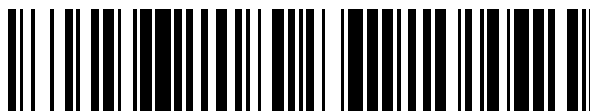


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 847 698**

51 Int. Cl.:

C07K 14/705 (2006.01)

C07K 16/28 (2006.01)

C12N 15/85 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.04.2016 PCT/GB2016/051164**

87 Fecha y número de publicación internacional: **03.11.2016 WO16174408**

96 Fecha de presentación y número de la solicitud europea: **26.04.2016 E 16720483 (3)**

97 Fecha y número de publicación de la concesión europea: **25.11.2020 EP 3288968**

54 Título: **Construcción de ácido nucleico**

30 Prioridad:

27.04.2015 GB 201507104

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

03.08.2021

73 Titular/es:

**AUTOLUS LIMITED (100.0%)
Forest House, 58 Wood Lane
London W12 7RZ, GB**

72 Inventor/es:

**PULÉ, MARTIN y
CORDOBA, SHAUN**

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 847 698 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

Construcción de ácido nucleico

5 Campo de la invención

La presente invención se refiere a construcciones y enfoques para modular la expresión relativa de polipéptidos coexpresados a partir de un solo vector. En particular, la invención se refiere a modular la expresión de una proteína transmembrana coexpresada a partir de un solo vector con un segundo polipéptido.

10 Antecedentes de la invención

Con frecuencia es deseable expresar diferentes proteínas del mismo vector, ya que la transducción múltiple de la misma célula es difícil, cara e impredecible. Por lo tanto, se han desarrollado diferentes métodos para permitir la coexpresión de dos proteínas a partir de un solo vector (véase figura 1).

Los intentos iniciales utilizaron dos promotores diferentes dentro del mismo casete. Esto da lugar a dos transcripciones separadas, cada una de las cuales codifica una proteína separada. Este es un enfoque difícil por varias razones. Un problema clave es la "interferencia del promotor" por la que un promotor domina y provoca el silenciamiento del segundo promotor. Además, diferentes promotores funcionan de manera diferente en diferentes contextos celulares y esto hace que sea difícil lograr el "ajuste" uniforme de la expresión relativa de cada transgén.

Un enfoque alternativo es utilizar una secuencia de entrada de ribosoma interna (IRES). En este caso, se genera un solo transcrito. La secuencia de IRES en el transcrito se coloca entre los marcos de lectura abiertos para los dos transgenes e imita una estructura de caperuza de ARNm. Por lo tanto, el ribosoma inicia la traducción en la caperuza 5' o la IRES dando lugar a la expresión de dos proteínas separadas. Una limitación clave de este enfoque es la incapacidad de controlar la expresión relativa. El transcrito 3' se expresa normalmente menos que el 5', pero la relación de expresión es difícil de predecir y ajustar.

Se ha proporcionado un enfoque adicional después de la caracterización de la función del péptido 2A del virus de la fiebre aftosa (FMDV) para permitir que el FMDV (y virus relacionados) expresen múltiples proteínas de un solo marco de lectura abierto (ORF) (Donnelly *et al.*; J. Gen. Virol.; 82, 1027-1041 (2001)). El péptido 2A (y homólogos) se escinde con una eficacia muy alta inmediatamente después de la traducción del ORF, permitiendo la expresión de múltiples péptidos a partir de un único ORF, como se describe, por ejemplo, en De Felipe P *et al.* (2004), Genetics Vaccines and Therapy, vol. 2, n.º 1, p. 13. Un problema con el uso del péptido 2A para escindir entre diferentes péptidos en el mismo ORF es que la expresión se limita a una relación de 1:1.

Por tanto, existe la necesidad de métodos alternativos para expresar más de una proteína a partir de un único vector que no estén asociados con las desventajas descritas anteriormente.

40 Sumario de la invención

La invención es como se define en las reivindicaciones. La presente invención se basa en la determinación de que, cuando una primera proteína transmembrana se coexpresa con una segunda proteína transmembrana como una poliproteína que después de la traducción se escinde posteriormente para separar ambas proteínas, la incorporación de una señal de retención intracelular en la primera o segunda proteína transmembrana permite modular la expresión en la superficie celular de esta proteína transmembrana en relación con la otra proteína transmembrana reduciendo su tráfico a la superficie celular y/o reduciendo su semivida en la superficie celular. No es necesario que esto se limite a un par de transgenes, pero puede usarse para permitir el control de la expresión relativa de múltiples proteínas traducidas inicialmente como una poliproteína, en donde al menos dos de las múltiples proteínas son proteínas transmembrana.

Como se emplea en esta memoria, 'poliproteína' se refiere a una secuencia polipeptídica traducida de una construcción de ácido nucleico como una entidad individual, pero que comprende secuencias polipeptídicas que se separan posteriormente y que actúan como entidades discretas (por ejemplo, proteínas separadas).

Por tanto, en un primer aspecto, la presente invención proporciona una construcción de ácido nucleico que codifica una poliproteína que comprende la siguiente estructura:

60 A-X-B

en la cual

65 A y B son primera y segunda proteínas transmembrana que, cuando la construcción de ácido nucleico se introduce en una célula, se expresan en la superficie celular; y

X es un sitio de escisión de proteasa o un péptido autoescindible,

en donde la primera o la segunda proteína transmembrana comprende una señal de retención intracelular heteróloga.

5 La señal de retención intracelular es heteróloga de la proteína transmembrana, es decir, no aparece en la versión de "tipo silvestre" de la proteína transmembrana (o parte de la misma) y se introduce por medios recombinantes. La señal de retención intracelular puede ser sintética u obtenible de otro organismo u otra proteína.

10 El endodominio de la proteína transmembrana puede comprender la señal de retención intracelular.

La señal de retención intracelular puede apartar la proteína transmembrana de la ruta secretora y/o dirigirla hacia un compartimento intracelular unido a membrana tal como un compartimento lisosómico, endosómico o del aparato de Golgi.

15 La señal de retención intracelular puede, por ejemplo, ser una señal de clasificación basada en tirosina, una señal de clasificación basada en dileucina, una señal de grupo ácido, una señal de evitación lisosomal, una señal tipo NPFX'(1,2)D, una señal KDEL, KKX'X' o KX'KX'X' (en la que X' es cualquier aminoácido).

20 La señal de retención intracelular puede comprender una secuencia seleccionada del grupo de: NPX'Y, YX'X'Z', [DE]X'X'X'L[LI], DX'X'LL, DP[FW], FX'DX'F, NPF, LZX'Z[DE], LLDLL, PWDLW, KDEL, KKX'X' o KX'KX'X'; en la que X' es cualquier aminoácido y Z' es un aminoácido con una cadena lateral hidrófoba voluminosa.

25 La señal de retención intracelular puede comprender cualquiera de las secuencias mostradas en las tablas hasta 6. La señal de retención intracelular puede comprender la señal de retención intracelular de proteína relacionada con tirosinasa (TYRP)-1. La señal de retención intracelular puede comprender el dominio intracelular TYRP-1. La señal de retención intracelular puede comprender la secuencia NQPLLED (SEQ ID NO: 35).

30 La señal de retención intracelular puede comprender la señal de retención intracelular E3/19K adenoviral. La señal de retención intracelular puede comprender el dominio citosólico E3/19K. La señal de retención intracelular puede comprender la secuencia KYKSRRSFIDEKKMP (SEQ ID NO: 36); o DEKKMP (SEQ ID NO: 37).

35 La señal de retención intracelular puede ser proximal o distal a un dominio transmembrana de la proteína transmembrana.

X puede ser una secuencia de ácido nucleico que codifica un péptido autoescindible, un sitio de escisión de furina o un sitio de escisión del virus del grabado del tabaco.

40 X puede ser una secuencia de ácido nucleico que codifica un péptido autoescindible 2A de un aftovirus o un cardiovirus o un péptido de tipo 2A.

La proteína transmembrana puede ser cualquier proteína transmembrana expresada transgénicamente que, cuando la construcción de ácido nucleico se introduce en una célula, se expresa en la superficie celular.

45 La proteína transmembrana se puede seleccionar de una lista de: receptores excitadores tales como 41BB, OX40, CD27, CD28 y moléculas relacionadas; o receptores inhibidores tales como PD1, CTLA4, LAIR1, CD22 y moléculas relacionadas; o moléculas de receptor de citocina tales como IL1R, IL2R, IL7R, IL15R y moléculas relacionadas; o moléculas de referencia tales como N-CAM, V-CAM, L1-CAM, LFA-1, CDH1-3, selectinas o integrinas;

50 La proteína transmembrana puede ser una proteína sintética tal como un gen suicida o un gen marcador.

La proteína transmembrana puede ser o comprender las cadenas α y/o β de un receptor de linfocitos T.

La proteína transmembrana puede ser un receptor quimérico de antígenos (CAR).

55 La primera o la segunda proteína transmembrana comprende una señal de retención intracelular heteróloga como se define en el presente documento.

60 Una o ambas de las proteínas transmembrana pueden ser una proteína de un solo dominio transmembrana, tal como una proteína transmembrana de tipo I.

65 La cantidad de una proteína transmembrana que comprende una señal de retención intracelular que se expresa en la superficie celular puede ser, por ejemplo, menos de 90 %, 70 %, 50 % o 30 % en comparación con una proteína transmembrana expresada a partir de la misma construcción de ácido nucleico que no comprende una señal de retención intracelular.

La invención también proporciona una construcción de ácido nucleico que codifica una poliproteína que comprende la siguiente estructura:

A-X-B-Y-C

en la cual

A, B y C son primero, segundo y tercer polipéptidos de interés (PDI); y X e Y pueden ser iguales o diferentes, siendo cada uno de los cuales un sitio de escisión de proteasa o un péptido autoescindible, en donde al menos dos de los PDI son proteínas transmembrana que, cuando la construcción de ácido nucleico se introduce en una célula, se expresan en la superficie celular y que comprenden una señal de retención intracelular heteróloga, y en donde los al menos dos PDI que son proteínas transmembrana y que comprenden una señal de retención intracelular heteróloga: (a) comprenden diferentes señales de retención intracelular; y/o (b) tienen la señal de retención intracelular localizada en una posición diferente en el PDI; de modo que cuando el ácido nucleico se expresa en una célula, existe una expresión relativa diferencial de las al menos dos proteínas transmembrana en la superficie celular.

En un segundo aspecto la presente invención proporciona un vector que comprende una construcción de ácido nucleico de acuerdo con el primer aspecto de la invención.

El vector puede ser un vector retroviral o un vector lentiviral o un transposón.

En un tercer aspecto la presente invención proporciona una célula que comprende una construcción de ácido nucleico de acuerdo con el primer aspecto de la invención o un vector de acuerdo con el segundo aspecto de la invención.

La divulgación se refiere además a un método para modular la expresión relativa en la superficie celular de una proteína transmembrana expresada como una poliproteína de un único ácido nucleico con una segunda proteína incluyendo una señal de retención intracelular en la secuencia de ácido nucleico que codifica la proteína transmembrana.

La inclusión de una señal de retención intracelular en una proteína transmembrana reduce la cantidad de proteína transmembrana expresada en la superficie celular. Como tal, se puede modular el nivel de expresión relativo de una proteína transmembrana derivada de una poliproteína que incluye un segundo polipéptido. Cuando la proteína transmembrana solo está activa en la superficie celular (o está predominantemente activa en la superficie celular), la reducción de la expresión relativa en la superficie celular de la proteína también reduce su actividad relativa.

Esta invención puede ampliarse para modular la expresión relativa de tres o más proteínas expresadas como un polipéptido concatenado, separadas por sitios de escisión y expresión relativa en la superficie dictada por señales de retención de actividad diferente, como se define en las reivindicaciones.

Descripción de las figuras

Figura 1: Métodos utilizados para expresar diferentes proteínas del mismo vector

(a) Dos promotores diferentes dentro del mismo casete dan como resultado dos transcritos diferentes, cada uno de los cuales da lugar a proteínas separadas. (b) Uso de una secuencia del sitio interno de entrada al ribosoma (IRES) que conduce a un transcrito único que se traduce en dos proteínas independientes. (c) Uso de los resultados del péptido FMDV 2A que da como resultado un único transcrito, y una única poliproteína que se escinde rápidamente en dos proteínas diferentes.

Figura 2: El endodominio TYRP1 es capaz de dirigir la retención de una proteína transmembrana con un endodominio complejo

Tyrp1 es una proteína transmembrana de tipo I, de 537aa de longitud. El motivo di-leucina, que retiene la proteína en el compartimiento intracelular, está indicado como un rectángulo negro en el dominio citoplasmático. (A) Tyrp1 (wt). Tyrp1 tipo silvestre consiste en un péptido de señal, un dominio luminal, un dominio transmembrana y un dominio citoplásmico. El dominio citoplasmático contiene la señal de retención de dileucina. (B) Tyrp1 (ts)-enlazador SG-eGFP. Esta construcción contiene el Tyrp1 de tipo silvestre fusionado simplemente a eGFP a través de un enlazador- serina-glicina-glicina-glicina-serina. El Tyrp1-L-eGFP representa el Tyrp1 proximal citoplasmático. (C) Tyrp1 luminal (LM)-Transmembrana (TM)-Enlazador SG-eGFP-Tyrp1 citoplasmático (CP). Esta construcción constituye el Tyrp1 distal citoplasmático, ya que el enlazador SG-eGFP se interpone entre los dominios transmembrana y citoplasmático. D: Tyrp1 luminal (LM)-Transmembrana (TM)-Enlazador SG-eGFP. Esta construcción sirve como el control positivo, ya que se ha excluido el dominio citoplasmático que contiene la señal de retención. Todas las construcciones se co-expresan con IRES.CD34. La tinción de las células SupT1 transducidas se muestra con tinción intracelular y superficial izquierda/derecha inferior respectivamente.

Figura 3: Funcionalidad de la señal de retención de TYRP1 en células primarias

Se generó una construcción que co-expresa un CAR anti-CD19 y un anti-CD33 usando un péptido de tipo FMD-

2A. También se generaron dos variantes de esta construcción: en la primera variante, el motivo di-leucina de TYRP1 se insertó en el endodominio de CAR anti-CD19 justo proximal al dominio TM; en la segunda variante, el mismo motivo di-leucina de TYRP1 se unió al extremo carboxilo del endodominio de CAR anti-CD19. Se aislaron PBMC de la sangre y se estimularon usando PHA e IL-2. Dos días más tarde las células fueron transducidas en placas recubiertas con retronectina con retrovirus que contienen las construcciones de CAR CD19:CD33 diferentes. El día 5 el nivel de expresión de los dos CAR traducidos por la construcción se evaluó mediante citometría de flujo usando las fusiones CD19-Fc y CD33-Fc recombinantes. A. Representación del gen sintético construido para permitir la co-expresión; B. Representación del posterior par de proteínas generado por las tres construcciones; C. Muestra la expresión de los dos receptores por citometría de flujo. En la construcción original, ambos CAR se expresan igualmente. Con la incorporación del motivo di-leucina distalmente en el endodominio del CAR anti-CD19, la expresión del CAR de CD33 sigue constante pero la expresión de CD19 baja hasta niveles intermedios. Con la incorporación del motivo di-leucina proximalmente en el endodominio del CAR anti-CD19, la expresión del CAR de CD33 sigue constante, pero la expresión de CD19 baja hasta niveles bajos.

Figura 4: Señal de retención de la cola citosólica de E3/19K

Se generó una construcción que co-expresa un CAR anti-CD19 y un anti-CD33 usando un péptido de tipo FMD-2A. También se generaron dos variantes de esta construcción: en la primera variante, los últimos 6aa de E3/19K (DEKKMP), que se vio que eran decisivos para su capacidad de retención en el aparato de Golgi/ER, se unieron al extremo carboxilo del endodominio del CAR anti-CD33; en la segunda variante, la cola citosólica completa de la proteína E3/19K del adenovirus se unió al extremo carboxilo del endodominio del CAR anti-CD33

Figura 5: Funcionalidad de la señal de retención de E3/19K

Las construcciones mostradas en la figura 4 se transfirieron a células 293T y se evaluó el nivel de expresión de los dos CAR traducidos por la construcción mediante citometría de flujo usando las fusiones CD19-Fc y CD33-Fc recombinantes. Se observó una clara retención cuando la proteína E3/19K del adenovirus de longitud completa, o el motivo DEKKMP se colocaba en el receptor anti-CD33. Los niveles de expresión del receptor anti-CD19 no se vieron afectados.

Descripción detallada

La invención es como se define en las reivindicaciones.

La presente invención proporciona una construcción de ácido nucleico que codifica una poliproteína Z que comprende la siguiente estructura:

A-X-B

en la que;

A y B son primera y segunda proteínas transmembrana que, cuando la construcción de ácido nucleico se introduce en una célula, se expresan en la superficie celular; y X es un sitio de escisión de proteasa o un péptido autoescindible,

en donde la primera o la segunda proteína transmembrana comprende una señal de retención intracelular heteróloga.

PROTEÍNA TRANSMEMBRANA

La presente invención permite la modulación de la expresión relativa de una proteína de superficie transmembrana. La proteína de superficie transmembrana es una proteína que, en ausencia de una señal de retención intracelular, se expresa en la superficie celular. Cuando se expresa en la superficie celular, al menos un dominio de la proteína transmembrana es exoplasmático (es decir, en el exterior de la célula).

La proteína transmembrana puede ser una proteína de un solo dominio transmembrana, es decir, puede comprender un único dominio transmembrana o puede comprender múltiples dominios transmembrana.

Las proteínas transmembrana pueden clasificarse por topología, es decir, con referencia a la posición de los dominios N y C-terminales. Las proteínas transmembrana de tipos I, II y III son moléculas de un solo dominio transmembrana, mientras que las proteínas transmembrana de tipo IV son moléculas de múltiples dominios transmembrana. Las proteínas transmembrana de tipo I están ancladas a la membrana lipídica con una secuencia de anclaje de detención-transferencia y tienen sus dominios N-terminales dirigidos a la luz del RE durante la síntesis (y al espacio extracelular, cuando la forma madura se localiza en la membrana plasmática). Los tipos II y III están anclados con una secuencia de señal-anclaje, dirigiéndose el tipo II a la luz del RE con su dominio C-terminal, mientras que el tipo III tiene sus dominios N-terminales dirigidos a la luz del RE. El tipo IV se subdivide en IV-A, con sus dominios N-terminales dirigidos al citosol y IV-B, con un dominio N-terminal dirigido a la luz.

La o las proteínas transmembrana de la presente invención pueden ser de cualquiera de los tipos I-IV.

El dominio transmembrana puede ser cualquier estructura de proteína que sea termodinámicamente estable en una membrana. Esta es normalmente una hélice alfa que comprende varios residuos hidrófobos.

- 5 El dominio transmembrana de cualquier proteína transmembrana se puede usar para suministrar la porción transmembrana. Los expertos en la materia pueden determinar la presencia y la extensión de un dominio transmembrana de una proteína utilizando el algoritmo TMHMM <http://www.cbs.dtu.dk/services/TMHMM-2.0/>). Adicionalmente, dado que el dominio transmembrana de una proteína es una estructura relativamente simple, es decir, una secuencia de polipéptidos prevista para formar una hélice alfa hidrófoba de longitud suficiente para abarcar la
- 10 membrana, también se puede usar un dominio TM diseñado artificialmente (el documento US7052906 B1 describe componentes transmembrana sintéticos).

El dominio transmembrana puede obtenerse de CD28, que proporciona buena estabilidad.

- 15 La estructura y el procesamiento de las proteínas transmembrana de tipo I son bien conocidos en la técnica. Dichas proteínas comprenden normalmente un dominio extracelular, un dominio transmembrana y un endodominio intracelular y son moléculas de un solo dominio transmembrana con una sola hélice α que atraviesa la membrana celular.

- 20 Las proteínas transmembrana de tipo I normalmente tienen un péptido señal que es rápidamente reconocido por el retículo endoplásmico (RE) y, por lo tanto, la proteína en traducción se redirige rápidamente al RE. Una hélice hidrófoba bloquea y después ancla la proteína en la membrana del RE.

- 25 Como se ha mencionado anteriormente, Las proteínas transmembrana de tipo I están ancladas a la membrana lipídica con una secuencia de anclaje de detención-transferencia. La secuencia de detención-transferencia detiene la translocación adicional del polipéptido y actúa como un anclaje transmembrana.

- 30 Como se emplea en esta memoria, la expresión proteína transmembrana de tipo I abarca cualquier proteína que comprenda un dominio transmembrana de tipo I y una secuencia de anclaje de detención-transferencia y se dirige, en ausencia de una señal de retención intracelular exógena, a la expresión en la superficie celular.

- 35 Se conocen en la técnica diversas proteínas transmembrana de tipo 1 que son adecuadas para su uso en la presente invención. Dichas proteínas incluyen, pero sin limitación, receptores inhibidores, receptores estimuladores, receptores de citocinas y proteínas G.

La o las proteínas transmembrana pueden ser una cadena α o β del receptor de linfocitos T.

La o las proteínas transmembrana pueden ser un receptor quimérico de antígenos (CAR).

- 40 Los CAR son proteínas que injertan un dominio de unión a antígeno en la función efectora de un linfocito T. Su forma habitual es la de una proteína de dominio transmembrana de tipo I con un extremo amino terminal que reconoce el antígeno, un espaciador, un dominio transmembrana, todo ello conectado a un endodominio compuesto que transmite señales de supervivencia y activación de linfocitos T.

- 45 El dominio de unión a antígeno puede proceder de un anticuerpo o mimético de anticuerpo, o puede ser otra entidad que se une específicamente al antígeno, tal como un ligando.

- 50 La forma más habitual de estas moléculas son fusiones de fragmentos variables de cadena sencilla (scFv) derivados de anticuerpos monoclonales que reconocen un antígeno diana, fusionados mediante un espaciador y un dominio transmembrana a un endodominio de señalización. Dichas moléculas dan como resultado la activación de los linfocitos T en respuesta al reconocimiento por el scFv de su diana. Cuando los linfocitos T expresan dicho CAR, reconocen y eliminan células diana que expresan el antígeno diana. Se han desarrollado varios CAR contra antígenos asociados al tumor, y las estrategias de transferencia adoptiva usando dichos linfocitos T que expresan CAR están actualmente en ensayo clínico para el tratamiento de diversos cánceres.

- 55 También es posible que el endodominio de señalización esté presente en una molécula separada. El término "CAR" en relación con la presente invención también abarca una molécula que comprende un dominio de unión a antígeno conectado a un dominio transmembrana. Dicho CAR puede ser capaz de interactuar con un dominio de señalización intracelular para estimular la activación de linfocitos T.

- 60 En la presente invención, cualquiera de A o B puede ser una proteína transmembrana que comprende una señal de retención intracelular heteróloga.

- 65 La mayoría de las proteínas transmembrana de interés solo están activas o están predominantemente activas cuando están en la membrana celular. Por lo tanto, hacer que una proporción de la proteína se retenga intracelularmente reduce la expresión relativa de la proteína en la superficie celular y, por lo tanto, reduce la actividad relativa de la

proteína.

SECUENCIA DE SEÑAL

- 5 La proteína transmembrana también puede comprender una secuencia señal de modo que cuando la proteína transmembrana se expresa dentro de una célula, la proteína nascente se dirige al retículo endoplásmico (RE).

10 El núcleo del péptido señal puede contener un largo tramo de aminoácidos hidrófobos que tiene tendencia a formar una única hélice alfa. El péptido señal puede empezar con un tramo corto de aminoácidos cargados positivamente, que ayuda a reforzar la topología adecuada del polipéptido durante la translocación. Al final del péptido señal normalmente hay un tramo de aminoácidos que reconoce y escinde la peptidasa señal. La peptidasa señal puede escindir durante o después de completarse la translocación, para generar un péptido señal libre y una proteína madura. Después, los péptidos señales libres se digieren por proteasas específicas.

- 15 El péptido señal puede estar en el extremo amino de la molécula.

SITIO DE ESCISIÓN

- 20 La presente construcción de ácido nucleico comprende una secuencia que codifica un sitio de escisión situado entre secuencias de ácido nucleico que codifican el primer y segundo polipéptidos, de modo que el primer y el segundo polipéptidos pueden expresarse como entidades separadas.

25 El sitio de escisión puede ser cualquier secuencia seleccionada de un sitio de escisión por proteasa o un péptido autoescindible, lo que permite que el polipéptido que comprende el primer y segundo PDI se separe.

- 30 El término "escisión" se usa en el presente documento por conveniencia, pero el sitio de escisión puede provocar que el primer y segundo PDI se separen en entidades individuales por un mecanismo que no sea la escisión clásica. Por ejemplo, para el péptido 2A autoescindible del virus de la fiebre aftosa (FMDV) (véase a continuación), se han propuesto diversos modelos para dar cuenta de la actividad de "escisión": proteólisis por una proteinasa de la célula hospedadora, autoproteólisis o un efecto traduccional (Donnelly et al (2001) J. Gen. Virol. 82:1027-1041). El mecanismo exacto de tal "escisión" no es importante para los fines de la presente invención, siempre y cuando el sitio de escisión, cuando se coloca entre el primer y segundo polipéptidos, haga que el primer y segundo polipéptidos se expresen como entidades separadas. El sitio de escisión puede ser un sitio de escisión de furina.

- 35 La furina es una enzima que pertenece a la familia de la proproteína convertasa de tipo subtilisina. Los miembros de esta familia son proproteína convertasas que procesan proteínas precursoras latentes en sus productos biológicamente activos. La furina es una serina endoproteasa dependiente de calcio que puede escindir de forma eficaz proteínas precursoras en sus sitios de procesamiento de aminoácidos básicos emparejados. Los ejemplos de sustratos de furina incluyen la hormona paratiroides, el precursor del factor de crecimiento transformante beta 1, la proalbúmina, la pro-beta-secretasa, la metaloproteínasa de matriz de membrana de tipo 1, la subunidad beta del profactor de crecimiento nervioso y el factor von Willebrand. La furina escinde proteínas justo cadena abajo de una secuencia diana de aminoácidos básicos (canónicamente, Arg-X-(Arg/Lys)-Arg) y es enriquecida en el aparato de Golgi.

- 45 El sitio de escisión puede ser un sitio de escisión del virus del grabado del tabaco (VGT).

50 La proteasa de VGT es una cisteína proteasa altamente específica de secuencia que es una proteasa de tipo quimotripsina. Es muy específica para su sitio de escisión diana y, por lo tanto, se usa con frecuencia para la escisión controlada de proteínas de fusión tanto *in vitro* como *in vivo*. El sitio de escisión del VGT consenso es ENLYFQ\S (donde 'I' denota el enlace peptídico escindido). Las células de mamífero, tales como las células humanas, no expresan la proteasa de VGT. Por tanto, en realizaciones en las que la presente construcción de ácido nucleico comprende un sitio de escisión de VGT y se expresa en una célula de mamífero, la proteasa de VGT exógena también debe expresarse en la célula de mamífero.

- 55 El sitio de escisión puede codificar un péptido autoescindible.

Un 'péptido autoescindible' se refiere a un péptido que funciona de tal modo que cuando se produce el polipéptido que comprende el primer y segundo PDI y el péptido autoescindible, se "escinde" inmediatamente o se separa en un primer y segundo polipéptidos distintos y discretos sin la necesidad de ninguna actividad de escisión externa.

- 60 El péptido autoescindible puede ser un péptido autoescindible 2A de un aftovirus o un cardiovirus. La escisión primaria 2A/2B de los aftovirus y cardiovirus está mediada por la "escisión" de 2A en su propio extremo C. En los aftovirus, tal como los virus de la fiebre aftosa (FMDV) y el virus de la rinitis equina A, la región 2A es una sección corta de aproximadamente 18 aminoácidos, que, junto con el residuo N-terminal de la proteína 2B (un resto de prolina conservado) representa un elemento autónomo capaz de mediar la "escisión" en su propio C-terminal.

- 65 Los 19 aminoácidos C-terminales de la proteína de cardiovirus más larga, junto con la prolina N-terminal de 2B median

la "escisión" con una eficacia aproximadamente igual a la secuencia del aftovirus FMDV 2a. Los cardiovirus incluyen el virus de la encefalomiocarditis (EMCV) y el virus de la encefalitis murina de Theiler (TMEV).

El análisis mutacional de EMCV y FMDV 2A ha revelado que el motivo DxExNPGP está íntimamente implicado en la actividad de "escisión" (Donnelly et al (2001) como anteriormente).

El sitio de escisión de la presente invención puede comprender la secuencia de aminoácidos: Dx₁Ex₂NPGP, donde x₁ y x₂ son cualquier aminoácido. X₁ puede seleccionarse del siguiente grupo: I, V, M y S. X₂ puede seleccionarse del siguiente grupo: T, M, S, L, E, Q y F.

Por ejemplo, el sitio de escisión puede comprender una de las secuencias de aminoácidos mostradas en la tabla 1.

Tabla 1

Motivo	Presente en:
DIETNPGP (SEQ ID NO: 1)	Picornavirus EMCB,
	EMCD, EMCPV21
DVETNPGP (SEQ ID NO: 2)	Picornavirus MENG0 y TMEBEAN; Virus de insectos DCV, ABPV
DVEMNPGP (SEQ ID NO: 3)	Picornavirus TMEGD7 y TMEBEAN
DVESNPGP (SEQ ID NO: 4)	Picornavirus FMDA10, FMDA12, FMDC1, FMD01K, FMDSAT3, FMDVSAT2, ERAV; Virus de insectos CrPV
DMESNPGP (SEQ ID NO: 5)	Picornavirus FMDV01G
DVELNPGP (SEQ ID NO: 6)	Picornavirus ERBV; Rotavirus porcino
DVEENPGP (SEQ ID NO: 7)	Picornavirus PTV-1; Virus de insectos TaV; Trypanosoma TSR1
DIELNPGP (SEQ ID NO: 8)	Rotavirus bovino, rotavirus humano
DIEQNPGP (SEQ ID NO: 9)	AP endonucleasa de Trypanosoma
DSEFNPGP (SEQ ID NO: 10)	Secuencia bacteriana <i>T. maritima</i>

El sitio de escisión, basado en una secuencia 2A puede tener, por ejemplo 15-22 aminoácidos de longitud. La secuencia puede comprender el extremo C-terminal de una proteína 2A, seguido de un resto de prolina (que corresponde a la prolina N-terminal de 2B).

Los estudios de mutaciones también han demostrado que, además de las secuencias 2A naturales, algunas variantes también son activas. El sitio de escisión puede corresponder a una secuencia variante de un polipéptido 2A de origen natural, que tiene una, dos o tres sustituciones de aminoácidos, que retiene la capacidad de inducir la "escisión" de una secuencia de poliproteína en dos o más proteínas separadas.

La secuencia de escisión puede seleccionarse de las siguientes que todas han demostrado en cierta medida ser activas (Donnelly et al (2001) como anteriormente):

LLNFDLLKLAGDVESNPGP (SEQ ID No. 11)
 LLNFDLLKLAGDVQSNPGP (SEQ ID No. 12)
 LLNFDLLKLAGDVEINPGP (SEQ ID No. 13)
 LLNFDLLKLAGDVEFNPGP (SEQ ID No. 14)
 LLNFDLLKLAGDVESHPPG (SEQ ID No. 15)
 LLNFDLLKLAGDVESEPPG (SEQ ID No. 16)
 LLNFDLLKLAGDVESQPPG (SEQ ID No. 17)
 LLNFDLLKLAGDVESNPGG (SEQ ID No. 18)

Basándose en la secuencia del motivo DxExNPGP, se han descubierto secuencias "similares a 2A" en picornavirus que no son aftovirus o cardiovirus, virus de insectos 'tipo picornavirus', rotavirus de tipo C y secuencias repetidas dentro de *Trypanosoma spp.*, y una secuencia bacteriana (Donnelly et al (2001), como anteriormente). El sitio de escisión puede comprender una de estas secuencias similar a 2A, tales como:

YHADYYKQRLIHDVEMNPGP (SEQ ID No. 19)
 HYAGYFADLLIHDVETNPGP (SEQ ID No. 20)
 QCTNYALLKLAGDVESNPGP (SEQ ID No. 21)

ATNFSLLKQAGDVEENPGP (SEQ ID No. 22)
 AARQMLLLSGDVETNPGP (SEQ ID No. 23)
 RAEGRGSLLTCGDVEENPGP (SEQ ID No. 24)
 TRAEIEDELIRAGIESNPGP (SEQ ID No. 25)
 TRAEIEDELIRADIESNPGP (SEQ ID No. 26)
 AKFQIDKILISGDVELNPGP (SEQ ID No. 27)
 SSIIRTKMLVSGDVEENPGP (SEQ ID No. 28)
 CDAQRQKLLLSGDIEQNPGP (SEQ ID No. 29)
 YPIDFGGFLVKADSEFNPGP (SEQ ID No. 30)

El sitio de escisión puede comprender la secuencia de tipo 2A mostrada como la SEQ ID NO: 24 (RAEGRGSLLTCGDVEENPGP).

Se ha demostrado que incluir una "extensión" N-terminal de entre 5 y 39 aminoácidos puede aumentar la actividad (Donnelly et al (2001) como anteriormente). En particular, la secuencia de escisión puede comprender una de las siguientes secuencias o una variante de la misma que tiene, por ejemplo, hasta 5 cambios de aminoácidos que retienen la actividad del sitio de escisión:

VTELLYRMKRAETYCPRLAIHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 31)
 LLAHPTEARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 32)
 EARHKQKIVAPVKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 33)
 APVKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 34)

SEÑAL DE RETENCIÓN INTRACELULAR

La construcción de ácido nucleico de la presente invención comprende una secuencia que codifica una proteína transmembrana que comprende una señal de retención intracelular heteróloga.

El direccionamiento de proteínas o la clasificación de proteínas es el mecanismo biológico por el cual las proteínas son transportadas a los destinos apropiados en la célula o fuera de ella. Las proteínas pueden dirigirse al espacio interior de un orgánulo, diferentes membranas intracelulares, membrana plasmática o al exterior de la célula mediante secreción. Este proceso de transporte se lleva a cabo basándose en la información de secuencia contenida en la propia proteína.

Las proteínas sintetizadas en el retículo endoplasmático (ER) rugoso de las células eucariotas utilizan la vía exocítica para el transporte a sus destinos finales. Las proteínas que carecen de señales de clasificación especiales son transportadas vectorialmente desde el ER a través del aparato de Golgi y la red trans-Golgi (TGN) a la membrana plasmática. Otras proteínas tienen señales de direccionamiento para su incorporación a orgánulos específicos de la vía exocítica, como los endosomas y los lisosomas.

Los lisosomas son orgánulos ácidos en los que las macromoléculas endógenas e internalizadas son degradadas por las hidrolasas lumenales. Las macromoléculas endógenas alcanzan el lisosoma al ser clasificadas en la TGN desde donde son transportadas a los endosomas y luego a los lisosomas.

La presente invención puede explotar las señales de direccionamiento utilizadas por una célula para clasificar proteínas en la ubicación intracelular correcta. Las señales pueden clasificarse ampliamente en los siguientes tipos:

- i) señales de endocitosis
- ii) señales de retención en el Golgi
- iii) señales de reciclaje en la TGN
- iv) señales de retención en el ER
- v) señales de clasificación lisosomal

La 'señal de retención intracelular' se refiere a una secuencia de aminoácidos que dirige la proteína en la que está incluida a un compartimento celular distinto de la membrana de la superficie celular o al exterior de la célula.

La señal de retención intracelular provoca una reducción en la cantidad de la proteína transmembrana expresada en la superficie de una célula en comparación con una proteína transmembrana de control equivalente, que no comprende una señal de retención intracelular.

En otras palabras, la proporción de proteína transmembrana traducida que comprende una señal de retención intracelular que se expresa en la superficie celular es menor que la proporción de una cantidad equivalente de una proteína transmembrana de control traducida equivalente que no comprende una señal de retención intracelular.

Por ejemplo, la cantidad de la proteína transmembrana que comprende una señal de retención intracelular que se expresa sobre la superficie de una célula puede ser menos de 75 %, menos de 50 %, menos de 25 % o menos de

10 % de la cantidad de una proteína transmembrana de control equivalente que no comprende una señal de retención intracelular.

Las construcciones que expresan una poliproteína que se escinde posteriormente por una proteasa generalmente están limitadas por el hecho de que la expresión de los péptidos a partir de la poliproteína se limita a una relación 1:1. Sin embargo, en la presente invención, la inclusión de una señal de retención intracelular en la proteína transmembrana significa que su expresión en la superficie celular puede ser modulada (por ejemplo, reducida en comparación con una proteína transmembrana de control equivalente que no comprende una señal de retención intracelular). Como tal, la relación de la proteína transmembrana que comprende la señal de retención intracelular expresada en la superficie celular en comparación con la expresión de la segunda proteína expresada en la poliproteína puede ser, por ejemplo aproximadamente 1:1,5, de 1:1,5-1:2, 1:2-1:3, 1:3-1:4, 1:4-1:5, o más de 1:5.

La cantidad de una proteína transmembrana expresada en la superficie de una célula puede determinarse usando métodos que son conocidos en la técnica, por ejemplo citometría de flujo o microscopia de fluorescencia.

La señal de retención intracelular puede apartar la proteína transmembrana de la ruta secretora durante la translocación desde la RE.

La señal de retención intracelular puede dirigir la proteína transmembrana a un compartimento o complejo intracelular. La señal de retención intracelular puede dirigir la proteína transmembrana a un compartimento intracelular unido a la membrana.

Por ejemplo, la señal de retención intracelular puede dirigir la proteína a un compartimento lisosómico, endosómico o de Golgi (red trans-Golgi, 'TGN').

Dentro de una célula normal, las proteínas que surgen de la biogénesis o la vía endocítica se clasifican en el compartimento intracelular apropiado después de un conjunto secuencial de decisiones de clasificación. En la membrana plasmática, las proteínas pueden permanecer en la superficie celular o ser internalizadas en los endosomas. En la TGN, la elección está entre ir a la membrana plasmática o ser desviado a los endosomas. En los endosomas, las proteínas pueden reciclarse a la membrana plasmática o ir a los lisosomas. Estas decisiones se rigen por la clasificación de las señales en las propias proteínas.

Los lisosomas son orgánulos celulares que contienen enzimas hidrolasas ácidas que descomponen los materiales de desecho y los desechos celulares. La membrana alrededor de un lisosoma permite que las enzimas digestivas trabajen al pH que requieren. Los lisosomas se fusionan con las vacuolas autofágicas (fagosomas) y dispensan sus enzimas en las vacuolas autofágicas, digiriendo su contenido.

Un endosoma es un compartimento limitado por membrana dentro de las células eucariotas. Es un compartimento de la vía de transporte de la membrana endocítica desde la membrana plasmática al lisosoma y proporciona un entorno para clasificar el material antes de que llegue al lisosoma degradante. Los endosomas pueden clasificarse como endosomas tempranos, endosomas tardíos o endosomas de reciclaje, dependiendo del tiempo que tarde el material endocitosado en llegar a ellos. La señal de retención intracelular usada en la presente invención puede dirigir la proteína a un compartimento endosomal tardío.

El aparato de Golgi es parte del sistema de endomembrana celular, el aparato de Golgi empaqueta proteínas dentro de la célula antes de enviarlas a su destino; es particularmente importante en el procesamiento de proteínas para la secreción.

Existe un considerable conjunto de conocimientos que ha surgido de los estudios que investigan las señales de clasificación presentes en proteínas conocidas y el efecto de alterar su secuencia y/o posición dentro de la molécula (Bonifacino and Traub (2003) Ann. Rev. Biochem. 72:395-447; Braulke and Bonifacino (2009) Biochimica and Biophysica Acta 1793:605-614; Griffith (2001) Current Biology 11:R226-R228; Mellman and Nelson (2008) Nat Rev Mol Cell Biol. 9:833-845; Dell'Angelica and Payne (2001) Cell 106:395-398; Schafer y col. (1995) EMBO J. 14:2424-2435; Trejo (2005) Mol. Pharmacol. 67:1388-1390). Numerosos estudios han demostrado que es posible insertar una o más señales de clasificación en una proteína de interés para alterar la ubicación intracelular de una proteína de interés (Pelham (2000) Meth. Enzymol. 327:279-283).

Por lo tanto, es perfectamente posible seleccionar una señal de clasificación que tenga una propiedad de localización deseada e incluirla dentro de una proteína de interés para dirigir la ubicación intracelular de esa proteína. En relación con la presente solicitud, por lo tanto, es posible seleccionar una señal de clasificación que tenga la cantidad deseada de reducción de la expresión en la membrana plasmática.

La posición óptima de la señal de clasificación en la proteína nascente de interés puede depender del tipo de proteína transmembrana (es decir, los tipos I-IV) y de si el extremo C está en el lado luminal o citoplasmático de la membrana (Goder and Spiess (2001) FEBS Lett 504:87-93). Esto puede determinarse fácilmente considerando la posición de la señal de clasificación en su proteína natural.

Los ejemplos de señales de endocitosis incluyen las del receptor de transferrina y el receptor de asialoglicoproteína.

5 Los ejemplos de señales que provocan el reciclaje TGN-endosoma incluyen aquellas que forman proteínas tales como los CI-MPR y CD-MPR, sortilina, las proteínas relacionadas con el receptor de LDL LRP3 y LRP10 y la β -secretasa, GGA1-3, LIMP-II, NCP1, mucolipn-1, sialina, GLUT8 y la cadena invariante.

10 Los ejemplos de señales de retención en la TGN incluyen las de las siguientes proteínas que están localizadas en la TGN: las enzimas de procesamiento de prohormonas furina, PC7, CPD y PAM; la glicoproteína E del herpesvirus 3 y TGN38.

15 Los ejemplos de señales de retención en el ER incluyen señales C-terminales tales como KDEL, KKXX o KXKXX y el motivo RXR(R) de los canales de potasio. Las proteínas del ER conocidas incluyen la proteína de adenovirus E19 y ERGIC53.

Los ejemplos de señales de clasificación lisosomal incluyen las que se encuentran en las proteínas de membrana lisosomal, tales como LAMP-1 y LAMP-2, CD63, CD68, endolina, DC-LAMP, cistinosina, intercambiador de fosfato de azúcar 2 y fosfatasa ácida.

20 La señal de retención intracelular puede ser de la proteína E19 de adenovirus. La señal de retención intracelular puede ser de la proteína E3/19K, que también se conoce como E3 gp 19 kDa; E19 o GP19K. La señal de retención intracelular puede comprender la cola citosólica completa de E3/19K, que se muestra como SEQ ID NO: 36; o los últimos 6 aminoácidos de esta cola, que se muestra como SEQ ID NO: 37. Los presentes inventores han mostrado que los últimos 6 aminoácidos son particularmente importantes para la retención (ejemplo 3 y figura 5)

25 SEQ ID NO: 36: KYKSRRSFIDEKKMP
SEQ ID NO: 37: DEKKMP

30 AJUSTABILIDAD

La expresión relativa en la superficie celular de una o más proteínas transmembrana se puede ajustar usando la invención mediante diversos métodos, tales como

- 35 a) alterar la posición de la señal de retención intracelular en la molécula de proteína; y/o
b) seleccionar una señal de retención intracelular particular.

La opción a) se analiza con más detalle más adelante.

40 Con respecto a la opción b), se encuentra disponible una variedad de señales de retención intracelular a partir de la gran cantidad de proteínas naturales que se clasifican a las distintas ubicaciones celulares dentro de las células eucariotas. También es posible usar señales de retención intracelular "sintéticas" que comprenden uno o más de los motivos que se encuentran en las proteínas de origen natural (véase la siguiente sección) y tienen una función de señal de clasificación similar.

45 Está disponible una cascada de intensidad de señal, dependiendo de la ubicación intracelular a la que la señal de clasificación envía la proteína relevante. En términos generales, cuanto más "intracelular" es la ubicación dirigida por la señal de clasificación, más "fuerte" es la señal en términos de disminución de la expresión relativa de la proteína.

50 Cuando una señal de clasificación dirige una proteína al compartimento lisosomal, la proteína es internalizada y degradada por la célula, lo que da como resultado un escape relativamente pequeño a la superficie celular. La proteína se degrada y se pierde del sistema una vez que entra en el lisosoma. Por lo tanto, las señales de clasificación lisosomal, tales como LAMP1, son las "más fuertes" en términos de reducción de la expresión relativa en la superficie celular.

55 Cuando una señal de clasificación dirige una proteína para ser retenida en el ER, solo una proporción muy pequeña de la proteína llega a la superficie celular. Por lo tanto, las señales de retención o reciclaje del ER, tales como la señal ERGIC-53 y KKFF son las siguientes más fuertes, en términos de reducción de la expresión relativa en la superficie celular.

60 Cuando una señal de clasificación dirige una proteína a los compartimentos endosomal, de Golgi o de la TGN, es probable que la proteína se recicle en cierta medida entre la TGN, el compartimento endosomal y la membrana plasmática. Estas señales proporcionan un nivel más limitado de reducción de la expresión ya que una proporción significativa de la proteína aún alcanzará la membrana plasmática.

65 En general, la reducción en la expresión vista con señales de clasificación conocidas se puede resumir de la siguiente manera:

Señales de clasificación lisosomal>señales de retención/reciclaje en el ER>señales de retención/reciclaje en la

TGN>señales de endocitosis.

La ajustabilidad usando diferentes señales de clasificación y/o diferentes posiciones de señales de clasificación dentro de la proteína es especialmente útil cuando se considera la expresión de múltiples proteínas, cada una con su propia expresión relativa. Por ejemplo, considérese una construcción de ácido nucleico que codifica una poliproteína que tiene la siguiente estructura:

A-X-B-Y-C

en la cual

A, B y C son polipéptidos, al menos dos de los cuales son proteínas transmembrana que, cuando la construcción de ácido nucleico se introduce en una célula, se expresan en la superficie celular y que comprenden una señal de retención intracelular heteróloga; y

X e Y son cada uno un sitio de escisión por proteasa o un péptido autoescindible.

Por ejemplo, B y C pueden ser proteínas transmembrana que comprenden una señal de retención intracelular heteróloga. Si se desea que A, B y C se expresen de tal manera que los niveles relativos sean $A > B > C$, entonces A puede no tener señal de retención intracelular, B puede tener una señal de retención intracelular heteróloga que provoque que una pequeña proporción de proteína B sea retenida en la célula (es decir, que no se exprese en la superficie celular) y C puede tener una señal de retención intracelular heteróloga que provoque que una gran proporción de proteína C sea retenida en la célula.

Como se explica a continuación, se pueden lograr cantidades diferentes de retención intracelular, que conducen a diferentes cantidades de expresión en la superficie celular:

- (a) usando diferentes señales de retención intracelular para las proteínas; y/o
- (b) teniendo la señal de retención intracelular localizada en una posición diferente en las proteínas.

TIPOS DE SEÑAL

En la técnica se conocen numerosas proteínas que incluyen una señal de retención intracelular y se dirigen a un compartimento intracelular.

La señal de retención intracelular puede ser una señal de retención de una proteína que reside en el compartimento lisosomal, endosomal o del aparato de Golgi.

Las señales de retención intracelular son bien conocidas en la técnica (véase, por ejemplo, Bonifacino y

Traub; Annu. Rev. Biochem.; 2003; 72; 395-447).

La señal de retención intracelular puede ser una señal de clasificación basada en tirosina, una señal de clasificación basada en dileucina, una señal de grupo ácido, una señal de evitación lisosomal, una señal tipo NPFX'(1,2)D, una señal KDEL, KKX'X' o KX'KX'X' (en la que X' es cualquier aminoácido).

Las señales de clasificación basadas en tirosina median la internalización rápida de las proteínas transmembrana de la membrana plasmática y el direccionamiento de las proteínas a los lisosomas (Bonifacino & Traub; como anteriormente). Dos tipos de señales de clasificación basadas en tirosina están representados por los motivos de consenso NPX'Y e YX'X'Z' (en el que Z' es un aminoácido con una cadena lateral hidrófoba voluminosa).

Las señales NPX'Y se han desarrollado para mediar la internalización rápida de las proteínas transmembrana tipo I, se producen en familias, tales como los miembros del receptor LDL, la integrina β , y las familias de la proteína precursora de β -amiloide.

Se proporcionan ejemplos de señales NPX'Y en la tabla 2.

Tabla 2 - Señales NPX'Y

Proteína	Especie	Secuencia
Receptor LDL	Humana	Tm-10-INFDNPVYQKTT-29
LRP1 (1)	Humana	Tm-21-VEIGNPTYKMYE-64
LRP1 (2)	Humana	Tm-55-TNFTNPVYATLY-33
LRP1	<i>Drosophila</i>	Tm-43-GNFANPVYESMY-38
LRP1 (1)	<i>C. elegans</i>	Tm-54-TTFTNPVYELED-91
LRP1 (2)	<i>C. elegans</i>	Tm-140-LRVDNPLYDPDS-4

(continuación)

Proteína	Especie	Secuencia
Megalina (1)	Humana	Tm-70-IIFENPMYSARD-125
Megalina (2)	Humana	Tm-144-TNFENPIYAQME-53
Integrina 13-1 (1)	Humana	Tm-18-DTGENPIYKSAV-11
Integrina 13-1 (2)	Humana	Tm-30-TTVVNPKEYGK
Integrina 13 (1)	<i>Drosophila</i>	Tm-26-WDTENPIYKQAT-11
Integrina 13 (2)	<i>Drosophila</i>	Tm-35-STFKNPMYAGK
APLP1	Humana	Tm-33-HGYENPTYRFLE-3
APP	Humana	Tm-32-NGYENPTYKFFE-4
Similar a APP	<i>Drosophila</i>	Tm-38-NGYENPTYKYFE-3
Receptor de insulina	Humana	Tm-36-YASSNPEYLSAS-379
Receptor EGR (1)	Humana	Tm-434-GSVQNPVYHNQP-96
Receptor EGR (2)	Humana	Tm-462-TAVGNPEYLNTV-68
Receptor EGR (3)	Humana	Tm-496-ISLDNPDYQQDF-34

Los números entre paréntesis indican motivos que están presentes en más de una copia dentro de la misma proteína. Las señales en esta y otras tablas deben considerarse ejemplos. Los restos clave se indican en negrita. Se indican los números de aminoácidos antes (es decir, amino terminal) y después (es decir, carboxilo terminal) de las señales Abreviaturas: Tm, transmembrana; LDL, lipoproteína de baja densidad; LRP1, proteína relacionada con el receptor LDL 1; APP, proteína precursora de amiloide 13; APLP1, Proteína similar a APP 1.

Las señales de tipo YX'X'Z' se encuentran en receptores endocíticos tales como el receptor de transferrina y el receptor de asialoglicoproteína, receptores de clasificación intracelular como los CI-MPR y CD-MPR, proteínas de membrana lisosomales tales como LAMP-1 y LAMP-2, y proteínas TGN tales como TGN38 y furina, así como en proteínas localizadas en orgánulos endosomales-lisosomales especializados tales como compartimentos de procesamiento de antígeno (por ejemplo, HLA-DM) y gránulos citotóxicos (por ejemplo, GMP-17). Las señales de tipo YX'X'Z' están involucradas en la internalización rápida de proteínas de la membrana plasmática. Sin embargo, su función no se limita a la endocitosis, ya que los mismos motivos se han implicado en el direccionamiento de las proteínas transmembrana a los lisosomas y a los orgánulos relacionados con los lisosomas.

Se proporcionan ejemplos de señales de tipo YX'X'Z' en la tabla 3.

Tabla 3 - Señales de tipo YX'X'Z'

Proteína	Especie	Secuencia
LAMP-1	Humana	Tm-RKRSHAGYQTI
LAMP-2a	Humana	Tm-KHHHAGYEQF
LAMP-2a	Pollo	Tm-KKHHNTGYEQF
LAMP-2b	Pollo	Tm-RRKSRTGYQSV
LAMP-2c	Pollo	Tm-RRKSYAGYQTL
LAMP	<i>Drosophila</i>	Tm-RRRSTSRGYMSF
LAMP	Lombriz de tierra	Tm-RKRSTRRGYESV
CD63	Humana	Tm-KSIRSGYEVN
GMP-17	Humana	Tm-HCGGPRPGYETL
GMP-17	Ratón	Tm-HCRTTRAEYETL
CD68	Humana	Tm-RRRPSAYQAL
CD1b	Humana	Tm-RRRSYQNIP
CD1c	Humana	Tm-KKHCSYQDIL
CD1d	Ratón	Tm-RRRSAYQDIR
CD1	Rata	Tm-RKRRRSYQDIM
Endolina	Rata	Tm-KFCKSKERNYHTL
Endolina	<i>Drosophila</i>	Tm-KFYKARNERNYHTL
TSC403	Humana	Tm-KIRLRCQSSGYQRI
TSC403	Ratón	Tm-KIRQRHQSSAYQRI
Cistinosina	Humana	Tm-HFCLYRKRPYDQLN
Portador de solutos putativo	Humana	Tm-12-SLSRSGSYKEI
TRP-2	Humana	Tm-RRLRKGYTPLMET-11
HLA-DM	Humana	Tm-RRAGHSSYTPLPGS-9
LmpA	Dictyostelium	Tm-KKLRQKQKQGYQAIINNE
Proteína lisosomal putativa	Dictyostelium	Tm-RSKSNQNNQSYNLIQL
LIMP-II	Dictyostelium	Tm-RKTFYNNNNQYNGYNIIN
Receptor de transferrina	Humana	16-PLSYTRFSLA-35-Tm
Receptor de asialoglicoproteína H1	Humana	MTKEYQDLQHL-29-Tm
CI-MPR	Humana	Tm-22-SYKYSKVNKE-132
CD-MPR	Humana	Tm-4C-PAAYRGVGGDD-16

(continuación)		
Proteína	Especie	Secuencia
CTLA-4	Humana	Tm-10-TGVYVKMPPT-16
Furina	Humana	Tm-17-LISYKGLPPE-29
TGN38	Rata	Tm-23-ASDYQRLNLKL
gp41	HIV-1	Tm-13-RQGYSPSFQT-144
Fosfatasa ácida	Humana	Tm-RMQAQPPGYRHVADGEDHA
Véase la leyenda de la tabla 1 para obtener una explicación del formato de señal		

Las señales de clasificación basadas en dileucina ([DE]X'X'X'LL[LI]) juegan papeles decisivos en la clasificación de muchas proteínas transmembrana de tipo I, tipo II y las proteínas transmembranales multipaso. Las señales de clasificación basadas en dileucina están involucradas en la internalización rápida y degradación lisosomal de las proteínas transmembrana y en el direccionamiento de proteínas a los compartimentos endosomales-lisomales tardíos. Las proteínas transmembrana que contienen formas constitutivamente activas de esta señal se localizan principalmente en los endosomas y lisosomas tardíos.

Se proporcionan ejemplos de señales de clasificación [DE]X'X'X'LL[LI] en la tabla 4.

Tabla 4 - Señales de clasificación [DE]X'X'X'LL[LI]

Proteína	Especie	Señal
CD3-Y	Humana	Tm-8-SDKQTLLPN-26
LIMP-II	Rata	Tm-11-DERAPLIRT
Nmb	Humana	Tm-37-QEKDPLLKN-7
QNR-71	Codorniz	Tm-37-TERNPLLKS-5
Pme117	Humana	Tm-33-GENSPLLSG-3
Tirosinasa	Humana	Tm-8-EEKQPLLME-12
Tirosinasa	Pez Medaka	Tm-16-GERQPLLQS-13
Tirosinasa	Pollo	Tm-8-PEIQPLLTE-13
TRP-1	Pez dorado	Tm-7-EGRQPLLGD-15
TRP-1	Humana	Tm-7-EANQPLLTD-20
TRP-1	Pollo	Tm-7-ELHQPLLTD-20
TRP-2	Pez cebra	Tm-5-REFEPLLNA-11
VMAT2	Humana	Tm-6-EEKMAILMD-29
VMAT1	Humana	Tm-6-EEKLAILSQ-32
VAchT	Ratón	Tm-10-SERDVLLDE-42
VAMP4	Humana	19-SERRNLLED-88-Tm
FcR neonatal	Rata	Tm-16-DDSGDLLPG-19
CD4	Humana	Tm-12-SQIKRLLSE-17
CD4	Gato	Tm-12-SHIKRLLSE-17
GLUT4	Ratón	Tm-17-RRTPSLLEQ-17
GLUT4	Humana	Tm-17-HRTPSLLEQ-17
IRAP	Rata	46-EPRGSRLVLR-53-Tm
Ii	Humana	MDDQRDLISNNEQLPMLGR-11-Tm
Ii	Ratón	MDDQRDLISNHEQLPILGN-10-Tm
Ii	Pollo	MAEEQRDLISSDGSSGVLP-12-Tm
Ii-1	Pez cebra	MEPDHQNESLIQRVPSAETILGR-12-Tm
Ii-2	Pez cebra	MSSEGNETPLISDQSSVNMGPQP-8-Tm
Lamp	Trypanosoma	Tm-RPRRRTEEDELPEEAEGIDPQN
Proteína de Menkes	Humana	Tm 74 PDKHSLLVGDFREDDDTAL
NPC1	Humana	Tm-13-TERERLLNF
AQP4	Humana	Tm-32-VETDDLIL-29

(continuación)

Proteína	Especie	Señal
RME-2	<i>C. elegans</i>	Tm-104-FENDSLL
Vam3p	<i>S. cerevisiae</i>	153-NEQSPLLHN-121-Tm
ALP	<i>S. cerevisiae</i>	7-SEQTRLVP-18-Tm
Gap1p	<i>S. cerevisiae</i>	Tm-23-EVDLDLLK-24
Véase la leyenda de la tabla 1 para obtener una explicación del formato de señal.		

Las señales DX'X'LL constituyen un tipo distinto de señales de clasificación basadas en dileucina. Estas señales están presentes en varios receptores transmembrana y otras proteínas que circulan entre la TGN y los endosomas, tales como los CI- y CD-MPR, sortilina, las proteínas relacionadas con el receptor de LDL LRP3 y LRP10, y la β -secretasa.

5

Se proporcionan ejemplos de señales de clasificación DX'X'LL en la tabla 5.

Tabla 5 - Señales de clasificación DX'X'LL

Proteína	Especie	Secuencia
CI-MPR	Humana	Tm-151-SFHDDSD ED LLHI
CI-MPR	Bovinos	Tm-150-TFHDDSD ED LLHV
CI-MPR	Conejo	Tm-151-SFHDDSD ED LLNI
CI-MPR	Pollo	Tm-148-SFHDDSD ED LLNV
CD-MPR	Humana	Tm-54-EESEERDDHLLPM
CD-MPR	Pollo	Tm-54-DESSERDDHLLPM
Sortilina	Humana	Tm-41-GYHDDSD ED LLLE
SorLA	Humana	Tm-41-ITGFSD ED VPMVIA
Activador de la cabeza BP	Hydra	Tm-41-INRFSD ED DEPLVVA
LRP3	Humana	Tm-237-MLEASD ED EALLVC
ST7	Humana	Tm-330-KNETSD ED EALLLC
LRP10	Ratón	Tm-235-WVVEAEDEPLLA
LRP10	Humana	Tm-237-WVAEAEDEPLLT
Beta-secretasa	Humana	Tm-9-HDDFADDISLLK
Mucolipina-1	Ratón	Tm-43-GRDSPEDHSLLVN
MHC-I no clásico	Ratón ciervo	Tm-6-VRCHPEDDRLLG
FLJ30532	Humana	Tm-83-HRVSQDDLDLLTS
GGA1	Humana	350-ASVSLLD ED ELMSL-275
GGA1	Humana	415-ASSGLDDLDLLGK-211
GGA2	Humana	408-VQNPSADRNLLDL-192
GGA3	Humana	384-NALSWLDEELLCL-326
GGA	<i>Drosophila</i>	447-TVDSIDDVPLLS ED -116

Véase la leyenda de la tabla 1 para obtener una explicación del formato de señal. Los restos de serina y treonina están subrayados.

- 10 Otra familia de motivos de clasificación es proporcionada por grupos de restos ácidos que contienen sitios para fosforilación por CKII. Este tipo de motivo a menudo se encuentra en las proteínas transmembrana que se localizan en la TGN en estado estacionario, incluidas las enzimas procesadoras de prohormonas furina, PC6B, PC7, CPD, y PAM, y la glicoproteína E del herpesvirus 3.

- 15 Se proporcionan ejemplos de señales de grupos ácidos en la tabla 6.

Tabla 6 - Señales de clasificación de grupos ácidos

Proteína	Especie	Secuencia
Furina	Ratón	Tm-31-QEECP <u>S</u> DSEED <u>E</u> G-14
PC6B (1) ^a	Ratón	Tm-39-RDRDYDEDEDDI-36
PC6B (2)	Ratón	Tm-69-LDE <u>T</u> EDDELEYDDES-4
PC7	Humana	Tm-38-KDPDEVE <u>T</u> ES-47
CPD	Humana	Tm-36-HEFQDE <u>T</u> DTEEE <u>T</u> -6
PAM	Humana	Tm-59-QEKEDDGSEEEY-12
VMAT2	Humana	Tm-35-GEDEESES <u>D</u>
VMAT1	Humana	Tm-35-GED <u>S</u> DEE <u>P</u> DHEE
VAMP4	Humana	25-LEDD <u>S</u> DEEEDF-81-Tm
Glicoproteína B	HCMV	Tm-125-KD <u>S</u> DEE <u>E</u> NV
Glicoproteína E	Herpesvirus 3	Tm-28-FED <u>S</u> ESTDTEEEF-21
Nef	HIV-1 (AAL65476)	55-LEAQEEEEV-139
Kex1p (1)	<i>S. cerevisiae</i>	Tm-29- ADDLES <u>G</u> LGAEDDLEQDEQLEG- 40
Kex1p (2)	<i>S. cerevisiae</i>	Tm-79-TEID <u>E</u> SFEMIDF
Kex2p	<i>S. cerevisiae</i>	Tm-36-TEPEEVEDFD <u>F</u> DL <u>S</u> DEDH- 61
Vps10p	<i>S. cerevisiae</i>	Tm-112-FEIEEDDVPTLEEEH-37

Véase la leyenda de la tabla 1 para obtener una explicación del formato de señal. Los restos de serina y treonina están subrayados.

^aEl número entre paréntesis es el número de motivo.

El receptor KDEL se une a la proteína en el compartimento intermedio ER-Golgi, o en el Golgi temprano y la devuelve al ER. Aunque la señal común de los mamíferos es KDEL, se ha demostrado que el receptor KDEL se une más fuertemente a la secuencia HDEL (Scheel et al; J. Biol. Chem. 268; 7465 (1993)). La señal de retención intracelular puede ser HDEL.

Las señales KX'X' y KX'KX'X' son señales de recuperación que se pueden colocar en el lado citoplasmático de una proteína de membrana tipo I. Los requisitos de secuencia de estas señales se proporcionan en detalle por Teasdale & Jackson (Annu. Rev. Cell Dev. Biol.; 12; 27 (1996)).

La señal de retención intracelular puede seleccionarse del grupo de: NPX'Y, YX'X'Z, [DE]X'X'X'L[LI], DX'X'LL, DP[FW], FX'DX'F, NPF, LZ'X'Z[DE], LLDLL, PWDLW, KDEL, HDEL, KX'X' o KX'KX'X'; en la que X' es cualquier aminoácido y Z' es un aminoácido con una cadena lateral hidrófoba voluminosa.

La señal de retención intracelular puede ser cualquier secuencia mostrada en las tablas 2 a 6.

La señal de retención intracelular puede comprender la señal de retención intracelular de proteína relacionada con tirosinasa (TYRP)-1. La señal de retención intracelular puede comprender el dominio intracelular TYRP-1. La señal de retención intracelular puede comprender la secuencia NQPLLT(D) (SEQ ID NO: 35).

TYRP1 es una proteína melanosómica bien caracterizada que se retiene en el melanosoma (un lisosoma especializado) con una eficacia >99 %. TYRP1 es una proteína transmembrana de 537 aminoácidos con un dominio luminal (1-477aa), un dominio transmembrana (478-501) y un dominio citoplasmático (502-537). Una señal de di-leucina que reside en el dominio citoplasmático provoca la retención de la proteína. Esta señal de di-leucina tiene la secuencia mostrada como SEQ ID NO: 35 (NQPLLT(D)).

La señal de retención intracelular puede estar en el endodominio de la proteína transmembrana. En otras palabras, la señal de retención intracelular puede estar en el dominio de la proteína transmembrana que estaría en el lado intracelular de la membrana celular si la proteína se expresara correctamente en la superficie celular.

El endodominio de la proteína transmembrana puede comprender al menos 100, al menos 150, al menos 200, al menos 300 o al menos 500 aminoácidos.

El endodominio que comprende la señal de retención intracelular puede estar localizado en el extremo carboxilo terminal de la proteína transmembrana. En particular, donde la proteína transmembrana comprende una secuencia señal en el extremo amino del péptido, el endodominio que comprende la señal de retención intracelular puede estar localizado en el extremo carboxilo.

La señal de retención intracelular puede ser proximal al dominio transmembrana, por ejemplo, puede estar inmediatamente conectado a él. La señal de retención intracelular puede ser distal al dominio transmembrana, por ejemplo, en el terminal carboxilo del endodominio. El posicionamiento de la señal de retención modula su actividad permitiendo "sintonizar" la expresión relativa de dos proteínas. Por ejemplo, en el caso del motivo TYRP1 di-leucina,

la colocación proximal da como resultado una expresión superficial de bajo nivel, mientras que la colocación distal da como resultado una expresión superficial intermedia, como se muestra en los Ejemplos.

POLIPÉPTIDO DE INTERÉS

5 Cualquiera de las proteínas transmembrana A o B en la poliproteína A-X-B codificada por una construcción de ácido nucleico de la invención puede ser la proteína transmembrana que comprende una señal de retención intracelular heteróloga.

10 Cualquiera de las proteínas A, B o C en la poliproteína A-X-B-Y-C codificada por una construcción de ácido nucleico de la invención puede ser una de las al menos dos proteínas transmembrana que comprenden una señal de retención intracelular heteróloga.

15 La otra secuencia de ácido nucleico puede codificar cualquier polipéptido de interés (PDI). Por ejemplo, los PDI pueden ser una proteína intracelular tal como una proteína nucleica, una proteína citoplasmática o una proteína localizada en un compartimento unido a la membrana; una proteína secretora o una proteína transmembrana.

20 Cualquiera o todos de A o B; o A, B o C de las poliproteínas codificadas por las secuencias de ácido nucleico en las construcciones definidas en el presente documento pueden ser un receptor de antígeno quimérico (CAR). Las construcciones de ácido nucleico descritas en los ejemplos codifican dos receptores de antígenos quiméricos.

Las construcciones de ácido nucleico descritas en los ejemplos codifican las siguientes poliproteínas que comprenden los diversos componentes en el orden en que se enumeran:

25 1. Poliproteína que comprende CAR anti-CD19 y CAR anti-CD33 con retención proximal de tyrp-1 en el CAR anti-CD19

Péptido de señal derivado de CD8a humana:
MSLPVTALLLPLALLLHAARP (SEQ ID No. 38)
scFv aCD19:

30 DIQMTQTTSSLSASLGDRVTISCRASQDISKYLNWYQQKPDGTVKLLIYHTSRLHSGVPSRF
SGSGSGTDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGKLEITKAGGGGSGGGGSGGG
GSGGGGSEVKLQESGPGVLVAPSQSLSVTCTVSGVSLPDYGVSWIRQPPRKGLEWLGVIW
GSETTYYN SALKSRLTIKDNSKSQVFLKMNSLQTDDTAIYYCAKHYYYGGSYAMDYWGQG
TSVTVS (SEQ ID No. 39)

35 Enlazador:
SD
CD8aSTK humana:
PTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACDI (SEQ ID No. 40)
CD28TM humana:
FWVLVVVGGVLACYSLLVTVAFIIFWV (SEQ ID No. 41)
40 Dominio intracelular de Tyrp-1 humano (señal de retención):
RARRSMDEANQPLLTDQYQCYAEEYEKLQNPNSV (SEQ ID No. 42)
Dominio intracelular de CD3zeta humana:

45 RRVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRRGRDPEMGGKPRRKNPQEG
LYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQALPPR (SEQ
ID No. 43)

45 Péptido 2A:
RAEGRGSLLTCGDVEENPGP (SEQ ID No. 24)
Péptido de señal derivado de la cadena kappa de la Ig de ratón:
MAVPTQVLGLLLLLTDA (SEQ ID No. 44)
50 scFv aCD33:

RCDIQMTQSPSSLSASVGDRVITICRASEDIYFNLVWYQQKPGKAPKLLIYDTNRLADGVPS
RFSGSGSGTQYTLTISSLQPEDFATYYCQHYKNYPLTFGGGTKEIKRSGGGGSGGGGSGG
GGGSGGGSRSEVQLVESGGGLVQPGGSLRLSCAASGFTLSNYGMHWIRQAPGKGLEW
VSSISLNGGSTYYRDSVKGRFTISRDNASTLYLQMNSLRAEDTAVYYCAAQDAYTGGYFD
YWGQGTLLTVSSM (SEQ ID No. 45)

Enlazador:
DPA

5 Bisagra y Fc derivados de la IgG1 humana con mutaciones para prevenir la asociación de FcRg (HCH2CH3pvaa):

EPKSPDKTHTCPPCPAPPVAGPSVFLFPPKPKDTLMIARTPEVTCVVDVSHEDPEVKFNW
YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVHLQDNLNGKEYCKVSNKALPAPIEKTISK
AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD
SDGSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTQKSLSLSPGK (SEQ ID No. 46)

Enlazador:
KDPK (SEQ ID No. 47)
CD148TM humana:
AVFGCIFGALVIVTVGGFIFW (SEQ ID No. 48)
Dominio intracelular de CD148 humana:

10

RKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQAADSNCGFAEEYEDLKLVGISQPKY
AAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYINANYMPGYHKKDFIATQGPLPNTL
KDFWRMVWEKNVYAIIMLTKEVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRD
FTVKNIQTSESHPLRQFHFTSWPDHGVPTDITLLINFRYLVRDYMKSPPESPILVHCSAGV
GRTGTGFIADRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSK
VDLIYQNTTAMTIYENLAPVTTFGKTNGYIA (SEQ ID No. 49)

15

2. Poliproteína que comprende CAR anti-CD19 y CAR anti-CD33 con retención distal de tyrp-1 en el CAR anti-CD19

Péptido de señal derivado de CD8a humana:
MSLPVTALLPLALLHAARP (SEQ ID No. 38)
scFv aCD19:

20

DIQMTQTSSLSASLGDRVITISCRASQDISKYLWYQQKPDGTVKLLIYHTSRLHSGVPSRF
SGSGSGTDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGTKEITKAGGGGSGGGGSGGG
GGGGGSEVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWIRQPPRKGLEWLGVIW
GSETTYNSALKSRLLTIKDNSKSQVFLKMNSLQTDITAIYYCAKHYYYGGSYAMDYWGQG
TSVTVS (SEQ ID No. 39)

25

Enlazador:
SD
CD8aSTK humana:
PTTTPAPRPPTPAPTIASQPLSLRPEACRPAAGGAVHTRGLDFACDI (SEQ ID No. 40)

30

CD28TM humana:
FWVLVVVGGVLACYSLLVTVAFIIFWV (SEQ ID No. 41)
Dominio intracelular de CD3zeta humana:

RRVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDRRRGRDPEMGGKPRRKNPQEG
LYNELQKDKMAEAYSEIGMKGERRRGKGHDGLYQGLSTATKDTYDALHMQALPPR (SEQ
ID No. 43)

35

Dominio intracelular de Tyrp-1 humano (señal de retención):
RARRSMDEANQPLLTDQYQCYAEEYEKLQNPNSV (SEQ ID No. 42)
Péptido 2A:
RAEGRGSLTTCGDVEENPGP (SEQ ID No. 24)

Péptido de señal derivado de la cadena kappa de la Ig de ratón:
MAVPTQVLGLLLLWLTA (SEQ ID No. 44)
scFv aCD33:

RCDIQMTQSPSSLSASVGDRTITCRASEDIYFNLVWYQQKPGKAPKLLIYDTNRLADGVPS
RFSGSGSGTQYTLTISSLQPEDFATYYCQHYKNYPLTFGGGTKEIKRSGGGSGGGSGG
GGSGGGSRSEVQLVESGGGLVQPGGSLRLSCAASGFTLSNYGMHWIRQAPGKGLEW
VSSISLNGGSTYYRDSVKGRFTISRDNASTLYLQMNSLRAEDTAVYYCAAQDAYTGGYFD
YWGQGTLLTVSSM (SEQ ID No. 45)

Enlazador:
DPA

Bisagra y Fc derivados de la IgG1 humana con mutaciones para prevenir la asociación de FcRg (HCH2CH3pvaa):

EPKSPDKTHTCPPCPAPPVAGPSVFLFPPKPKDTLMIARTPEVTCVVDVSHEDPEVKFNW
YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK
AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTTPVLD
SDGSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTQKSLSLSPGK (SEQ ID No. 46)

Enlazador:
KDPK (SEQ ID No. 47)

CD148TM humana:
AVFGCIFGALVIVTVGGFIFW (SEQ ID No. 48)
Dominio intracelular de CD148 humana:

RKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQAADSNCGFAEEYEDLKLVGISQPKY
AAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYINANYMPGYHKKDFIATQGPLPNTL
KDFWRMVWEKNVYAIIMLTKEVEQGRTKCEEYWPSKQAQDYGDITVAMTSEIVLPEWTIRD
FTVKNIQTSESHPLRQFHFTSWPDHGVPTDLDLLINFRYLVRDYMKQSPPEPILVHCSAGV
GRTGTFFIAIDRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIRSQKDSK
VDLIYQNTTAMTIYENLAPVTTFGKTNGYIA (SEQ ID No. 49)

3. Poliproteína que comprende CAR anti-CD19 y CAR anti-CD33 con retención de E3/19K en el CAR anti-CD33

Péptido de señal derivado de CD8a humana:
MSLPVTALLLPLALLLHAARP (SEQ ID No. 38)
scFv aCD19:

DIQMTQTTSSLSASLGDRVTISCRASQDISKYLWYQQKPDGTVKLLIYHTSRLHSGVPSRF
SGSGSGTDYSLTISNLEQEDIATYFCQQGNTLPYTFGGGTKEITKAGGGSGGGSGGG
GGGGGSEVKLQESGPGLVAPSQSLSVTCTVSGVSLPDYGVSWIRQPPRKGLEWLGVIW
GSETTYNSALKSRLTIKDNSKSQVFLKMNSLQTDITAIYYCAKHYYYGGSYAMDYWGQG
TSVTVS (SEQ ID No. 39)

Enlazador:
SD

CD8aSTK humana:
PTTTPAPRPPTPAPTASQPLSLRPEACRPAAGGAVHTRGLDFACDI (SEQ ID No. 40)

CD28TM humana:
FWVLVVGGVLACYSLLVTVAFIIFWV (SEQ ID No. 41)
Dominio intracelular de CD3zeta humana:

RRVKFSRSADAPAYQQGQNQLYNELNLGRREEYDVLDKRRGRDPEMGGKPRRKNPQEG
LYNELQKDKMAEAYSEIGMKGERRRGKGDGLYQGLSTATKDYDALHMQALPPR (SEQ
ID No. 43)

Péptido 2A:

RAEGRGSLLTCGDVEENPGP (SEQ ID No. 24)

Péptido de señal derivado de la cadena kappa de la Ig de ratón:

MAVPTQVLGLLLLWLTDA (SEQ ID No. 44)

scFv aCD33:

RCDIQMTQSPSSLSASVGDRTITCRASEDIYFNLVWYQQKPGKAPKLLIYDTNRLADGVPS
RFGSGSGGTQYTLTISSLQPEDFATYYCQHYKNYPLTFGGGTKLEIKRSGGGGSGGGGSG
GGSGGGGSRSEVQLVESGGGLVQPGGSLRLSCAASGFTLSNYGMHWIRQAPGKGLEW
VSSISLNGGSTYYRDSVKGRFTISRDNASTLYLQMNSLRAEDTAVYYCAAQDAYTGGYFD
YWGQGTLLTVSSM (SEQ ID No. 45)

Enlazador:

DPA

Bisagra y Fc derivados de la IgG1 humana con mutaciones para prevenir la asociación de FcRg (HCH2CH3pvaa):

EPKSPDKTHTCPPCPAPPVAGPSVFLFPPKPKDTLMIARTPEVTCVVDVSHEDPEVKFNW
YVDGVEVHNAKTKPREEQYNSTYRVVSVLTVLHQDWLNGKEYCKVSNKALPAPIEKTISK
AKGQPREPQVYTLPPSRDELTKNQVSLTCLVKGFYPSDIAVEWESNGQPENNYKTTPPVLD
SDGSFFLYSKLTVDKSRWQQGNVFCFSVMHEALHNHYTQKSLSLSPGK (SEQ ID No. 46)

Enlazador:

KDPK (SEQ ID No. 47)

CD148TM humana:

AVFGCIFGALVIVTVGGFIFW (SEQ ID No. 48)

Dominio intracelular de CD148 humana:

RKKRKDAKNNEVSFSQIKPKKSKLIRVENFEAYFKKQQAADSNCGFAEEYEDLKLVGISQPKY
AAELAENRGKNRYNNVLPYDISRVKLSVQTHSTDDYINANYMPGYHSHKDFIATQGPLPNTL
KDFWRMVWEKNVYAIIMLTKEVEQGRKCEEWPSKQAQDYGDITVAMTSEIVLPEWTIRD
FTVKNIQTSSESHPLRQFHFTSWPDHGVPTDILLINFRYLVRDYMKSPPESPILVHCSAGV
GRTGTGFIADRLIYQIENENTVDVYGIVYDLRMHRPLMVQTEDQYVFLNQCVLDIVRSQKDSK
VDLIYQNTTAMTIYENLAPVTTFGKTNGYIA (SEQ ID No. 49)

Cola citosólica adenovírica de E3/19K:

KYKSRRSFIDEKKMP (SEQ ID No. 36)

En la poliproteína 3 anterior, la cola citosólica de E3/19K se puede reemplazar con una versión truncada que tiene la secuencia: DEKKMP

Como se muestra en los ejemplos, se encontró que, para la señal de retención de tyrp-1, se podrían lograr bajos niveles de expresión colocando la señal de retención entre "CD28TM humana" y el "dominio intracelular de CD3zeta" en la secuencia proporcionada anteriormente, mientras que se podrían lograr niveles medios de expresión colocando la señal de retención entre el "dominio intracelular de CD3zeta humano" y el "péptido 2A".

La señal de retención de E3/19K provocó una reducción en la expresión del CAR anti-CD33 cuando se colocó en el extremo C del CAR anti-CD33.

VECTOR

La presente invención también proporciona un vector que comprende una construcción de ácido nucleico según el primer aspecto de la invención.

Dicho vector se puede usar para introducir la construcción de ácido nucleico en una célula hospedadora de modo que exprese el primer y el segundo polipéptido.

El vector puede, por ejemplo, ser un plásmido o un vector viral, tal como un vector retroviral o un vector lentiviral, o un vector basado en transposón o ARNm sintético.

El vector puede tener la capacidad de transfectar o transducir una célula de mamífero, por ejemplo, un linfocito T.

CÉLULA

5 La presente invención proporciona además una célula que comprende una construcción de ácido nucleico o vector de la presente invención que expresa el primer y segundo polipéptido codificado por la secuencia de ácido nucleico.

La célula puede ser cualquier célula eucariota capaz de expresar una proteína transmembrana en la superficie celular, tal como una célula inmunológica.

PROTEÍNA

10 La presente divulgación también proporciona una proteína transmembrana que comprende un dominio extracelular, un dominio transmembrana y un endodominio en donde el endodominio comprende una señal de retención intracelular exógena como se define en el presente documento.

'Exógena' significa que la señal de retención intracelular no forma parte de la secuencia de tipo silvestre de la proteína transmembrana. La secuencia de tipo silvestre se refiere a la secuencia de aminoácidos de la proteína que aparece habitualmente en la naturaleza.

20 Los presentes inventores han demostrado que la inclusión de una señal de retención intracelular exógena en el endodominio de una proteína transmembrana provoca que la proteína transmembrana se dirija a un compartimento intracelular. Como tal, la cantidad de proteína transmembrana expresada en la superficie celular se reduce en comparación con una proteína transmembrana equivalente que no incluye una señal de retención intracelular en el dominio transmembrana.

MÉTODO

30 En un aspecto adicional, la presente divulgación se refiere a un método para modular la expresión relativa en la superficie celular de una proteína transmembrana expresada a partir de una única construcción de ácido nucleico como una poliproteína con una segunda proteína; incluyendo una señal de retención intracelular en la secuencia de ácido nucleico que codifica la proteína transmembrana.

35 La invención se describirá adicionalmente a continuación por medio de ejemplos, que pretenden servir para ayudar a un experto en la materia a poner en práctica la invención y que de ninguna manera pretenden limitar el alcance de la invención.

Ejemplos

40 Ejemplo 1 - Análisis de las señales de retención lisosómica de TYRP1

La capacidad de la señal de retención de la proteína relacionada con la tirosinasa 1 (TYRP1) para provocar la retención de un polipéptido cuando se encuentra en el contexto de un endodominio más complejo se determinó usando varias construcciones (figura 2). La construcción de tipo silvestre se comparó con las construcciones en las que se añadió 45 proteína verde fluorescente mejorada (eGFP) o se reemplazó el endodominio de TYRP1. Cuando se agregó eGFP, se colocó después o antes del endodominio nativo, por lo que la señal de retención estaba en su ubicación nativa (justo debajo de la membrana) o distal a ella.

50 Todas las construcciones se co-expresan con IRES.CD34. La tinción de células SupT1 transducidas se muestra con tinción intracelular y superficial en la figura 2.

Se encontró que el reemplazo del endodominio dio como resultado una expresión de superficie muy brillante, la introducción de eGFP después de la señal de retención casi no produjo expresión de superficie y la introducción antes de la señal de retención dio lugar a una expresión de superficie intermedia (figura 2).

55 Ejemplo 2 - Modulación de la expresión relativa de una proteína transmembrana coexpresada a partir de un casete de expresión único con una proteína separada

60 Se modificó un casete de expresión que codifica dos proteínas transmembrana CAR para que una de las proteínas CAR tuviera la señal de retención lisosomal de TYRP1 introducida proximal o distal a la membrana. La expresión de cada una de estas dos nuevas variantes en la superficie celular se comparó con la de la proteína CAR original no modificada.

65 Se aislaron PBMC de la sangre y se estimularon usando PHA e IL-2. Dos días más tarde las células fueron transducidas en placas recubiertas con retronectina con retrovirus que contienen la construcción de CAR CD19:CD33. El día 5, el nivel de expresión de los dos CAR traducidos por la construcción se evaluó mediante citometría de flujo y

se agotaron los linfocitos CD56+ (predominantemente los linfocitos NK). El día 6 las PBMC se colocaron en un co-cultivo con células diana en una relación efector:célula diana de 1:2. El día 8 el sobrenadante se recogió y se analizó mediante ELISA para determinar la secreción de IFN-gamma.

5 El patrón observado con las fusiones Tyrp1-eGFP fue de cierta reducción de la expresión de la proteína transmembrana modificada con la señal de retención distal y una reducción marcada en el caso de la señal de retención proximal. Como cabía esperar, la expresión de la segunda proteína transmembrana del casete no se alteró (figura 3).

10 Ejemplo 3 - Modulación de la expresión usando una señal de retención de la proteína E3/19K adenovírica

La proteína E3/19K del adenovirus humano es una glucoproteína transmembrana tipo I del retículo endoplásmico/Golgi que anula el transporte de la superficie celular del complejo mayor de histocompatibilidad clase I (MHC-I) y las moléculas de las cadenas A y B relacionadas con el MHC-I (MICA/B). Se identificó que el motivo de retención dependía de la cola citosólica de la proteína E3/19K de adenovirus. De manera más específica, se encontró que los últimos 6aa DEKKMP eran los más importantes para la retención. Se encontró que el posicionamiento óptimo era en el extremo c de la proteína.

20 Se modificó un casete de expresión que codifica dos proteínas transmembrana CAR, como se describe en el ejemplo 2, de modo que una de las proteínas CAR tenía el motivo de retención de la proteína E3/19K de adenovirus. En este experimento, el motivo de retención en el segundo CAR en el casete de expresión (el CAR inhibidor anti-CD33).

25 Se generaron construcciones que comprenden la cola citosólica completa de la proteína E3/19K de adenovirus o solo los últimos 6aa de E3/19K (DEKKMP), que se descubrió que eran decisivas por su capacidad de retención en el aparato de Golgi/RE (figura 4). Estas construcciones se transfectaron en células 293T y se tiñeron principalmente con una proteína CD19-Fc de conejo soluble quimérica y una proteína CD33-Fc de ratón. Posteriormente, estas células se tiñeron con un anti-Fc de conejo-FITC y un anti-Fc de ratón-APC (figura 5). Estas células muestran una retención clara cuando la proteína E3/19K de adenovirus de longitud completa, o el motivo DEKKMP, se colocaba en el receptor anti-CD33 pero no tuvo efecto sobre los niveles de expresión del receptor anti-CD19.

30 Aunque la presente invención se ha descrito en relación con realizaciones preferidas específicas, debe entenderse que la invención según se reivindica no debe considerarse limitada indebidamente a dichas realizaciones específicas.

LISTADO DE SECUENCIAS

35 <110> UCL Business PLC

<120> Construcción de ácido nucleico

40 <130> P106298PCT

<150> GB1507104,6

<151> 27/04/2015

45 <160> 197

<170> PatentIn versión 3.5

50 <210> 1

<211> 8

<212> PRT

<213> Secuencia artificial

55 <220>

<223> Sitio de escisión

<400> 1

Asp Ile Glu Thr Asn Pro Gly Pro
1 5

60 <210> 2

<211> 8

<212> PRT

<213> Secuencia artificial

65

ES 2 847 698 T3

	<220>								
	<223> Sitio de escisión								
	<400> 2								
5		Asp	Val	Glu	Thr	Asn	Pro	Gly	Pro
		1				5			
	<210> 3								
	<211> 8								
10	<212> PRT								
	<213> Secuencia artificial								
	<220>								
	<223> Sitio de escisión								
15	<400> 3								
		Asp	Val	Glu	Met	Asn	Pro	Gly	Pro
		1				5			
	<210> 4								
	<211> 8								
	<212> PRT								
	<213> Secuencia artificial								
20									
	<220>								
	<223> Sitio de escisión								
25	<400> 4								
		Asp	Val	Glu	Ser	Asn	Pro	Gly	Pro
		1				5			
30	<210> 5								
	<211> 8								
	<212> PRT								
	<213> Secuencia artificial								
35	<220>								
	<223> Sitio de escisión								
40	<400> 5								
		Asp	Met	Glu	Ser	Asn	Pro	Gly	Pro
		1				5			
	<210> 6								
	<211> 8								
	<212> PRT								
	<213> Secuencia artificial								
45	<220>								
	<223> Sitio de escisión								
50	<400> 6								
		Asp	Val	Glu	Leu	Asn	Pro	Gly	Pro
		1				5			
55	<210> 7								
	<211> 8								
	<212> PRT								
	<213> Secuencia artificial								
60	<220>								

<223> Sitio de escisión

<400> 7

5 Asp Val Glu Glu Asn Pro Gly Pro
1 5

<210> 8
<211> 8
<212> PRT
10 <213> Secuencia artificial

<220>
<223> Sitio de escisión

15 <400> 8

Asp Ile Glu Leu Asn Pro Gly Pro
1 5

<210> 9
20 <211> 8
<212> PRT
<213> Secuencia artificial

<220>
25 <223> Sitio de escisión

<400> 9

Asp Ile Glu Gln Asn Pro Gly Pro
1 5

30 <210> 10
<211> 8
<212> PRT
35 <213> Secuencia artificial

<220>
<223> Sitio de escisión

<400> 10

40 Asp Ser Glu Phe Asn Pro Gly Pro
1 5

<210> 11
45 <211> 19
<212> PRT
<213> Secuencia artificial

<220>
50 <223> Secuencia de escisión

<400> 11

Leu Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn
1 5 10 15

Pro Gly Pro

55 <210> 12
<211> 19
<212> PRT
<213> Secuencia artificial

<220>
<223> Secuencia de escisión

5 <400> 12

Leu	Leu	Asn	Phe	Asp	Leu	Leu	Lys	Leu	Ala	Gly	Asp	Val	Gln	Ser	Asn
1				5					10					15	

Pro Gly Pro

10 <210> 13
<211> 19
<212> PRT
<213> Secuencia artificial

15 <220>
<223> Secuencia de escisión

<400> 13

Leu	Leu	Asn	Phe	Asp	Leu	Leu	Lys	Leu	Ala	Gly	Asp	Val	Glu	Ile	Asn
1				5					10					15	

Pro Gly Pro

20 <210> 14
<211> 19
<212> PRT
<213> Secuencia artificial

25 <220>
<223> Secuencia de escisión

<400> 14

30

Leu	Leu	Asn	Phe	Asp	Leu	Leu	Lys	Leu	Ala	Gly	Asp	Val	Glu	Phe	Asn
1				5					10					15	

Pro Gly Pro

35 <210> 15
<211> 19
<212> PRT
<213> Secuencia artificial

<220>
<223> Secuencia de escisión

40 <400> 15

Leu	Leu	Asn	Phe	Asp	Leu	Leu	Lys	Leu	Ala	Gly	Asp	Val	Glu	Ser	His
1				5					10					15	

Pro Gly Pro

45 <210> 16
<211> 19
<212> PRT
<213> Secuencia artificial

50 <220>
<223> Secuencia de escisión

<400> 16

	Leu	Leu	Asn	Phe	Asp	Leu	Leu	Lys	Leu	Ala	Gly	Asp	Val	Glu	Ser	Glu
	1				5					10					15	

Pro Gly Pro

5

<210> 17
 <211> 19
 <212> PRT
 <213> Secuencia artificial

10

<220>
 <223> Secuencia de escisión

<400> 17

15

	Leu	Leu	Asn	Phe	Asp	Leu	Leu	Lys	Leu	Ala	Gly	Asp	Val	Glu	Ser	Gln
	1				5					10					15	

Pro Gly Pro

<210> 18
 <211> 19
 <212> PRT
 <213> Secuencia artificial

20

<220>
 <223> Secuencia de escisión

25

<400> 18

	Leu	Leu	Asn	Phe	Asp	Leu	Leu	Lys	Leu	Ala	Gly	Asp	Val	Glu	Ser	Asn
	1				5					10					15	

Pro Gly Gly

30

<210> 19
 <211> 20
 <212> PRT
 <213> Secuencia artificial

35

<220>
 <223> Sitio de escisión, secuencia de tipo 2A

<400> 19

	Tyr	His	Ala	Asp	Tyr	Tyr	Lys	Gln	Arg	Leu	Ile	His	Asp	Val	Glu	Met
	1				5					10					15	

Asn Pro Gly Pro

40

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

45

<220>
 <223> Sitio de escisión, secuencia de tipo 2A

50

<400> 20

ES 2 847 698 T3

His Tyr Ala Gly Tyr Phe Ala Asp Leu Leu Ile His Asp Ile Glu Thr
1 5 10 15

Asn Pro Gly Pro
20

5 <210> 21
<211> 20
<212> PRT
<213> Secuencia artificial

10 <220>
<223> Sitio de escisión, secuencia de tipo 2A

<400> 21

Gln Cys Thr Asn Tyr Ala Leu Leu Lys Leu Ala Gly Asp Val Glu Ser
1 5 10 15

Asn Pro Gly Pro
20

15 <210> 22
<211> 19
<212> PRT
<213> Secuencia artificial

20 <220>
<223> Sitio de escisión, secuencia de tipo 2A

<400> 22

25 Ala Thr Asn Phe Ser Leu Leu Lys Gln Ala Gly Asp Val Glu Glu Asn
1 5 10 15

Pro Gly Pro

30 <210> 23
<211> 19
<212> PRT
<213> Secuencia artificial

35 <220>
<223> Sitio de escisión, secuencia de tipo 2A

<400> 23

Ala Ala Arg Gln Met Leu Leu Leu Leu Ser Gly Asp Val Glu Thr Asn
1 5 10 15

Pro Gly Pro

40 <210> 24
<211> 20
<212> PRT
<213> Secuencia artificial

45 <220>
<223> Sitio de escisión, secuencia de tipo 2A

<400> 24

ES 2 847 698 T3

Arg Ala Glu Gly Arg Gly Ser Leu Leu Thr Cys Gly Asp Val Glu Glu
1 5 10 15

Asn Pro Gly Pro
20

<210> 25

<211> 20

<212> PRT

<213> Secuencia artificial

<220>

<223> Sitio de escisión, secuencia de tipo 2A

<400> 25

Thr Arg Ala Glu Ile Glu Asp Glu Leu Ile Arg Ala Gly Ile Glu Ser
1 5 10 15

Asn Pro Gly Pro
20

<210> 26

<211> 20

<212> PRT

<213> Secuencia artificial

<220>

<223> Sitio de escisión, secuencia de tipo 2A

<400> 26

Thr Arg Ala Glu Ile Glu Asp Glu Leu Ile Arg Ala Asp Ile Glu Ser
1 5 10 15

Asn Pro Gly Pro
20

<210> 27

<211> 20

<212> PRT

<213> Secuencia artificial

<220>

<223> Sitio de escisión, secuencia de tipo 2A

<400> 27

Ala Lys Phe Gln Ile Asp Lys Ile Leu Ile Ser Gly Asp Val Glu Leu
1 5 10 15

Asn Pro Gly Pro
20

<210> 28

<211> 20

<212> PRT

<213> Secuencia artificial

<220>

<223> Sitio de escisión, secuencia de tipo 2A

<400> 28

ES 2 847 698 T3

		Ser Ser Ile Ile Arg Thr Lys Met Leu Val Ser Gly Asp Val Glu Glu
		1 5 10 15
		Asn Pro Gly Pro
		20
5	<210> 29	
	<211> 20	
	<212> PRT	
	<213> Secuencia artificial	
	<220>	
10	<223> Sitio de escisión, secuencia de tipo 2A	
	<400> 29	
		Cys Asp Ala Gln Arg Gln Lys Leu Leu Leu Ser Gly Asp Ile Glu Gln
		1 5 10 15
		Asn Pro Gly Pro
		20
15	<210> 30	
	<211> 20	
	<212> PRT	
	<213> Secuencia artificial	
20	<220>	
	<223> Sitio de escisión, secuencia de tipo 2A	
	<400> 30	
		Tyr Pro Ile Asp Phe Gly Gly Phe Leu Val Lys Ala Asp Ser Glu Phe
		1 5 10 15
		Asn Pro Gly Pro
		20
25	<210> 31	
	<211> 57	
	<212> PRT	
30	<213> Secuencia artificial	
	<220>	
	<223> Secuencia de escisión	
35	<400> 31	
		Val Thr Glu Leu Leu Tyr Arg Met Lys Arg Ala Glu Thr Tyr Cys Pro
		1 5 10 15
		Arg Pro Leu Ala Ile His Pro Thr Glu Ala Arg His Lys Gln Lys Ile
		20 25 30
		Val Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala
		35 40 45
		Gly Asp Val Glu Ser Asn Pro Gly Pro
		50 55
40	<210> 32	
	<211> 40	
	<212> PRT	

<213> Secuencia artificial
 <220>
 <223> Secuencia de escisión
 5 <400> 32
 Leu Leu Ala Ile His Pro Thr Glu Ala Arg His Lys Gln Lys Ile Val
 1 5 10 15
 Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly
 20 25 30
 Asp Val Glu Ser Asn Pro Gly Pro
 35 40
 10 <210> 33
 <211> 33
 <212> PRT
 <213> Secuencia artificial
 15 <220>
 <223> Secuencia de escisión
 <400> 33
 Glu Ala Arg His Lys Gln Lys Ile Val Ala Pro Val Lys Gln Thr Leu
 1 5 10 15
 Asn Phe Asp Leu Leu Lys Leu Ala Gly Asp Val Glu Ser Asn Pro Gly
 20 25 30
 20 Pro
 <210> 34
 <211> 24
 <212> PRT
 25 <213> Secuencia artificial
 <220>
 <223> Secuencia de escisión
 30 <400> 34
 Ala Pro Val Lys Gln Thr Leu Asn Phe Asp Leu Leu Lys Leu Ala Gly
 1 5 10 15
 Asp Val Glu Ser Asn Pro Gly Pro
 20
 35 <210> 35
 <211> 7
 <212> PRT
 <213> Secuencia artificial
 <220>
 40 <223> Señal de retención intracelular
 <400> 35
 Asn Gln Pro Leu Leu Thr Asp
 1 5
 45

<210> 36
 <211> 15
 <212> PRT
 <213> Secuencia artificial
 5
 <220>
 <223> Señal de retención intracelular
 <400> 36
 10
 Lys Tyr Lys Ser Arg Arg Ser Phe Ile Asp Glu Lys Lys Met Pro
 1 5 10 15
 <210> 37
 <211> 6
 <212> PRT
 <213> Secuencia artificial
 15
 <220>
 <223> Señal de retención intracelular
 20
 <400> 37
 Asp Glu Lys Lys Met Pro
 1 5
 25
 <210> 38
 <211> 21
 <212> PRT
 <213> Secuencia artificial
 30
 <220>
 <223> Péptido de señal derivado de CD8a humana
 <400> 38
 Met Ser Leu Pro Val Thr Ala Leu Leu Leu Pro Leu Ala Leu Leu Leu
 1 5 10 15
 His Ala Ala Arg Pro
 20
 35
 <210> 39
 <211> 248
 <212> PRT
 <213> Secuencia artificial
 40
 <220>
 <223> scFv aCD19
 45
 <400> 39

ES 2 847 698 T3

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

5 <210> 40
 <211> 47
 <212> PRT
 <213> Secuencia artificial

10 <220>
 <223> CD8aSTK humano
 <400> 40

ES 2 847 698 T3

	Pro Thr Thr Thr Pro Ala Pro Arg Pro Pro Thr Pro Ala Pro Thr Ile
	1 5 10 15
	Ala Ser Gln Pro Leu Ser Leu Arg Pro Glu Ala Cys Arg Pro Ala Ala
	20 25 30
	Gly Gly Ala Val His Thr Arg Gly Leu Asp Phe Ala Cys Asp Ile
	35 40 45
5	
	<210> 41
	<211> 27
	<212> PRT
	<213> Secuencia artificial
10	
	<220>
	<223> CD28TM humano
	<400> 41
15	
	Phe Trp Val Leu Val Val Val Gly Gly Val Leu Ala Cys Tyr Ser Leu
	1 5 10 15
	Leu Val Thr Val Ala Phe Ile Ile Phe Trp Val
	20 25
	<210> 42
	<211> 36
	<212> PRT
	<213> Secuencia artificial
20	
	<220>
	<223> Dominio intracelular humano Typr-1 (señal de retención)
25	
	<400> 42
	Arg Ala Arg Arg Ser Met Asp Glu Ala Asn Gln Pro Leu Leu Thr Asp
	1 5 10 15
	Gln Tyr Gln Cys Tyr Ala Glu Glu Tyr Glu Lys Leu Gln Asn Pro Asn
	20 25 30
	Gln Ser Val Val
	35
30	
	<210> 43
	<211> 113
	<212> PRT
	<213> Secuencia artificial
35	
	<220>
	<223> Dominio intracelular CD3zeta humano
	<400> 43

ES 2 847 698 T3

Arg Arg Val Lys Phe Ser Arg Ser Ala Asp Ala Pro Ala Tyr Gln Gln
1 5 10 15

Gly Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg Arg Glu Glu
20 25 30

Tyr Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu Met Gly Gly
35 40 45

Lys Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn Glu Leu Gln
50 55 60

Lys Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met Lys Gly Glu
65 70 75 80

Arg Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly Leu Ser Thr
85 90 95

Ala Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala Leu Pro Pro
100 105 110

Arg

5 <210> 44
<211> 18
<212> PRT
<213> Secuencia artificial

10 <220>
<223> Péptido de señal derivado de la cadena kappa de la Ig de ratón

<400> 44

Met Ala Val Pro Thr Gln Val Leu Gly Leu Leu Leu Leu Trp Leu Thr
1 5 10 15

15 Asp Ala

<210> 45
<211> 254
<212> PRT
20 <213> Secuencia artificial

<220>
<223> scFv aCD33

25 <400> 45

Arg Cys Asp Ile Gln Met Thr Gln Ser Pro Ser Ser Leu Ser Ala Ser
1 5 10 15

Val Gly Asp Arg Val Thr Ile Thr Cys Arg Ala Ser Glu Asp Ile Tyr
20 25 30

Phe Asn Leu Val Trp Tyr Gln Gln Lys Pro Gly Lys Ala Pro Lys Leu
35 40 45

Leu Ile Tyr Asp Thr Asn Arg Leu Ala Asp Gly Val Pro Ser Arg Phe

ES 2 847 698 T3

50

55

60

Ser Gly Ser Gly Ser Gly Thr Gln Tyr Thr Leu Thr Ile Ser Ser Leu
65 70 75 80

Gln Pro Glu Asp Phe Ala Thr Tyr Tyr Cys Gln His Tyr Lys Asn Tyr
85 90 95

Pro Leu Thr Phe Gly Gln Gly Thr Lys Leu Glu Ile Lys Arg Ser Gly
100 105 110

Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser Gly Gly
115 120 125

Gly Gly Ser Arg Ser Glu Val Gln Leu Val Glu Ser Gly Gly Gly Leu
130 135 140

Val Gln Pro Gly Gly Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Phe
145 150 155 160

Thr Leu Ser Asn Tyr Gly Met His Trp Ile Arg Gln Ala Pro Gly Lys
165 170 175

Gly Leu Glu Trp Val Ser Ser Ile Ser Leu Asn Gly Gly Ser Thr Tyr
180 185 190

Tyr Arg Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala
195 200 205

Lys Ser Thr Leu Tyr Leu Gln Met Asn Ser Leu Arg Ala Glu Asp Thr
210 215 220

Ala Val Tyr Tyr Cys Ala Ala Gln Asp Ala Tyr Thr Gly Gly Tyr Phe
225 230 235 240

Asp Tyr Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Met
245 250

<210> 46

<211> 231

<212> PRT

<213> Secuencia artificial

<220>

<223> Bisagra y Fc derivados de la IgG1 humana con mutaciones para prevenir la asociación de FcRg (HCH2CH3pvaa)

<400> 46

Glu Pro Lys Ser Pro Asp Lys Thr His Thr Cys Pro Pro Cys Pro Ala

ES 2 847 698 T3

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

<210> 47
 <211> 4
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Enlazador

<400> 47

Lys Asp Pro Lys
 1

<210> 48
 <211> 21
 <212> PRT
 <213> Secuencia artificial
 5
 <220>
 <223> CD148TM humano
 <400> 48
 10
 Ala Val Phe Gly Cys Ile Phe Gly Ala Leu Val Ile Val Thr Val Gly
 1 5 10 15
 Gly Phe Ile Phe Trp
 20
 <210> 49
 <211> 341
 <212> PRT
 <213> Secuencia artificial
 15
 <220>
 <223> Dominio intracelular CD148 humano
 20
 <400> 49
 Arg Lys Lys Arg Lys Asp Ala Lys Asn Asn Glu Val Ser Phe Ser Gln
 1 5 10 15
 Ile Lys Pro Lys Lys Ser Lys Leu Ile Arg Val Glu Asn Phe Glu Ala
 20 25 30
 Tyr Phe Lys Lys Gln Gln Ala Asp Ser Asn Cys Gly Phe Ala Glu Glu
 35 40 45
 Tyr Glu Asp Leu Lys Leu Val Gly Ile Ser Gln Pro Lys Tyr Ala Ala
 50 55 60
 Glu Leu Ala Glu Asn Arg Gly Lys Asn Arg Tyr Asn Asn Val Leu Pro
 65 70 75 80
 Tyr Asp Ile Ser Arg Val Lys Leu Ser Val Gln Thr His Ser Thr Asp
 85 90 95
 Asp Tyr Ile Asn Ala Asn Tyr Met Pro Gly Tyr His Ser Lys Lys Asp

ES 2 847 698 T3

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

<210> 50

<211> 4

<212> PRT

<213> Secuencia artificial

<220>

<223> Señal de retención intracelular

<220>

<221> misc_feature

	<222> (4)..(4)				
	<223> Xaa puede ser cualquier aminoácido de origen natural				
5	<400> 50				
			Asn Pro Phe Xaa		
			1		
	<210> 51				
	<211> 4				
10	<212> PRT				
	<213> Secuencia artificial				
	<220>				
15	<223> Señal de retención intracelular				
	<400> 51				
			Lys Asp Glu Leu		
			1		
20	<210> 52				
	<211> 4				
	<212> PRT				
	<213> Secuencia artificial				
25	<220>				
	<223> Señal de retención intracelular				
	<220>				
	<221> misc_feature				
30	<222> (3)..(4)				
	<223> Xaa puede ser cualquier aminoácido de origen natural				
	<400> 52				
			Lys Lys Xaa Xaa		
35			1		
	<210> 53				
	<211> 5				
	<212> PRT				
40	<213> Secuencia artificial				
	<220>				
	<223> Señal de retención intracelular				
45	<220>				
	<221> misc_feature				
	<222> (2)..(2)				
	<223> Xaa puede ser cualquier aminoácido de origen natural				
50	<220>				
	<221> misc_feature				
	<222> (4)..(5)				
	<223> Xaa puede ser cualquier aminoácido de origen natural				
55	<400> 53				
			Lys Xaa Lys Xaa Xaa		
			1		5
60	<210> 54				
	<211> 4				
	<212> PRT				
	<213> Secuencia artificial				

	<220>	
	<223> Señal de retención intracelular	
5	<220>	
	<221> misc_feature	
	<222> (3)..(3)	
	<223> Xaa puede ser cualquier aminoácido de origen natural	
10	<400> 54	
		Asn Pro Xaa Tyr
		1
	<210> 55	
15	<211> 4	
	<212> PRT	
	<213> Secuencia artificial	
	<220>	
20	<223> Señal de retención intracelular	
	<220>	
	<221> MISC_FEATURE	
	<222> (2)..(2)	
25	<223> Xaa puede ser cualquier aminoácido de origen natural.	
	<220>	
	<221> MISC_FEATURE	
	<222> (3)..(3)	
30	<223> Xaa puede ser cualquier aminoácido de origen natural.	
	<220>	
	<221> MISC_FEATURE	
	<222> (4)..(4)	
35	<223> Xaa es un aminoácido con una cadena lateral hidrófoba voluminosa.	
	<400> 55	
		Tyr Xaa Xaa Xaa
40		1
	<210> 56	
	<211> 8	
45	<212> PRT	
	<213> Secuencia artificial	
	<220>	
	<223> Señal de retención intracelular	
50	<220>	
	<221> misc_feature	
	<222> (3)..(5)	
	<223> Xaa puede ser cualquier aminoácido de origen natural	
55	<400> 56	
		Asp Glu Xaa Xaa Xaa Leu Leu Ile
		1 5
60	<210> 57	
	<211> 5	
	<212> PRT	
	<213> Secuencia artificial	

	<220>				
	<223>	Señal de retención intracelular			
5	<220>				
	<221>	misc_feature			
	<222>	(2)..(3)			
	<223>	Xaa puede ser cualquier aminoácido de origen natural			
10	<400>	57			
			Asp	Xaa	Xaa Leu Leu
			1		5
15	<210>	58			
	<211>	4			
	<212>	PRT			
	<213>	Secuencia artificial			
20	<220>				
	<223>	Señal de retención intracelular			
	<400>	58			
25			Asp	Pro	Phe Trp
			1		
	<210>	59			
	<211>	5			
	<212>	PRT			
	<213>	Secuencia artificial			
30	<220>				
	<223>	Señal de retención intracelular			
35	<220>				
	<221>	misc_feature			
	<222>	(2)..(2)			
	<223>	Xaa puede ser cualquier aminoácido de origen natural			
40	<220>				
	<221>	misc_feature			
	<222>	(4)..(4)			
	<223>	Xaa puede ser cualquier aminoácido de origen natural			
	<400>	59			
45			Phe	Xaa	Asp Xaa Phe
			1		5
50	<210>	60			
	<211>	6			
	<212>	PRT			
	<213>	Secuencia artificial			
55	<220>				
	<223>	Señal de retención intracelular			
60	<220>				
	<221>	MISC_FEATURE			
	<222>	(2)..(2)			
	<223>	Xaa es un aminoácido con una cadena lateral hidrófoba voluminosa.			
	<220>				
	<221>	MISC_FEATURE			
	<222>	(3)..(3)			
	<223>	Xaa puede ser cualquier aminoácido de origen natural.			

<220>
 <221> MISC_FEATURE
 <222> (4)..(4)
 5 <223> Xaa es un aminoácido con una cadena lateral hidrófoba voluminosa.

 <400> 60

 Leu Xaa Xaa Xaa Asp Glu
 1 5
 10
 <210> 61
 <211> 5
 <212> PRT
 <213> Secuencia artificial
 15
 <220>
 <223> Señal de retención intracelular

 <400> 61
 20
 Leu Leu Asp Leu Leu
 1 5

 <210> 62
 <211> 5
 25 <212> PRT
 <213> Secuencia artificial

 <220>
 <223> Señal de retención intracelular
 30
 <400> 62

 Pro Trp Asp Leu Trp
 1 5
 35
 <210> 63
 <211> 4
 <212> PRT
 <213> Secuencia artificial
 40
 <220>
 <223> Sitio de escisión de furina

 <220>
 <221> MISC_FEATURE
 45 <222> (2)..(2)
 <223> Xaa puede ser cualquier aminoácido de origen natural.

 <220>
 <221> MISC_FEATURE
 50 <222> (3)..(3)
 <223> Xaa puede ser ARG o LYS

 <400> 63

 Arg Xaa Xaa Arg
 1
 55
 <210> 64
 <211> 7
 <212> PRT
 60 <213> Secuencia artificial

 <220>
 <223> Sitio de escisión del virus del grabado del tabaco (VGT)

<400> 64

Glu Asn Leu Tyr Phe Gln Ser
1 5

5

<210> 65
<211> 8
<212> PRT
<213> Secuencia artificial

10

<220>
<223> Sitio de escisión

15

<220>
<221> misc_feature
<222> (2)..(2)
<223> Xaa puede ser cualquier aminoácido de origen natural

20

<220>
<221> misc_feature
<222> (4)..(4)
<223> Xaa puede ser cualquier aminoácido de origen natural

25

<400> 65

Asp Xaa Glu Xaa Asn Pro Gly Pro
1 5

30

<210> 66
<211> 8
<212> PRT
<213> Secuencia artificial

35

<220>
<223> Sitio de escisión

40

<220>
<221> MISC_FEATURE
<222> (2)..(2)
<223> Xaa puede seleccionarse del siguiente grupo: ILE, VAL, MET y SER

45

<220>
<221> MISC_FEATURE
<222> (4)..(4)
<223> Xaa puede seleccionarse del siguiente grupo: THR, MET, SER, LEU, GLU, GLN y PHE.

50

<400> 66

Asp Xaa Glu Xaa Asn Pro Gly Pro
1 5

55

<210> 67
<211> 4
<212> PRT
<213> Secuencia artificial

60

<220>
<223> Señal de retención en el retículo endoplásmico (RE)

<220>
<221> misc_feature
<222> (2)..(2)
<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 67

ES 2 847 698 T3

Arg Xaa Arg Arg
1

5 <210> 68
<211> 12
<212> PRT
<213> *Homo sapiens*

<400> 68
10 Ile Asn Phe Asp Asn Pro Val Tyr Gln Lys Thr Thr
1 5 10

<210> 69
<211> 12
<212> PRT
<213> *Homo sapiens*

<400> 69
20 Val Glu Ile Gly Asn Pro Thr Tyr Lys Met Tyr Glu
1 5 10

<210> 70
<211> 12
<212> PRT
<213> *Homo sapiens*

<400> 70
30 Thr Asn Phe Thr Asn Pro Val Tyr Ala Thr Leu Tyr
1 5 10

<210> 71
<211> 12
<212> PRT
<213> *Drosophila sp.*
35 <400> 71

Gly Asn Phe Ala Asn Pro Val Tyr Glu Ser Met Tyr
1 5 10
40 <210> 72
<211> 12
<212> PRT
<213> *Caenorhabditis elegans*
45 <400> 72

Thr Thr Phe Thr Asn Pro Val Tyr Glu Leu Glu Asp
1 5 10

<210> 73
50 <211> 12
<212> PRT
<213> *Caenorhabditis elegans*

<400> 73
55 Leu Arg Val Asp Asn Pro Leu Tyr Asp Pro Asp Ser
1 5 10

<210> 74
60 <211> 12
<212> PRT

ES 2 847 698 T3

<213> *Homo sapiens*
 <400> 74
 5 Ile Ile Phe Glu Asn Pro Met Tyr Ser Ala Arg Asp
 1 5 10
 <210> 75
 <211> 12
 <212> PRT
 10 <213> *Homo sapiens*
 <400> 75
 15 Thr Asn Phe Glu Asn Pro Ile Tyr Ala Gln Met Glu
 1 5 10
 <210> 76
 <211> 12
 <212> PRT
 20 <213> *Homo sapiens*
 <400> 76
 25 Asp Thr Gly Glu Asn Pro Ile Tyr Lys Ser Ala Val
 1 5 10
 <210> 77
 <211> 11
 <212> PRT
 30 <213> *Homo sapiens*
 <400> 77
 35 Thr Thr Val Val Asn Pro Lys Tyr Glu Gly Lys
 1 5 10
 <210> 78
 <211> 12
 <212> PRT
 40 <213> *Drosophila* sp.
 <400> 78
 45 Trp Asp Thr Glu Asn Pro Ile Tyr Lys Gln Ala Thr
 1 5 10
 <210> 79
 <211> 11
 <212> PRT
 50 <213> *Drosophila* sp.
 <400> 79
 55 Ser Thr Phe Lys Asn Pro Met Tyr Ala Gly Lys
 1 5 10
 <210> 80
 <211> 12
 <212> PRT
 60 <213> *Homo sapiens*
 <400> 80
 65 His Gly Tyr Glu Asn Pro Thr Tyr Arg Phe Leu Glu
 1 5 10

ES 2 847 698 T3

<210> 81
 <211> 12
 <212> PRT
 <213> *Homo sapiens*
 5
 <400> 81
 Asn Gly Tyr Glu Asn Pro Thr Tyr Lys Phe Phe Glu
 1 5 10
 10
 <210> 82
 <211> 12
 <212> PRT
 <213> *Drosophila* sp.
 15
 <400> 82
 Asn Gly Tyr Glu Asn Pro Thr Tyr Lys Tyr Phe Glu
 1 5 10
 20
 <210> 83
 <211> 12
 <212> PRT
 <213> *Homo sapiens*
 25
 <400> 83
 Tyr Ala Ser Ser Asn Pro Glu Tyr Leu Ser Ala Ser
 1 5 10
 30
 <210> 84
 <211> 12
 <212> PRT
 <213> *Homo sapiens*
 <400> 84
 Gly Ser Val Gln Asn Pro Val Tyr His Asn Gln Pro
 1 5 10
 35
 <210> 85
 <211> 12
 <212> PRT
 <213> *Homo sapiens*
 40
 <400> 85
 Thr Ala Val Gly Asn Pro Glu Tyr Leu Asn Thr Val
 1 5 10
 45
 <210> 86
 <211> 12
 <212> PRT
 <213> *Homo sapiens*
 50
 <400> 86
 Ile Ser Leu Asp Asn Pro Asp Tyr Gln Gln Asp Phe
 1 5 10
 55
 <210> 87
 <211> 11
 <212> PRT
 <213> *Homo sapiens*
 60
 <400> 87

ES 2 847 698 T3

	Arg Lys Arg Ser His Ala Gly Tyr Gln Thr Ile
	1 5 10
5	<210> 88
	<211> 10
	<212> PRT
	<213> <i>Homo sapiens</i>
10	<400> 88
	Lys His His His Ala Gly Tyr Glu Gln Phe
	1 5 10
15	<210> 89
	<211> 11
	<212> PRT
	<213> <i>Gallus gallus</i>
	<400> 89
20	Lys Lys His His Asn Thr Gly Tyr Glu Gln Phe
	1 5 10
25	<210> 90
	<211> 11
	<212> PRT
	<213> <i>Gallus gallus</i>
	<400> 90
30	Arg Arg Lys Ser Arg Thr Gly Tyr Gln Ser Val
	1 5 10
35	<210> 91
	<211> 11
	<212> PRT
	<213> <i>Gallus gallus</i>
	<400> 91
40	Arg Arg Lys Ser Tyr Ala Gly Tyr Gln Thr Leu
	1 5 10
45	<210> 92
	<211> 12
	<212> PRT
	<213> <i>Drosophila sp.</i>
	<400> 92
50	Arg Arg Arg Ser Thr Ser Arg Gly Tyr Met Ser Phe
	1 5 10
55	<210> 93
	<211> 11
	<212> PRT
	<213> <i>Lumbricus terrestris</i>
	<400> 93
60	Arg Lys Arg Ser Arg Arg Gly Tyr Glu Ser Val
	1 5 10
	<210> 94
	<211> 10
	<212> PRT

ES 2 847 698 T3

<213> *Homo sapiens*

<400> 94

5 Lys Ser Ile Arg Ser Gly Tyr Glu Val Met
1 5 10

<210> 95
<211> 12
<212> PRT
10 <213> *Homo sapiens*

<400> 95

15 His Cys Gly Gly Pro Arg Pro Gly Tyr Glu Thr Leu
1 5 10

<210> 96
<211> 12
<212> PRT
20 <213> *Mus musculus*

<400> 96

25 His Cys Arg Thr Arg Arg Ala Glu Tyr Glu Thr Leu
1 5 10

<210> 97
<211> 10
<212> PRT
30 <213> *Homo sapiens*

<400> 97

35 Arg Arg Arg Pro Ser Ala Tyr Gln Ala Leu
1 5 10

<210> 98
<211> 9
<212> PRT
40 <213> *Homo sapiens*

<400> 98

45 Arg Arg Arg Ser Tyr Gln Asn Ile Pro
1 5

<210> 99
<211> 10
<212> PRT
50 <213> *Homo sapiens*

<400> 99

55 Lys Lys His Cys Ser Tyr Gln Asp Ile Leu
1 5 10

<210> 100
<211> 10
<212> PRT
60 <213> *Mus musculus*

<400> 100

ES 2 847 698 T3

		Arg Arg Arg Ser Ala Tyr Gln Asp Ile Arg
		1 5 10
5	<210> 101	
	<211> 11	
	<212> PRT	
	<213> <i>Rattus norvegicus</i>	
	<400> 101	
10		Arg Lys Arg Arg Arg Ser Tyr Gln Asp Ile Met
		1 5 10
	<210> 102	
	<211> 13	
15	<212> PRT	
	<213> <i>Rattus norvegicus</i>	
	<400> 102	
20		Lys Phe Cys Lys Ser Lys Glu Arg Asn Tyr His Thr Leu
		1 5 10
	<210> 103	
	<211> 14	
	<212> PRT	
25	<213> <i>Drosophila</i> sp.	
	<400> 103	
30		Lys Phe Tyr Lys Ala Arg Asn Glu Arg Asn Tyr His Thr Leu
		1 5 10
	<210> 104	
	<211> 14	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
35	<400> 104	
40		Lys Ile Arg Leu Arg Cys Gln Ser Ser Gly Tyr Gln Arg Ile
		1 5 10
	<210> 105	
	<211> 14	
	<212> PRT	
	<213> <i>Mus musculus</i>	
	<400> 105	
45		Lys Ile Arg Gln Arg His Gln Ser Ser Ala Tyr Gln Arg Ile
		1 5 10
	<210> 106	
	<211> 15	
50	<212> PRT	
	<213> <i>Homo sapiens</i>	
	<400> 106	
55		His Phe Cys Leu Tyr Arg Lys Arg Pro Gly Tyr Asp Gln Leu Asn
		1 5 10 15
	<210> 107	
	<211> 11	
	<212> PRT	
60	<213> <i>Homo sapiens</i>	

ES 2 847 698 T3

<400> 107

Ser Leu Ser Arg Gly Ser Gly Tyr Lys Glu Ile
1 5 10

5 <210> 108
<211> 13
<212> PRT
<213> *Homo sapiens*

10 <400> 108

Arg Arg Leu Arg Lys Gly Tyr Thr Pro Leu Met Glu Thr
1 5 10

15 <210> 109
<211> 14
<212> PRT
<213> *Homo sapiens*

20 <400> 109

Arg Arg Ala Gly His Ser Ser Tyr Thr Pro Leu Pro Gly Ser
1 5 10

25 <210> 110
<211> 18
<212> PRT
<213> *Dictyostelium discoideum*

30 <400> 110

Lys Lys Leu Arg Gln Gln Lys Gln Gln Gly Tyr Gln Ala Ile Ile Asn
1 5 10 15

Asn Glu

35 <210> 111
<211> 15
<212> PRT
<213> *Dictyostelium discoideum*

40 <400> 111

Arg Ser Lys Ser Asn Gln Asn Gln Ser Tyr Asn Leu Ile Gln Leu
1 5 10 15

45 <210> 112
<211> 17
<212> PRT
<213> *Dictyostelium discoideum*

50 <400> 112

Arg Lys Thr Phe Tyr Asn Asn Asn Gln Tyr Asn Gly Tyr Asn Ile Ile
1 5 10 15

Asn

55 <210> 113
<211> 10
<212> PRT
<213> *Homo sapiens*

60 <400> 113

ES 2 847 698 T3

		Pro Leu Ser Tyr Thr Arg Phe Ser Leu Ala
		1 5 10
5	<210> 114	
	<211> 11	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
10	<400> 114	
	Met Thr Lys Glu Tyr Gln Asp Leu Gln His Leu	
	1 5 10	
15	<210> 115	
	<211> 10	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
	<400> 115	
20		
	Ser Tyr Lys Tyr Ser Lys Val Asn Lys Glu	
	1 5 10	
25	<210> 116	
	<211> 10	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
	<400> 116	
30		
	Pro Ala Ala Tyr Arg Gly Val Gly Asp Asp	
	1 5 10	
35	<210> 117	
	<211> 10	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
	<400> 117	
40		
	Thr Gly Val Tyr Val Lys Met Pro Pro Thr	
	1 5 10	
45	<210> 118	
	<211> 10	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
	<400> 118	
50		
	Leu Ile Ser Tyr Lys Gly Leu Pro Pro Glu	
	1 5 10	
55	<210> 119	
	<211> 11	
	<212> PRT	
	<213> <i>Rattus norvegicus</i>	
	<400> 119	
60		
	Ala Ser Asp Tyr Gln Arg Leu Asn Leu Lys Leu	
	1 5 10	
	<210> 120	
	<211> 11	
	<212> PRT	

<213> Virus de la inmunodeficiencia humana tipo 1

<400> 120

5 Arg Gln Gly Tyr Ser Pro Leu Ser Phe Gln Thr
1 5 10

<210> 121

<211> 19

<212> PRT

10 <213> *Homo sapiens*

<400> 121

Arg Met Gln Ala Gln Pro Pro Gly Tyr Arg His Val Ala Asp Gly Glu
1 5 10 15

Asp His Ala

15

<210> 122

<211> 9

<212> PRT

<213> Secuencia artificial

20

<220>

<223> Señal de clasificación basada en dileucina

<220>

25 <221> misc_feature

<222> (3)..(5)

<223> Xaa puede ser cualquier aminoácido de origen natural

<400> 122

30

Asp Glu Xaa Xaa Xaa Leu Leu Leu Ile
1 5

<210> 123

<211> 9

35 <212> PRT

<213> *Homo sapiens*

<400> 123

Ser Asp Lys Gln Thr Leu Leu Pro Asn
1 5

40

<210> 124

<211> 9

<212> PRT

45 <213> *Rattus norvegicus*

<400> 124

Asp Glu Arg Ala Pro Leu Ile Arg Thr
1 5

50

<210> 125

<211> 9

<212> PRT

<213> *Homo sapiens*

55

<400> 125

Gln Glu Lys Asp Pro Leu Leu Lys Asn
1 5

ES 2 847 698 T3

5	<p><210> 126 <211> 9 <212> PRT <213> <i>Coturnix coturnix</i></p>	
	<400> 126	
10		Thr Glu Arg Asn Pro Leu Leu Lys Ser 1 5
15	<p><210> 127 <211> 9 <212> PRT <213> <i>Homo sapiens</i></p>	
	<400> 127	
20		Gly Glu Asn Ser Pro Leu Leu Ser Gly 1 5
25	<p><210> 128 <211> 9 <212> PRT <213> <i>Homo sapiens</i></p>	
	<400> 128	
30		Glu Glu Lys Gln Pro Leu Leu Met Glu 1 5
35	<p><210> 129 <211> 9 <212> PRT <213> <i>Oryzias latipes</i></p>	
	<400> 129	
40		Gly Glu Arg Gln Pro Leu Leu Gln Ser 1 5
45	<p><210> 130 <211> 9 <212> PRT <213> <i>Gallus gallus</i></p>	
	<400> 130	
50		Pro Glu Ile Gln Pro Leu Leu Thr Glu 1 5
55	<p><210> 131 <211> 9 <212> PRT <213> <i>Carassius auratus</i></p>	
	<400> 131	
60		Glu Gly Arg Gln Pro Leu Leu Gly Asp 1 5
	<p><210> 132 <211> 9 <212> PRT <213> <i>Homo sapiens</i></p>	

ES 2 847 698 T3

	<400> 132		
		Glu Ala Asn Gln Pro Leu Leu Thr Asp	
		1	5
5	<210> 133		
	<211> 9		
	<212> PRT		
	<213> <i>Gallus gallus</i>		
10	<400> 133		
		Glu Leu His Gln Pro Leu Leu Thr Asp	
		1	5
15	<210> 134		
	<211> 9		
	<212> PRT		
	<213> <i>Danio rerio</i>		
20	<400> 134		
		Arg Glu Phe Glu Pro Leu Leu Asn Ala	
		1	5
25	<210> 135		
	<211> 9		
	<212> PRT		
	<213> <i>Homo sapiens</i>		
	<400> 135		
30		Glu Glu Lys Met Ala Ile Leu Met Asp	
		1	5
35	<210> 136		
	<211> 9		
	<212> PRT		
	<213> <i>Homo sapiens</i>		
	<400> 136		
40		Glu Glu Lys Leu Ala Ile Leu Ser Gln	
		1	5
45	<210> 137		
	<211> 9		
	<212> PRT		
	<213> <i>Mus musculus</i>		
	<400> 137		
50		Ser Glu Arg Asp Val Leu Leu Asp Glu	
		1	5
55	<210> 138		
	<211> 9		
	<212> PRT		
	<213> <i>Homo sapiens</i>		
60	<400> 138		
		Ser Glu Arg Arg Asn Leu Leu Glu Asp	
		1	5
60	<210> 139		
	<211> 9		

ES 2 847 698 T3

	<212> PRT	
	<213> <i>Rattus norvegicus</i>	
5	<400> 139	
		Asp Asp Ser Gly Asp Leu Leu Pro Gly
		1 5
10	<210> 140	
	<211> 9	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
	<400> 140	
15		Ser Gln Ile Lys Arg Leu Leu Ser Glu
		1 5
20	<210> 141	
	<211> 9	
	<212> PRT	
	<213> <i>Felis catus</i>	
	<400> 141	
25		Ser His Ile Lys Arg Leu Leu Ser Glu
		1 5
30	<210> 142	
	<211> 9	
	<212> PRT	
	<213> <i>Mus musculus</i>	
	<400> 142	
35		Arg Arg Thr Pro Ser Leu Leu Glu Gln
		1 5
40	<210> 143	
	<211> 9	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
	<400> 143	
45		His Arg Thr Pro Ser Leu Leu Glu Gln
		1 5
50	<210> 144	
	<211> 10	
	<212> PRT	
	<213> <i>Rattus norvegicus</i>	
	<400> 144	
		Glu Pro Arg Gly Ser Arg Leu Leu Val Arg
		1 5 10
55	<210> 145	
	<211> 19	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
	<400> 145	

ES 2 847 698 T3

	Met Asp Asp Gln Arg Asp Leu Ile Ser Asn Asn Glu Gln Leu Pro Met	
	1 5 10 15	
	Leu Gly Arg	
5	<210> 146 <211> 19 <212> PRT <213> <i>Mus musculus</i>	
	<400> 146	
	Met Asp Asp Gln Arg Asp Leu Ile Ser Asn His Glu Gln Leu Pro Ile	
	1 5 10 15	
10	Leu Gly Asn	
	<210> 147 <211> 20 <212> PRT <213> <i>Gallus gallus</i>	
15	<400> 147	
	Met Ala Glu Glu Gln Arg Asp Leu Ile Ser Ser Asp Gly Ser Ser Gly	
	1 5 10 15	
	Val Leu Pro Ile	
	20	
20	<210> 148 <211> 23 <212> PRT <213> <i>Danio rerio</i>	
25	<400> 148	
	Met Glu Pro Asp His Gln Asn Glu Ser Leu Ile Gln Arg Val Pro Ser	
	1 5 10 15	
	Ala Glu Thr Ile Leu Gly Arg	
	20	
30	<210> 149 <211> 23 <212> PRT <213> <i>Danio rerio</i>	
35	<400> 149	
	Met Ser Ser Glu Gly Asn Glu Thr Pro Leu Ile Ser Asp Gln Ser Ser	
	1 5 10 15	
	Val Asn Met Gly Pro Gln Pro	
	20	
40	<210> 150 <211> 24 <212> PRT <213> <i>Trypanosoma brucei</i>	
	<400> 150	

ES 2 847 698 T3

	Arg Pro Arg Arg Arg Thr Glu Glu Asp Glu Leu Leu Pro Glu Glu Ala
	1 5 10 15
	Glu Gly Leu Ile Asp Pro Gln Asn
	20
5	<210> 151
	<211> 19
	<212> PRT
	<213> <i>Homo sapiens</i>
10	<400> 151
	Pro Asp Lys His Ser Leu Leu Val Gly Asp Phe Arg Glu Asp Asp Asp
	1 5 10 15
	Thr Ala Leu
15	<210> 152
	<211> 9
	<212> PRT
	<213> <i>Homo sapiens</i>
	<400> 152
20	Thr Glu Arg Glu Arg Leu Leu Asn Phe
	1 5
25	<210> 153
	<211> 8
	<212> PRT
	<213> <i>Homo sapiens</i>
	<400> 153
30	Val Glu Thr Asp Asp Leu Ile Leu
	1 5
35	<210> 154
	<211> 7
	<212> PRT
	<213> <i>Caenorhabditis elegans</i>
	<400> 154
40	Phe Glu Asn Asp Ser Leu Leu
	1 5
45	<210> 155
	<211> 9
	<212> PRT
	<213> <i>Saccharomyces cerevisiae</i>
	<400> 155
50	Asn Glu Gln Ser Pro Leu Leu His Asn
	1 5
	<210> 156
	<211> 8
	<212> PRT
	<213> <i>Saccharomyces cerevisiae</i>
	<400> 156

ES 2 847 698 T3

		Ser Glu Gln Thr Arg Leu Val Pro
		1 5
5	<210> 157	
	<211> 8	
	<212> PRT	
	<213> <i>Saccharomyces cerevisiae</i>	
10	<400> 157	
		Glu Val Asp Leu Asp Leu Leu Lys
		1 5
15	<210> 158	
	<211> 5	
	<212> PRT	
	<213> Secuencia artificial	
20	<220>	
	<223> Señal de clasificación basada en dileucina	
	<220>	
	<221> misc_feature	
	<222> (2)..(3)	
25	<223> Xaa puede ser cualquier aminoácido de origen natural	
	<400> 158	
		Asp Xaa Xaa Leu Leu
		1 5
30	<210> 159	
	<211> 13	
	<212> PRT	
	<213> <i>Homo sapiens</i>	
35	<400> 159	
		Ser Phe His Asp Asp Ser Asp Glu Asp Leu Leu His Ile
		1 5 10
40	<210> 160	
	<211> 13	
	<212> PRT	
	<213> <i>Bos taurus</i>	
45	<400> 160	
		Thr Phe His Asp Asp Ser Asp Glu Asp Leu Leu His Val
		1 5 10
50	<210> 161	
	<211> 13	
	<212> PRT	
	<213> <i>Oryctolagus cuniculus</i>	
	<400> 161	
55		Ser Phe His Asp Asp Ser Asp Glu Asp Leu Leu Asn Ile
		1 5 10
60	<210> 162	
	<211> 13	
	<212> PRT	
	<213> <i>Gallus gallus</i>	

ES 2 847 698 T3

<400> 162

5 Ser Phe His Asp Asp Ser Asp Glu Asp Leu Leu Asn Val
1 5 10

<210> 163
<211> 13
<212> PRT
<213> *Homo sapiens*

10 <400> 163

Glu Glu Ser Glu Glu Arg Asp Asp His Leu Leu Pro Met
1 5 10

15 <210> 164
<211> 13
<212> PRT
<213> *Gallus gallus*

20 <400> 164

Asp Glu Ser Glu Glu Arg Asp Asp His Leu Leu Pro Met
1 5 10

25 <210> 165
<211> 12
<212> PRT
<213> *Homo sapiens*

30 <400> 165

Gly Tyr His Asp Asp Ser Asp Glu Asp Leu Leu Glu
1 5 10

35 <210> 166
<211> 13
<212> PRT
<213> *Homo sapiens*

<400> 166

Ile Thr Gly Phe Ser Asp Asp Val Pro Met Val Ile Ala
1 5 10

40 <210> 167
<211> 13
<212> PRT
<213> *Hydra vulgaris*

45 <400> 167

Ile Asn Arg Phe Ser Asp Asp Glu Pro Leu Val Val Ala
1 5 10

50 <210> 168
<211> 13
<212> PRT
<213> *Homo sapiens*

55 <400> 168

Met Leu Glu Ala Ser Asp Asp Glu Ala Leu Leu Val Cys
1 5 10

60 <210> 169

ES 2 847 698 T3

<211> 13
 <212> PRT
 <213> *Homo sapiens*
 5 <400> 169
 Lys Asn Glu Thr Ser Asp Asp Glu Ala Leu Leu Leu Cys
 1 5 10
 <210> 170
 10 <211> 12
 <212> PRT
 <213> *Mus musculus*
 <400> 170
 15 Trp Val Val Glu Ala Glu Asp Glu Pro Leu Leu Ala
 1 5 10
 <210> 171
 <211> 12
 20 <212> PRT
 <213> *Homo sapiens*
 <400> 171
 Trp Val Ala Glu Ala Glu Asp Glu Pro Leu Leu Thr
 25 1 5 10
 <210> 172
 <211> 12
 30 <212> PRT
 <213> *Homo sapiens*
 <400> 172
 His Asp Asp Phe Ala Asp Asp Ile Ser Leu Leu Lys
 35 1 5 10
 <210> 173
 <211> 13
 <212> PRT
 <213> *Mus musculus*
 40 <400> 173
 Gly Arg Asp Ser Pro Glu Asp His Ser Leu Leu Val Asn
 1 5 10
 45 <210> 174
 <211> 12
 <212> PRT
 <213> *Peromyscus maniculatus*
 50 <400> 174
 Val Arg Cys His Pro Glu Asp Asp Arg Leu Leu Gly
 1 5 10
 55 <210> 175
 <211> 13
 <212> PRT
 <213> *Homo sapiens*
 60 <400> 175

ES 2 847 698 T3

	His Arg Val Ser Gln Asp Asp Leu Asp Leu Leu Thr Ser
	1 5 10
5	<210> 176 <211> 13 <212> PRT <213> <i>Homo sapiens</i> <400> 176
10	Ala Ser Val Ser Leu Leu Asp Asp Glu Leu Met Ser Leu
	1 5 10
15	<210> 177 <211> 13 <212> PRT <213> <i>Homo sapiens</i> <400> 177
20	Ala Ser Ser Gly Leu Asp Asp Leu Asp Leu Leu Gly Lys
	1 5 10
25	<210> 178 <211> 13 <212> PRT <213> <i>Homo sapiens</i> <400> 178
30	Val Gln Asn Pro Ser Ala Asp Arg Asn Leu Leu Asp Leu
	1 5 10
35	<210> 179 <211> 13 <212> PRT <213> <i>Homo sapiens</i> <400> 179
40	Asn Ala Leu Ser Trp Leu Asp Glu Glu Leu Leu Cys Leu
	1 5 10
45	<210> 180 <211> 13 <212> PRT <213> <i>Drosophila sp.</i> <400> 180
50	Thr Val Asp Ser Ile Asp Asp Val Pro Leu Leu Ser Asp
	1 5 10
55	<210> 181 <211> 13 <212> PRT <213> <i>Mus musculus</i> <400> 181
60	Gln Glu Glu Cys Pro Ser Asp Ser Glu Glu Asp Glu Gly
	1 5 10
60	<210> 182 <211> 13 <212> PRT <213> <i>Mus musculus</i>

ES 2 847 698 T3

<400> 182

Arg Asp Arg Asp Tyr Asp Glu Asp Asp Glu Asp Asp Ile
1 5 10

5

<210> 183
<211> 15
<212> PRT
<213> *Mus musculus*

10

<400> 183

Leu Asp Glu Thr Glu Asp Asp Glu Leu Glu Tyr Asp Asp Glu Ser
1 5 10 15

15

<210> 184
<211> 10
<212> PRT
<213> *Homo sapiens*

20

<400> 184

Lys Asp Pro Asp Glu Val Glu Thr Glu Ser
1 5 10

25

<210> 185
<211> 13
<212> PRT
<213> *Homo sapiens*

30

<400> 185

His Glu Phe Gln Asp Glu Thr Asp Thr Glu Glu Glu Thr
1 5 10

35

<210> 186
<211> 14
<212> PRT
<213> *Homo sapiens*

<400> 186

Gln Glu Lys Glu Asp Asp Gly Ser Glu Ser Glu Glu Glu Tyr
1 5 10

40

<210> 187
<211> 9
<212> PRT
<213> *Homo sapiens*

45

<400> 187

Gly Glu Asp Glu Glu Ser Glu Ser Asp
1 5

50

<210> 188
<211> 12
<212> PRT
<213> *Homo sapiens*

55

<400> 188

Gly Glu Asp Ser Asp Glu Glu Pro Asp His Glu Glu
1 5 10

60

<210> 189

<211> 11
 <212> PRT
 <213> *Homo sapiens*

5 <400> 189

Leu Glu Asp Asp Ser Asp Glu Glu Glu Asp Phe
 1 5 10

<210> 190
 <211> 9
 <212> PRT
 <213> Citomegalovirus humano

10 <400> 190

15

Lys Asp Ser Asp Glu Glu Glu Asn Val
 1 5

<210> 191
 <211> 13
 <212> PRT
 <213> Herpesvirus humano 3

20 <400> 191

25

Phe Glu Asp Ser Glu Ser Thr Asp Thr Glu Glu Glu Phe
 1 5 10

<210> 192
 <211> 9
 <212> PRT
 <213> Virus de la inmunodeficiencia humana tipo 1

30 <400> 192

35

Leu Glu Ala Gln Glu Glu Glu Glu Val
 1 5

<210> 193
 <211> 22
 <212> PRT
 <213> *Saccharomyces cerevisiae*

40 <400> 193

Ala Asp Asp Leu Glu Ser Gly Leu Gly Ala Glu Asp Asp Leu Glu Gln
 1 5 10 15

Asp Glu Gln Leu Glu Gly
 20

45

<210> 194
 <211> 12
 <212> PRT
 <213> *Saccharomyces cerevisiae*

50 <400> 194

Thr Glu Ile Asp Glu Ser Phe Glu Met Thr Asp Phe
 1 5 10

55

<210> 195
 <211> 18
 <212> PRT
 <213> *Saccharomyces cerevisiae*

<400> 195

Thr Glu Pro Glu Glu Val Glu Asp Phe Asp Phe Asp Leu Ser Asp Glu
1 5 10 15

Asp His

5

<210> 196

<211> 15

<212> PRT

<213> *Saccharomyces cerevisiae*

10

<400> 196

Phe Glu Ile Glu Glu Asp Asp Val Pro Thr Leu Glu Glu Glu His
1 5 10 15

15

<210> 197

<211> 4

<212> PRT

<213> Secuencia artificial

20

<220>

<223> Señal de retención intracelular

<400> 197

His Asp Glu Leu
1

25

REIVINDICACIONES

1. Una construcción de ácido nucleico que codifica una poliproteína que comprende la siguiente estructura:

5 A-X-B

en la cual

A y B son primera y segunda proteínas transmembrana que, cuando la construcción de ácido nucleico se introduce en una célula, se expresan en la superficie celular; y

10 X es un sitio de escisión de proteasa o un péptido autoescindible, en donde la primera o la segunda proteína transmembrana comprende una señal de retención intracelular heteróloga.

15 2. Una construcción de ácido nucleico según la reivindicación 1, en donde la señal de retención intracelular dirige la proteína transmembrana a un compartimento lisosómico, endosómico o del aparato de Golgi.

3. Una construcción de ácido nucleico según la reivindicación 1 o 2, en donde la señal de retención intracelular se selecciona del siguiente grupo: una señal de endocitosis; una señal de retención en el aparato de Golgi; una señal de reciclaje en la red trans-Golgi (TGN); una señal de retención en el retículo endoplásmico (ER); y una señal de clasificación lisosomal.

4. Una construcción de ácido nucleico según cualquiera de las reivindicaciones anteriores, en donde la señal de retención intracelular es próxima a un dominio transmembrana de la proteína transmembrana.

25 5. Una construcción de ácido nucleico según cualquiera de las reivindicaciones 1 a 3, en donde la señal de retención intracelular es distal a un dominio transmembrana de la proteína transmembrana.

6. Una construcción de ácido nucleico según cualquier reivindicación anterior, en donde X es un péptido autoescindible, un sitio de escisión de furina o un sitio de escisión del virus del grabado del tabaco.

30 7. Una construcción de ácido nucleico según cualquiera de las reivindicaciones anteriores, en donde una proteína transmembrana es un receptor de antígeno quimérico (CAR).

8. Una construcción de ácido nucleico según cualquiera de las reivindicaciones anteriores, en donde una o ambas de la primera y segunda proteínas transmembrana es una proteína de un solo dominio transmembrana.

9. Una construcción de ácido nucleico según la reivindicación 8, en donde una o ambas de la primera y segunda proteínas transmembrana es una proteína transmembrana de tipo I.

40 10. Una construcción de ácido nucleico que codifica una poliproteína que comprende la siguiente estructura:

A - X - B - Y - C

en la cual

45 A, B y C son primero, segundo y tercer polipéptidos de interés (PDI); y X e Y pueden ser iguales o diferentes, siendo cada uno de los cuales un sitio de escisión de proteasa o un péptido autoescindible, en donde al menos dos de los PDI son proteínas transmembrana que, cuando la construcción de ácido nucleico se introduce en una célula, se expresan en la superficie celular y que comprenden una señal de retención intracelular heteróloga, y en donde los al menos dos PDI Z que son proteínas transmembrana y que comprenden una señal de retención intracelular heteróloga:

55 (a) comprenden diferentes señales de retención intracelular; y/o (b) tienen la señal de retención intracelular localizada en una posición diferente en el PDI; de modo que cuando el ácido nucleico se expresa en una célula, existe una expresión relativa diferencial de las al menos dos proteínas transmembrana en la superficie celular.

11. Un vector que comprende una construcción de ácido nucleico según cualquiera de las reivindicaciones anteriores.

60 12. Una célula que comprende una construcción de ácido nucleico según una cualquiera de las reivindicaciones 1 a 10 o un vector según la reivindicación 11.

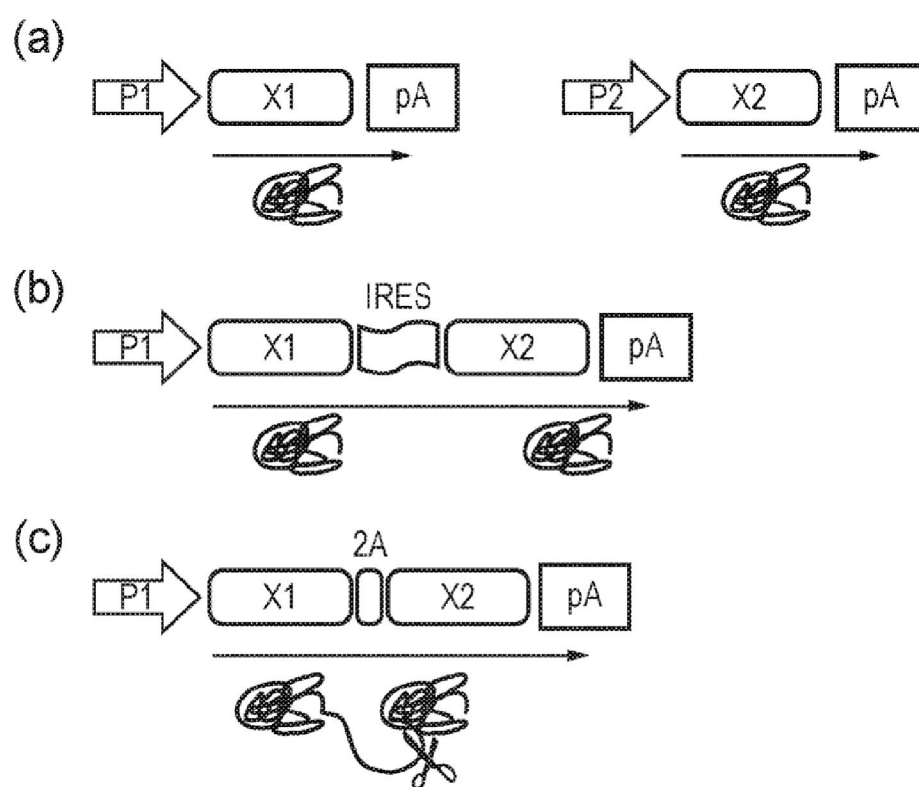


FIG. 1

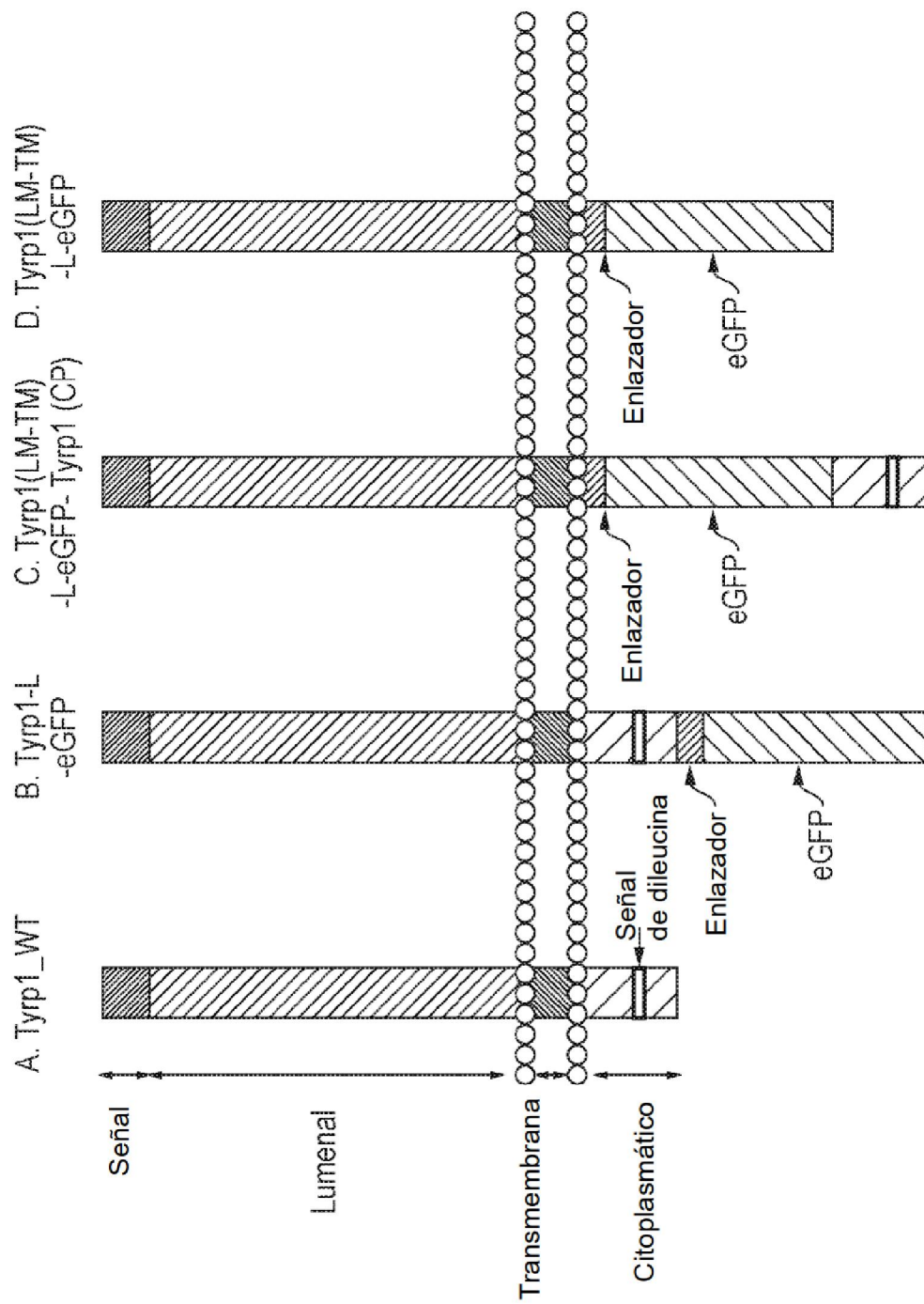


FIG. 2

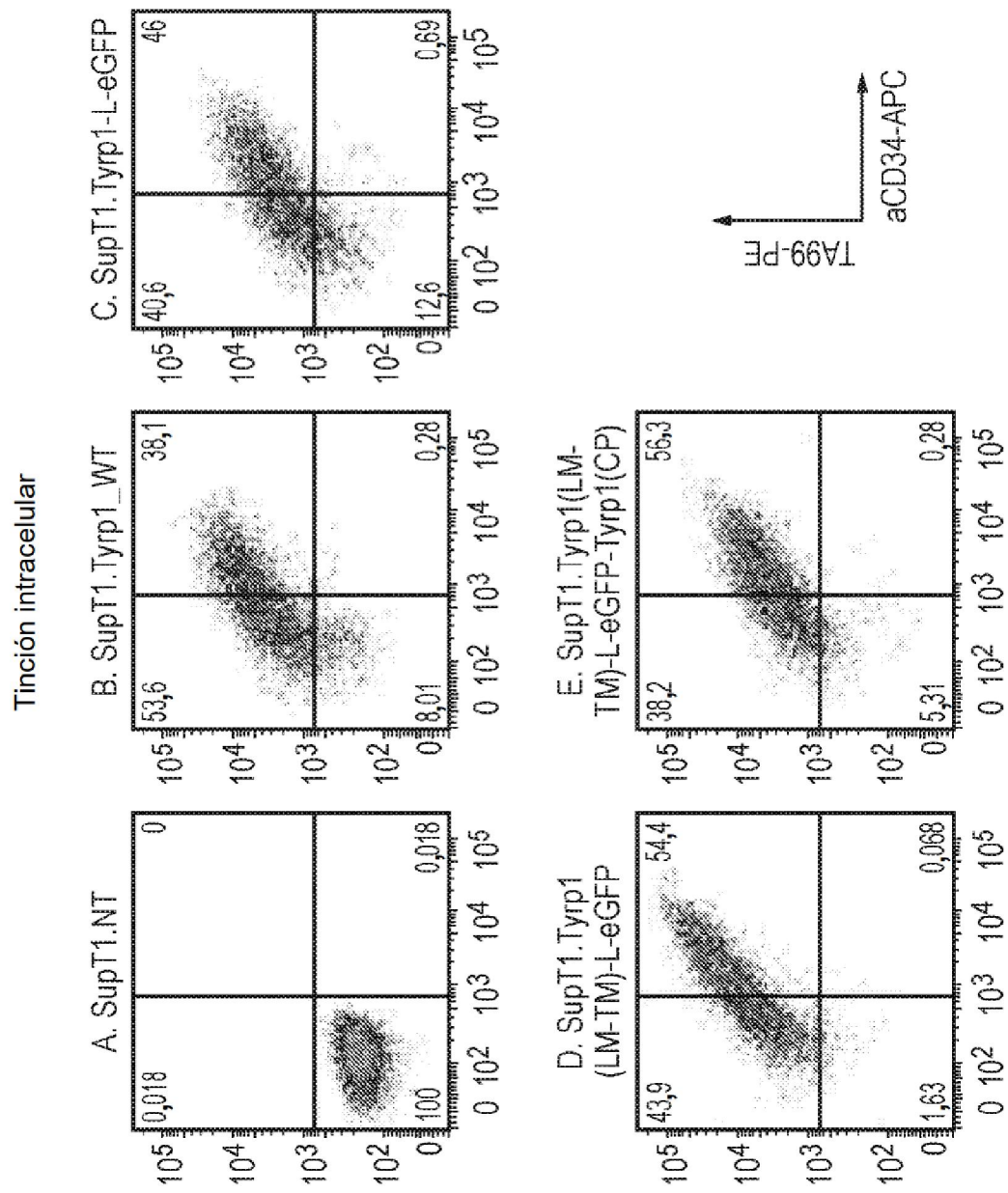


FIG 2 (Continuación)

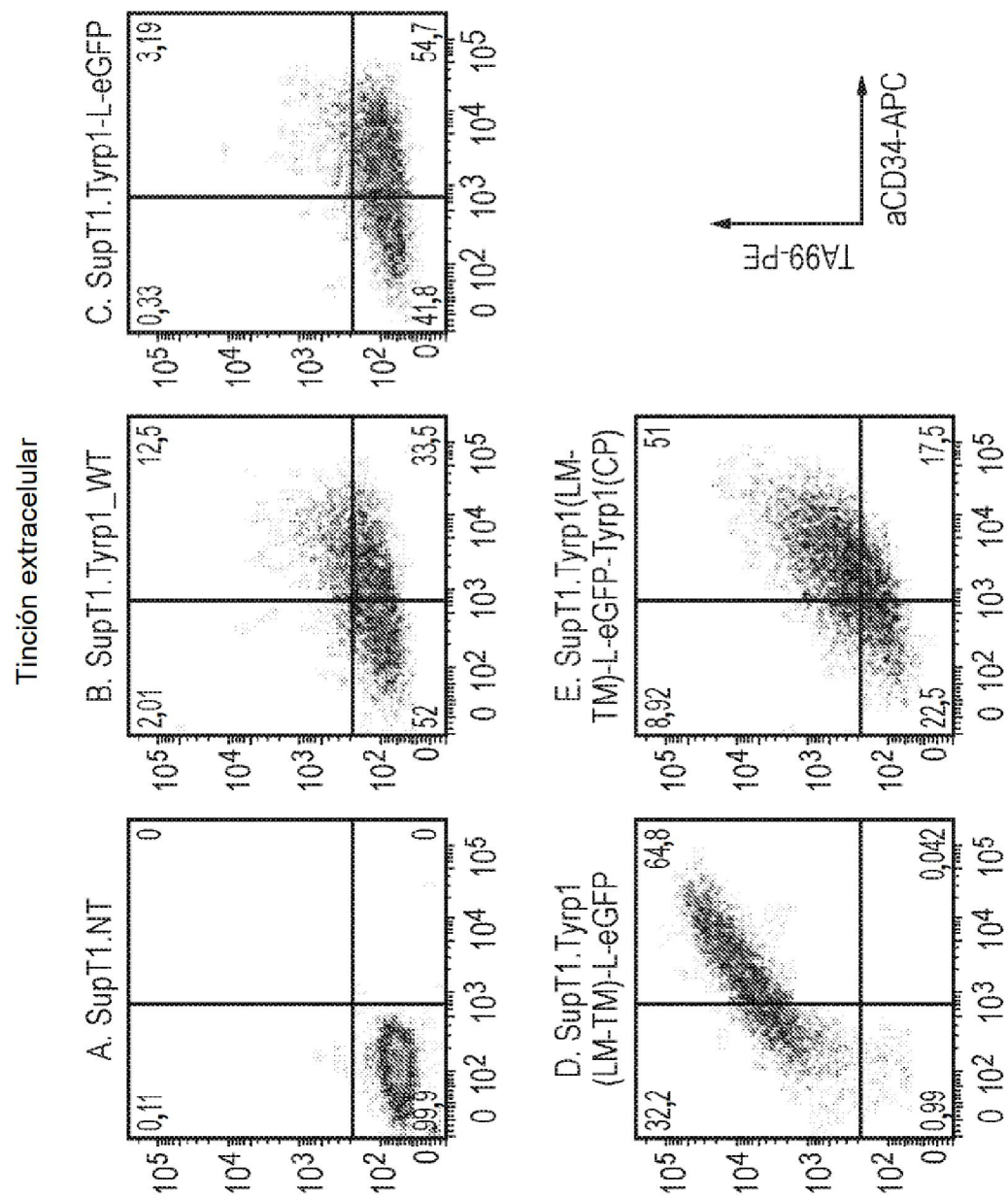


FIG 2 (Continuación)

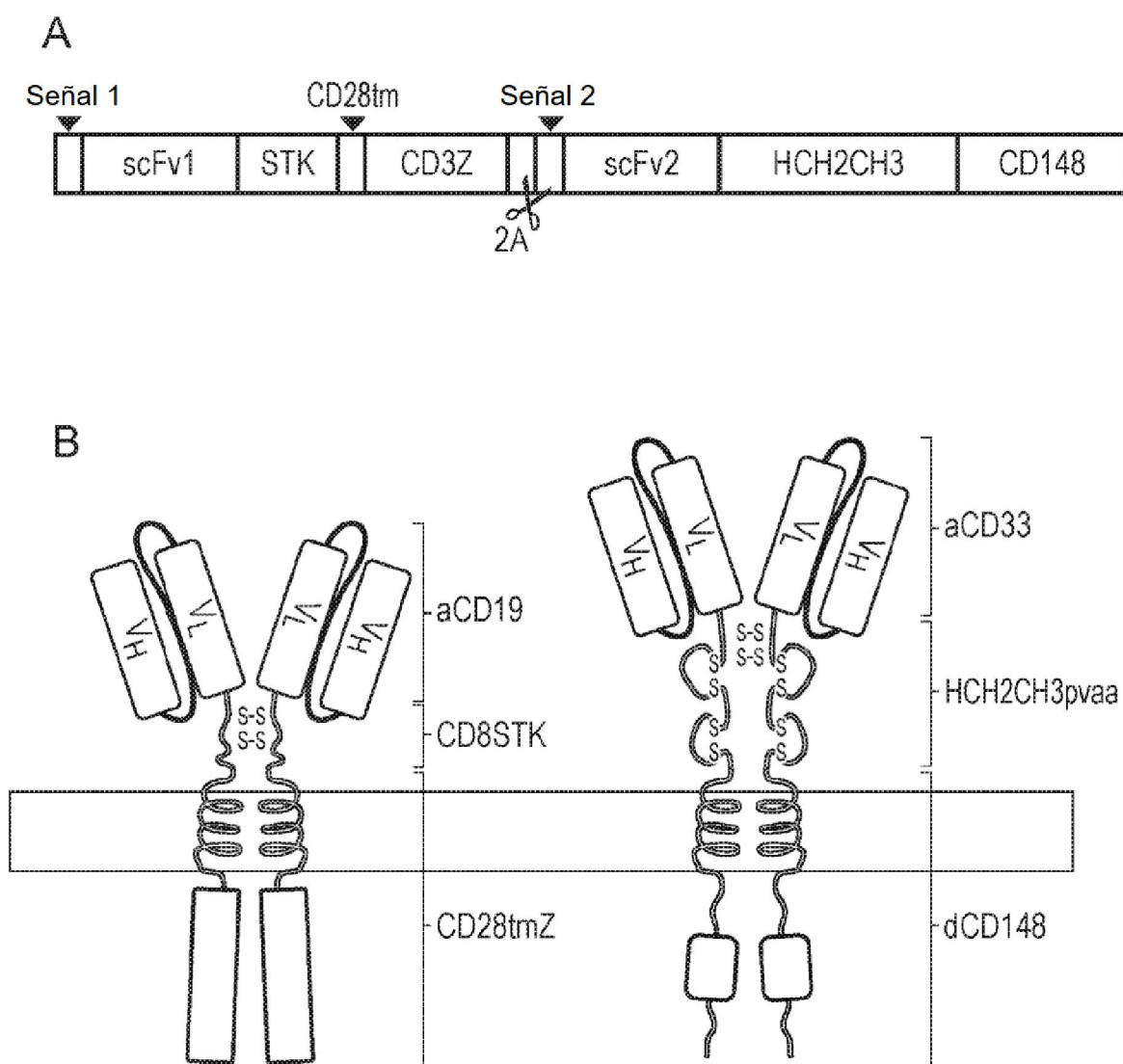


FIG. 3

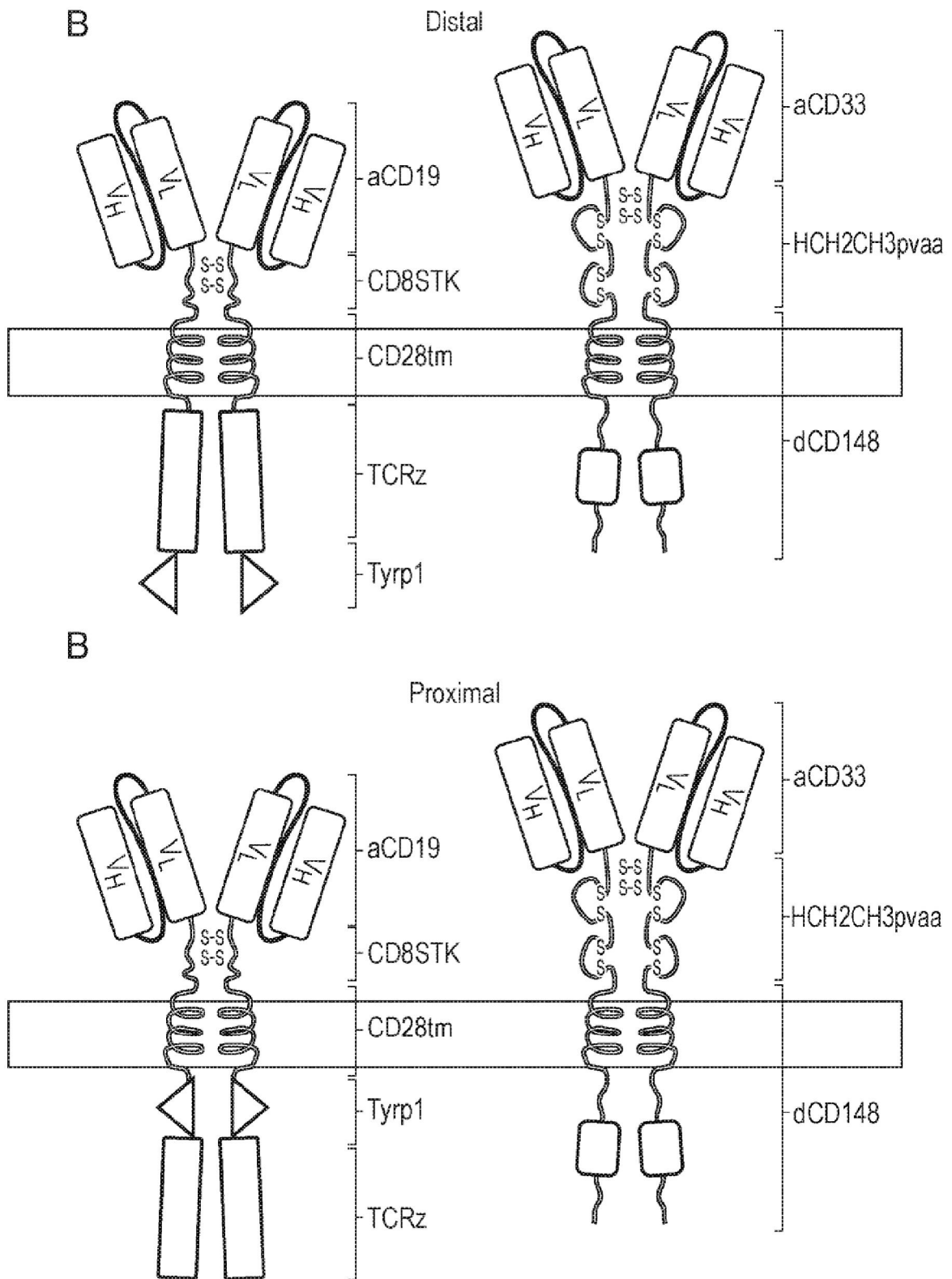


FIG 3 (Continuación)

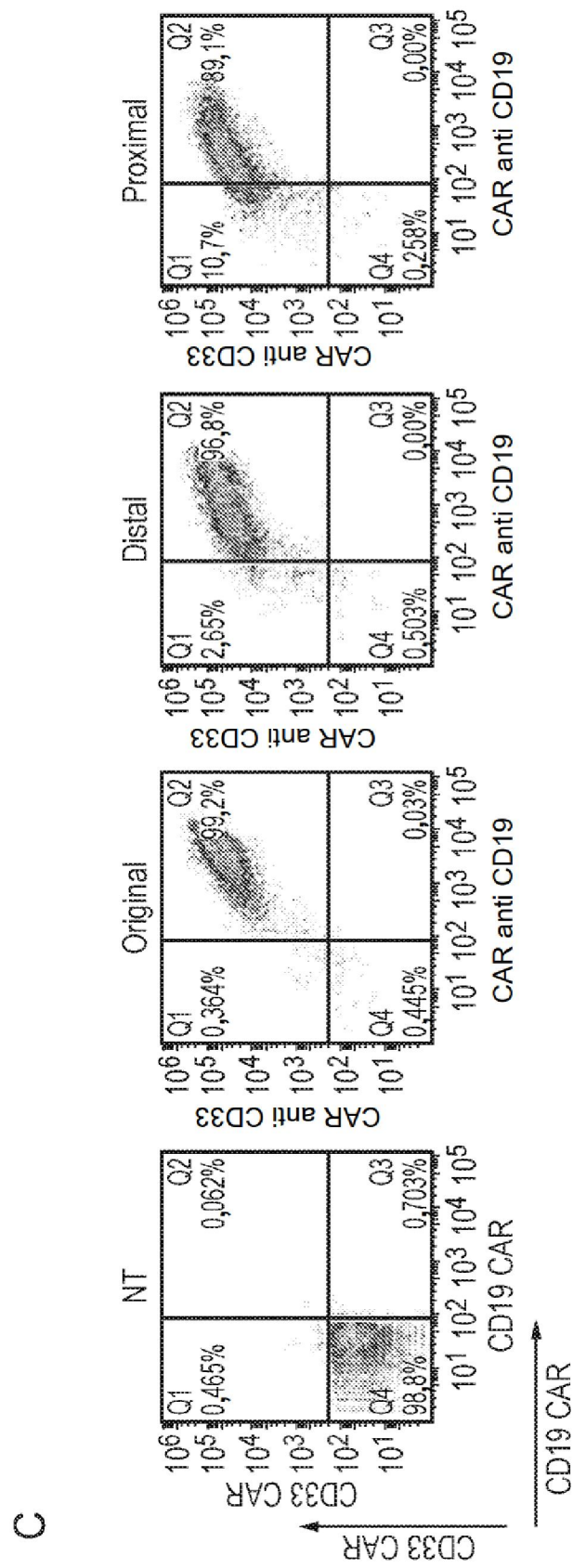


FIG 3 (Continuación)

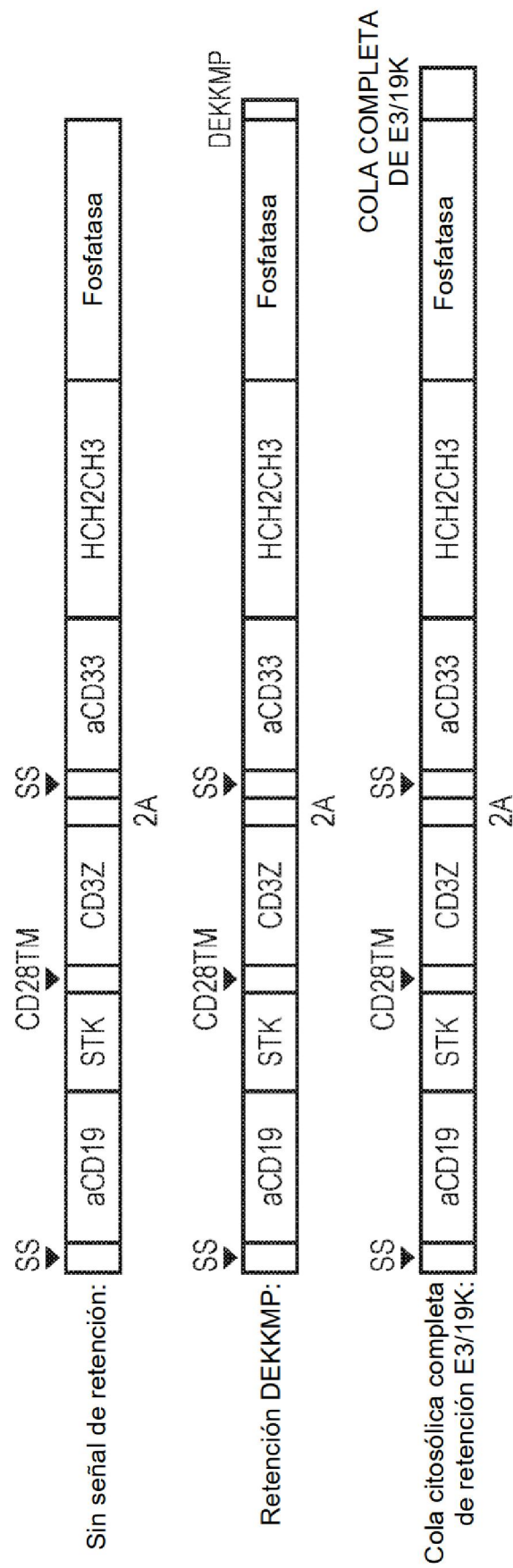


FIG. 4

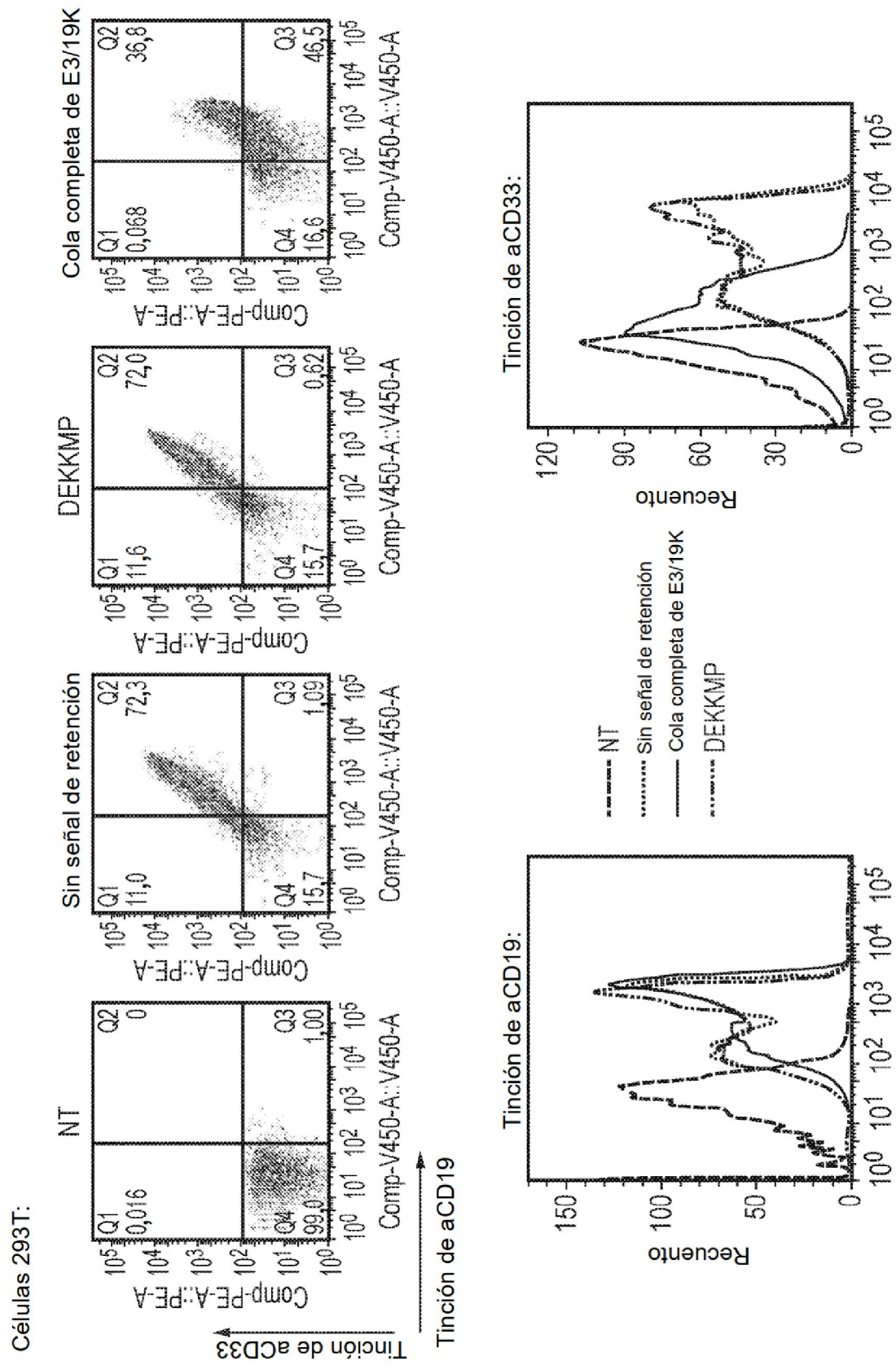


FIG. 5