundo polipéptido en el polipéptido de fusión puede comprender un péptido de señal. El péptido de señal puede ser un dominio de transducción de proteínas, tal como el péptido TAT, o un péptido de la tercera hélice del homeodominio de Antennapedia (Derossi et al., 1994, 1996; Vives et al., 1997; Vives, 2003; Richard et al., 2005) o la SLN (secuencia de localización nuclear) para administrar el polipéptido de fusión en el núcleo de una célula eucariótica (Macara, 2001).

Un aspecto adicional de la presente invención se refiere a la célula troncal adulta tal como puede obtenerse con el procedimiento anteriormente indicado para preparar una célula troncal adulta transformada de la presente invención. Las células madre preferentemente se infectan o se transfectan con el vector de expresión según la invención. En una forma de realización preferida, la célula troncal adulta es una célula troncal del linaje hematopoyético que expresa la recombinasa elaborada a medida, el polipéptido de fusión anteriormente indicado o que comprende el vector de expresión anteriormente indicado. Las células madre hematopoyéticas (CMH) son células CD34+ derivadas de la médula ósea que pueden, por ejemplo, purificarse a partir de sangre periférica movilizada con G-CSF de donantes (por ejemplo pacientes infectados por el VIH) mediante leucaféresis rutinaria (Scherr y Eder, 2002). Las células genéticamente modificadas *in vitro* a continuación pueden formularse para la reinfusión en los pacientes.

En el estado de la técnica, la expresión "células madre" se refiere a células que (a) presentan la capacidad de autorrenovación y (b) la capacidad de formar por lo menos un y con frecuencia varios tipos de célula especializada debido a su capacidad de división asimétrica (Donovan y Gearhart, 2001). Las células madre adultas pueden aislarse a partir de tejidos diferentes del adulto, es decir, de individuos diferenciados. Dichas células madre se denominan en el estado de la técnica, "células madre adultas multipotentes". La diferencia esencial entre las células madre pluripotentes embrionarias y las células madre multipotentes adultas radica en el número de tejidos diferenciados que pueden obtenerse a partir de cada tipo de célula.

En una forma de realización adicional, el vector de expresión de la presente invención se utiliza para transformar células T, por ejemplo células primarias CD4+ (células sanguíneas) de pacientes infectados por el VIH-1.

En una etapa adicional del procedimiento de la presente invención, el vector de expresión que comprende la secuencia de ácidos nucleicos codificante de una recombinasa elaborada a medida de la invención, la proteína recombinasa, la proteína de fusión o la célula troncal adulta obtenida mediante los procedimientos de la presente invención se formulan como composición farmacéutica para la utilización en la prevención y/o el tratamiento de una infección por VIH-1 y/o para la reducción de la carga vírica en un sujeto infectado por VIH-1.

Un objeto adicional de la invención es la composición farmacéutica obtenida mediante el procedimiento anteriormente indicado. La composición farmacéutica se encuentra presente preferentemente en forma de una

solución adecuada para la aplicación intravenosa (infusión).

La preparación farmacéutica puede comprender además uno o más portadores, excipientes y/o adyuvantes farmacéuticamente aceptables. Los portadores, excipientes y adyuvantes adecuados para la utilización en la composición farmacéutica son conocidos en la técnica.

La composición farmacéutica de la presente invención preferentemente reduce la carga vírica en un sujeto infectado por un retrovirus a menos de 5.000 equivalentes genómicos/ml de plasma, preferentemente a menos de 500 equivalentes genómicos/ml de más y más preferentemente a menos de 50 equivalentes genómicos/ml de plasma al administrarla en el sujeto. De esta manera, la composición farmacéutica de la presente invención que comprende un vector de expresión codificante de una recombinasa elaborada a medida (o la recombinasa elaborada a medida en forma de una proteína o polipéptido de fusión o una célula troncal que comprende el vector de expresión) es capaz de reducir la carga vírica en un sujeto infectado por un retrovirus mediante la erradicación del reservorio genético de retrovirus dentro de las células hospedadoras, evitando de esta manera ciclos vitales adicionales del virus.

La expresión "carga vírica" tal como se utiliza en la presente memoria se refiere a, por ejemplo, los equivalentes de ARN del VIH (es decir, los genomas) que se asocian a 1 ml de plasma del paciente (Dybul et al., 2002). De esta manera, la carga vírica se determina mediante la medición del contenido de ADN vírico en una muestra obtenida del paciente. Actualmente existen tres tipos principales de ensayos de carga vírica:

- 1) la reacción en cadena de la polimerasa-transcripción inversa del ARN del VIH (RT-PCR): Amplicor VIH-1 Monitor Test; Roche Diagnostics.
- 2) ADN de cadena ramificada (ADNr): ensayo de ARN del VIH Versant™; Bayer Diagnostics, y
- 3) amplificación basada en la secuencia de ácidos nucleicos (NASBA, por sus siglas en inglés): ensayo NucliSens™; bioMérieux.

En una forma de realización preferida, la composición farmacéutica de la presente invención es capaz de reducir la carga vírica en un sujeto infectado por un retrovirus a menos de 5.000 equivalentes genómicos/ml de plasma, preferentemente a menos de 500 equivalentes genómicos/ml de plasma y más preferentemente a menos de 50 equivalentes genómicos/ml de plasma. Los pacientes con una carga vírica inferior a 5.000 equivalentes genómicos/ml de plasma se considera que se encuentran relativamente bien adaptados al tratamiento médico. Sin embargo, el objetivo en la terapia actual del SIDA es una reducción de la carga vírica a un nivel inferior al límite de detección de los ensayos de carga vírica, que actualmente es inferior a aproximadamente 50 equivalentes genómicos/ml de plasma.

El sujeto infectado por VIH-1, en el que debe administrarse la composición farmacéutica, se selecciona de entre el grupo que consiste en seres humanos y primates. Sin embargo, el sujeto preferentemente es un ser humano.

En general, debe administrarse en el sujeto una cantidad eficaz del vector de expresión, la recombinasa elaborada a medida o la célula transformada de la invención. La administración puede ser, por ejemplo, administración intravenosa o intramuscular.

En una forma de realización, la composición farmacéutica se formula para la administración concomitante con otros agentes activos de la terapia antiretroviral altamente activa (TARAA). La terapia antiretroviral altamente activa, TARAA, es una terapia de combinación con diana en la transcriptasa inversa, proteasa y fusión víricas (Gulick et al., 1997; Lalezari et al., 2003).

En otra forma de realización, la composición farmacéutica se formula para la administración concomitante o posterior a la terapia de activación inmunológica global o a la activación específica de la expresión génica de provirus. La terapia de activación inmunológica se basa sobre la hipótesis de que la activación deliberada de células infectadas latentemente por el VIH puede acelerar la erradicación de los reservorios víricos persistentes. La erradicación ocurriría mediante eliminación inmunitaria por muerte programada de las células que expresan activamente productos (proapoptóticos) del VIH-1 (Kulkosky y Bray, 2006). La activación inmunológica global (activación de células inmunitarias, incluyendo células en reposo) habitualmente se consigue mediante, por ejemplo, la administración de inmunotoxinas, citocinas (por ejemplo IL-2) o anticuerpos activadores de las células T (por ejemplo OKT3).

A partir de que la activación inmunitaria realizada para activar deliberadamente los reservorios latentes resistentes a la TARAA desafortunadamente no ha conseguido eliminar permanentemente el VIH-1 y el rebote vírico (para una revisión ver Kulkosky y Bray, 2006; Marcello, 2006; Shehu-Xhilaga et al., 2005) debido al hecho de que la activación global de las células T aparentemente también induce la replicación vírica e incrementa el número de potenciales células diana con VIH-1 más allá del nivel que puede ser contenido por la TARAA (Fraser et al., 2000), resultan necesarios tratamientos específicos adicionales para tratar el VIH. Un enfoque es la activación de la transcripción de genomas víricos que de otro modo estarían inactivos. Puede conseguirse una activación específica de la expresión

de genes províricos latentes mediante la administración del éster de forbol Prostratin o la citocina humana IL-7, en la que ambos aparentemente reactivan el VIH-1 latente en ausencia de proliferación celular (Marcello, 2006). Además, la activación transcripcional selectiva del VIH-1 también puede conseguirse con inhibidores de la histona desacetilasa (HDAC1) tales como, por ejemplo, el ácido valproico, que finalmente induce la multiplicación expansiva del VIH-1 a partir de células en reposo en ausencia de activación celular (Marcello, 2006; Lehrman et al., 2005).

Sin embargo, la terapia de activación inmunitaria global o la activación específica de la expresión génica provírica o estrategias de terapia similares se beneficia mucho de la eliminación concurrente del ADN provírico, reduciendo de esta manera en el paciente la reserva de células infectadas.

La presente invención proporciona además un procedimiento de tratamiento y/o prevención de la infección por VIH-1 en un sujeto. En una forma de realización, la secuencia del retrovirus que infecta el sujeto se analiza en una muestra obtenida del mismo y debe administrarse en el sujeto por lo menos un vector de expresión codificante de una recombinasa elaborada a medida, por lo menos una recombinasa elaborada a medida o por lo menos una célula transformada con dicho vector de expresión, por ejemplo una célula troncal adulta, en el caso de que el ADN provírico del sujeto comprenda la secuencia diana asimétrica de SEC ID nº 1. La muestra obtenida del sujeto puede ser una muestra de sangre, por ejemplo que comprenda células CD4+ infectadas.

Breve descripción de las figuras

Figura 1: la figura 1 proporciona los sitios diana de recombinasas específicas, loxP (reconocida por Cre), loxH (reconocida por Fre), rox (reconocida por Dre), zox (reconocida por Zre) y loxLTR (reconocida por Tre) (SEC ID nº 3 a nº 7), así como sitios diana asimétricos presentes en una pluralidad de cepas del VIH-1 (SEC ID nº 1 y nº 2). La secuencia espaciadora se presenta en negrita. Las posiciones de nucleótido en el hemisitio izquierdo y el hemisitio derecho de SEC ID nº 1 o SEC ID nº 2 que no son homólogas respecto a los sitios diana conocidos se encuentran subrayadas. La SEC ID nº 1 presenta una homología de 54% respecto a un grupo de sitios diana de recombinasa conocidos; SEC ID nº 2 presenta una homología de 42% respecto a un grupo de sitios diana de recombinasa conocidos.

Figura 2: la figura 2 proporciona las secuencias de los subsitios utilizados para la generación de las recombinasas elaboradas a medida TRE 3.0 y TRE 4.0. Las posiciones de nucleótidos en el hemisitio izquierdo y en el hemisitio derecho de la SEC ID nº 1 o nº 2 que no son homólogas a uno de los sitios diana conocidos se encuentran subrayados. Los subsitios 1 y 2 se generan en la etapa (d) del procedimiento de la invención y los subsitios 1A-C y los subsitios 2A-C se generan en la etapa (e) del procedimiento de la invención. Los subsitios 1a/b y los subsitios 2a/b, así como los subsitios intermediarios 1b+1, 1b +2, 2a+1 y 2a+2 se generan en etapas adicionales del procedimiento de la invención.

Figura 3: la figura 3 proporciona un diagrama general de la estrategia de evolución y muestra los resultados de las primeras etapas de evolución para la SEC ID nº 1. La colección inicial de recombinasas se generó mediante agrupación y mezcla de familias de Cre y varias recombinasas de tipo Cre conocidas (A). Se llevó a cabo un enriquecimiento en las recombinasas deseadas mediante la realización de varias rondas de evolución de proteínas ligada a sustrato en los subsitios del segundo subconjunto. En (B), se sometió a ensayo la actividad de recombinación tras el primer ciclo de evolución, en (E) tras el sexto ciclo de evolución. El producto recombinado es más corto, ya que la secuencia entre los sitios diana ha sido extraída, la cual se encuentra marcada con un triángulo. El producto no recombinado se encuentra marcado con dos triángulos. Se amplificaron las recombinasas deseadas mediante PCR, resultando en un producto de 1,7 kb (C y D). Una de las no correspondencias en los subsitios 1C y 2C se trató separadamente (mediante un sitio diana loxP mutado denominado ΔloxP, con el fin de obtener una actividad de recombinación residual más elevada en los subsitios 1C y 2C(D)).

Figura 4: la figura 4 proporciona un diagrama general de las etapas posteriores de la estrategia de evolución y muestra los resultados de las etapas finales de la evolución para la SEC ID nº 1. La parte (A) de la figura corresponde a las etapas de los primeros ciclos de evolución descritos en detalle en la figura 3.

(B) Se llevó a cabo un enriquecimiento adicional en las recombinasas deseadas mediante la realización de varias rondas de evolución de proteínas ligada a sustrato en los subsitios del tercer subconjunto, seguido de varias rondas de ciclos de evolución ligada a sustrato en algunos sustratos intermediarios (C) y finalmente en los primeros subsitios (D) con el fin de obtener recombinasas específicas para el sitio diana asimétrico SEC ID nº 1. En (E), se sometió a ensayo la actividad de recombinación tras el 8º ciclo de evolución con cantidades relativamente elevadas del activador transcripcional L-arabina (L-ara, 200 µg/ml), en (F) se midió la actividad de recombinación tras el 16º ciclo de evolución con inducción transcripcional más baja, con sólo 50 µg/ml de L-ara. El producto recombinado se ha marcado con un triángulo, el no recombinado se ha marcado con dos triángulos. (G) y (H) muestran la actividad de recombinación respectiva tras los ciclos de evolución indicados, las concentraciones de L-ara indicadas y en los subsitios indicados. Tras el ciclo final de evolución (ciclo

43), se aisló y se enriqueció la población de recombinasas con la actividad máxima, resultando en una colección activa. SE seleccionaron los clones individuales de recombinasas y se sometieron a análisis adicionales.

- 5 Figura 5: la figura 5 proporciona una alineación de las secuencias de proteína de (a) la SEC ID nº 36 de la recombinasa Cre, (b) la SEC ID nº 37 de la secuencia de consenso común de las recombinasas Tre (específica para los sitios diana asimétricos dentro de la LTR del VIH-1), (c) la SEC ID nº 38 de las secuencia de consenso obligatoria/restrictiva de las recombinasas Tre específicas para la SEC ID nº 1, y (d) la SEC ID nº 39 de la secuencia de consenso relajada (con una obligación de 95%) de las recombinasas Tre específicas para la SEC ID nº 1. Los tramos con letras en negrita indican los aminoácidos conservados; las posiciones no especificadas se indican con una X.

Ejemplos

15 Ejemplo 1:

identificación de secuencias de diana asimétrica presentes en una pluralidad de cepas

- 20 Se identificaron las secuencias de diana asimétrica mediante un procedimiento de 3 etapas: generación de una matriz de pesos de posición basado en los sitios de reconocimiento conocidos de una recombinasa, provisión de secuencias genómicas en las que deben realizarse las búsquedas para sitios de reconocimiento, y búsqueda binaria de sitios diana potenciales en las secuencias proporcionadas y puntuación de los aciertos resultantes basándose en la matriz inicial de pesos de posición, en la que los nucleótidos se transforman en espacio binario.

- 25 Se encontró que la SEC ID nº 1 estaba presente en 348 de 379 cepas de VIH-1 de subtipo B. Además, la SEC ID nº 1 se encontró que estaba presente en 32 de 40 cepas de VIH-1 del subtipo A. Se encuentra localizada en la región R de la LTR.

- 30 Se encontró que la SEC ID nº 2 estaba presente en 288 de 379 cepas de VIH-1 de subtipo B. No se encontró la SEC ID nº 2 en ninguna de las 40 cepas de VIH-1 de subtipo A en las que se realizó la búsqueda. Se encuentra localizada en la región U3 de la LTR.

Ejemplo 2:

- 35 Generación de una recombinasa elaborada a medida que reconoce y se recombina con una secuencia diana asimétrica dentro de la LTR del VIH-1

- 40 Para iniciar el procedimiento de evolución, se seleccionaron las secuencias de LTR del VIH-1 que se encontraban altamente conservadas entre las cepas de VIH-1. Se encontró que la SEC ID nº 1 y la SEC ID nº 2 satisfacían estos criterios y representaban secuencias diana asimétricas para las que puede seleccionarse una recombinasa elaborada a medida.

- 45 A continuación se describe en detalle la generación de una recombinasa elaborada a medida que reconoce la SEC ID nº 1.

- 50 En la figura 2 se proporciona de manera general un diagrama simplificado de la estrategia de evolución y los avances realizados hasta ahora. Globalmente, la evolución de una nueva recombinasa Tre que reconoce específicamente la secuencia diana deseada se llevó a cabo tal como se da a conocer en el documento nº WO 2008/083931 y se indica en Sarkar et al. (Sarkar et al., Science, 2007).

- 55 La colección inicial de recombinasas se generó mediante agrupación y mezcla de familias de Cre y varias recombinasas de tipo Cre conocidas, a saber una colección de mutantes de Cre (Cre/Fre) (Buchholz y Stewart, Nature, 2001), una colección de Tre generada previamente (Sarkar et al., Science, 2007), Dre (Sauer et al., Nucl. Acids Res., 2004), Zre (aislada a partir de *Salmonella enterica*, GenBank nº de acceso NZ_ABEW01000015) y Shew (*Shewanella* sp., cepa ANA-3, GenBank nº de acceso CP000470) (figura 2A). Con el fin de incrementar la eficiencia de la mezcla de familias, se modificó Dre y Shew ("tipo Cre") mediante la modificación de la secuencia de nucleótidos de manera que fuese similar al máximo a Cre sin resultar en una secuencia de proteínas modificada. Para llevar a cabo la evolución de proteínas ligada a sustrato, se clonó la colección inicial de recombinasas en los vectores de evolución (Sarkar et al., Science, 2007) que contenía el subsitio diana respectivo.

- 60 Se generó un primer subconjunto de dos ácidos nucleicos diana que comprendían secuencias diana, en el que la primera secuencia diana se denominó subsitio 1 y comprendía la SEC ID nº 8 y en el que la segunda secuencia diana se denominó subsitio 2 y comprendía la SEC ID nº 12.

- 65 A continuación, se generó un segundo subconjunto de ácidos nucleicos diana que comprendía secuencias diana modificadas, basándose en los subsitios 1 y 2. Tal como se muestra en la figura 2, comprenden las SEC ID nº 9 a nº

11 y nº 13 a nº 15, respectivamente. Se subrayan en gris las no correspondencias de las secuencias respecto a la reserva de secuencias diana conocidas reconocidas por las recombinasas contenidas en la colección inicial. Las colecciones de subsitio se cultivaron independientemente en *E. coli*. En los primeros ciclos de evolución (1 a 4) se indujo la expresión de recombinasa con 200 µg/ml de L-arabinosa; en los ciclos siguientes, con 100 µg/ml (ciclos 5 a 6). Los cultivos se cultivaron durante 13 horas a 37°C, bajo agitación. Para someter a ensayo la actividad de recombinación, el ADN plasmídico aislado de los cultivos se digirió con BsrGI y XbaI. La digestión del ADN plasmídico recombinado resulta en un fragmento más corto que con el plásmido no recombinado (Sarkar et al., Science, 2007). Tras el primer ciclo de evolución, se observaron bandas muy débiles o ninguna banda de producto recombinado para todas las colecciones de subsitio 1A, 1B, 1C, 2A, 2B y 2C (figura 3B). Con el fin de aislar las recombinasas deseadas, se digirió el ADN plasmídico con NdeI y se amplificó por PCR antes de clonarlo nuevamente dentro de los vectores de evolución respectivos para el siguiente ciclo de evolución (Sarkar et al., Science, 2007). Tras el primer ciclo de evolución, se obtuvo una actividad residual suficiente para las colecciones de subsitio 1A, 1B, 2A y 2B, pero no para las colecciones de subsitio 1C y 2C (figura 3C). Se encontró que la no correspondencia en la posición 11 de la secuencia diana de 34 pb era la más crítica para la actividad de recombinación. Por lo tanto, se generó un sitio loxP modificado, denominado ΔloxP, que contenía únicamente dicha mutación individual (T) y la repetición invertida respectiva en la posición 24 (A). Al someter a ensayo la colección inicial de recombinasas en el vector de evolución que contenía el sitio diana ΔloxP, se obtuvo una actividad residual suficiente (figura 3D). La colección amplificada de las recombinasas deseadas se clonó nuevamente introduciéndola en los vectores de evolución con los subsitios 1C y 2C, respectivamente, y se sometió a ensayo para la actividad de recombinación. Se obtuvo también actividad residual suficiente para las colecciones de subsitio 1C y 2D (figura 3D), de manera que pudo iniciarse el procedimiento de evolución en todos los subsitios. Se llevaron a cabo 6 ciclos de evolución en cada uno de los subsitios 1A, 1B, 1C, 2A, 2B y 2C, enriqueciendo significativamente las colecciones de recombinasas deseadas, tal como puede observarse en el gel de ensayo mostrado en la figura 3E (comparar con la figura 3B).

Las colecciones de subsitio enriquecidas se agruparon y se mezclaron con el fin de obtener las colecciones para continuar el procedimiento de ciclado evolutivo con el 2º subconjunto de secuencias diana, es decir el subsitio 1 y el subsitio 2. A continuación, se enriquecieron las colecciones de recombinasas deseadas tal como se ha indicado anteriormente antes de agruparlas y mezclarlas para la etapa final del ciclado evolutivo, es decir la evolución del sitio diana que debía ser reconocido específicamente por la recombinasa Tre. El ácido nucleico codificante de la recombinasa puede obtenerse a partir de la colección y clonarse en un vector de expresión.

Se utilizaron los materiales y procedimientos indicados en el documento nº WO 2008/083931 y en Buchholz y Stewart, 2001, a menos que se indique lo contrario.

Las recombinasas TRE obtenidas adaptadas para reconocer y recombinarse con las secuencias diana asimétricas SEC ID nº 1 y SEC ID nº 2 se secuenciaron y se confirmó y su capacidad de recombinar ADN provírico de diferentes cepas de VIH-1 diana.

Listado de referencias

Abremski K, Hoess RH, Sternberg N (1983) "Studies on the properties of P1 site-specific recombination: evidence for topologically unlinked products following recombination." Cell 32, 1301-1311.

Abremski K, Hoess R (1983) "Bacteriófago P1 site-specific recombination. Purification and properties of the Cre recombinase protein." J. Biol. Chem. 259, 1509-1514.

Adachi A, Gendelman HE, Koenig S, Folks T, Willey R, Rabson A, Martin MA (1986) "Production of acquired immunodeficiency syndrome-associated retrovirus in human and nonhuman cells transfected with an infectious molecular clone." J. Virol. 59, 284-291.

Alper H, Fischer C, Nevoigt E, Stephanopoulos G (2006) "Tuning genetic control through promoter engineering" Proc. Natl. Acad. Sci. USA 102, 12678-12683.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic acids research, 25, 3389-3402.

Beyer WR, Westphal M, Ostertag W, von Laer D (2002) "Oncoretrovirus and lentivirus vectors pseudotyped with lymphocytic choriomeningitis virus glycoprotein: generation, concentration and broad host range." J. Virol. 76, 1488-1495.

Blackard JT, Renjifo BR, Mwakagile D, Montano MA, Fawzi WW, Essex M (1999) "Transmission of human immunodeficiency type 1 viruses with intersubtype recombinant long terminal repeat sequences." Virology 254, 220-225.

Bloom JD, Meyer MM, Meinhold P, Otey CR, MacMillan D, Arnold FH (2005) "Evolving strategies for enzyme

- engineering." *Curr. Opin. Struct. Biol.* 15, 447-452.
- Buchholz F, Ringrose L, Angrand PO, Rossi F, Stewart AF (1996) "Different thermostabilities of FLP and Cre recombinases: implications for applied site-specific recombination." *Nucl. Acids Res.* 24, 4256-4262.
- 5 Buchholz F, Angrand PO, Stewart AF (1998) "Improved properties of FLP recombinase evolved by cycling mutagenesis." *Nat. Biotechnol.* 16, 657-662.
- Buchholz F, Stewart AF (2001) "Alteration of Cre recombinase site specificity by substrate-linked protein evolution." *Nat. Biotechnol.* 19, 1047-1052.
- 10 Chiu YL, Soros VB, Kreisberg JF, Stopak K, Yonemoto W, Greene WC (2005) "Cellular APOBEC3G restricts HIV-1 infection in resting CD4+ T cells." *Nature* 435, 108-114
- 15 Chun T-W, Engel D, Berrey MM, Shea T, Corey L, Fauci AS (1998) "Early establishment of a pool of latently infected, resting CD4+ T cells during primary HIV-1 infection." *Proc. Natl. Acad. Sci. USA* 95, 8869-8873.
- Coates CJ, Kaminski JM, Summers JB, Segal DJ, Miller AD, Kolb AF (2005) "Site-directed genome modification: derivatives of DNA-modifying enzymes as targeting tools." *Trends Biotechnol.* 23, 407-419.
- 20 Collins CH, Yokobayashi Y, Umeno D, Arnold FH, (2003) "Engineering proteins that bind, move, make and break DNA." *Curr. Opin. Biotechnol.* 14, 665.
- Combes P, Till R, Bee S, Smith MC (2002) "The streptomyces genome contains multiple pseudo-attB sites for the (phi)C31-encoded site-specific recombination system." *J. Bacteriol.* 184, 5746-5752.
- 25 Crameri A, Raillard SA, Bermudez E, Stemmer WP (1998) "DNA shuffling of a family of genes from diverse species accelerates directed evolution." *Nature* 391, 288-291.
- 30 Derossi D, Joliot AH, Chassaing G, Prochiantz A (1994) "The third helix of the Antennapedia homeodomain translocates through biological membranes." *J. Biol. Chem.* 269, 10444-10450.
- Derossi D, Calvet S, Trembleau A, Chassaing G, Prochiantz A (1996) "Cell internalization of the third helix of the Antennapedia homeodomain is receptor-independent." *J Biol Chem* 271, 18188-18193.
- 35 Donovan, P.J., Gearhart, J. (2001) "The end of the beginning for pluripotent stem cells." *Nature* 414, 92-97.
- Donzella GA, Schols D, Lin SW, Este JA, Nagashima KA, Maddon PJ, Allaway GP, Sakmar TP, Henson G, De Clercq E, Moore JP (1998) "AMD3100, a small molecule inhibitor of HIV-1 entry via the CXCR4 co-receptor." *Nature Medicine* 4, 72-77.
- 40 Dybul M, Fauci AS, Bartlett JG, Kaplan JE, Pau AK (2002) "Guidelines for using antiretroviral agents among HIV-infected adults and adolescents." *Annals of Internal Medicine* 137, 381-433.
- 45 Eddy, S.R. (1998) *Profile hidden Markov models*. Bioinformatics (Oxford, England), 14, 755-763.
- Edelman GM, Meech R, Owens GC, Jones FS (2000) "Synthetic promoter elements obtained by nucleotide sequence variation and selection for activity." *Proc. Natl. Acad. Sci. USA* 97, 3038-3043.
- 50 Emerman M, Malim MH (1998) "HIV-1 regulatory/accessory genes: keys to unraveling viral and host cell biology." *Science* 280, 1880-1884.
- Finzi D, Hemankova M, Pierson T, Carruth LM, Buck C, Chaisson RE, Quinn TC, Chadwick K, Margolick J, Brookmeyer R, Gallant J, Markowitz M, Ho DD, Richman DD, Siliciano RF (1997) "Identification of a reservoir for HIV-1 in patients on highly active antiretroviral therapy." *Science* 278, 1295-1300.
- 55 Flowers CC, Woffendin C, Petryniak J, Yang S, Nabel GJ (1997) "Inhibition of recombinant human immunodeficiency virus type 1 replication by a site-specific recombinase." *J. Virol.* 71, 2685-2692.
- 60 Gulick RM, Mellors JW, Havlir D, Eron JJ, Gonzalez C, McMahon D, Richman DD, Valentine FT, Jonas L, Meibohm A, Emini EA, Chodakewitz JA (1997) "Treatment with indinavir, zidovudine, and lamivudine in adults with human immunodeficiency virus infection and prior antiretroviral therapy." *N. Engl. J. Med.* 337, 734-739.
- Guzman LM, Belin D, Carson MJ, Beckwith J (1995) "Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter." *J. Bacteriol.* 177, 4121-4130.
- 65

- Hartenbach S, Fussenegger M (2006) "A novel synthetic mammalian promoter derived from an internal ribosome entry site." *Biotechnology and Bioengineering* 95, 547-559.
- 5 Hauber I, Bevec D, Heukeshoven J, Krätzer F, Horn F, Choidas A, Harrer T, Hauber J (2005) "Identification of cellular deoxyhypusine synthase as a novel target for antiretroviral therapy." *J. Clin. Invest.* 115, 76-85.
- 10 Hazuda DJ, Young SD, Guare JP, Anthony NJ, Gomez RP, Wai JS, Vacca JP, Handt L, Motzel SL, Klein HJ, Dornadula G, Danovich RM, Witmer MV, Wilson KA, Tussey L, Schleif WA, Gabryelski LS, Jin L, Miller MD, Casimiro DR, Emini EA, Shiver JW (2004) "Integrase inhibitors and cellular immunity suppress retroviral replication in rhesus macaques." *Science* 305, 528-532.
- 15 Hoess RH, Abremski K (1985) "Mechanism of strand cleavage and exchange in the Cre-lox site-specific recombination system." *J. Mol. Biol.* 181, 351-362.
- 20 Johannes TW, Zhao H (2006) "Directed evolution of enzymes and biosynthetic pathways." *Curr. Opin. Microbiol.* 9, 261-267.
- 25 Krasnow MA, Cozzarelli NR (1983) "Site-specific relaxation and recombination by the Tn3 resolvase: Recognition of the DNA path between oriented res sites." *Cell* 32, 1313-1324.
- 30 Kulkosky J, Bray S (2006) "HAART-persistent HIV-1 latent reservoirs: their origin, mechanisms of stability and potential strategies for eradication." *Curr. HIV Res.* 4, 199-208.
- 35 Lalezari JP, Henry K, O'Hearn M, Montaner JS, Piliero PJ, Trottier B, Walmsley S, Cohen C, Kuritzkes DR, Eron Jr. JJ, Chung J, DeMasi R, Donatacci L, Drobnos C, Delehanty J, Salgo M (2003) "Enfuvirtide, an HIV-1 fusion inhibitor, for drug-resistant HIV infection in North and South America." *N. Engl. J. Med.* 348, 2175-2185.
- 40 Lee YS, Park JS (1998) "A novel mutant loxP containing part of long terminal repeat of HIV-1 in spacer region: presentation of possible target site for antiviral strategy using site-specific recombinase." *Biochem. Biophys. Res. Comm.* 253, 588-593.
- 45 Lee YS, Kim ST, Kim GW, Lee M, Park JS (2000) "An engineered lox sequence containing part of a long terminal repeat of HIV-1 permits Cre recombinase-mediated DNA excision." *Biochem. Cell Biol.* 78, 653-658.
- 50 Lehrman G, Hogue IB, Palmer S, Jennings C, Spina CA, Wiegand A, Landay AL, Coombs RW, Richman DD, Mellors JW, Coffin JM, Bosch RJ, Margolis DM (2005) "Depletion of latent HIV-1 infection in vivo: a proof-of-concept study" *Lancet* 366, 549-555.
- 55 Lewandoski, M. (2001) "Conditional control of gene expression in the mouse." *Nat. Rev. Genet.* 2, 743-755.
- 60 Lin Q, Jo D, Gebre-Amlak KD, Ruley HE (2004) "Enhanced cell-permeant Cre protein for site-specific recombination in cultured cells." *BMC Biotechnol.* 4, 25.
- 65 Little SJ, Holte S, Routy JP, Daar ES, Markowitz M, Collier AC, Koup RA, Mellors JW, Connick E, Conway B, Kilby M, Wang L, Whitcomb JM, Hellmann NS, Richman DD (2002) "Antiretroviral-drug resistance among patients recently infected with HIV." *N. Engl. J. Med.* 347, 385-394.
- Macara IG (2001) "Transport into and out of the nucleus." *Microbiology and molecular biology reviews* 65, 570-594.
- 50 Malim MH, Hauber J, Fenrick R, Cullen BR (1988) "Immunodeficiency virus rev transactivator modulates the expression of the viral regulatory genes." *Nature* 335, 181-183. Marcello A (2006) "Latency: the hidden HIV-1 challenge." *Retrovirology* 3, 7.
- 55 Matsumura I, Ellington AD (2001) "In vitro evolution of beta-glucuronidase into a beta-galactosidase proceeds through non-specific intermediates." *J. Mol. Biol.* 305, 331-339. Minshull J, Stemmer WP. (1999) "Protein evolution by molecular breeding." *Curr. Opin. Chem. Biol.* 3, 284-290.
- 60 Nagy A (2000) "Cre recombinase: the universal reagent for genome tailoring." *Genesis* 26, 99-109.
- 65 Needleman SB, Wunsch CD (1970) "A general method applicable to the search for similarities in the amino acid sequence of two proteins." *J. Mol. Biol.* 48, 443-453.
- Nolden L, Edenhofer F, Haupt S, Koch P, Wunderlich FT, Siemen H, Brustle O. (2006) "Site-specific recombination in human embryonic stem cells induced by cell-permeant Cre recombinase." *Nat. Methods* 3, 461-467.
- O'Doherty U, Swiggard WJ, Malim MH (2000) "Human immunodeficiency virus type 1 spinoculation enhances

infection through virus binding." J. Virol. 74, 10074-10080.

Pearson WR, Lipman DJ (1988) "Improved tools for biological sequence comparison." Proc Natl Acad Sci USA 85, 2444-2448.

Peitz M, Pfannkuche K, Rajewsky K, Edenhofer F. (2002) "Ability of the hydrophobic FGF and basic TAT peptides to promote cellular uptake of recombinant Cre recombinase: A tool for efficient genetic engineering of mammalian genomes." Proc. Natl. Acad. Sci. USA 99, 4489-4494.

Ratner L, Starcich B, Josephs SF, Hahn BH, Reddy EP, Livak KJ, Petteway SR, Jr., Pearson ML, Haseltine WA, Arya SK, (1985) "Polymorphism of the 3' open reading frame of the virus associated with the acquired immune deficiency syndrome, human T-lymphotropic virus type III." Nucl. Acids Res. 13, 8219-8229.

Richard JP, Melikov K, Brooks H, Prevot P, Lebleu B, Chernomordik LV (2005) "Cellular uptake of the unconjugated TAT peptide involves clathrin-dependent endocytosis and heparin sulfate receptors." J. Biol. Chem. 280, 15300-15306.

Rüfer AW, Sauer B (2002) "Non-contact positions impose site selectivity on Cre recombinase." Nucl. Acids Res. 30, 2764-2771.

Ruhl M, Himmelsbach M, Bahr GM, Hammerschmid F, Jaksche H, Wolff B, Aschauer H, Farrington GK, Probst H, Bevec D, Hauber J (1993) "Eukaryotic initiation factor 5A is a cellular target of the human immunodeficiency virus type 1 Rev activation domain mediating trans-activation" J. Cell Biol. 123, 1309-1320.

Sanger F, Nickler S, Coulson AR (1977) "DNA sequencing with chain-terminating inhibitors." Proc. Natl. Acad. Sci. USA 74, 5463-5467.

Santoro SW, Schultz PG (2002) "Directed evolution of the site specificity of Cre recombinase." Proc. Natl. Acad. Sci. USA 99, 4185-4190.

Saraf-Levy T, Santoro SW, Volpin H, Kushnirsky T, Eyal Y, Schultz PG, Gidoni D, Carmi N (2006) "Site-specific recombination of asymmetric lox sites mediated by a heterotetrameric Cre recombinase complex." Bioorg. Med. Chem. 14, 3081-3089.

Sauer B, McDermott J (2004) "DNA recombination with a heterospecific Cre homolog identified from comparison of the pac-c1 regions of P1-related phages." Nucl. Acids. Res. 32, 6086-6095.

Schambach A, Bohne J, Chandra S, Will E, Margison GP, Williams DA, Baum C (2006) "Equal potency of gammaretroviral and lentiviral SIN vectors for expression of O6-methylguanine-DNA methyltransferase in hematopoietic cells." Molecular Therapy 13, 391-400.

Scherr M, Eder M (2002) "Gene Transfer into Hematopoietic Stem Cells Using Lentiviral Vectors." Current Gene Therapy 2, 45-55.

Shehu-Xhilaga M, Tachedjian G, Crowe SM, Kedzierska K. (2005) "Antiretroviral compounds: mechanisms underlying failure of HAART to eradicate HIV-1." Curr. Med. Chem. 12, 1705-1719.

Shimshek DR, Kim J, Hubner MR, Spergel DJ, Buchholz F, Casanova E, Stewart AF, See-burg PH, Sprengel R (2002) "Codon-improved Cre recombinase (iCre) expression in the mouse." Genesis 32(1), 19-26.

Smith Tf, Waterman MS (1981) "Overlapping genes and information theory." J. Theor. Biol. 91, 379-380.

Stark WM, Boocock MR, Sherratt DJ (1992) "Catalysis by site-specific recombinases." Trends Genet. 8, 432-439.

Stemmer WPC (1994) "Rapid evolution of a protein in vitro by DNA shuffling." Nature 370, 389-391.

Sternberg N, Hamilton D (1981) "Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites." J. Mol. Biol. 150, 467-486.

Van Duyne GD (2001) "A structural view of cre-loxp site-specific recombination." Annu. Rev. Biophys. Biomol. Struct. 30, 87-104.

Vives E, Brodin P, Lebleu B (1997) "A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus." J. Biol. Chem. 272, 16010-16017.

Vives E (2003) "Cellular uptake of the TAT peptide: an endocytosis mechanism following ionic interactions." J. Mol.

Recognit. 16, 265-271.

Volkert FC, Broach JR (1986) "Site-specific recombination promotes plasmid amplification in yeast." Cell 46, 541-550.

5 Voziyanov Y, Konieczka JH, Stewart AF, Jayaram M (2003) "Stepwise manipulation of DNA specificity in FLP recombinase: progressively adapting FLP to individual and combinatorial mutations in its target site." J. Mol. Biol. 326, 65-76.

10 Yuan L, Kurek I, English J, Keenan R (2005) "Laboratory-directed protein evolution" Microbiol. Mol. Biol. Rev. 69, 373-92.

WO 2002/44409

WO 2008/083931

15 Fawell S, Seery J, Daikh Y, Moore C, Chen LL, Pepinsky B, Barsoum J., Tat-mediated delivery of heterologous proteins into cells., Proc Natl Acad Sci USA. 1994 Jan 18;91(2):664-8.

20 Elliott G, O'Hare P., Intercellular trafficking and protein delivery by a herpesvirus structural protein., Cell. 1997 Jan 24;88(2):223-33

Oess S, Hildt E., Novel cell permeable motif derived from the PreS2-domain of hepatitis-B virus surface antigens., Gene Ther. 2000 May;7(9):750-8.

25 Listado de secuencias

<110> Heinrich-Pette-Institut Leibniz-Institut für Experimentelle Virologie Max-Planck-Gesellschaft zur Foerderung der Wissenschaften e.V.

30 <120> Recombinasa elaborada a medida para la recombinación de sitios diana asimétricos en una pluralidad de cepas de retrovirus

<130> P 85349

35 <160> 65

<170> PatentIn versión 3.5

<210> 1

40 <211> 34

<212> ADN

<213> Virus de la inmunodeficiencia humana de tipo 1

<400> 1

45 aaccactgc ttaagcctca ataaagcttg cctt 34

<210> 2

<211> 34

<212> ADN

50 <213> Virus de la inmunodeficiencia humana de tipo 1

<400> 2

ctgggcggga ctggggagtg gcgagccctc agat 34

55 <210> 3

<211> 34

<212> ADN

<213> Secuencia artificial

60 <220>

<223> loxP

<400> 3

ataacttctg ataattgatg ctatacgaag ttat 34

65 <210> 4

	<211> 34	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> loxH	
	<400> 4	
10	atatatacgt atatagacat atatacgtat atat	34
	<210> 5	
	<211> 34	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> rox	
	<400> 5	
20	ctaactttaa ataatgccaa ttatttaaag ttat	34
	<210> 6	
	<211> 34	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> zox	
30	<400> 6	
	ataacttctg ataacacaca ttatgcgaag ttat	34
	<210> 7	
	<211> 34	
35	<212> ADN	
	<213> Human immunodeficiency virus	
	<400> 7	
40	acaacatcct attacaccct atagccaac atgg	34
	<210> 8	
	<211> 34	
	<212> ADN	
	<213> Secuencia artificial	
45	<220>	
	<223> sub sitio 1	
	<400> 8	
50	aaccactgc ttaagcctca ataagcagtg gggt	34
	<210> 9	
	<211> 34	
	<212> ADN	
55	<213> Secuencia artificial	
	<220>	
	<223> sub sitio 1A	
60	<400> 9	
	aacacattgt ataagcctca atatacaatg tggt	34
	<210> 10	
	<211> 34	
65	<212> ADN	
	<213> Secuencia artificial	

	<220>		
	<223> subsitio 1B		
5	<400> 10 ataccactgt ataagcctca atatacagtg gtat	34	
	<210> 11		
	<211> 34		
10	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> subsitio 1C		
15	<400> 11 ataacattgc ttaagcctca ataagcaatg ttat	34	
	<210> 12		
20	<211> 34		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
25	<223> subsitio 2		
	<400> 12 aaggcaagct ttaagcctca ataaagcttg cctt	34	
30	<210> 13		
	<211> 34		
	<212> ADN		
	<213> Secuencia artificial		
35	<220>		
	<223> subsitio 2A		
	<400> 13 aagactaggt ataagcctca atatacctag tctt	34	
40	<210> 14		
	<211> 34		
	<212> ADN		
	<213> Secuencia artificial		
45	<220>		
	<223> subsitio 2B		
	<400> 14 atagcaaggt ataagcctca atataccttg ctat	34	
50	<210> 15		
	<211> 34		
	<212> ADN		
55	<213> Secuencia artificial		
	<220>		
	<223> subsitio 2C		
60	<400> 15 ataactagct ttaagcctca ataaagctag ttat	34	
	<210> 16		
	<211> 34		
65	<212> ADN		
	<213> Secuencia artificial		

	<220>		
	<223> subsitio 1		
5	<400> 16 ctgggcggga ctggggagtg gcagtcccg cag	34	
	<210> 17		
	<211> 34		
10	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> subsitio 1A		
15	<400> 17 ctgacttcga ctagggagtg gtagtcgaag tcag	34	
	<210> 18		
20	<211> 34		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
25	<223> subsitio 1B		
	<400> 18 ataggttcga atggggagtg gcattcgaac ctat	34	
30	<210> 19		
	<211> 34		
	<212> ADN		
	<213> Secuencia artificial		
35	<220>		
	<223> subsitio 1C		
	<400> 19 ataaccggga ataggagtg gtattccgg ttat	34	
40	<210> 20		
	<211> 34		
	<212> ADN		
	<213> Secuencia artificial		
45	<220>		
	<223> subsitio 2		
	<400> 20 atctgagggc tcggggagtg gcgagccctc agat	34	
50	<210> 21		
	<211> 34		
	<212> ADN		
55	<213> Secuencia artificial		
	<220>		
	<223> subsitio 2A		
60	<400> 21 atctgttggc atggggagtg gcatgccaac agat	34	
	<210> 22		
	<211> 34		
65	<212> ADN		
	<213> Secuencia artificial		

<400> 25

ES 2 577 929 T3

atgtccttac	tgaccacaaa	caaccattcg	gttgccattat	cgtaacggtga	accgcctagt	60
acgcttaacg	acagcctgaa	ggacagctac	cagcgttcta	ccgacgaact	gcaggcgctg	120
ctgtctaaac	cgctggcgca	gctgaccgac	gcgacaaaac	tgcgatatccg	tgaaatcacc	180
caggcgaaac	tgaaacactt	cctggacaac	ggtcaccgga	cgcgctcgtgc	gaatacctgg	240
agagcgctca	tgtcccgttg	ggcgaagtcc	gagtcattgt	gcttgacgaa	taacctgaca	300
ccgttaccgg	caacacctga	agttgttgcc	acttttattg	aatattatca	ggcgagctca	360
tatacaacac	tctcgcaata	tgogtgggag	ataaactcgt	ttcatgttga	gtgcgggctg	420
ctaagcccag	ttagcagcaa	aactgttcaa	gacaaacaga	acgagatccg	aatagtaaaag	480
cttgaatccg	gtggtctagc	tcaagaacag	gctacaccgt	tccgaactcca	tcatctccag	540
atgcttattg	aaagctacgg	agaaagcgaa	cgcttgctgg	ataaacgtaa	tctggcatta	600
ttgaacattg	cttatgagtc	cctgttacgt	gaatccgaac	ttctcaggat	caagggttga	660
catctcaaaa	gtactttcga	aggggattac	gtactcagtg	ttccctatac	gaaaacgaac	720
gatagcgggc	aagaagaagt	ggtgaatatt	accccgctgg	gatttaaact	gatccagcga	780
tacattcagg	gcgctgggct	aactaaagaa	gactacctgt	ttcagccgat	cggaagaagt	840
aataaagtct	cggtacaagc	caagccgatg	tcaactcgca	ccgtggatcg	ggtttttcta	900
tgggcttttg	aatcgcttgg	cattgacaga	cactcggcct	ggtctggaca	cagtgcccg	960
atcgagcgcg	cgcaagattt	gctcgccgca	ggatattcaa	tagcgagat	ccaggaaaat	1020
ggtcgctgga	agagtcgaat	gatggtcttg	aggtatggca	aagacatcaa	agctaaagaa	1080
agcgcaatgg	cgaagatgct	ggcagaacgc	cgtag			1116

<210> 26

<211> 1029

5 <212> ADN

<213> Salmonella enterica

<400> 26

atgaccaatt	cggttaaccgt	tcatcaaaaat	ttaccggcgt	tacctgttga	gtctgtcaac	60
gaagaggttc	gtcgaaacct	gaatgcgatg	ttccgtgata	aagaggcgtt	ttccgaacat	120
acctggaaaa	tgctgatgtc	agtttgccgt	tcattgggctg	aatgggtgtc	ggttaacggt	180
tatacctggt	tcccggcaca	cccgaaaaat	gtgcgtgact	atttactcgg	attgcaagcg	240
cgcggaactg	cggttaagtc	gatagagcaa	catattgtct	aactgaacat	gctgcataaa	300
cgatctggct	tgccacgccc	aagcgacagt	atggccgtct	cccttggtat	gcgcccggata	360
cgaaaggaaa	acgtggatgc	cgccgagcgg	gctaaacagg	cgctagcatt	tgagcgacgg	420
gatttttcaa	cggtccgctc	tctcctcata	gaaagtgagc	gtatcattga	tattcgcaac	480
ctggcattcc	tggggctggc	atataacacc	ctgttacgca	tatctgaaat	taccgcac	540
cgcgttaaaag	atatttcgcg	taccgaaggt	ggcggaatgc	tgatccgcat	tggacggaca	600
aaaagcgttg	tcagcaccgc	cgccatagag	aaagcgctaa	gccttgatat	aacgtatctg	660
gtagaccgat	ggatcacggc	ctccggcgta	ggagaagacc	cggaataa	cctgttctgc	720
cgcgtacgta	aaaatggcgt	agcagtccca	tcatccacca	gcaaactgac	aacgcgatcg	780
ctggaacata	tcttcgaggc	agcacactac	aaagtttatg	gtgcaaaaaga	tgcattccggc	840

gaacgctatc	tggcgtggtc	aggacacagt	gcccggtgtag	gcgctgcgcg	tgacatggcc	900
cgcgcgggtg	ttggcgtaac	ggaaattatg	caagctgggtg	gctggactaa	cgtcgaaata	960
gttatgagct	atatccgtaa	tctggacagt	gaaacaggcg	ctatggtccg	ccttctggag	1020
gaaacctaa						1029

10

<210> 27

<211> 34

<212> ADN

15 <213> Secuencia artificial

<220>

<223> Subsitio 1b+2

20

<400> 27

aacacactgc ttaagcctca ataagcagtg tggt 34

<210> 28

<211> 34

25 <212> ADN

<213> Secuencia artificial

<220>

<223> Subsitio 1b+1

30

<400> 28

	atcacactgc ttaagcctca ataagcagtg tgat	34
	<210> 29	
	<211> 34	
5	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Subsitio 1b	
10	<400> 29	
	ataacactgc ttaagcctca ataagcagtg ttat	34
	<210> 30	
15	<211> 34	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
20	<223> Subsitio 1a	
	<400> 30	
	aaccattgt ataagcctca atatacaatg gggt	34
25	<210> 31	
	<211> 34	
	<212> ADN	
	<213> Secuencia artificial	
30	<220>	
	<223> Subsitio 2a+2	
	<400> 31	
35	aaggcaaggt ttaagcctca ataaaccttg cctt	34
	<210> 32	
	<211> 34	
	<212> ADN	
	<213> Secuencia artificial	
40	<220>	
	<223> Subsitio 2a+1	
	<400> 32	
45	aaggctaggt ttaagcctca ataaacctag cctt	34
	<210> 33	
	<211> 34	
	<212> ADN	
50	<213> Secuencia artificial	
	<220>	
	<223> Subsitio 2a	
55	<400> 33	
	aaggctaggt ataagcctca atatacctag cctt	34
	<210> 34	
	<211> 34	
60	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> Subsitio 2b	
65	<400> 34	

ataacaagct ttaagcctca ataaagcttg ttat 34

<210> 35
 <211> 34
 5 <212> ADN
 <213> Secuencia artificial

<220>
 <223> delta loxP

10 <400> 35
 ataacttcgt ttaatgtatg ctaaacgaag ttat 34

<210> 36
 15 <211> 343
 <212> PRT
 <213> Bacteriófago P1

<400> 36

20 Met Ser Asn Leu Leu Thr Val His Gln Asn Leu Pro Ala Leu Pro Val
 1 5 10 15
 Asp Ala Thr Ser Asp Glu Val Arg Lys Asn Leu Met Asp Met Phe Arg
 20 25 30
 Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Met Leu Leu Ser Val
 35 40 45
 Cys Arg Ser Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
 50 55 60
 Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu Tyr Leu Gln Ala
 65 70 75 80
 Arg Gly Leu Ala Val Lys Thr Ile Gln Gln His Leu Gly Gln Leu Asn
 85 90 95
 Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Ser Asp Ser Asn Ala
 100 105 110
 Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
 115 120 125
 Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
 130 135 140
 Val Arg Ser Leu Met Glu Asn Ser Asp Arg Cys Gln Asp Ile Arg Asn
 145 150 155 160
 Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ala Glu
 165 170 175
 Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg
 180 185 190
 Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
 195 200 205
 Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
 210 215 220
 Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
 225 230 235 240
 Arg Val Arg Lys Asn Gly Val Ala Ala Pro Ser Ala Thr Ser Gln Leu
 245 250 255
 Ser Thr Arg Ala Leu Glu Gly Ile Phe Glu Ala Thr His Arg Leu Ile
 260 265 270
 Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
 275 280 285
 His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
 290 295 300
 Ser Ile Pro Glu Ile Met Gln Ala Gly Gly Trp Thr Asn Val Asn Ile
 305 310 315 320
 Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
 325 330 335
 Arg Leu Leu Glu Asp Gly Asp
 340

<210> 37
 <211> 343
 25 <212> PRT

<213> Secuencia artificial

<220>

<223> secuencia común de consenso de Tre al100%

5

<220>

<221> misc_feature

<222> (3)..(5)

<223> Xaa puede ser cualquier aminoácido natural

10

<220>

<221> misc_feature

<222> (7)..(10)

<223> Xaa puede ser cualquier aminoácido natural

15

<220>

<221> misc_feature

<222> (12)..(12)

<223> Xaa puede ser cualquier aminoácido natural

20

<220>

<221> misc_feature

<222> (15)..(16)

<223> Xaa puede ser cualquier aminoácido natural

25

<220>

<221> misc_feature

<222> (18)..(19)

<223> Xaa puede ser cualquier aminoácido natural

30

<220>

<221> misc_feature

<222> (22)..(26)

<223> Xaa puede ser cualquier aminoácido natural

35

<220>

<221> misc_feature

<222> (28)..(31)

<223> Xaa puede ser cualquier aminoácido natural

40

<220>

<221> misc_feature

<222> (34)..(35)

<223> Xaa puede ser cualquier aminoácido natural

45

<220>

<221> misc_feature

<222> (37)..(37)

<223> Xaa puede ser cualquier aminoácido natural

50

<220>

<221> misc_feature

<222> (39)..(40)

<223> Xaa puede ser cualquier aminoácido natural

55

<220>

<221> misc_feature

<222> (43)..(44)

<223> Xaa puede ser cualquier aminoácido natural

60

<220>

<221> misc_feature

<222> (48)..(48)

<223> Xaa puede ser cualquier aminoácido natural

65

<220>

	<221> misc_feature
	<222> (51)..(51)
	<223> Xaa puede ser cualquier aminoácido natural
5	<220>
	<221> misc_feature
	<222> (53)..(53)
	<223> Xaa puede ser cualquier aminoácido natural
10	<220>
	<221> misc_feature
	<222> (57)..(60)
	<223> Xaa puede ser cualquier aminoácido natural
15	<220>
	<221> misc_feature
	<222> (62)..(63)
	<223> Xaa puede ser cualquier aminoácido natural
20	<220>
	<221> misc_feature
	<222> (66)..(67)
	<223> Xaa puede ser cualquier aminoácido natural
25	<220>
	<221> misc_feature
	<222> (69)..(70)
	<223> Xaa puede ser cualquier aminoácido natural
30	<220>
	<221> misc_feature
	<222> (73)..(73)
	<223> Xaa puede ser cualquier aminoácido natural
35	<220>
	<221> misc_feature
	<222> (77)..(77)
	<223> Xaa puede ser cualquier aminoácido natural
40	<220>
	<221> misc_feature
	<222> (80)..(80)
	<223> Xaa puede ser cualquier aminoácido natural
45	<220>
	<221> misc_feature
	<222> (84)..(84)
	<223> Xaa puede ser cualquier aminoácido natural
50	<220>
	<221> misc_feature
	<222> (86)..(86)
	<223> Xaa puede ser cualquier aminoácido natural
55	<220>
	<221> misc_feature
	<222> (88)..(88)
	<223> Xaa puede ser cualquier aminoácido natural
60	<220>
	<221> misc_feature
	<222> (93)..(94)
	<223> Xaa puede ser cualquier aminoácido natural
65	<220>
	<221> misc_feature

<222> (98)..(98)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 5 <221> misc_feature
 <222> (102)..(102)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 10 <221> misc_feature
 <222> (105)..(105)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 15 <221> misc_feature
 <222> (107)..(108)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 20 <221> misc_feature
 <222> (110)..(112)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 25 <221> misc_feature
 <222> (116)..(117)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 30 <221> misc_feature
 <222> (122)..(122)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 35 <221> misc_feature
 <222> (131)..(131)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 40 <221> misc_feature
 <222> (138)..(138)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 45 <221> misc_feature
 <222> (140)..(140)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 50 <221> misc_feature
 <222> (142)..(147)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 55 <221> misc_feature
 <222> (149)..(151)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 60 <221> misc_feature
 <222> (153)..(156)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 65 <221> misc_feature
 <222> (158)..(158)

<223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 5 <222> (160)..(160)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 10 <222> (163)..(163)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 15 <222> (166)..(166)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 20 <222> (174)..(175)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 25 <222> (177)..(180)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 30 <222> (182)..(183)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 35 <222> (185)..(185)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 40 <222> (187)..(187)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 45 <222> (189)..(189)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 50 <222> (193)..(193)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 55 <222> (197)..(199)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 60 <222> (207)..(207)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 65 <222> (216)..(217)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (219)..(219)
 5 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (221)..(221)
 10 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (225)..(225)
 15 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (227)..(227)
 20 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (232)..(236)
 25 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (241)..(245)
 30 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (247)..(247)
 35 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (249)..(249)
 40 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (251)..(251)
 45 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (253)..(255)
 50 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (258)..(260)
 55 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (262)..(263)
 60 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (266)..(268)
 65 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (270)..(273)
 <223> Xaa puede ser cualquier aminoácido natural
 5
 <220>
 <221> misc_feature
 <222> (277)..(278)
 <223> Xaa puede ser cualquier aminoácido natural
 10
 <220>
 <221> misc_feature
 <222> (281)..(281)
 <223> Xaa puede ser cualquier aminoácido natural
 15
 <220>
 <221> misc_feature
 <222> (284)..(285)
 <223> Xaa puede ser cualquier aminoácido natural
 20
 <220>
 <221> misc_feature
 <222> (305)..(305)
 <223> Xaa puede ser cualquier aminoácido natural
 25
 <220>
 <221> misc_feature
 <222> (307)..(307)
 <223> Xaa puede ser cualquier aminoácido natural
 30
 <220>
 <221> misc_feature
 <222> (316)..(316)
 <223> Xaa puede ser cualquier aminoácido natural
 35
 <220>
 <221> misc_feature
 <222> (319)..(320)
 <223> Xaa puede ser cualquier aminoácido natural
 40
 <220>
 <221> misc_feature
 <222> (332)..(332)
 <223> Xaa puede ser cualquier aminoácido natural
 45
 <220>
 <221> misc_feature
 <222> (341)..(343)
 <223> Xaa puede ser cualquier aminoácido natural
 50
 <400> 37

Met Ser Xaa Xaa Xaa Thr Xaa Xaa Xaa Xaa Leu Xaa Ala Leu Xaa Xaa
1 5 10 15
Asp Xaa Xaa Ser Asp Xaa Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Arg
20 25 30
Asp Xaa Xaa Ala Xaa Ser Xaa Xaa Thr Trp Xaa Xaa Leu Leu Ser Xaa
35 40 45
Cys Arg Xaa Trp Xaa Ala Trp Cys Xaa Xaa Xaa Xaa Arg Xaa Xaa Phe
50 55 60
Pro Xaa Xaa Pro Xaa Xaa Val Arg Xaa Tyr Leu Xaa Leu Gln Xaa
65 70 75 80
Arg Gly Leu Xaa Val Xaa Thr Xaa Gln Gln His Leu Xaa Xaa Leu Asn
85 90 95
Met Xaa His Arg Arg Xaa Gly Leu Xaa Arg Xaa Xaa Asp Xaa Xaa Xaa
100 105 110
Val Ser Leu Xaa Xaa Arg Arg Ile Arg Xaa Glu Asn Val Asp Ala Gly
115 120 125
Glu Arg Xaa Lys Gln Ala Leu Ala Phe Xaa Arg Xaa Asp Xaa Xaa Xaa
130 135 140
Xaa Xaa Xaa Leu Xaa Xaa Xaa Ser Xaa Xaa Xaa Xaa Asp Xaa Arg Xaa
145 150 155 160
Leu Ala Xaa Leu Gly Xaa Ala Tyr Asn Thr Leu Leu Arg Xaa Xaa Glu
165 170 175
Xaa Xaa Xaa Xaa Arg Xaa Xaa Asp Xaa Ser Xaa Thr Xaa Gly Gly Arg
180 185 190
Xaa Leu Ile His Xaa Xaa Xaa Thr Lys Thr Leu Val Ser Thr Xaa Gly
195 200 205
Val Glu Lys Ala Leu Ser Leu Xaa Xaa Thr Xaa Leu Xaa Glu Arg Trp
210 215 220
Xaa Ser Xaa Ser Gly Val Ala Xaa Xaa Xaa Xaa Xaa Tyr Leu Phe Cys
225 230 235 240
Xaa Xaa Xaa Xaa Xaa Gly Xaa Ala Xaa Pro Xaa Ala Xaa Xaa Xaa Leu
245 250 255
Ser Xaa Xaa Xaa Leu Xaa Xaa Ile Phe Xaa Xaa Xaa His Xaa Xaa Xaa
260 265 270
Xaa Gly Ala Lys Xaa Xaa Ser Gly Xaa Arg Tyr Xaa Xaa Trp Ser Gly
275 280 285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290 295 300
Xaa Ile Xaa Glu Ile Met Gln Ala Gly Gly Trp Xaa Thr Val Xaa Xaa
305 310 315 320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Xaa Gly Ala Met Val
325 330 335
Arg Leu Leu Glu Xaa Xaa Xaa
340

<210> 38
<211> 343
5 <212> PRT
<213> Secuencia artificial

<220>
10 <223> secuencia de consenso de Tre 3.0 al 100%

<220>
<221> misc_feature
<222> (3)..(5)
15 <223> Xaa puede ser cualquier aminoácido natural

<220>
<221> misc_feature
<222> (7)..(10)
20 <223> Xaa puede ser cualquier aminoácido natural

<220>
<221> misc_feature
<222> (16)..(16)
25 <223> Xaa puede ser cualquier aminoácido natural

<220>

	<221> misc_feature
	<222> (18)..(19)
	<223> Xaa puede ser cualquier aminoácido natural
5	<220>
	<221> misc_feature
	<222> (22)..(26)
	<223> Xaa puede ser cualquier aminoácido natural
10	<220>
	<221> misc_feature
	<222> (28)..(31)
	<223> Xaa puede ser cualquier aminoácido natural
15	<220>
	<221> misc_feature
	<222> (34)..(35)
	<223> Xaa puede ser cualquier aminoácido natural
20	<220>
	<221> misc_feature
	<222> (37)..(37)
	<223> Xaa puede ser cualquier aminoácido natural
25	<220>
	<221> misc_feature
	<222> (39)..(40)
	<223> Xaa puede ser cualquier aminoácido natural
30	<220>
	<221> misc_feature
	<222> (43)..(43)
	<223> Xaa puede ser cualquier aminoácido natural
35	<220>
	<221> misc_feature
	<222> (48)..(48)
	<223> Xaa puede ser cualquier aminoácido natural
40	<220>
	<221> misc_feature
	<222> (51)..(51)
	<223> Xaa puede ser cualquier aminoácido natural
45	<220>
	<221> misc_feature
	<222> (53)..(53)
	<223> Xaa puede ser cualquier aminoácido natural
50	<220>
	<221> misc_feature
	<222> (57)..(60)
	<223> Xaa puede ser cualquier aminoácido natural
55	<220>
	<221> misc_feature
	<222> (62)..(63)
	<223> Xaa puede ser cualquier aminoácido natural
60	<220>
	<221> misc_feature
	<222> (66)..(67)
	<223> Xaa puede ser cualquier aminoácido natural
65	<220>
	<221> misc_feature

- <222> (69)..(70)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
5 <221> misc_feature
<222> (73)..(73)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
10 <221> misc_feature
<222> (77)..(77)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
15 <221> misc_feature
<222> (80)..(80)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
20 <221> misc_feature
<222> (84)..(84)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
25 <221> misc_feature
<222> (88)..(88)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
30 <221> misc_feature
<222> (94)..(94)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
35 <221> misc_feature
<222> (98)..(98)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
40 <221> misc_feature
<222> (102)..(102)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
45 <221> misc_feature
<222> (105)..(105)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
50 <221> misc_feature
<222> (107)..(108)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
55 <221> misc_feature
<222> (111)..(112)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
60 <221> misc_feature
<222> (116)..(117)
<223> Xaa puede ser cualquier aminoácido natural
- <220>
65 <221> misc_feature
<222> (122)..(122)

<223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 5 <222> (131)..(131)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 10 <222> (138)..(138)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 15 <222> (140)..(140)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 20 <222> (142)..(147)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 25 <222> (149)..(151)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 30 <222> (153)..(156)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 35 <222> (158)..(158)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 40 <222> (160)..(160)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 45 <222> (163)..(163)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 50 <222> (166)..(166)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 55 <222> (174)..(174)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 60 <222> (177)..(180)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 65 <222> (182)..(183)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (185)..(185)
 5 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (187)..(187)
 10 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (189)..(189)
 15 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (193)..(193)
 20 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (197)..(199)
 25 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (207)..(207)
 30 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (216)..(217)
 35 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (219)..(219)
 40 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (221)..(221)
 45 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (225)..(225)
 50 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (227)..(227)
 55 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (232)..(236)
 60 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (241)..(245)
 65 <223> Xaa puede ser cualquier aminoácido natural

5 <220>
 <221> misc_feature
 <222> (247)..(247)
 <223> Xaa puede ser cualquier aminoácido natural
 10 <220>
 <221> misc_feature
 <222> (249)..(249)
 <223> Xaa puede ser cualquier aminoácido natural
 15 <220>
 <221> misc_feature
 <222> (251)..(251)
 <223> Xaa puede ser cualquier aminoácido natural
 20 <220>
 <221> misc_feature
 <222> (253)..(255)
 <223> Xaa puede ser cualquier aminoácido natural
 25 <220>
 <221> misc_feature
 <222> (258)..(260)
 <223> Xaa puede ser cualquier aminoácido natural
 30 <220>
 <221> misc_feature
 <222> (262)..(263)
 <223> Xaa puede ser cualquier aminoácido natural
 35 <220>
 <221> misc_feature
 <222> (266)..(268)
 <223> Xaa puede ser cualquier aminoácido natural
 40 <220>
 <221> misc_feature
 <222> (270)..(273)
 <223> Xaa puede ser cualquier aminoácido natural
 45 <220>
 <221> misc_feature
 <222> (277)..(278)
 <223> Xaa puede ser cualquier aminoácido natural
 50 <220>
 <221> misc_feature
 <222> (281)..(281)
 <223> Xaa puede ser cualquier aminoácido natural
 55 <220>
 <221> misc_feature
 <222> (284)..(285)
 <223> Xaa puede ser cualquier aminoácido natural
 60 <220>
 <221> misc_feature
 <222> (305)..(305)
 <223> Xaa puede ser cualquier aminoácido natural
 65 <220>
 <221> misc_feature
 <222> (316)..(316)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>

<221> misc_feature
 <222> (319)..(320)
 <223> Xaa puede ser cualquier aminoácido natural

5

<220>
 <221> misc_feature
 <222> (332)..(332)
 <223> Xaa puede ser cualquier aminoácido natural

10

<220>
 <221> misc_feature
 <222> (341)..(343)
 <223> Xaa puede ser cualquier aminoácido natural

15

<400> 38
 Met Ser Xaa Xaa Xaa Thr Xaa Xaa Xaa Xaa Leu Ser Ala Leu Leu Xaa
 1 5 10 15
 Asp Xaa Xaa Ser Asp Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Arg
 20 25 30
 Asp Xaa Xaa Ala Xaa Ser Xaa Xaa Thr Trp Xaa Val Leu Leu Ser Xaa
 35 40 45
 Cys Arg Xaa Trp Xaa Ala Trp Cys Xaa Xaa Xaa Xaa Arg Xaa Xaa Phe
 50 55 60
 Pro Xaa Xaa Pro Xaa Xaa Val Arg Xaa Tyr Leu Xaa Leu Gln Xaa
 65 70 75 80
 Arg Gly Leu Xaa Val Asn Thr Xaa Gln Gln His Leu Ala Xaa Leu Asn
 85 90 95
 Met Xaa His Arg Arg Xaa Gly Leu Xaa Arg Xaa Xaa Asp Ser Xaa Xaa
 100 105 110
 Val Ser Leu Xaa Xaa Arg Arg Ile Arg Xaa Glu Asn Val Asp Ala Gly
 115 120 125
 Glu Arg Xaa Lys Gln Ala Leu Ala Phe Xaa Arg Xaa Asp Xaa Xaa Xaa
 130 135 140
 Xaa Xaa Xaa Leu Xaa Xaa Ser Xaa Xaa Xaa Xaa Asp Xaa Arg Xaa
 145 150 155 160
 Leu Ala Xaa Leu Gly Xaa Ala Tyr Asn Thr Leu Leu Arg Xaa Ser Glu
 165 170 175
 Xaa Xaa Xaa Xaa Arg Xaa Xaa Asp Xaa Ser Xaa Thr Xaa Gly Gly Arg
 180 185 190
 Xaa Leu Ile His Xaa Xaa Xaa Thr Lys Thr Leu Val Ser Thr Xaa Gly
 195 200 205
 Val Glu Lys Ala Leu Ser Leu Xaa Xaa Thr Xaa Leu Xaa Glu Arg Trp
 210 215 220
 Xaa Ser Xaa Ser Gly Val Ala Xaa Xaa Xaa Xaa Tyr Leu Phe Cys
 225 230 235 240
 Xaa Xaa Xaa Xaa Xaa Gly Xaa Ala Xaa Pro Xaa Ala Xaa Xaa Xaa Leu
 245 250 255
 Ser Xaa Xaa Xaa Leu Xaa Xaa Ile Phe Xaa Xaa Xaa His Xaa Xaa Xaa
 260 265 270
 Xaa Gly Ala Lys Xaa Xaa Ser Gly Xaa Arg Tyr Xaa Xaa Trp Ser Gly
 275 280 285
 His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
 290 295 300
 Xaa Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Xaa Thr Val Xaa Xaa
 305 310 315 320
 Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Xaa Gly Ala Met Val
 325 330 335
 Arg Leu Leu Glu Xaa Xaa Xaa
 340

<210> 39
 <211> 343
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> secuencia de consenso de Tre 3.0 al 95%

25

<220>

	<221> misc_feature
	<222> (3)..(3)
	<223> Xaa puede ser cualquier aminoácido natural
5	<220>
	<221> misc_feature
	<222> (5)..(5)
	<223> Xaa puede ser cualquier aminoácido natural
10	<220>
	<221> misc_feature
	<222> (7)..(71)
	<223> Xaa puede ser cualquier aminoácido natural
15	<220>
	<221> misc_feature
	<222> (9)..(10)
	<223> Xaa puede ser cualquier aminoácido natural
20	<220>
	<221> misc_feature
	<222> (16)..(16)
	<223> Xaa puede ser cualquier aminoácido natural
25	<220>
	<221> misc_feature
	<222> (18)..(18)
	<223> Xaa puede ser cualquier aminoácido natural
30	<220>
	<221> misc_feature
	<222> (23)..(23)
	<223> Xaa puede ser cualquier aminoácido natural
35	<220>
	<221> misc_feature
	<222> (25)..(25)
	<223> Xaa puede ser cualquier aminoácido natural
40	<220>
	<221> misc_feature
	<222> (31)..(31)
	<223> Xaa puede ser cualquier aminoácido natural
45	<220>
	<221> misc_feature
	<222> (34)..(35)
	<223> Xaa puede ser cualquier aminoácido natural
50	<220>
	<221> misc_feature
	<222> (37)..(37)
	<223> Xaa puede ser cualquier aminoácido natural
55	<220>
	<221> misc_feature
	<222> (51)..(51)
	<223> Xaa puede ser cualquier aminoácido natural
60	<220>
	<221> misc_feature
	<222> (57)..(57)
	<223> Xaa puede ser cualquier aminoácido natural
65	<220>
	<221> misc_feature

<222> (63)..(63)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 5 <221> misc_feature
 <222> (77)..(77)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 10 <221> misc_feature
 <222> (80)..(80)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 15 <221> misc_feature
 <222> (94)..(94)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 20 <221> misc_feature
 <222> (102)..(102)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 25 <221> misc_feature
 <222> (107)..(108)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 30 <221> misc_feature
 <222> (111)..(111)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 35 <221> misc_feature
 <222> (117)..(117)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 40 <221> misc_feature
 <222> (122)..(122)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 45 <221> misc_feature
 <222> (131)..(131)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 50 <221> misc_feature
 <222> (142)..(143)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 55 <221> misc_feature
 <222> (147)..(147)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 60 <221> misc_feature
 <222> (150)..(151)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 65 <221> misc_feature
 <222> (153)..(153)

<223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 5 <222> (155)..(155)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 10 <222> (160)..(160)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 15 <222> (163)..(163)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 20 <222> (166)..(166)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 25 <222> (182)..(183)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 30 <222> (187)..(187)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 35 <222> (189)..(189)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 40 <222> (198)..(198)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 45 <222> (216)..(216)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 50 <222> (219)..(219)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 55 <222> (227)..(227)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 60 <222> (232)..(232)
 <223> Xaa puede ser cualquier aminoácido natural
 <220>
 <221> misc_feature
 65 <222> (241)..(241)
 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (244)..(245)
 5 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (249)..(249)
 10 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (254)..(254)
 15 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (266)..(266)
 20 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (268)..(268)
 25 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (270)..(270)
 30 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (277)..(278)
 35 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (281)..(281)
 40 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (305)..(305)
 45 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (319)..(319)
 50 <223> Xaa puede ser cualquier aminoácido natural

<220>
 <221> misc_feature
 <222> (341)..(342)
 55 <223> Xaa puede ser cualquier aminoácido natural

<400> 39

ES 2 577 929 T3

```

Met Ser Xaa Leu Xaa Thr Xaa His Xaa Xaa Leu Ser Ala Leu Leu Xaa
1      5      10      15
Asp Xaa Thr Ser Asp Glu Xaa Arg Xaa Asn Leu Met Asp Val Xaa Arg
20      25      30
Asp Xaa Xaa Ala Xaa Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35      40      45
Cys Arg Xaa Trp Ala Ala Trp Cys Xaa Leu Asn Asn Arg Lys Xaa Phe
50      55      60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu Xaa Leu Gln Xaa
65      70      75      80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Xaa Leu Asn
85      90      95
Met Leu His Arg Arg Xaa Gly Leu Pro Arg Xaa Xaa Asp Ser Xaa Ala
100      105      110
Val Ser Leu Val Xaa Arg Arg Ile Arg Xaa Glu Asn Val Asp Ala Gly
115      120      125
Glu Arg Xaa Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Xaa Xaa Gln
130      135      140
Val Arg Xaa Leu Met Xaa Xaa Ser Xaa Arg Xaa Gln Asp Ile Arg Xaa
145      150      155      160
Leu Ala Xaa Leu Gly Xaa Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165      170      175
Ile Ala Arg Ile Arg Xaa Xaa Asp Ile Ser Xaa Thr Xaa Gly Gly Arg
180      185      190
Met Leu Ile His Ile Xaa Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195      200      205
Val Glu Lys Ala Leu Ser Leu Xaa Val Thr Xaa Leu Val Glu Arg Trp
210      215      220
Ile Ser Xaa Ser Gly Val Ala Xaa Asp Pro Asn Asn Tyr Leu Phe Cys
225      230      235      240
Xaa Val Arg Xaa Xaa Gly Val Ala Xaa Pro Ser Ala Thr Xaa Gln Leu
245      250      255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Xaa Ala Xaa His Xaa Leu Ile
260      265      270
Tyr Gly Ala Lys Xaa Xaa Ser Gly Xaa Arg Tyr Leu Ala Trp Ser Gly
275      280      285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290      295      300
Xaa Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Xaa Ser
305      310      315      320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325      330      335
Arg Leu Leu Glu Xaa Xaa Asp
340

```

<210> 40
 5 <211> 343
 <212> PRT
 <213> Secuencia artificial

<220>
 10 <223> Tre3-5

<400> 40

Met Ser Glu Leu Gln Thr Leu His Arg Asn Leu Ser Ala Leu Leu Val
1 5 10 15
Asp Ala Thr Ser Asp Glu Thr Arg Lys Asn Leu Met Asp Val Phe Arg
20 25 30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Ile
35 40 45
Cys Arg Ser Trp Ala Ala Trp Cys Lys Leu Asn Asp Arg Lys Trp Phe
50 55 60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Thr
65 70 75 80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85 90 95
Met Leu His Arg Arg Ser Gly Leu Pro Arg Leu Ser Asp Ser Asp Ala
100 105 110
Val Ser Leu Val Met Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
115 120 125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Leu Asp Gln
130 135 140
Val Arg Ser Leu Thr Glu Asp Ser Asp Arg Cys Arg Asp Ile Arg Asn
145 150 155 160
Leu Ala Leu Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165 170 175
Ile Ala Arg Val Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
180 185 190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195 200 205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210 215 220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225 230 235 240
Pro Val Arg Lys Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245 250 255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
260 265 270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275 280 285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290 295 300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
305 310 315 320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325 330 335
Arg Leu Leu Glu Asp Gly Asp
340

<210> 41
5 <211> 343
<212> PRT
<213> Secuencia artificial

<220>
10 <223> Tre 3-18

<400> 41

ES 2 577 929 T3

```

Met Ser Asp Leu Gln Thr Leu His Arg Asn Leu Ser Ala Leu Leu Val
1      5      10      15
Asp Ala Thr Ser Asp Glu Ala Arg Lys Asn Leu Met Asp Val Leu Arg
20      25      30
Asp Arg Gln Ala Tyr Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35      40      45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
50      55      60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Thr
65      70      75      80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85      90      95
Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Gly Asp Ser Asp Ala
100      105      110
Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
115      120      125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Ala Asp Phe Asp Gln
130      135      140
Val Arg Ala Leu Ile Glu Asn Ser Asn Arg Cys Gln Asp Ile Arg Thr
145      150      155      160
Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165      170      175
Val Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
180      185      190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195      200      205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210      215      220
Ile Ser Val Ser Gly Val Ala Gly Asp Pro Asn Asn Tyr Leu Phe Cys
225      230      235      240
Gln Val Arg Arg Ser Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245      250      255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
260      265      270
Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275      280      285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290      295      300
Pro Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Glu Ser
305      310      315      320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325      330      335
Arg Leu Leu Glu Asp Ser Asp
340

```

<210> 42
 <211> 343
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Tre3-23

10 <400> 42

ES 2 577 929 T3

```

Met Ser Lys Leu Leu Thr Val His Gln Asn Leu Ser Ala Leu Leu Val
1      5      10      15
Asp Ala Thr Ser Asp Glu Thr Arg Lys Asn Leu Met Asp Val Leu Arg
20      25      30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35      40      45
Cys Arg Ser Trp Ala Ala Trp Cys Asp Leu Asn Asn Arg Lys Arg Phe
50      55      60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu His Leu Gln Val
65      70      75      80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85      90      95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asn Ala
100      105      110
Val Ser Leu Val Met Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
115      120      125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Leu Asp Gln
130      135      140
Val His Ser Leu Met Gly Asn Ser Asp Arg Gly Gln Asp Ile Arg Asn
145      150      155      160
Leu Ala Leu Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165      170      175
Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg
180      185      190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195      200      205
Val Glu Lys Ala Leu Ser Leu Arg Val Thr Lys Leu Val Glu Arg Trp
210      215      220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225      230      235      240
Pro Val Arg Ile Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245      250      255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
260      265      270
Tyr Gly Ala Lys Gly Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275      280      285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290      295      300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Glu Ser
305      310      315      320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325      330      335
Arg Leu Leu Glu Asp Gly Asp
340

```

<210> 43

<211> 343

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Tre3-27

10

<400> 43

Met Ser Glu Leu Leu Thr Leu His Gln Asn Leu Ser Ala Leu Leu Val
1 5 10 15
Asp Ala Thr Ser Asp Glu Ala Arg Lys Asn Leu Met Asp Val Leu Arg
20 25 30
Asp Arg Gln Ala Phe Ser Lys His Thr Trp Lys Val Leu Leu Ser Val
35 40 45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
50 55 60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
65 70 75 80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85 90 95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asp Ala
100 105 110
Val Ser Leu Val Ile Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
115 120 125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
130 135 140
Val Arg Ala Leu Met Glu Asn Ser Asp Arg Gly Gln Asp Ile Arg Ser
145 150 155 160
Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165 170 175
Ile Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
180 185 190
Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195 200 205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210 215 220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225 230 235 240
Gln Val Arg Lys Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245 250 255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
260 265 270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Gln Ala Trp Ser Gly
275 280 285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290 295 300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
305 310 315 320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325 330 335
Arg Leu Leu Glu Asp Gly Asp
340

<210> 44
<211> 343
5 <212> PRT
<213> Secuencia artificial

<220>
<223> Tre3-31

10 <400> 44

Met Ser Asp Leu Gln Thr Leu His Arg Asn Leu Ser Ala Leu Leu Val
1 5 10 15
Asp Ala Ser Ser Asp Glu Ala Arg Lys Asn Leu Met Asp Val Leu Arg
20 25 30
Asp Arg Gln Ala Tyr Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35 40 45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asp Asn Arg Lys Trp Phe
50 55 60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Ala
65 70 75 80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85 90 95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asn Ala
100 105 110
Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
115 120 125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
130 135 140
Val Arg Ala Leu Met Glu Asn Ser Asp Arg Gly Gln Asp Ile Arg Ser
145 150 155 160
Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165 170 175
Ile Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
180 185 190
Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195 200 205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Glu Leu Val Glu Arg Trp
210 215 220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225 230 235 240
Pro Val Arg Ile Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245 250 255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Ala His Arg Leu Ile
260 265 270
His Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275 280 285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290 295 300
Pro Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Asn
305 310 315 320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325 330 335
Arg Leu Leu Glu Asp Asp
340

<210> 45

<211> 343

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Tre3-37

10

<400> 45

Met Ser Glu Leu Leu Thr Val His Gln Asn Leu Ser Ala Leu Leu Val
1 5 10 15
Asp Ala Thr Ser Asp Glu Thr Cys Lys Asn Leu Met Asp Val Phe Arg

```

      20      25      30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Arg Val Leu Leu Ser Val
      35      40      45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
      50      55      60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
      65      70      75      80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
      85      90      95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asp Ala
      100      105      110
Val Ser Leu Val Ile Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
      115      120      125
Glu Arg Thr Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
      130      135      140
Val Arg Ala Leu Met Glu Asn Ser Asp Arg Gly Gln Asp Ile Arg Thr
      145      150      155      160
Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
      165      170      175
Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg
      180      185      190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
      195      200      205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
      210      215      220
Ile Ser Val Ser Gly Val Ala Asn Asp Pro Asn Asn Tyr Leu Phe Cys
      225      230      235      240
Gln Val Arg Lys Ser Gly Val Ala Val Pro Ser Ala Thr Ser Arg Leu
      245      250      255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Gly Ala Ala His Cys Leu Ile
      260      265      270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
      275      280      285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
      290      295      300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
      305      310      315      320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
      325      330      335
Arg Leu Leu Glu Asp Gly Asp
      340

```

<210> 46
 <211> 343
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Tre3-43

10

```

<400> 46
Met Ser Glu Leu Gln Thr Leu His Gln Asn Leu Ser Ala Leu Leu Val
1      5      10      15
Asp Ala Thr Ser Asp Glu Thr Arg Lys Asn Leu Met Asp Val Leu Arg
      20      25      30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
      35      40      45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Ser Asn Asn Arg Lys Trp Phe

```

ES 2 577 929 T3

```

50          55          60
Pro Ala Asp Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
65          70          75          80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85          90          95
Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Gly Asp Ser Asn Ala
100          105          110
Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
115          120          125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
130          135          140
Ala Arg Ala Leu Met Glu Asn Ser Asp Arg Gly Gln Asp Ile Arg Ser
145          150          155          160
Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165          170          175
Ile Ala Arg Ile Arg Val Lys Asp Ile Ser His Thr Asp Gly Gly Arg
180          185          190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Thr Gly
195          200          205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210          215          220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225          230          235          240
Gln Val Arg Arg Asn Gly Val Ala Val Pro Ser Ala Thr Cys Gln Leu
245          250          255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Gly Ala Ala His Cys Leu Ile
260          265          270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275          280          285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290          295          300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
305          310          315          320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325          330          335
Arg Leu Leu Glu Asp Gly Asp
340

```

<210> 47

<211> 343

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Tre3-46

10

<400> 47

```

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

```

[illegible]

<210> 48
<211> 343
<212> PRT
<213> Secuencia artificial

<220>
<223> Tre3-48

<400>	48															
Met	Ser	Lys	Leu	Gln	Thr	Val	His	Arg	Asn	Leu	Ser	Ala	Leu	Leu	Val	
1				5					10					15		
Asp	Ala	Thr	Ser	Asp	Glu	Thr	Arg	Lys	Asn	Leu	Met	Asp	Val	Phe	Arg	
			20					25					30			
Asp	Arg	Gln	Ala	Phe	Ser	Glu	His	Thr	Trp	Lys	Val	Leu	Leu	Ser	Val	
		35					40					45				
Cys	Arg	Thr	Trp	Ala	Ala	Trp	Cys	Lys	Leu	Asn	Asn	Arg	Lys	Trp	Phe	
	50					55					60					
Pro	Ala	Glu	Pro	Glu	Asp	Val	Arg	Asp	Tyr	Leu	Leu	His	Leu	Gln	Val	
65				70					75					80		
Arg	Gly	Leu	Ala	Val	Asn	Thr	Ile	Gln	Gln	His	Leu	Ala	Gln	Leu	Asn	
				85					90					95		
Met	Leu	His	Arg	Arg	Ser	Gly	Leu	Pro	Arg	Leu	Ser	Asp	Ser	Asn	Ala	
			100					105					110			
Val	Ser	Leu	Val	Met	Arg	Arg	Ile	Arg	Lys	Glu	Asn	Val	Asp	Ala	Gly	

ES 2 577 929 T3

```

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

```

<210> 49

<211> 343

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Tre3-63

10

<400> 49

```

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

```

```

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

```

<210> 50
 <211> 343
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Tre3-74

```

10 <400> 50
Met Ser Asp Leu Gln Thr Val His Arg Asn Leu Ser Ala Leu Leu Val
1      5      10      15
Asp Ala Thr Ser Asp Glu Thr Arg Arg Asn Leu Met Asp Val Phe Arg
20     25     30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35     40     45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
50     55     60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
65     70     75     80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85     90     95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asn Ala
100    105    110
Val Ser Leu Val Met Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
115    120    125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Arg
130    135    140
Val Arg Ala Leu Met Glu Asn Ser Asp Arg Cys Gln Asp Ile Arg Ser
145    150    155    160
Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165    170    175
Ile Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg

```

```

180      185      190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195      200      205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210      215      220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225      230      235      240
Gln Val Arg Lys Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245      250      255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
260      265      270
Tyr Gly Ala Lys Asp Gly Ser Gly Arg Arg Tyr Leu Ala Trp Ser Gly
275      280      285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290      295      300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Glu Ser
305      310      315      320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325      330      335
Arg Leu Leu Glu Gly Asp Asp
340

```

<210> 51
 <211> 343
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Tre3-78

```

10 <400> 51
Met Ser Asp Leu Gln Thr Val His Arg Asn Leu Ser Ala Leu Leu Ile
1      5      10      15
Asp Ala Thr Ser Asp Glu Thr Arg Lys Asn Leu Met Asp Val Phe Arg
20      25      30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35      40      45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
50      55      60
Pro Thr Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
65      70      75      80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85      90      95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asp Ala
100      105      110
Val Ser Leu Val Ile Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
115      120      125
Glu Arg Thr Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
130      135      140
Val Arg Ala Leu Met Glu Asn Ser Asp Arg Gly Gln Asp Ile Arg Asn
145      150      155      160
Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165      170      175
Ile Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
180      185      190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195      200      205
Val Glu Lys Ala Leu Ser Leu Gly Ile Thr Lys Leu Val Glu Arg Trp

```

210		215		220
Ile Ser Val Ser Gly Val	Ala Asp Asp Pro	Asn Asn Tyr Leu Phe Cys		
225	230	235	240	
Pro Val Arg Ile Asn Gly Val	Ala Val Pro Ser	Ala Thr Ser Gln Leu		
	245	250	255	
Ser Thr Ser Ala Leu Gln Gly	Ile Phe Ala Ala His Arg	Leu Ile		
	260	265	270	
Tyr Gly Ala Lys Asp Gly Ser	Gly Gln Arg Tyr Leu	Ala Trp Ser Gly		
	275	280	285	
His Ser Ala Arg Val Gly	Ala Ala Arg Asp Met	Ala Arg Ala Gly Val		
	290	295	300	
Ser Ile Ala Glu Ile Met	Gln Ala Gly Gly Trp	Thr Thr Val Asp Ser		
305	310	315	320	
Val Met Asn Tyr Ile Arg	Asn Leu Asp Ser	Glu Thr Gly Ala Met Val		
	325	330	335	
Arg Leu Leu Glu Asp Gly Gly				
	340			

<210> 52
 <211> 343
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Tre3-89

<400> 52
Met Ser Asp Leu Leu Thr Val His Arg Asn Leu Ser Ala Leu Leu Val
1 5 10 15
Asp Ala Thr Ser Asp Glu Ala Arg Lys Asn Leu Met Asp Val Leu Arg
20 25 30
Asp Arg Gln Ala Tyr Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35 40 45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
50 55 60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
65 70 75 80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85 90 95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asn Ala
100 105 110
Val Ser Leu Val Met Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
115 120 125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Leu Asp Gln
130 135 140
Val Arg Ser Leu Met Glu Asp Ser Asp Arg Cys Gln Asp Ile Arg Ser
145 150 155 160
Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165 170 175
Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg
180 185 190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195 200 205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210 215 220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225 230 235 240
Gln Val Arg Arg Asn Gly Val Ala Ala Pro Ser Ala Thr Ser Gln Leu

245 250 255
 Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
 260 265 270
 Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
 275 280 285
 His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
 290 295 300
 Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Glu Ser
 305 310 315 320
 Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
 325 330 335
 Arg Leu Leu Glu Asp Asp Asp
 340

<210> 53
 <211> 343
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Tre3-94

10 <400> 53
 Met Ser Lys Leu Leu Thr Val His Arg Asn Leu Ser Ala Leu Leu Val
 1 5 10 15
 Asp Ala Thr Ser Asp Glu Thr Arg Lys Asn Leu Met Asp Val Leu Arg
 20 25 30
 Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
 35 40 45
 Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
 50 55 60
 Pro Ala Glu Pro Glu Gly Val Arg Asp Tyr Leu Leu Tyr Leu Gln Ala
 65 70 75 80
 Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
 85 90 95
 Met Leu His Arg Arg Ser Gly Leu Pro Arg Leu Ser Asp Ser Asp Ala
 100 105 110
 Val Ser Leu Val Met Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
 115 120 125
 Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Leu Asp Gln
 130 135 140
 Val Arg Ser Leu Met Gly Ser Ser Glu Arg Cys Gln Asp Ile Arg Ser
 145 150 155 160
 Leu Ala Leu Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
 165 170 175
 Ile Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
 180 185 190
 Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
 195 200 205
 Val Glu Lys Ala Leu Ser Leu Gly Val Thr Glu Leu Val Glu Arg Trp
 210 215 220
 Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
 225 230 235 240
 Pro Val Arg Lys Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
 245 250 255
 Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
 260 265 270
 Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
 275 280 285
 His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
 290 295 300
 Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
 305 310 315 320
 Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
 325 330 335
 Arg Leu Leu Glu Asp Asp Asn
 340

15 <210> 54

<211> 343
 <212> PRT
 <213> Secuencia artificial

5 <220>
 <223> Tre3-96

<400> 54
 Met Ser Lys Leu Gln Thr Val His Arg Asn Leu Ser Ala Leu Leu Val
 1 5 10 15
 Asp Ala Thr Ser Asp Glu Thr Arg Lys Asn Leu Met Asp Val Phe Arg
 20 25 30
 Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
 35 40 45
 Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
 50 55 60
 Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Ala
 65 70 75 80
 Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
 85 90 95
 Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asn Ala
 100 105 110
 Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
 115 120 125
 Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
 130 135 140
 Val Arg Ser Leu Met Glu Asn Ser Asp Arg Gly Gln Asp Ile Arg Thr
 145 150 155 160
 Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
 165 170 175
 Ile Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
 180 185 190
 Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
 195 200 205
 Val Glu Lys Ala Leu Ser Leu Arg Val Thr Lys Leu Val Glu Arg Trp
 210 215 220
 Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
 225 230 235 240
 Pro Val Arg Ile Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
 245 250 255
 Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
 260 265 270
 Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
 275 280 285
 His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
 290 295 300
 Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
 305 310 315 320
 Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
 325 330 335
 Arg Leu Leu Glu Asp Gly Asp
 340

10

<210> 55
 <211> 343
 <212> PRT
 <213> Secuencia artificial

15

<220>
 <223> Tre3-101

20 <400> 55

```

Met Ser Lys Leu Leu Thr Val His Gln Asn Leu Ser Ala Leu Leu Val
1      5      10      15
Asp Ala Thr Ser Asp Glu Ala Arg Lys Asn Leu Ala Asp Val Phe Arg
      20      25      30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
      35      40      45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asp Asn Arg Lys Trp Phe
      50      55      60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Thr
      65      70      75      80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
      85      90      95
Met Leu His Arg Arg Ser Gly Leu Ala Arg Pro Gly Asp Ser Asn Ala
      100      105      110
Val Ser Leu Val Ile Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
      115      120      125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
      130      135      140
Val Arg Ser Leu Met Glu Asn Ser Asp Arg Gly Gln Asp Ile Arg Thr
      145      150      155      160
Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
      165      170      175
Ile Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
      180      185      190
Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
      195      200      205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
      210      215      220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
      225      230      235      240
Pro Val Arg Ile Asn Gly Val Ala Ala Pro Ser Ala Ala Ser Gln Leu
      245      250      255
Ser Thr Tyr Thr Leu Gln Gly Ile Phe Lys Ala Ala His Arg Leu Ile
      260      265      270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
      275      280      285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
      290      295      300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
      305      310      315      320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
      325      330      335
Arg Leu Leu Glu Asp Gly Asp
      340

```

<210> 56
5 <211> 343
<212> PRT
<213> Secuencia artificial

<220>
10 <223> Tre3-103

<400> 56

ES 2 577 929 T3

```

Met Ser Glu Leu Gln Thr Val His His Ser Leu Ser Ala Leu Leu Val
1      5      10      15
Asp Ala Thr Ser Asp Glu Ala Arg Lys Asn Leu Ala Asp Val Phe Arg
20      25      30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35      40      45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
50      55      60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Thr
65      70      75      80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85      90      95
Met Leu His Arg Arg Ser Gly Leu Pro Arg Pro Gly Asp Ser Asn Ala
100      105      110
Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
115      120      125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Glu Gln
130      135      140
Val Arg Ser Leu Met Glu Asn Ser Asp Arg Cys Gln Asp Ile Arg Ser
145      150      155      160
Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165      170      175
Ile Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
180      185      190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Thr Gly
195      200      205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210      215      220
Ile Ser Ala Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225      230      235      240
Pro Val Arg Ile Asn Gly Val Ala Ala Pro Ser Ala Ala Ser Gln Leu
245      250      255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Ala Ala Ala His Arg Leu Ile
260      265      270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275      280      285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290      295      300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Glu Ser
305      310      315      320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325      330      335
Arg Leu Leu Glu Asp Asp Asp
340

```

<210> 57
 <211> 343
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Tre3-106

10 <400> 57

Met Ser Glu Leu Gln Thr Leu His Arg Asn Leu Ser Ala Leu Leu Val
1 5 10 15
Asp Ala Thr Ser Asp Glu Thr Arg Lys Asn Leu Met Asp Val Phe Arg
20 25 30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35 40 45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
50 55 60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
65 70 75 80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85 90 95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asn Ala
100 105 110
Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
115 120 125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
130 135 140
Val Arg Ala Leu Met Glu Asn Ser Asp Arg Gly Gln Asp Ile Arg Ser
145 150 155 160
Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165 170 175
Ile Ala Arg Ile Arg Val Arg Asp Ile Ser Arg Thr Asp Gly Gly Arg
180 185 190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195 200 205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210 215 220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225 230 235 240
Gln Val Arg Lys Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245 250 255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
260 265 270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275 280 285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290 295 300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
305 310 315 320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325 330 335
Arg Leu Leu Glu Asp Gly Asp
340

<210> 58
<211> 343
5 <212> PRT
<213> Secuencia artificial

<220>
<223> Tre3-135

10 <400> 58

Met Ser Lys Leu Leu Thr Val His Gln Asn Leu Ser Ala Leu Leu Val
1 5 10 15
Asp Ala Thr Ser Asp Glu Thr Arg Lys Asn Leu Met Gly Val Leu Arg
20 25 30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35 40 45
Cys Arg Thr Trp Val Ala Trp Cys Lys Leu Asn Asn Arg Arg Trp Phe
50 55 60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
65 70 75 80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85 90 95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asp Ala
100 105 110
Val Ser Leu Val Ile Arg Arg Ile Arg Arg Glu Asn Val Asp Ala Gly
115 120 125
Glu Arg Thr Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
130 135 140
Val Arg Ala Leu Met Gly Asn Ser Asp Arg Gly Gln Asp Ile Arg Ser
145 150 155 160
Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165 170 175
Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asn Gly Gly Arg
180 185 190
Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195 200 205
Val Glu Lys Ala Leu Ser Leu Arg Val Thr Lys Leu Val Glu Arg Trp
210 215 220
Ile Ser Val Ser Gly Val Ala Gly Asp Pro Asn Asn Tyr Leu Phe Cys
225 230 235 240
Gln Val Arg Lys Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245 250 255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Gly Ala Ala His Cys Leu Ile
260 265 270
Tyr Gly Ala Lys Asp Asp Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275 280 285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290 295 300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Glu Ser
305 310 315 320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325 330 335
Arg Leu Leu Glu Asp Asp Asp
340

<210> 59
<211> 343
5 <212> PRT
<213> Secuencia artificial

<220>
<223> Tre3-002

10 <400> 59

Met Ser Asn Leu Leu Thr Leu His Gln Asn Leu Ser Ala Leu Leu Val
1 5 10 15
Asp Val Thr Ser Asp Glu Val Arg Lys Asn Leu Met Asp Val Phe Arg
20 25 30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
35 40 45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
50 55 60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
65 70 75 80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85 90 95
Met Leu His Arg Arg Ser Gly Leu Pro Arg Leu Ser Asp Ser Asn Ala
100 105 110
Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
115 120 125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
130 135 140
Val Arg Ser Leu Met Glu Asp Ser Asp Arg Gly Gln Asp Ile Arg Thr
145 150 155 160
Leu Ala Phe Leu Gly Val Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165 170 175
Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg
180 185 190
Ile Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195 200 205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210 215 220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225 230 235 240
Arg Val Arg Lys Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245 250 255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Ala Ala Thr His Arg Gln Ile
260 265 270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275 280 285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290 295 300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
305 310 315 320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325 330 335
Arg Leu Leu Glu Asp Gly Asn
340

<210> 60
<211> 343
5 <212> PRT
<213> Secuencia artificial

<220>
10 <223> Tre3-023

<400> 60
Met Ser Asn Leu Gln Thr Leu His Arg Asn Leu Ser Ala Leu Leu Ile
1 5 10 15
Asp Ala Thr Ser Asp Glu Ala Arg Lys Asn Leu Met Asp Val Phe Arg

ES 2 577 929 T3

```

      20      25      30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Arg Val Leu Leu Ser Val
      35      40      45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe
      50      55      60
Pro Ala Glu Pro Glu Asp Val Arg Gly Tyr Leu Leu His Leu Gln Ala
65      70      75      80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
      85      90      95
Met Leu His Arg Arg Ser Gly Leu Pro Arg Leu Ser Asp Ser Asp Ala
      100      105      110
Val Ser Leu Val Val Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
      115      120      125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
      130      135      140
Val Arg Ala Leu Met Glu Asn Ser Asp Arg Gly Gln Asp Ile Arg Thr
145      150      155      160
Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
      165      170      175
Ile Ala Arg Ile Arg Ile Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg
      180      185      190
Met Leu Ile His Ile Ser Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
      195      200      205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
      210      215      220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225      230      235      240
Gln Val Arg Arg Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
      245      250      255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Ala His His Leu Ile
      260      265      270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
      275      280      285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
      290      295      300
Pro Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
305      310      315      320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
      325      330      335
Arg Leu Leu Glu Asp Gly Asp
      340

```

<210> 61
 <211> 343
 5 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Tre3-030

10

```

<400> 61
Met Ser Asn Leu Leu Thr Leu His Gln Asn Leu Ser Ala Leu Leu Val
1      5      10      15
Asp Val Thr Ser Asp Glu Val Arg Glu Asn Leu Met Asp Val Phe Arg
      20      25      30
Asp Arg Gln Ala Phe Ser Glu His Thr Trp Lys Val Leu Leu Ser Val
      35      40      45
Cys Arg Thr Trp Ala Ala Trp Cys Lys Leu Asn Asn Arg Lys Trp Phe

```

ES 2 577 929 T3

```

50          55          60
Pro Ala Glu Pro Glu Asp Val Arg Asp Tyr Leu Leu His Leu Gln Val
65          70          75          80
Arg Gly Leu Ala Val Asn Thr Ile Gln Gln His Leu Ala Gln Leu Asn
85          90          95
Met Leu His Arg Arg Phe Gly Leu Pro Arg Pro Gly Asp Ser Asp Ala
100        105        110
Val Ser Leu Val Met Arg Arg Ile Arg Lys Glu Asn Val Asp Ala Gly
115        120        125
Glu Arg Ala Lys Gln Ala Leu Ala Phe Glu Arg Thr Asp Phe Asp Gln
130        135        140
Val Arg Ala Leu Met Gly Asn Ser Asn Arg Cys Gln Asp Ile Arg Thr
145        150        155        160
Leu Ala Phe Leu Gly Ile Ala Tyr Asn Thr Leu Leu Arg Ile Ser Glu
165        170        175
Ile Ala Arg Ile Arg Val Lys Asp Ile Ser Arg Thr Asp Gly Gly Arg
180        185        190
Met Leu Ile His Ile Gly Arg Thr Lys Thr Leu Val Ser Thr Ala Gly
195        200        205
Val Glu Lys Ala Leu Ser Leu Gly Val Thr Lys Leu Val Glu Arg Trp
210        215        220
Ile Ser Val Ser Gly Val Ala Asp Asp Pro Asn Asn Tyr Leu Phe Cys
225        230        235        240
Gln Val Arg Lys Asn Gly Val Ala Val Pro Ser Ala Thr Ser Gln Leu
245        250        255
Ser Thr Ser Ala Leu Gln Gly Ile Phe Glu Ala Thr His Arg Leu Ile
260        265        270
Tyr Gly Ala Lys Asp Gly Ser Gly Gln Arg Tyr Leu Ala Trp Ser Gly
275        280        285
His Ser Ala Arg Val Gly Ala Ala Arg Asp Met Ala Arg Ala Gly Val
290        295        300
Ser Ile Ala Glu Ile Met Gln Ala Gly Gly Trp Thr Thr Val Asp Ser
305        310        315        320
Val Met Asn Tyr Ile Arg Asn Leu Asp Ser Glu Thr Gly Ala Met Val
325        330        335
Arg Leu Leu Glu Asp Asp Asp
340

```

<210> 62

<211> 343

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Tre3-031

10

<400> 62

```

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

```

[illegible]

<210> 63
<211> 343
<212> PRT
<213> Secuencia artificial

<220>
<223> Tre3-033

<400>	63															
Met	Ser	Asn	Leu	Leu	Thr	Leu	His	Gln	Asn	Leu	Ser	Ala	Leu	Leu	Val	
1				5					10					15		
Asp	Ala	Thr	Ser	Asp	Glu	Ala	Arg	Lys	Asn	Leu	Met	Asp	Val	Leu	Arg	
			20					25					30			
Asp	Arg	Gln	Ala	Phe	Ser	Glu	His	Thr	Trp	Lys	Val	Leu	Leu	Ser	Val	
			35				40					45				
Cys	Arg	Thr	Trp	Ala	Ala	Trp	Cys	Lys	Leu	Asn	Asn	Arg	Lys	Trp	Phe	
	50					55					60					
Pro	Ala	Glu	Pro	Glu	Asp	Val	Arg	Asp	Tyr	Leu	Leu	His	Leu	Gln	Val	
65					70					75					80	
Arg	Gly	Leu	Ala	Val	Asn	Thr	Ile	Gln	Gln	His	Leu	Ala	Gln	Leu	Asn	
				85					90					95		
Met	Leu	His	Arg	Arg	Phe	Gly	Leu	Pro	Arg	Pro	Gly	Asp	Ser	Asp	Ala	
			100					105					110			
Val	Ser	Leu	Val	Ile	Arg	Arg	Ile	Arg	Arg	Glu	Asn	Val	Asp	Ala	Gly	

[illegible]

<210> 64
<211> 343
<212> PRT
<213> Secuencia artificial

<220>
<223> Tre3-039

<400>	64															
Met	Ser	Asn	Leu	Leu	Thr	Ser	His	Gln	Asn	Leu	Ser	Ala	Leu	Leu	Val	
1				5					10					15		
Asp	Val	Thr	Ser	Asp	Glu	Val	Arg	Lys	Asn	Leu	Met	Asp	Val	Phe	Arg	
			20					25					30			
Asp	Arg	Gln	Ala	Phe	Ser	Glu	His	Thr	Trp	Lys	Val	Leu	Leu	Ser	Val	
		35					40					45				
Cys	Arg	Thr	Trp	Ala	Ala	Trp	Cys	Lys	Leu	Asn	Asn	Arg	Lys	Trp	Phe	
	50					55				60						
Pro	Ala	Glu	Pro	Glu	Asp	Val	Arg	Asp	Tyr	Leu	Leu	His	Leu	Gln	Val	
65					70					75				80		
Arg	Gly	Leu	Ala	Val	Asn	Thr	Ile	Gln	Gln	His	Leu	Ala	Gln	Leu	Asn	
				85					90					95		
Met	Leu	His	Arg	Arg	Ser	Gly	Leu	Pro	Arg	Pro	Gly	Asp	Ser	Asp	Ala	
			100					105					110			
Val	Ser	Leu	Val	Ile	Arg	Arg	Ile	Arg	Arg	Glu	Asn	Val	Asp	Ala	Gly	
		115					120					125				
Glu	Arg	Ala	Lys	Gln	Ala	Leu	Ala	Phe	Glu	Arg	Thr	Asp	Phe	Glu	Gln	
	130					135					140					
Val	Arg	Ser	Leu	Met	Glu	Asn	Ser	Asp	Arg	Gly	Gln	Asp	Ile	Arg	Asn	

```

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

```

<210> 65

<211> 351

5 <212> PRT

<213> Secuencia artificial

<220>

<223> Tre 1.0

10

<220>

<221> Señal

<222> (2)..(9)

<223> Secuencia de localización nuclear (SLN) para dirigir la proteína al núcleo de las células de mamífero

15

<400> 65

```

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

```

ES 2 577 929 T3

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

REIVINDICACIONES

1. Procedimiento para preparar un vector de expresión que codifica una recombinasa elaborada a medida, pudiendo dicha recombinasa elaborada a medida recombinar unas secuencias diana asimétricas dentro de la LTR del ADN provírico de una pluralidad de cepas de retrovirus que son cepas del VIH-1, que comprende las etapas siguientes:

(a) identificar unas secuencias con una homología de por lo menos 30% respecto a la secuencia de hemisio izquierdo y secuencia de hemisio derecho de por lo menos un sitio diana de recombinasa conocido en la secuencia de la LTR del ADN provírico de una pluralidad de cepas de retrovirus, en el que las secuencias homólogas se encuentran separadas por un espaciador de 5 a 12 nucleótidos y en el que la secuencia diana asimétrica se encuentra en una pluralidad de cepas de retrovirus, en el que la secuencia diana asimétrica presenta la secuencia expuesta como SEC ID nº 1,

(b) identificar dos secuencias, en el que la primera secuencia corresponde a la secuencia de la secuencia diana asimétrica de la etapa (a) homóloga al hemisio izquierdo de dicho sitio diana conocido y a la que se hace referencia como "secuencia 1 de hemisio" y en el que la segunda secuencia corresponde a la secuencia de la secuencia diana asimétrica de la etapa (a) homóloga al hemisio derecho y a la que se hace referencia como "secuencia 2 de hemisio";

(c) determinar los nucleótidos dentro de las secuencias de la etapa (b) que se diferencian de las secuencias de hemisio izquierdo y de hemisio derecho homólogas correspondientes de dicho por lo menos un sitio diana homólogo conocido de la etapa (a);

(d) generar un primer subconjunto de dos ácidos nucleicos diana que comprende unas secuencias diana, en el que la primera secuencia diana se denomina subsitio 1 y comprende, adyacentes entre sí y en orden 5' a 3', la secuencia 1 de hemisio de la etapa (b), la secuencia espaciadora de la secuencia diana asimétrica y una repetición invertida de la secuencia 1 de hemisio, y en el que la segunda secuencia diana se denomina subsitio 2 y comprende, adyacentes entre sí y en orden 5' a 3', una repetición invertida de la secuencia 2 de hemisio, la secuencia espaciadora de la secuencia diana asimétrica y la secuencia 2 de hemisio de la etapa (b),

(e) generar un segundo subconjunto de ácidos nucleicos diana que comprende unas secuencias diana modificadas sobre la base de las secuencias diana del primer subconjunto de la etapa (d),

en el que, en las secuencias basadas en el subsitio 1, en la secuencia de hemisio izquierdo, una parte de los nucleótidos que se diferencian de la secuencia de hemisio homóloga correspondiente de por lo menos un sitio diana conocido de la etapa (a) se sustituye por los nucleótidos naturales hallados en dicho sitio diana conocido, hasta que dicha secuencia de hemisio contenga uno, dos o tres nucleótidos que se diferencian de dicho sitio diana conocido, en el que el hemisio derecho de dicha secuencia diana modificada está formado por una repetición invertida de dicha secuencia modificada de hemisio izquierdo, que se encuentra separada de dicha secuencia modificada de hemisio izquierdo por la secuencia espaciadora de la secuencia diana asimétrica, y

en el que, en las secuencias basadas en el subsitio 2, en la secuencia de hemisio derecho, una parte de los nucleótidos que se diferencian de la secuencia de hemisio homóloga correspondiente de dicho por lo menos un sitio diana conocido de la etapa (a) se sustituye por los nucleótidos naturales hallados en dicho sitio diana conocido, hasta que dicha secuencia de hemisio contenga uno, dos o tres nucleótidos que se diferencian de dicho sitio diana conocido, en el que el hemisio izquierdo de dicha secuencia diana modificada está formado por una repetición invertida de dicha secuencia modificada de hemisio derecho, que se encuentra separada de dicha secuencia modificada de hemisio derecho por la secuencia espaciadora de la secuencia diana asimétrica,

de manera que en todas las secuencias modificadas de hemisio originadas a partir de una secuencia diana del primer subconjunto de la etapa (d) tomadas conjuntamente, pueden encontrarse todos los nucleótidos que se diferencian, mientras que ninguna de dichas secuencias modificadas de hemisio sola comprende todos los nucleótidos que se diferencian,

(f) aplicar por separado la evolución molecular dirigida a por lo menos una recombinasa que reconoce un sitio diana homólogo conocido según la etapa (a) utilizando cada ácido nucleico del segundo subconjunto obtenido en la etapa (e) como sustrato,

(g) mezclar las colecciones de recombinasas evolucionadas en la etapa (f), en el que todas las colecciones de recombinasas evolucionadas sobre las secuencias basadas en el subsitio 1 se combinan y mezclan, y en el que todas las colecciones de recombinasas evolucionadas sobre las secuencias basadas en el subsitio 2 se combinan y mezclan;

- (h) aplicar la evolución molecular dirigida a las colecciones mezcladas obtenidas en la etapa (g) utilizando cada ácido nucleico del subconjunto según la etapa (d) como un sustrato;
- 5 (i) mezclar las colecciones de recombinasas evolucionadas en la etapa (h);
- (j) aplicar la evolución de proteína ligada a sustrato de la colección mezclada obtenida en la etapa (g) utilizando un ácido nucleico que comprende la secuencia diana asimétrica de la etapa (a) como un sustrato, hasta que se obtiene por lo menos una recombinasa que resulta activa sobre la secuencia diana asimétrica dentro de la LTR del ADN retrovítico de la etapa (a);
- 10 (k) aislar el ácido nucleico que codifica dicha por lo menos una recombinasa obtenida en la etapa (j) a partir de la colección; y
- 15 (l) clonar el ácido nucleico obtenido en la etapa (k) en un vector de expresión adecuado.
2. Procedimiento según la reivindicación 1, en el que dicha por lo menos una recombinasa conocida cuya secuencia diana es utilizada en la etapa (a) y sobre la que es aplicada la evolución molecular dirigida en la etapa (f), es parte de una colección de recombinasas.
- 20 3. Procedimiento según la reivindicación 1 o 2, en el que el vector de expresión en la etapa (1) es seleccionado de entre el grupo que consiste en vectores retrovíticos, vectores lentivíticos, vectores *spumavirus* y vectores adenovíticos.
- 25 4. Procedimiento para preparar una recombinasa elaborada a medida, en el que dicho procedimiento comprende el procedimiento según cualquiera de las reivindicaciones 1 a 3 y la etapa adicional de expresar la recombinasa elaborada a medida a partir del ácido nucleico que codifica la recombinasa insertada en el vector de expresión en una célula hospedadora adecuada, en el que la recombinasa se expresa opcionalmente como un polipéptido de fusión que comprende la secuencia de aminoácidos de la recombinasa elaborada a medida.
- 30 5. Procedimiento para preparar una célula transformada, en el que dicho procedimiento comprende el procedimiento según cualquiera de las reivindicaciones 1 a 3 y la etapa adicional de introducir el vector de expresión *in vitro* en una célula, en el que la célula es preferentemente una célula troncal adulta.
- 35 6. Procedimiento según cualquiera de las reivindicaciones 1 a 5, que comprende además la etapa de preparar el vector de expresión, la recombinasa o la célula como una composición farmacéutica.
- 40 7. Ácido nucleico que codifica una recombinasa elaborada a medida, pudiendo dicha recombinasa elaborada a medida recombinar unas secuencias diana asimétricas que comprenden la SEC ID nº 1 dentro de la LTR del ADN provírico de una pluralidad de cepas de VIH-1, en el que la secuencia de aminoácidos de la recombinasa elaborada a medida presenta una identidad de secuencia de por lo menos 99% respecto a una secuencia según la SEC ID nº 38, en el que dicha recombinasa elaborada a medida comprende todos los intercambios de aminoácido en comparación con la secuencia de Cre (SEC ID nº 36) de P12S, P15L, M44V, K86N, G93A, A175S y P307A, en el que el ácido nucleico es opcionalmente un vector de expresión.
- 45 8. Ácido nucleico según la reivindicación 7, como puede obtenerse a partir de la etapa (k) del procedimiento según cualquiera de las reivindicaciones 1 a 3 o vector de expresión como puede obtenerse a partir del procedimiento según cualquiera de las reivindicaciones 1 a 3.
- 50 9. Ácido nucleico según cualquiera de las reivindicaciones 7 a 8, en el que la secuencia de aminoácidos de la recombinasa elaborada a medida comprende la secuencia según SEC ID nº 38.
10. Ácido nucleico según la reivindicación 9, en el que la secuencia de aminoácidos de la recombinasa elaborada a medida comprende la secuencia según SEC ID nº 39.
- 55 11. Ácido nucleico según la reivindicación 10, en el que la secuencia de aminoácidos de la recombinasa elaborada a medida es seleccionada de entre el grupo que consiste en SEC ID nº 40 a 64 y una combinación de secuencias de entre el grupo que consiste en SEC ID nº 40 a 64.
- 60 12. Recombinasa elaborada a medida codificada por el ácido nucleico según cualquiera de las reivindicaciones 7 a 11, que se expresa opcionalmente como una proteína de fusión.
13. Célula transformada que comprende un ácido nucleico según cualquiera de las reivindicaciones 7 a 11, en la que dicha célula es preferentemente una célula troncal del linaje hematopoyético.
- 65 14. Composición farmacéutica que comprende un ácido nucleico según cualquiera de las reivindicaciones 7 a 11,

una recombinasa elaborada a medida según la reivindicación 12 y/o una célula transformada según la reivindicación 13.

5 15. Composición farmacéutica según la reivindicación 14, en la que la composición farmacéutica está destinada a la utilización en el tratamiento o la prevención de la infección por VIH-1 en un sujeto.

16. Composición farmacéutica según la reivindicación 15, en la que el ADN provírico hallado en una muestra obtenida del sujeto comprende la secuencia diana asimétrica de SEC ID nº 1 sobre la que se ha seleccionado la recombinasa.

Fig. 1

loxP (Cre)	ATAACTTCGTATA ATGTATGCT TATACGAAGTTAT	SEC ID no°: 3
loxH (Fre)	ATATATACGTATAT TAGACAT ATATACGTATATAT	SEC ID no°: 4
rox (Dre)	CTAACTTTAAATA ATGCCAAT TATTTAAAGTTAT	SEC ID no°: 5
zox (Zre)	ATAACTTCGTATA ACACACAT TATGCGAAGTTAT	SEC ID no°: 6
loxLTR (Tre)	ACAACATCCTATT ACACCCT TATATGCCAACATGG	SEC ID no°: 7
SEC ID no°: 1	<u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>T</u>	
SEC ID no°: 2	<u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>G</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u>	

Fig. 2

SEC ID no°: 1	<u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>T</u>	
Subsitio 1	<u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>T</u> <u>T</u>	SEC ID no°: 8
Subsitio 1b+2	<u>A</u> <u>A</u> <u>C</u> <u>A</u> <u>C</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>T</u> <u>G</u> <u>T</u> <u>T</u>	SEC ID no°: 27
Subsitio 1b+1	<u>A</u> <u>T</u> <u>C</u> <u>A</u> <u>C</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>T</u> <u>G</u> <u>A</u> <u>T</u>	SEC ID no°: 28
Subsitio 1b	<u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>C</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>T</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 29
Subsitio 1a	<u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>T</u> <u>T</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>T</u> <u>T</u>	SEC ID no°: 30
Subsitio 1A	<u>A</u> <u>A</u> <u>C</u> <u>A</u> <u>C</u> <u>A</u> <u>T</u> <u>T</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>G</u> <u>T</u> <u>G</u> <u>T</u> <u>T</u>	SEC ID no°: 9
Subsitio 1B	<u>A</u> <u>T</u> <u>A</u> <u>C</u> <u>C</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>C</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 10
Subsitio 1C	<u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>C</u> <u>A</u> <u>T</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>G</u> <u>T</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 11
Subsitio 2	<u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>T</u>	SEC ID no°: 12
Subsitio 2a+2	<u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>T</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>T</u>	SEC ID no°: 31
Subsitio 2a+1	<u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>T</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>T</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>T</u>	SEC ID no°: 32
Subsitio 2a	<u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>C</u> <u>T</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>C</u> <u>C</u> <u>T</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>T</u>	SEC ID no°: 33
Subsitio 2b	<u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>G</u> <u>T</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 34
Subsitio 2A	<u>A</u> <u>A</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>C</u> <u>C</u> <u>T</u> <u>A</u> <u>G</u> <u>T</u> <u>C</u> <u>T</u> <u>T</u>	SEC ID no°: 13
Subsitio 2B	<u>A</u> <u>T</u> <u>A</u> <u>G</u> <u>C</u> <u>A</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>A</u> <u>C</u> <u>C</u> <u>T</u> <u>T</u> <u>G</u> <u>C</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 14
Subsitio 2C	<u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>C</u> <u>T</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>T</u> <u>A</u> <u>G</u> <u>T</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 15
Δ loxP	ATAACTTCGTTT TAATGTATGCT AAACGAAGTTAT	SEC ID no°: 35
SEC ID no°: 2	<u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>G</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u>	
Subsitio 1	<u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>G</u> <u>T</u> <u>C</u> <u>C</u> <u>C</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>G</u>	SEC ID no°: 16
Subsitio 1A	<u>C</u> <u>T</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>T</u> <u>C</u> <u>G</u> <u>A</u> <u>C</u> <u>T</u> <u>A</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>G</u> <u>T</u> <u>C</u> <u>G</u> <u>A</u> <u>A</u> <u>G</u> <u>T</u> <u>C</u> <u>A</u> <u>G</u>	SEC ID no°: 17
Subsitio 1B	<u>A</u> <u>T</u> <u>A</u> <u>G</u> <u>G</u> <u>T</u> <u>T</u> <u>C</u> <u>G</u> <u>A</u> <u>A</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>T</u> <u>T</u> <u>C</u> <u>G</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 18
Subsitio 1C	<u>A</u> <u>T</u> <u>A</u> <u>A</u> <u>C</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>A</u> <u>T</u> <u>A</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>T</u> <u>T</u> <u>C</u> <u>C</u> <u>C</u> <u>G</u> <u>G</u> <u>T</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 19
Subsitio 2	<u>A</u> <u>T</u> <u>C</u> <u>T</u> <u>G</u> <u>A</u> <u>G</u> <u>G</u> <u>G</u> <u>C</u> <u>T</u> <u>C</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>G</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>T</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u>	SEC ID no°: 20
Subsitio 2A	<u>A</u> <u>T</u> <u>C</u> <u>T</u> <u>G</u> <u>T</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>A</u> <u>C</u> <u>A</u> <u>G</u> <u>A</u> <u>T</u>	SEC ID no°: 21
Subsitio 2B	<u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>G</u> <u>A</u> <u>T</u> <u>G</u> <u>G</u> <u>C</u> <u>T</u> <u>T</u> <u>A</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>T</u> <u>A</u> <u>A</u> <u>G</u> <u>C</u> <u>C</u> <u>A</u> <u>T</u> <u>C</u> <u>A</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 22
Subsitio 2C	<u>A</u> <u>T</u> <u>A</u> <u>T</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>G</u> <u>C</u> <u>A</u> <u>C</u> <u>A</u> <u>G</u> <u>G</u> <u>G</u> <u>A</u> <u>G</u> <u>T</u> <u>G</u> <u>G</u> <u>T</u> <u>G</u> <u>T</u> <u>G</u> <u>C</u> <u>C</u> <u>C</u> <u>A</u> <u>C</u> <u>A</u> <u>T</u> <u>A</u> <u>T</u>	SEC ID no°: 23

Fig. 3

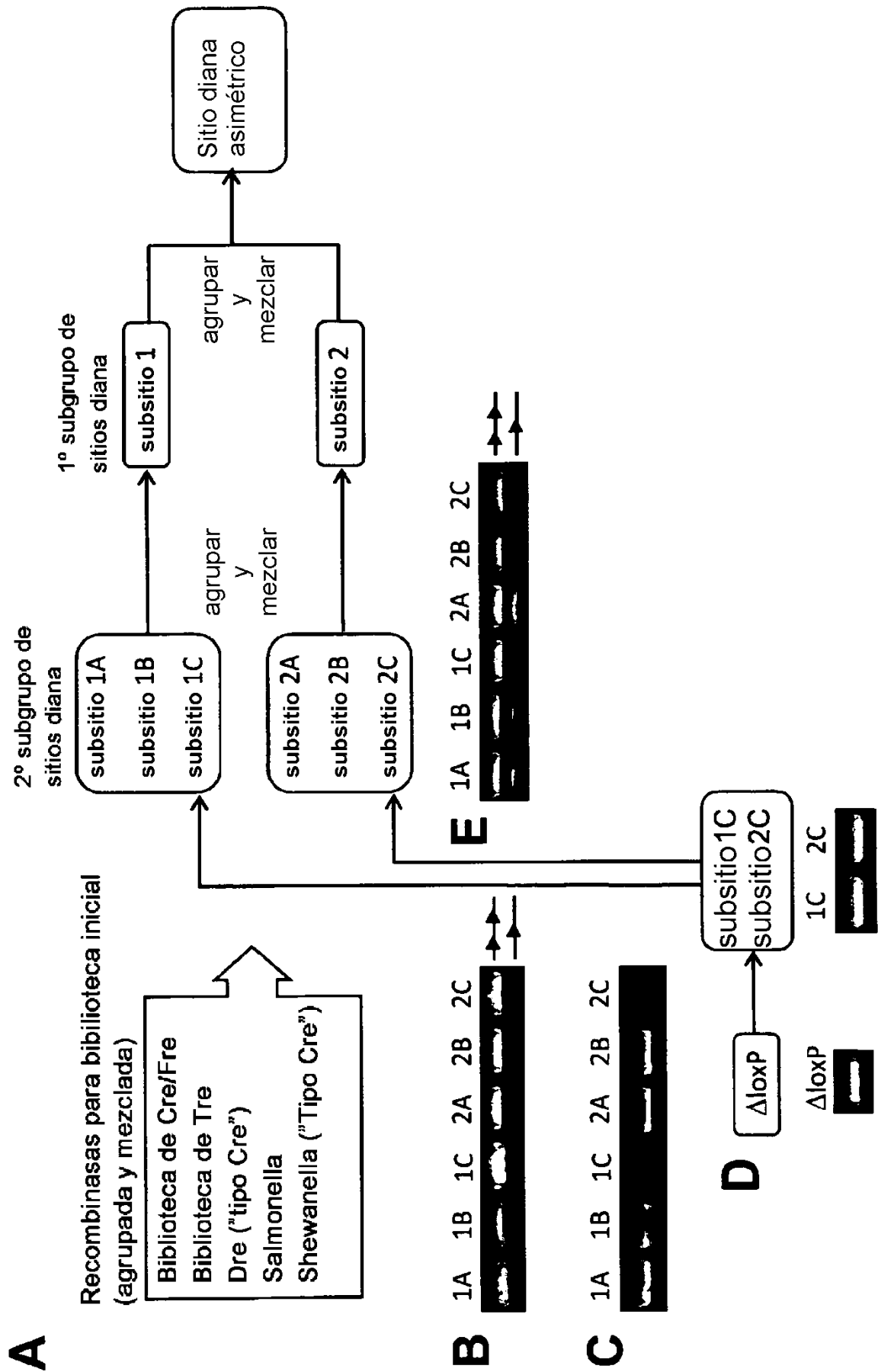


Fig. 4

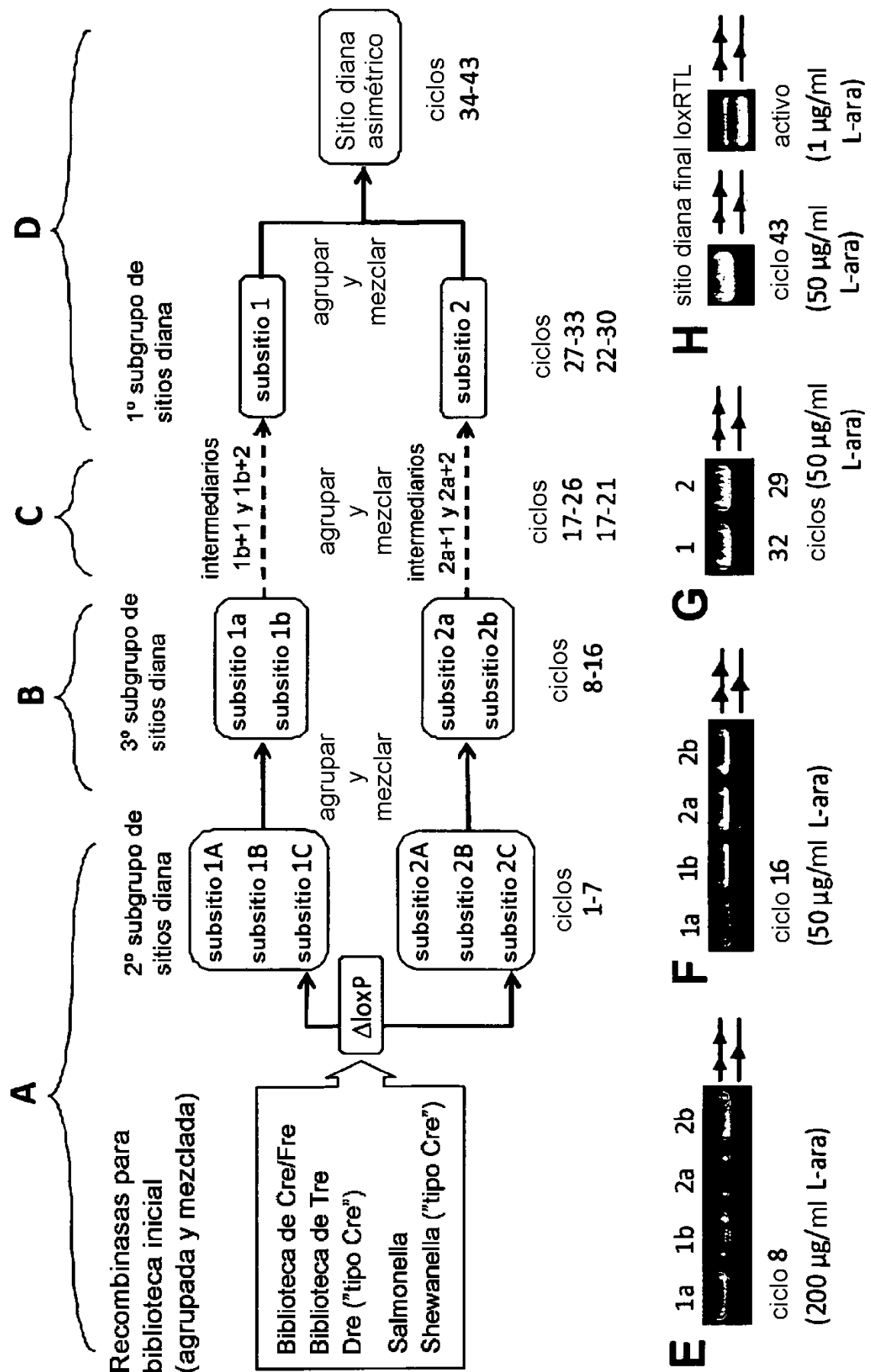


Fig. 5

- a) Cre SEC ID no^o: 36
 b) secuencia de consenso común de Tree 100% SEC ID no^o: 37
 c) secuencia de consenso de Tre 3.0 100% SEC ID no^o: 38
 d) secuencia de consenso de Tre 3.0 95% SEC ID no^o: 39

```

1      12 15      30      44
MSNLLTVHQNLPALPVDATSDEVVRKNLMDMFRDRQAFSEHTWKMLLSVCRSWAAWCKLN a
MSXXXTXXXXLXALXDXSDXXXXXXLXXXRDXXAXSXXTWXXLLSXCXRXWAWCXXX b
MSXXXTXXXXLSALLXDXSDXXXXXXLXXXRDXXAXSXXTWXVLLSXCXRXWAWCXXX c
MSXLXTXHXSLSALLXDXTSDEXRXNLMDVXRDXAXSEHTWKVLLSVCRXWAAWCXLN d

60      86      93
NRKWFPAEPEDVRDYLLYLQARGLAVKTIQQHLGQLNMLHRRSGLPRPSDSNAVSLVMRR a
XXXXFPXXPXXVRXYLLXLQXRGLXVTXQQHLXXLNMXHRRXGLRXXDXXXVSLXXRR b
XXXXFPXXPXXVRXYLLXLQXRGLXVNTXQQHLAXLNMXHRRXGLRXXDXXXVSLXXRR c
NRKXFPAPEDVRDYLLXLQXRGLAVNTIQQHLAXLNMLHRRXGLPRXXDSXAVSLVXRR d

120      175
IRKENVDAGERAKQALAFERTDFDQVRSLMENS DRCQDIRNLAFGLGIAYNTLLRIAEIAR a
IRXENVDA GERXKQALAFXRDXXXXXXXLXXXXSXXXXDXRXLAXLGXAYNTLLRXXEXXX b
IRXENVDA GERXKQALAFXRDXXXXXXXLXXXXSXXXXDXRXLAXLGXAYNTLLRXSEXXX c
IRXENVDA GERXKQALAFERTDXXQVRXLMXXSXRQDIRXLAXLGXAYNTLLRISEIAR d

180
IRVKDISRTDGGRMLIHIGRTKTLVSTAGVEKALS LGVTKLVERWISVSGVADDPNNYLF a
XXXXDXSXTXGGRXLIHXXXTKTLVSTXGVEKALS LXXTXLXERWXSXSGVAXXXXXXYLF b
XXXXDXSXTXGGRXLIHXXXTKTLVSTXGVEKALS LXXTXLXERWXSXSGVAXXXXXXYLF c
IRXXDISXTXGGRMLIHIXRTKTLVSTAGVEKALS LXVTXLVERWISXSGVAXDPNNYLF d

240      259 262
CRVRKNGVAAPSATSQ LSTRALEGIFEATHRLIYGAKDDSGQRYLAWSGHSARVGAARDM a
CXXXXXGXAXPXAXXXLSXXXLXXIFXXXHXXXXGAKXXSGXRYXXWSGHSARVGAARDM b
CXXXXXGXAXPXAXXXLSXXXLXXIFXXXHXXXXGAKXXSGXRYXXWSGHSARVGAARDM c
CXVRXXGVAXPSATXQLSTSALQGIFXAXHXL IYGAKXXSGXRYLAWSGHSARVGAARDM d

300      307      317 320
ARAGVSIPEIMQAGGWTNVNIVMNYIRNLDSETGAMVRLLEDGD* a
ARAGVXIXEIMQAGGWXTVXXVMNYIRNLDSEXGAMVRLLEXXX* b
ARAGVXIAEIMQAGGWXTVXXVMNYIRNLDSEXGAMVRLLEXXX* c
ARAGVXIAEIMQAGGWTTVXSMNYIRNLDSETGAMVRLLEXXD* d

```