

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 549 488**

21 Número de solicitud: 201431066

51 Int. Cl.:

**C07K 19/00** (2006.01)

**C12N 15/62** (2006.01)

**G01N 33/533** (2006.01)

12

SOLICITUD DE PATENTE

A1

22 Fecha de presentación:

**14.07.2014**

43 Fecha de publicación de la solicitud:

**28.10.2015**

71 Solicitantes:

**UNIVERSIDAD DEL PAÍS VASCO (100.0%)  
Barrio Sarriena, s/n  
48940 Leioa (Bizkaia) ES**

72 Inventor/es:

**RODRÍGUEZ PÉREZ, Jose Antonio y  
GARCÍA-SANTISTEBAN, Iraia**

74 Agente/Representante:

**ARIAS SANZ, Juan**

54 Título: **BIOSENSOR PARA DETERMINAR LA FUNCIÓN EXPORTADORA DE CRM1**

57 Resumen:

Biosensor para determinar la función exportadora de CRM1.

La presente invención se refiere a un método para determinar si la función de una variante de CRM1 está alterada en una muestra y a un método para identificar un compuesto con capacidad de modular la función de CRM1 o de una variante equivalente de la misma. La presente invención también se refiere a una proteína de fusión y al uso de dicha proteína de fusión para determinar si la función de una variante de CRM1 está alterada en una muestra y para identificar un compuesto con capacidad de modular la función de CRM1.

ES 2 549 488 A1

## DESCRIPCIÓN

Biosensor para determinar la función exportadora de CRM1

### 5 CAMPO DE LA INVENCION

La presente invención se refiere a un método para determinar si la función de una variante de CRM1 está alterada en una muestra y a un método para identificar un compuesto con capacidad de modular la función de CRM1 o de una variante equivalente de la misma. La presente invención también se refiere a una proteína de fusión y al uso de dicha proteína de fusión para determinar si la función de una variante de CRM1 está alterada en una muestra y para identificar un compuesto con capacidad de modular la función de CRM1.

### 15 ANTECEDENTES DE LA INVENCION

La proteína CRM1/exportina1, es un miembro de la familia de las carioferinas, un grupo de proteínas que funcionan como receptores en el transporte de macromoléculas entre el núcleo y el citoplasma a través de los poros nucleares. Los receptores de transporte nuclear se pueden clasificar funcionalmente en dos grandes grupos: los receptores de importación (importinas) y los receptores de exportación (exportinas). En el citoplasma celular, las importinas reconocen y se unen a secuencias específicas de aminoácidos, denominadas señales de localización nuclear (NLS, *nuclear localization signal*) en las proteínas "carga" que han de ser transportadas al núcleo. La importina facilita la entrada del cargo al núcleo a través del poro nuclear y, una vez en el núcleo, el complejo importina/cargo se disocia por la unión de RanGTP a la importina. El proceso de transporte inverso, la exportación de proteínas al citoplasma, esta mediada principalmente por el receptor de exportación CRM1. En el núcleo, y en presencia de RanGTP, CRM1 reconoce y se une a secuencias de aminoácidos denominadas señales de exportación nuclear (NES, *nuclear export signal*) en los cargos que han de ser exportados. El complejo trimérico CRM1/cargo/RanGTP atraviesa el poro nuclear y se disocia en el citoplasma tras la hidrólisis del GTP. El transporte nucleocitoplásmico de proteínas es un proceso esencial para la homeostasis celular, y se encuentra alterado en varios procesos patológicos, incluido el cancer.

Numerosas proteínas celulares dependen de la función de CRM1 para ser exportadas al citoplasma y así desarrollar de modo correcto su función. Entre estas proteínas se encuentran importantes reguladores del ciclo celular y de la apoptosis como p53 o survivina

(Stommel et al., 1999, EMBO J. 18:1660-1672; Rodríguez et al., 2002, Exp. Cell Res. 275: 44–53), cuya alteración es causa frecuente del desarrollo tumoral. Recientemente, se han identificado mutaciones puntuales del gen XPO1, que codifica CRM1, en pacientes con leucemia linfocítica crónica (CLL) (Puente et al., 2011, Nature. 475: 101-105). Debido a su importante función en la homeostasis celular, y a su papel central como regulador de proteínas relacionadas con el desarrollo de tumores, CRM1 constituye una interesante diana terapéutica para el tratamiento del cancer (Turner et al., 2012, Biochem. Pharmacol., 83:1021-1032).

10 La función de CRM1 se ha analizado en numerosos estudios utilizando diferentes métodos, dirigidos fundamentalmente a determinar su capacidad de unión a una secuencia NES (aislada o en el contexto de una proteína). Estos métodos pueden clasificarse en dos grandes grupos según se lleven a cabo en un tampón de reacción bien definido bioquímicamente (ensayos *in vitro*), o en células intactas o permeabilizadas (ensayos  
15 celulares).

Los ensayos *in vitro* se basan en la utilización de proteínas recombinantes producidas *in vitro* o purificadas a partir de sistemas de expresión procarióticos o eucarióticos. A la secuencia de aminoácidos de estas proteínas recombinantes se añaden a menudo epítopos artificiales (por ejemplo GST o 6xHistidine) que facilitan su aislamiento o detección. Usualmente, las proteínas recombinantes (CRM1 y NES) purificadas se añaden a un  
20 tampón, que contiene RanGTP, el tercer componente del complejo de exportación. Utilizando diversas variantes de ensayos de co-purificación (por ejemplo GST-pull down) (Grespin et al., 2008, J. Biol. Chem. 283: 25576-25588; Güttler et al., 2010, Nat. Struct. Mol. Biol. 17:1367-1376) o ensayos basados en la cuantificación de la hidrólisis de GTP (RanGAP protection assay) (Askjaer et al., 1999, Mol. Cell. Biol., 19: 6276–6285) se  
25 determina si se produce la interacción CRM1/NES. Los ensayos *in vitro* presentan la limitación de que se llevan a cabo en un medio distinto al fisiológico.

30 Los ensayos celulares pueden estar dirigidos a evaluar la función de la proteína CRM1 producida de modo natural por las propias células en las que se realiza el ensayo (CRM1 endógena), por ejemplo mediante la transfección o microinyección de sustratos de CRM1 fluorescentes (Ossareh-Nazari et al., 1997, Science 278: 141-144; Cabot et al. 2002, Biol. Reprod. 67: 814-819; Fetz et al., 2009, Sensors 9: 5423-5445). Sin embargo, los ensayos  
35 que analizan la función de CRM1 endógena no permiten comparar directamente la función de distintas variantes mutadas de CRM1. Dicha comparación sí es posible utilizando

ensayos celulares en los que se induce la expresión ectópica de CRM1 mediante transfección. La interacción de CRM1 ectópica con una NES se ha investigado mediante co-inmunoprecipitación (Kanai et al., 2007, Nat. Cell Biol., 9: 1175 – 1183), mediante ensayos de doble híbrido (Yang et al., 2001, Mol. Cell 8: 397-406) y también  
5 mediante ensayos de relocalización. En estos últimos, se co-transfectan células con plásmidos de expresión que codifican CRM1 y una proteína con NES fusionadas a epítopos que faciliten su detección mediante microscopía de fluorescencia (por ejemplo GFP o Flag). La proteína con NES se utiliza como “cebo” y la interacción CRM1/NES se detecta como un cambio en la localización subcelular de CRM1  
10 (Daelemans et al., 2002, Proc. Natl. Acad. Sci. USA 99: 14440-14445). Estos ensayos evalúan la capacidad de CRM1 para unirse a una NES, pero no directamente su capacidad para exportar una proteína.

Existe por tanto, una necesidad en el estado de la técnica de desarrollar biosensores  
15 que permitan la determinación de la funcionalidad en un entorno celular de variantes de CRM1 así como la identificación de compuestos capaces de modular la función de CRM1 basados en el análisis de la capacidad de exportación de dicha exportina.

## **COMPENDIO DE LA INVENCION**

20

Los autores de la presente invención han desarrollado un método para evaluar la función de la exportina CRM1 basado en la detección de la localización subcelular de un biosensor construido a partir de un cargo natural de CRM1, survivina. Survivina es una proteína constitutivamente citoplasmática. Los autores de la presente invención  
25 han diseñado una proteína de fusión que comprende survivina cuya localización es constitutivamente nuclear y que relocaliza al citoplasma en respuesta a la sobreexpresión de una proteína CRM1 funcional.

En un primer aspecto, la invención se refiere a un método para determinar si la función  
30 de una variante de CRM1 en una muestra que comprende al menos una célula está alterada, que comprende:

(i) proporcionar una muestra que comprende al menos una célula en donde dicha al menos una célula comprende una proteína de fusión y en donde dicha proteína de fusión comprende:

35

a) un primer dominio polipeptídico que comprende una variante de survivina seleccionada entre:

1) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional; y

5

2) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende la señal de exportación nuclear (NES) de survivina;

b) al menos, un segundo dominio polipeptídico que comprende una señal de localización nuclear (NLS); y

10

c) al menos, un tercer dominio polipeptídico reportero;

en donde si la variante de survivina carece de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además

d) un cuarto dominio polipeptídico que comprende una NES;

15

(ii) introducir en dicha al menos una célula un casete de expresión o un plásmido que codifica para dicha variante de CRM1;

(iii) crecer dicha al menos una célula bajo condiciones que permitan la correcta expresión de dicha variante de CRM1;

(iv) detectar la señal del dominio reportero de dicha proteína de fusión;

(v) determinar la localización subcelular de dicha proteína de fusión;

20

(vi) comparar la localización subcelular de dicha proteína de fusión determinada en la etapa (v) con dicha localización subcelular obtenida para dicha proteína de fusión en una muestra de referencia;

en donde una alteración de la localización subcelular de dicha proteína de fusión es indicativa de una función alterada de CRM1.

25

En un segundo aspecto, la invención se relaciona con un método para identificar un compuesto con capacidad de modular la función de CRM1 o de una variante funcionalmente equivalente de la misma en una muestra que comprende al menos una célula, comprendiendo dicho método:

30

(i) poner en contacto un compuesto candidato con una muestra que comprende al menos una célula en donde dicha al menos una célula comprende:

a) una proteína de fusión que comprende:

1) un primer dominio polipeptídico que comprende una variante de survivina seleccionada entre:

- una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional; y

5

- una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende una señal de exportación nuclear (NES) de survivina;

2) al menos, un segundo dominio polipeptídico que comprende una señal de localización nuclear (NLS); y

10

3) al menos, un tercer dominio polipeptídico reportero;

en donde si la variante de survivina carece de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además

4) un cuarto dominio polipeptídico que comprende una NES;

15

b) un casete de expresión o un plásmido que codifica para dicha CRM1 o dicha variante funcionalmente equivalente de la misma;

(ii) detectar la señal del dominio reportero de dicha proteína de fusión;

(iii) detectar la localización subcelular de dicha proteína de fusión; y

20

(iv) comparar la localización subcelular de dicha proteína de fusión obtenida en la etapa (iii) con la localización subcelular de dicha proteína de fusión obtenida en una muestra que comprende al menos una célula que no ha sido tratada con dicho compuesto;

en donde una alteración de la localización subcelular de la proteína de fusión detectada en la etapa (iii) con respecto a la localización subcelular en dicha muestra no tratada es indicativa de que dicho compuesto es capaz de modular la función de CRM1 o de una variante funcionalmente equivalente de la misma.

25

En un tercer aspecto, la invención se relaciona con una proteína de fusión que comprende:

(i) un primer dominio polipeptídico que comprende una variante de survivina seleccionada entre:

30

a) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional; y

35

b) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende una señal de exportación nuclear (NES) de survivina;

(ii) al menos, un segundo dominio polipeptídico que comprende una señal de localización nuclear (NLS);

(iii) al menos, un tercer dominio polipeptídico reportero;

5 En donde si la variante de survivina carece de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además

(iv) un cuarto dominio polipeptídico que comprende una NES.

En un cuarto aspecto, la invención se refiere a un ácido nucleico que codifica para dicha proteína de fusión.

10

En un quinto aspecto, la invención se refiere a un casete de expresión que comprende dicho ácido nucleico operativamente unido a un sistema de transcripción y/o traducción apropiado.

En un sexto aspecto, la invención se refiere a un plásmido que comprende dicho ácido nucleico o dicho casete de expresión.

15

En un séptimo aspecto, la invención se refiere a una célula hospedadora que comprende dicho ácido nucleico, dicho casete de expresión o dicho plásmido.

20 En un octavo aspecto, la invención se refiere al uso de dicha proteína de fusión para determinar si la función de una variante de CRM1 está alterada en una muestra o para identificar un compuesto con capacidad de modular la función de CRM1 o de una variante funcionalmente equivalente de la misma en una muestra.

## 25 **BREVE DESCRIPCIÓN DE LAS FIGURAS**

**Figura 1.** Elementos funcionales y secuencia de la proteína de fusión (SEQ ID NO: 52). El panel A muestra una representación simplificada de la proteína survivina humana y la proteína de fusión derivada de ella. En la representación de survivina se indica su tamaño (142 aminoácidos), la posición de la NES, el dominio de dimerización (dim) y el segundo dominio de exportación carboxiterminal (C-exp). En la representación de la proteína de fusión se indica la posición de la NES, las NLS y el epitopo 3xFlag. El panel B muestra la secuencia de aminoácidos de la proteína de fusión indicando la secuencia de la NES (en recuadro), las NLS (subrayadas) y el epitopo 3xFlag (subrayado doble). El panel C muestra la secuencia de nucleótidos que codifica la proteína de fusión, indicando la secuencia que codifica la NES (en recuadro), las NLS (subrayadas), el epitopo 3xFlag (subrayado doble) y

35

el codón de parada (negrita). En mayúscula, se indican los sitios de restricción BgIII (AGATCT), HindIII (AAGCTT) y NotI (CGGCCGCG).

5

## DESCRIPCIÓN DETALLADA DE LA INVENCION

### Método para determinar si la función de una variante de CRM1 está alterada

10 En un primer aspecto, la invención se refiere a un método, de aquí en adelante “primer método de la invención”, para determinar si la función de una variante de CRM1 en una muestra que comprende al menos una célula está alterada, que comprende:

(i) proporcionar una muestra que comprende al menos una célula en donde dicha al menos una célula comprende una proteína de fusión y en donde dicha proteína de fusión comprende:

15

a) un primer dominio polipeptídico que comprende una variante de survivina seleccionada entre:

1) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional; y

20

2) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende la señal de exportación nuclear (NES) de survivina;

25

b) al menos, un segundo dominio polipeptídico que comprende una señal de localización nuclear (NLS); y

c) al menos, un tercer dominio polipeptídico reportero;

en donde si la variante de survivina carece de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además

30

d) un cuarto dominio polipeptídico que comprende una NES;

(ii) introducir en dicha al menos una célula un casete de expresión o un plásmido que codifica para dicha variante de CRM1;

(iii) crecer las células bajo condiciones que permitan la correcta expresión de dicha variante de CRM1;

35

(iv) detectar la señal del dominio reportero de dicha proteína de fusión;

(v) determinar la localización subcelular de dicha proteína de fusión;



(vi) comparar la localización subcelular de dicha proteína de fusión determinada en la etapa (v) con dicha localización subcelular obtenida para dicha proteína de fusión en una muestra de referencia;

5 en donde una alteración de la localización subcelular de dicha proteína de fusión es indicativa de una función alterada de CRM1.

La etapa (i) del primer método de la invención comprende proporcionar una muestra que comprende al menos una célula en donde dicha al menos una célula comprende una proteína de fusión y en donde dicha proteína de fusión comprende:

10 a) un primer dominio polipeptídico que comprende una variante de survivina seleccionada entre:

1) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y

15 2) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende la señal de exportación nuclear (NES) de survivina;

20 b) al menos, un segundo dominio polipeptídico que comprende una señal de localización nuclear (NLS); y

c) al menos, un tercer dominio polipeptídico reportero;

en donde si la variante de survivina carece de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además

d) un cuarto dominio polipeptídico que comprende una NES;

25

El primer método de la invención está dirigido a determinar si la función de una variante de CRM1 está alterada.

30 El término "CRM1" (del inglés, *Chromosomal Maintenance 1*) o "exportina 1" según se usa en la presente descripción, se refiere a una proteína que media el transporte nuclear de macromoléculas (incluyendo ARN y proteínas) desde el núcleo celular hacia el citoplasma mediante el reconocimiento en dichas macromoléculas de una señal de exportación nuclear (NES) rica en residuos de leucina. La proteína CRM1 está formada por varios dominios HEAT (del inglés, *Huntingtin, elongation factor 3, protein phosphatase 2A and the yeast*  
35 *kinase Tor1*), cuyo número es variable dependiendo de la especie (por ejemplo, CRM1 humana contiene 20 dominios HEAT), que a su vez están constituidos por una cadena alfa y

una cadena beta antiparalelas que dan lugar a “ranuras hidrofóbicas” en las que se acomodan los residuos hidrofóbicos de la secuencia NES. El transporte de dichas macromoléculas lleva asociado un consumo de energía en forma de GTP y se lleva a cabo mediante el acoplamiento de CRM1 con un complejo proteico formado por la proteína RANBP3 y una GTPasa. La proteína CRM1 es el producto del gen *crm1* o *xpo1*. CRM1 incluye la proteína CRM1 de cualquier mamífero incluyendo, pero sin estar limitado a, animales domésticos y de granja (vacas, caballos, cerdos, ovejas, cabras, perros, gatos o roedores), primates y humanos.

El término “variante de CRM1” según se utiliza en la presente descripción, se refiere a toda aquella proteína o polipéptido, o a la secuencia de nucleótidos que codifica la misma, que resulta de la inserción, delección o sustitución de uno o varios aminoácidos de la secuencia de CRM1. Variantes de CRM1 según la invención incluyen polipéptidos que son sustancialmente homólogos a CRM1. Un polipéptido es sustancialmente homólogo a otro polipéptido cuando su secuencia de aminoácidos tiene un grado de identidad de al menos el 50%, al menos el 60%, al menos el 70%, al menos el 80%, al menos el 90%, al menos el 91%, al menos el 92%, al menos el 93%, al menos el 94%, al menos el 95%, al menos el 96%, al menos el 97%, al menos el 98% o al menos el 99%. El grado de identidad entre dos secuencias de aminoácidos puede determinarse por métodos convencionales, por ejemplo, mediante algoritmos estándar de alineamiento de secuencias conocidos en el estado de la técnica, tales como, por ejemplo BLAST (Altschul S.F. *et al.* Basic Local Alignment Search Tool. *J Mol Biol.* 1990 Oct 5; 215(3):403-10). En una realización preferida de la invención, dicha variante de CRM1 es una variante de la proteína CRM1 humana. En una realización aún más preferida de la invención, dicha variante de CRM1 es CRM1 A541K en donde el residuo de alanina en la posición 541 es sustituido por una lisina (SEQ ID NO: 1).

La expresión “función de una variante de CRM1” se refiere a la capacidad de una variante de CRM1 de exportar macromoléculas, tal como, por ejemplo, la proteína de fusión definida en la presente invención, desde el núcleo al citoplasma.

El término “muestra”, según se usa en el presente documento, se refiere a una pequeña parte de un sujeto que comprende al menos una célula, y puede estar constituida por una única célula, 2 células, 3 células, 10 células, 20 células, 100 células, 500 células, 1000 células, 2000 células, 5000 células, 10000 células o más. La muestra puede ser fresca o congelada. Dicha célula debe ser una célula eucariota. Ejemplos de células eucariotas para llevar a cabo el primer método de la invención incluyen células de mamíferos, plantas,

insectos y hongos. Células de hongos incluyen, preferiblemente, células de levaduras tales como *Saccharomyces*, *Pichia pastoris* y *Hansenula polymorpha*. Células de insectos incluyen, sin limitación, células de *Drosophila* y células Sf9. Células de plantas incluyen, entre otras, células de plantas de cultivos tales como cereales, plantas medicinales, ornamentales o de bulbos. Células de mamíferos adecuadas para llevar a cabo el primer método de la presente invención incluyen líneas celulares epiteliales (porcinas, humanas, etc.), líneas celulares tumorales como por ejemplo, líneas celulares de osteosarcoma (humanas, etc.), líneas celulares de neuroblastoma (humanas, etc.), carcinomas epiteliales (humanos, etc.), células gliales (murinas, etc.), líneas celulares hepáticas (de mono, etc.), células CHO (Chinese Hamster Ovary), células COS, células BHK, células HeLa, 911, AT1080, A549, 293 o PER.C6, células ECCs (Embryonal Cancer Cells) humana NTERA-2, células D3 de la línea de mESCs (Mouse Embryonic Stem Cells), células troncales embrionarias humanas tales como HS293 y BGV01, SHEF1, SHEF2 y HS181, células NIH3T3, 293T, REH y MCF-7 y células hMSCs (human Mesenchymal Stem Cells). En una realización preferida de la invención, dicha muestra comprende una población de células 293T.

El término “proteína de fusión” tal y como se usa en el presente documento, se refiere a proteínas generadas mediante ingeniería genética que consisten en dos o más dominios funcionales derivados de distintas proteínas. Una proteína de fusión puede ser obtenida mediante cualquier técnica convencional conocida en el estado de la técnica (por ejemplo, mediante la expresión de una secuencia nucleotídica que codifica para dicha proteína de fusión en una célula adecuada). A modo ilustrativo, la proteína de fusión puede generarse como se muestra en el Ejemplo 1 de la presente solicitud.

25

De acuerdo con el primer método de la invención, la proteína de fusión comprende:

a) un primer dominio polipeptídico que comprende una variante de survivina seleccionada entre:

1) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional; y

30

2) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende la señal de exportación nuclear (NES) de survivina;

35

b) al menos, un segundo dominio polipeptídico que comprende una señal de localización nuclear (NLS); y

c) al menos, un tercer dominio polipeptídico reportero;

en donde si la variante de survivina carece de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además

d) un cuarto dominio polipeptídico que comprende una NES;

El término “dominio polipeptídico o polipéptido”, según se usa en la presente descripción, usado aquí indistintamente con proteína, se refiere a una cadena de aminoácidos de cualquier longitud en donde los distintos aminoácidos se encuentran unidos entre sí mediante enlaces peptídicos. Cada dominio polipeptídico comprende una secuencia de aminoácidos y tiene una estructura de una proteína activa funcionalmente y que puede existir independientemente del resto de la cadena proteica. Cada dominio polipeptídico forma una estructura tridimensional compacta y, a menudo, puede adquirir su conformación estable de manera independiente.

#### *Primer dominio polipeptídico*

De acuerdo con el primer método de la invención, el primer dominio polipeptídico de la proteína de fusión comprende una variante de survivina seleccionada entre:

1) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional ; y

2) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende la señal de exportación nuclear (NES) de survivina.

El término “survivina”, también denominada “BIRC5” (del inglés *Baculoviral Inhibitor of apoptosis Repeat-Containing 5*), según se utiliza en el presente documento, se refiere a una proteína miembro de la familia de inhibidores de apoptosis (IAP). La función de survivina es inhibir la activación de las caspasas, lo cual tiene como consecuencia la regulación negativa (inhibición) del proceso apoptótico o muerte celular programada. Survivina también es una proteína reguladora del ciclo celular e interviene en la separación de las cromátidas y en la citoquinesis. Survivina es un “carga” natural de CRM1. Estructuralmente, dicha proteína comprende en el extremo amino terminal un dominio BIR, una señal de exportación nuclear

(NES) y un dominio de homodimerización localizado en el extremo carboxilo terminal de survivina que está poco caracterizado. El dominio BIR (del inglés *Baculovirus Inhibitor of Apoptosis protein Repeat*) es un motivo estructural muy conservado de aproximadamente, 70 residuos de aminoácidos que comprende 3 cisteínas y 1 histidina queladas con un ion de zinc. Típicamente, el dominio BIR está compuesto por 5 alfa hélices y 3 láminas beta. La señal de exportación nuclear (NES) de survivina está formada por una secuencia de aminoácidos similar a la secuencia NES consenso ( $\Phi$ -X<sub>2-3</sub>- $\Phi$ -X<sub>2-3</sub>- $\Phi$ -X- $\Phi$ , en donde  $\Phi$  representa un aminoácido hidrofóbico). El dominio de homodimerización de survivina permite la formación de dímeros de survivina. Dicho dominio, estructuralmente poco caracterizado, es rico en residuos hidrofóbicos y solapa parcialmente con la señal NES de survivina. La señal NES de survivina humana está comprendida entre las posiciones 84 y 109 de la secuencia aminoacídica. La proteína survivina es producto del gen *survivina* que da lugar a cuatro transcritos alternativos según se lleve a cabo el procesamiento del ARNm: survivina (constituida por 3 intrones y 4 exones), survivina 2B (con un exón 2 alternativo al exón 2 de survivina), survivina-delta-ex3 (en donde el exón 3 es eliminado) y survivina 3B (con un exón 3 alternativo al exón 4 de survivina). Survivina incluye la proteína survivina de cualquier mamífero incluyendo, pero sin estar limitado a, animales domésticos y de granja (vacas, caballos, cerdos, ovejas, cabras, perros, gatos o roedores), primates y humanos. La secuencia de la proteína survivina humana se encuentra depositada en la base de datos de GenBank bajo el número de acceso ACC51660 (versión del 2 de septiembre de 2004)

El término “variante de survivina” según se utiliza en la presente descripción, se refiere a toda aquella proteína que resulta de la inserción, delección o sustitución de uno o varios aminoácidos de la secuencia de survivina. La expresión “variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional” tal y como se usa en la presente descripción, se refiere a un polipéptido sustancialmente homólogo a survivina que no es capaz de dimerizar. El término “capacidad de dimerizar” o “capacidad de dimerización” tal y como se usa en la presente descripción, se refiere a la capacidad que tiene una proteína, en particular una variante de survivina, de formar un dímero mediante la asociación con otra molécula similar. La determinación de la capacidad de dimerización de las variantes de survivina contempladas en la presente invención puede llevarse a cabo mediante cualquier técnica conocida en el estado de la técnica (Scopes, Protein Purification: principles and practice, 1897, Springer-Verlag, Nueva York). A modo ilustrativo se pueden citar técnicas de separación bioquímica para resolver proteínas en base a su tamaño molecular (por ejemplo, electroforesis en gel nativo en condiciones no reductoras, cromatografía de filtración en gel, ultracentrifugación

analítica, etc), y/o comparación de propiedades físico-químicas antes y después de la inducción de la ruptura de los enlaces entre ambos monómeros mediante el empleo de agentes activos con sulfhidrilo (por ejemplo, yodoacetamida, N-etilmaleimida) o reductores de disulfuro (por ejemplo 2-mercaptoetanol, ditioneitol), u otras metodologías equivalentes.

5 Dicha variante de survivina incluye variantes de la proteína survivina de cualquier mamífero incluyendo, pero sin estar limitado a, animales domésticos y de granja (vacas, caballos, cerdos, ovejas, cabras, perros, gatos o roedores), primates y humanos. En una realización particular de la invención, dicha variante de survivina es una variante de survivina humana.

10 En una realización particular, dicha variante de survivina es una forma mutada de survivina que comprende un dominio de homodimerización truncado obtenido mediante la delección parcial de dicho dominio. El experto en la materia entenderá que dicha variante podrá presentar delecciones en cualquier región de la secuencia de aminoácidos de dicho dominio de homodimerización, siempre y cuando la variante de survivina resultante carezca de dicha  
 15 capacidad de dimerización con otro monómero de survivina. La generación de las variantes de survivina contempladas en la presente invención puede llevarse a cabo mediante técnicas de ingeniería genética bien conocidas en el estado de la técnica. Véase, por ejemplo, Alberts B, et al, "Molecular Biology of the Cell" (Garland Publishing Inc., Nueva York, NY, E.E.U.U., 2008).

20 En una realización particular, dicha variante de survivina carece de la totalidad del dominio de homodimerización, y, en consecuencia, presenta la señal NES truncada. En una realización preferida, dicha variante de survivina carece de los residuos de aminoácidos comprendidos entre las posiciones 96 y 104 de la secuencia de survivina humana.

25 En una realización todavía más particular, dicha variante de survivina carece de la totalidad del dominio de homodimerización y de la totalidad del dominio NES. En una realización preferida, dicha variante de survivina carece de los aminoácidos comprendidos entre las posiciones 84 y 142 de la secuencia de survivina humana (SEQ ID NO: 2).

30 Alternativamente, la variante de survivina contemplada en la presente invención carece de parte del dominio de homodimerización, es decir de uno o más aminoácidos de dicho dominio de homodimerización, en donde dichos uno o más aminoácidos forman parte de los residuos que codifican la señal NES de survivina. Por lo tanto, en una realización particular  
 35 una variante de survivina comprende una survivina con el dominio de homodimerización y la señal NES truncados. Alternativamente, la variante de survivina contemplada en la presente

invención carece de parte del dominio de homodimerización, es decir, de uno o más aminoácidos de dicho dominio de homodimerización en donde dichos uno o más aminoácidos no forman parte de los residuos que codifican la señal NES de survivina. Por lo tanto, en otra realización particular, la variante de survivina comprende una  
5 survivina con el dominio de homodimerización truncado pero en donde la señal NES no carece de ningún residuo aminoacídico, es decir, en donde la señal NES está completa. En otra realización particular de la invención, dicha variante de survivina carece de los aminoácidos comprendidos entre las posiciones 101 y 142 de la secuencia de survivina humana.

10

*Segundo dominio polipeptídico*

De acuerdo con el primer método de la invención, la proteína de fusión comprende, al menos, un segundo dominio polipeptídico que comprende, una señal de localización  
15 nuclear (NLS). El término "al menos uno" tal y como se usa en la presente descripción, se refiere a uno o más. Es decir, en una realización particular, dicha proteína de fusión comprende un segundo dominio polipeptídico que comprende una NLS; en otra realización particular, dicha proteína de fusión comprende un segundo dominio polipeptídico que comprende dos NLS. La presente invención también contempla  
20 realizaciones en donde dicha proteína de fusión comprende un segundo dominio polipeptídico que comprende tres NLS, cuatro NLS, cinco NLS o más.

Tal como se usa en la presente descripción, el término "señal de localización nuclear" o "NLS" se refiere a una secuencia que dirige el polipéptido al que está unido al núcleo.  
25 Secuencias de señalización nuclear adecuadas para el transporte de la proteína de fusión de la invención al núcleo comprenden, tanto secuencias que actúan independientemente de activadores externos como secuencias que son externamente activables en cuyo caso la célula que comprende el polipéptido se pone en contacto con un compuesto capaz de activar la señal de localización nuclear. Secuencias NLS que no  
30 requieren de un activador para promover la translocación de las proteínas son ampliamente conocidas para el experto en la materia e incluyen todas aquellas secuencias capaces de ser reconocidas por un receptor de importación formando un complejo que se transloca a través del poro nuclear mediante un proceso en el que ocurre hidrólisis de GTP. Un tipo de NLS son aquellas formadas por una región corta de  
35 aminoácidos básicos, tales como la NLS del antígeno T

grande del virus SV40 formada por la secuencia PKKKRKV (SEQ ID NO: 3), que también ha demostrado ser efectiva en células de mamíferos. Otro tipo de NLSs son las que están formadas por una secuencia bipartita formada por dos regiones de aminoácidos básicos separados por un espaciador de 10 a 12 aminoácidos, tales como la NLS de la nucleoplasmina formada por la secuencia KRPAATKKAGQAKKKK (SEQ ID NO: 4). Un tercer tipo de NLS son aquellas que se asemejan al homeodominio de la proteína c-myc formada por la secuencia RQRRNELKRSF (SEQ ID NO: 5) en donde no se observa una acumulación particular de aminoácidos básicos. Otros ejemplos de NLS incluyen, la secuencia RMRKFKNKGKDTAELRRRRVEVSVELRKAKKDEQILKRRNV (SEQ ID NO: 6) del dominio IBB de la importina alfa, las secuencias VSRKRPRP (SEQ ID NO: 7) y PPKKARED (SEQ ID NO: 8) de la proteína T de mioma, la secuencia NLS de la proteína p53 humana, la secuencia SALIKKKKKMAP (SEQ ID NO: 9) de la proteína c-abl IV de ratón, las secuencias DRLRR (SEQ ID NO: 10) y PKQKKRK (SEQ ID NO: 11) del virus de la gripe, la secuencia RKLKKIKKL (SEQ ID NO: 12) del antígeno delta del virus de la hepatitis, la secuencia REKKKFLKRR (SEQ ID NO: 13) de la proteína Mx1 de ratón, la secuencia KRKGDEV DGVDEVAKKSKK (SEQ ID NO: 14) de la poli(ADP-ribosa) polimerasa, la secuencia RKCLQAGMNLEARKTKK (SEQ ID NO: 15) de los receptores de hormonas esteroideas humanos y la secuencia NLS de la proteína p17 del reovirus aviar (IAAKRGRQLD- SEQ ID NO: 16). Adicionalmente, la invención contempla el uso de NLS tales como las descritas en US2006121513 y WO05120588 así como de sustituyentes del tipo de los descritos en EP1695717 que permiten el transporte de proteínas que los contienen a través de la membrana nuclear.

En una realización particular, dicho segundo polipéptido comprende la secuencia NLS del antígeno T grande del virus SV40 (SEQ ID NO: 3).

En otra realización particular, dicha proteína de fusión comprende dos segundos dominios polipeptídicos en donde dichos dos segundos dominios polipeptídicos son iguales o diferentes. Es decir, la proteína de fusión puede presentar dos cualesquiera de las secuencias NLS mencionadas anteriormente. En una realización particular, dichos dos segundos dominios polipeptídicos son diferentes. En otra realización particular, dicha proteína de fusión comprende dos segundos dominios polipeptídicos idénticos. En una realización aún más particular, dicha proteína de fusión comprende dos segundos dominios polipeptídicos idénticos en donde cada uno de dichos dos segundos dominios polipeptídicos comprende la secuencia NLS del antígeno T grande del virus SV40 (SEQ ID NO: 3).



*Tercer dominio polipeptídico*

La proteína de fusión comprende al menos, un tercer dominio polipeptídico reportero. El término “dominio polipeptídico reportero” o “polipéptido reportero” tal y como se usa en el presente documento, se refiere a un polipéptido que permite la detección y/o purificación de dicha proteína de fusión. Dicho polipéptido reportero es capaz de unirse con alta afinidad a uno o más ligandos, como por ejemplo anticuerpos específicos o sustratos inmovilizados y en ningún caso interferirá significativamente con la actividad del polipéptido al que está unido. Polipéptidos reporteros adecuados para detectar y/o purificar la proteína de fusión de acuerdo a la presente invención incluyen el epítipo flag, el epítipo HA, el epítipo Myc, los epítipos AU1 y AU5, el epítipo Glu-Glu, el epítipo KT3, el epítipo IRS, el epítipo Btag, el epítipo de proteína quinasa C, el epítipo de VSV, polipéptido de hexahistidinas y péptido de unión a calmodulina.

En una realización particular, dicho tercer dominio polipeptídico reportero comprende el epítipo flag.

En otra realización particular, dicha proteína de fusión comprende dos terceros dominios polipeptídicos en donde dichos dos terceros dominios polipeptídicos son iguales o diferentes. Es decir, la proteína de fusión puede presentar dos cualesquiera de los polipéptidos reporteros mencionados anteriormente. En una realización particular, dichos dos terceros dominios polipeptídicos son diferentes. En otra realización particular, dicha proteína de fusión comprende dos terceros dominios polipeptídicos idénticos. En otra realización particular, dicha proteína de fusión comprende tres terceros dominios polipeptídicos en donde dichos tres terceros dominios polipeptídicos son iguales o diferentes. Es decir, la proteína de fusión puede presentar tres cualesquiera de los polipéptidos reporteros definidos anteriormente. En una realización particular, dicha proteína de fusión comprende tres terceros dominios polipeptídicos diferentes. En otra realización particular, dicha proteína de fusión comprende tres terceros dominios polipeptídicos idénticos. En una realización aún más particular, dicha proteína de fusión comprende tres terceros dominios polipeptídicos idénticos en donde cada uno de dichos tres terceros dominios polipeptídicos comprende el epítipo flag.

*Cuarto dominio polipeptídico*

35

De acuerdo con el primer método de la invención, tal y como se ha mencionado en el contexto del primer dominio polipeptídico de la proteína de fusión, en caso de que la variante de survivina carezca de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y carezca de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además un cuarto dominio polipeptídico que comprende una NES.

El término “señal de exportación nuclear” o NES” según se usa en la presente descripción, se refiere a una secuencia que dirige al péptido al que está unido al citoplasma. Secuencias de exportación nuclear adecuadas para el transporte de la proteína de fusión de la invención al citoplasma comprenden la secuencia consenso  $\Phi$ -X<sub>2-3</sub>-  $\Phi$ -X<sub>2-3</sub>-  $\Phi$ -X- $\Phi$ , en donde  $\Phi$  representa un aminoácido hidrofóbico. Secuencias NES adecuadas para la presente invención incluyen pero no están limitadas a: la secuencia NES de survivina humana formada por la secuencia CAFLSVKKQFEELTL (SEQ ID NO: 17), la secuencia NES del inhibidor de la proteína quinasa dependiente de AMPc humano formada por la secuencia LALKLAGLDI (SEQ ID NO: 18), la secuencia NES de p53 humana formada por la secuencia EMFRELNEALELKD (SEQ ID NO: 19), la secuencia NES de ubiquitin-proteín ligasa Mdm2 humana formada por la secuencia SISLSFDESLALCVI (SEQ ID NO: 20), la secuencia NES de Xsmad4a de *Xenopus spp* formada por la secuencia GIDL SGLTLQ (SEQ ID NO: 21), la secuencia NES de catenina p120 humana (isoforma 3AB) formada por la secuencia ESLEEELDVLVLDDEGG (SEQ ID NO: 22), la secuencia NES de MAP quinasa 2 murina formada por la secuencia DKERWEDVKEEMTSALATMRVDYE (SEQ ID NO: 23), la secuencia NES de p150 murina formada por la secuencia EAINKLESNLRELQICPAT (SEQ ID NO: 24) secuencia NES de ciclina B1 humana formada por la secuencia DLCQAFSDVILAV (SEQ ID NO: 25), la secuencia NES de actina de *Saccharomyces spp* formada por la secuencia SLPHAILRIDLA (SEQ ID NO: 26), la secuencia NES de la proteína Rev del virus HIV-1 formada por la secuencia LPPLERLTL (SEQ ID NO: 27), la secuencia NES de la proteína Rex del virus T-linfotrópico tipo 1 formada por la secuencia LSLDS (SEQ ID NO: 28), la secuencia NES de MAP quinasa quinasa específica dual 1 de *Xenopus spp* formada por la secuencia LQKKLEELELDEQ (SEQ ID NO: 29), la secuencia NES de la proteína sustrato 15 del receptor del factor de crecimiento epidérmico humano formada por la secuencia LELAIALSKSEI (SEQ ID NO: 30), la secuencia NES de BRCA1 humana formada por la secuencia LECPICLEL (SEQ ID NO: 31), la secuencia NES de p65 humana formada por la secuencia LSEALLQLQF (SEQ ID NO: 32), la secuencia NES del factor de crecimiento similar a AP-1 de *Schizosaccharomyces spp* formada por la secuencia

ESFDIDDLCSKLNKAKCS (SEQ ID NO: 33) y la secuencia NES del factor de transcripción STAT-1 humano formada por la secuencia LLLKKMYLM (SEQ ID NO: 34).

5 En una realización particular, dicho cuarto polipéptido comprende la secuencia NES de survivina humana (SEQ ID NO: 17).

10 El experto en la materia entenderá que puede ser deseable que la proteína de fusión comprenda adicionalmente un péptido flexible que una dos dominios de manera covalente. El término “péptido flexible”, “péptido espaciador”, “péptido *linke*” o “péptido conector”, según se usa en la presente descripción, se refiere a un péptido que une de manera covalente dos dominios, que no forma parte de dichos dominios, permitiendo el movimiento de un dominio con respecto del otro, sin causar sustancialmente un detrimento en la función de uno de los dominios unidos. La proteína de fusión puede comprender un péptido espaciador entre el primer y segundo dominio polipeptídico y/o un péptido espaciador entre 15 el segundo y tercer dominio polipeptídico y/o un péptido espaciador entre el tercer y cuarto dominio polipeptídico.

20 El péptido flexible comprende al menos un aminoácido, al menos dos aminoácidos, al menos tres aminoácidos, al menos cuatro aminoácidos, al menos cinco aminoácidos, al menos seis aminoácidos, al menos siete aminoácidos, al menos ocho aminoácidos, al menos nueve aminoácidos, al menos 10 aminoácidos, al menos 12 aminoácidos, al menos 14 aminoácidos, al menos 16 aminoácidos, al menos 18 aminoácidos, al menos 20 aminoácidos, al menos 25 aminoácidos, al menos 30 aminoácidos, al menos 35 aminoácidos, al menos 40 aminoácidos, al menos 45 aminoácidos o aproximadamente 50 25 aminoácidos.

Péptidos flexibles adecuados para su uso en la presente invención se describen en el documento WO2009150284. Péptidos enlazadores adecuados para su uso en la presente invención incluyen péptidos que comprenden las secuencias  $(\text{Gly-Ser})_n$ ,  $(\text{Gly}_m\text{Ser})_n$  o  $(\text{Ser}_m\text{Gly})_n$ , en donde m es 1 a 6, en particular 1 a 4 y típicamente 2 a 4 y n es 1 a 30 o 1 a 10 y, típicamente, 1 a 4 y que, opcionalmente, comprenden algunos restos de glutámico (Glu) o lisina (Lys) repartidos a lo largo de la secuencia para mejorar la solubilidad (véase, por ejemplo, WO 96/06641, que proporciona ejemplos de péptidos enlazadores). Péptidos enlazadores ilustrativos incluyen, sin limitación, péptidos que comprenden la secuencia 30 GGSSRSSSSGGGGSGGGG (SEQ ID NO: 35), GSGRSGGGSGGGGS (SEQ ID NO: 36), EGSSGSGSESKST (SEQ ID NO: 37), EGKSSGSGSESKSTQ (SEQ ID NO: 38),

EGKSSGSGSESKVD (SEQ ID NO: 39), GSTSGSGKSSEGKG (SEQ ID NO: 40), KESGSVSSEQLAQFRSLD (SEQ ID NO: 41) y ESGSVSSEELAFRSLD (SEQ ID NO: 42). Otros ejemplos no limitativos de péptidos flexibles incluyen los siguientes: el péptido de secuencia TATPATTPTTAPTAGT (SEQ ID NO: 43); el péptido de secuencia TATPATTPTTAPTAGTTATPATTPTTAPTAGT (SEQ ID NO: 44); el péptido de secuencia GTKVHMK (SEQ ID NO: 45) formado por los residuos 53-56 y 57-59 de tetranectina (Nielsen et al., 1997, "Crystal structure of tetranectin, a trimeric plasminogen-binding protein with an alpha-helical coiled coil." FEBS Lett 412:388-396); la hebra conectora 3 de la fibronectina humana; la subsecuencia PGTSGQQPSVGQQ (SEQ ID NO: 46) correspondiente al número de aminoácidos 2037-2049 de la fibronectina, y dentro de esa subsecuencia el fragmento GTSGQ (SEQ ID NO: 47) correspondiente a los aminoácidos 2038-2042; la secuencia de 10 aminoácidos de la región de la bisagra superior de la IgG3 murina PKPSTPPGSS (SEQ ID NO: 48); el péptido de secuencia APAETKAEPMT (SEQ ID NO: 49); el péptido de secuencia GAP; el péptido de secuencia SGGSGSGGQ (SEQ ID NO: 50); y el péptido de secuencia GGSSRSSH (SEQ ID NO: 51).

En una realización particular dicha proteína de fusión comprende la secuencia mostrada en SEQ ID NO: 52.

Los dominios que forman la proteína de fusión no han de estar necesariamente dispuestos en dicho orden. Por lo tanto, realizaciones particulares de la invención incluyen cualquier combinación en el orden de la disposición relativa de dichos primero, segundo y tercer dominio polipeptídico; es decir, la invención contempla proteínas de fusión en las que el primer dominio polipeptídico está situado en posición amino-terminal respecto al segundo dominio polipeptídico y el segundo dominio polipeptídico está situado en posición amino-terminal respecto al tercer dominio polipeptídico. En otra realización particular, la invención contempla proteínas de fusión en las que el primer dominio polipeptídico está situado en posición carboxilo-terminal con respecto al segundo dominio polipeptídico y el segundo dominio polipeptídico está situado en posición carboxilo-terminal con respecto al tercer dominio polipeptídico. En otra realización particular, la invención contempla proteínas de fusión en las que el segundo dominio polipeptídico está situado en posición amino-terminal respecto al primer dominio polipeptídico y el primer dominio polipeptídico está situado en posición amino-terminal con respecto al tercer dominio polipeptídico.

En aquellas realizaciones en donde dicha proteína de fusión comprende adicionalmente un cuarto dominio polipeptídico tal y como se ha definido anteriormente, dichos primero,

segundo, tercero y cuarto dominios polipeptídicos pueden estar dispuestos en cualquier orden relativo. En una realización particular, la invención contempla proteínas de fusión en las que el primer dominio polipeptídico está situado en posición amino-terminal respecto al segundo dominio polipeptídico, el segundo dominio polipeptídico está situado en posición amino-terminal respecto al tercer dominio polipeptídico y el tercer dominio polipeptídico está situado en posición amino-terminal con respecto al cuarto dominio polipeptídico. En otra realización particular, la invención contempla proteínas de fusión en las que el primer dominio polipeptídico está situado en posición carboxilo-terminal con respecto al segundo dominio polipeptídico, el segundo dominio polipeptídico está situado en posición carboxilo-terminal con respecto al tercer dominio polipeptídico y el tercer dominio polipeptídico está situado en posición carboxilo-terminal con respecto al cuarto dominio polipeptídico.

En una realización aún más particular, el orden relativo de dichos dominios polipeptídicos es el mostrado en la secuencia SEQ ID NO: 52.

De acuerdo con la etapa (i) del primer método de la invención se proporciona una muestra que comprende al menos una célula que comprende una proteína de fusión según se ha definido anteriormente. Según esta etapa (i), dicha al menos una célula se ha modificado de forma que exprese dicha proteína de fusión. Para ello, dicha célula puede haber sido modificada genéticamente mediante la introducción de un ácido nucleico que codifique la proteína de fusión. La proteína de fusión puede ser introducida en dicha célula de forma que se exprese de manera estable o de manera transitoria. En una realización preferida de la invención, dicha célula expresa de forma transitoria dicha proteína de fusión. Con el fin de obtener una expresión estable es necesario incluir en la transfección un gen que codifique para resistencia a un determinado antibiótico o un gen reportero que permita la identificación de las células que han incorporado el ácido nucleico que codifica la proteína de fusión al genoma y diferenciarlas de las células en las que el ácido nucleico se encuentra en posición extracromosómica. El gen que permite seleccionar las células se puede aportar formando parte del mismo vector que contiene el ácido nucleico que codifica dicha proteína de fusión o, alternativamente, se puede aportar separadamente mediante co-transfección con un segundo plásmido que contiene dicho gen de resistencia. En este último caso, el vector que contiene el ácido nucleico que codifica dicha proteína de fusión se aporta a la mezcla de transfección en un exceso molar con respecto al gen de resistencia de forma que por cada evento de integración del gen de resistencia exista una alta probabilidad de integración del gen que contiene el promotor objeto de estudio. La selección de las células se puede llevar a cabo mediante cualquier proceso de selección convencional (véase por ejemplo Ausubel,

F.M. *et al.*, *Current Protocols in Molecular Biology* (1997) 9.5.1-9.5.19). Métodos adecuados para la introducción de material genético en la célula o células incluyen, sin limitación, precipitación con fosfato de calcio, lipofección, bombardeo de partículas, microinyección, electroporación, sistemas de dispersión coloidal (es decir, complejos de macromoléculas, nanocápsulas, microesferas, perlas y sistemas a base de lípidos incluyendo emulsiones de aceite en agua, micelas, micelas mixtas y liposomas), etc. Estos métodos son conocidos y se describen en la bibliografía publicada de manera que permita a un experto en la técnica realizar estos métodos. En una realización preferida de la invención, la modificación de las células para que expresen la proteína de fusión se lleva a cabo mediante la técnica de lipofección. Dicha técnica está basada en la introducción del ácido nucleico que codifica la proteína de fusión en un liposoma catiónico que entra en la célula por endocitosis.

Alternativamente, la célula se puede haber obtenido mediante la introducción directa de la proteína de fusión bien mediante microinyección o bien mediante modificación de la proteína de fusión con una región polipeptídica que permite la translocación de dicho polipéptido a través de membranas biológicas. Como el experto en la materia entenderá, la obtención de la célula que expresa la proteína de fusión de la invención mediante el empleo de la técnica de microinyección de dicha proteína de fusión comprende un paso previo de expresión de dicha proteína de fusión en un sistema adecuado (por ejemplo, en una célula bacteriana) y su posterior purificación mediante cualquier método apropiado para ello conocido en el estado de la técnica. Estas secuencias son conocidas de forma genérica como “dominios de transducción de proteína” (Protein-transducing domains) o “PTDs”. PTDs adecuados para su uso en la presente invención incluyen, sin limitación, polipéptidos que comprenden la región mínima de la proteína TAT de HIV formada por la secuencia de aminoácidos RKKRRQRR (residuos 49-57 de TAT) (SEQ ID NO: 53), variantes sintéticas de dicha secuencia tales como YARKARRQARR (SEQ ID NO: 54); YARAARRAARR (SEQ ID NO: 55); YARAARRAARA (SEQ ID NO: 56); YARAAARQARA (SEQ ID NO: 57), la proteína VP22 de HSV-1, polipéptidos que comprenden la secuencia RQIKIWFQNRRMKWKK (SEQ ID NO: 58), derivada de la tercera hélice del homeodominio de la proteína Antennapedia, homeodominios derivados de las proteínas Fushi tarazu (Ftz) y Engrailed (En), poliilisina, poliarginina (por ejemplo, Arg9), secuencias formadas por lisina y arginina, Transportan, MAP, MTS, o PEP-1.

De acuerdo con la etapa (ii) del primer método de la invención se introduce en dicha al menos una célula un casete de expresión o un plásmido que codifica para dicha variante de CRM1 cuya función se desea analizar. La expresión “introducir en dicha al menos una célula

un casete de expresión o un plásmido que codifica para dicha variante de CRM1”, tal y como se usa en la presente descripción, se refiere a un proceso para transportar dicho casete de expresión o plásmido a la célula. Dichos casete de expresión o plásmido que codifica para dicha variante de CRM1 pueden ser introducidos en dicha célula de forma que la variante de CRM1 se exprese de manera estable o de manera transitoria. Métodos apropiados para introducir un casete de expresión o un plásmido (es decir, un ácido nucleico) en una célula se han detallado en el contexto de la etapa (i) del primer método de la invención. En una realización preferida de la invención, dicha célula expresa de forma transitoria dicha variante de CRM1. En una realización más preferida de la invención, la introducción de una variante de CRM1 en dicha célula se realiza mediante lipofección.

Las características de la variante de CRM1 han sido definidas anteriormente. En una realización particular, dicha variante de CRM1 está fusionada a una secuencia que codifica para un polipéptido reportero que permite la detección y/o purificación de dicha variante. Es decir, opcionalmente dicho plásmido o dicho casete de expresión codifica para una proteína de fusión que comprende dicha variante de CRM1 y un polipéptido reportero. En una realización preferida, dicho polipéptido reportero es un grupo detectable. Los grupos detectables útiles incluyen fluoróforos. Por “fluoróforo” o “fluorocromo” o “cromóforo” se entiende un compuesto fluorescente que puede reemitir luz tras excitación lumínica. Los fluoróforos que se pueden usar incluyen fluoróforos biológicos (por ejemplo, proteínas) y químicos. Fluoróforos biológicos ilustrativos comprenden T-zafiro, Cerulean, mCFPm, CyPet, EGFP, PA-EGFP, Emerald, YFP, Venus, mCitrine, mKO, mOrange, DSRed, JRed, mStrawberry, mCherry, PA-mCherry, mRuby, Tomato, mPlum, mKate, mKatushka, Kaede, Halotag y flúor supereclíptico. Fluoróforos químicos ilustrativos comprenden Alexafluor, rodamina, BODIPY, tetrametilrodamina, colorantes de cianina, fluoresceína, puntos cuánticos, colorantes IR, colorantes FM, colorante ATTO.

En una realización todavía más preferida de la invención, dicho grupo detectable es un fluoróforo. Ejemplos ilustrativos no limitativos de fluoróforos han sido definidos anteriormente en la descripción. En una realización preferida de la invención, dicha variante de CRM1 está fusionada al fluoróforo biológico YFP. Como el experto en la materia entenderá, la proteína de fusión que comprende dicha variante de CRM1 y la proteína fluorescente YFP, puede comprender adicionalmente, un péptido enlazador entre ambas proteínas. El término “péptido enlazador” ha sido definido anteriormente y se usa de la misma manera.

35

En una realización todavía más preferida de la invención, dicha variante de CRM1 es CRM1 A541K fusionada al fluoróforo biológico YFP (SEQ ID NO: 59).

5 En una realización preferida, las etapas (i) y (ii) del primer método de la invención se llevan a cabo de manera simultánea. Es decir, de acuerdo con la presente invención la modificación de dicha al menos una célula para que exprese dicha proteína de fusión definida en la etapa (i) del primer método de la invención y la introducción en dicha al menos una célula de un casete de expresión o un plásmido que codifica para dicha variante de CRM1 tal y como se ha explicado en la etapa (ii) del primer método de la invención se llevan  
10 a cabo al mismo tiempo tal y como se muestra en el Ejemplo 2 de la presente solicitud. La modificación de dicha al menos una célula para que exprese dicha proteína de fusión y dicha variante de CRM1 se puede llevar a cabo mediante cualquiera de las técnicas detalladas en el contexto de la etapa (i) del primer método de la invención.

15 En una realización preferida, la transformación de dicha al menos una célula con dicha proteína de fusión y dicha variante de CRM1 se lleva a cabo mediante lipofección. En otra realización preferida, dicha al menos una célula expresa de manera transitoria ambas proteínas, la proteína de fusión y la variante de CRM1, definidas anteriormente. La expresión de proteínas simultáneamente puede llevarse a cabo mediante la utilización de un  
20 único vector de expresión que permita la expresión de dichas proteínas bajo un único promotor o bajo promotores diferentes. La expresión de proteínas simultáneamente también puede llevarse a cabo mediante la co-transfección de dos vectores de expresión independientes. En este último caso es deseable que cada vector de expresión constituya aproximadamente la mitad de la cantidad de ADN total transfectado; por ejemplo, se puede  
25 utilizar entre aproximadamente ente 0,1 y 3  $\mu\text{g}$  de cada vector, preferiblemente 500 ng del vector de expresión que codifica para dicha proteína de fusión y 500 ng del vector que codifica para dicha variante de CRM1.

La etapa (iii) del primer método de la invención comprende crecer dicha al menos una célula  
30 bajo condiciones que permitan la correcta expresión de dicha variante de CRM1. Dichas condiciones adecuadas para la expresión de una variante de CRM1 según la invención incluyen condiciones que permiten un crecimiento óptimo de dicha al menos una célula y las que permiten la expresión de la variante de CRM1 en dicha al menos una célula. El experto en la materia sabrá qué tipo de condiciones son las óptimas para cada tipo celular donde se  
35 exprese la variante de CRM1. Se elegirán los medios y condiciones apropiadas para el crecimiento óptimo de dicha al menos una célula y para la correcta expresión de la variante



de CRM1; dichos medios y condiciones de cultivo son ampliamente conocidos por los expertos en la materia. La elección de dichos medios y condiciones de cultivo dependerá del tipo celular elegido para la expresión de la variante de CRM1. En aquellas realizaciones en donde las etapas (i) y (ii) se llevan a cabo simultáneamente, las condiciones de esta etapa (iii) también se referirán a condiciones apropiadas para la expresión de la proteína de fusión.

En una forma de realización preferida, las células transformadas con dicha variante de CRM1 son de la línea celular 293T. Dichas células son transformadas con un plásmido que codifica para dicha variante de CRM1, preferiblemente para una proteína de fusión que comprende la variante de CRM1 fusionada a YFP. Las condiciones óptimas para el crecimiento celular y la expresión de dicha variante de CRM1 son, por ejemplo, esas en donde las células crecen en condiciones de saturación de CO<sub>2</sub> de aproximadamente el 5% y a una temperatura aproximada de 37°C durante un periodo de tiempo entre 24 y 72 horas. Estas condiciones también son óptimas para la expresión de la proteína de fusión.

Con el fin de comprobar si dicha variante de CRM1 se expresa correctamente, el experto en la materia puede utilizar cualquier método conocido que permita determinar la expresión de una proteína. La determinación de los niveles de expresión de las proteínas se puede llevar a cabo mediante técnicas inmunológicas tales como por ejemplo, ELISA, inmunotransferencia o inmunofluorescencia. La inmunotransferencia se basa en la detección de proteínas previamente separadas mediante electroforesis en gel en condiciones desnaturizantes e inmovilizadas en una membrana, generalmente nitrocelulosa, mediante incubación con un anticuerpo específico y un sistema de revelado (por ejemplo, quimioluminiscencia). El análisis mediante inmunofluorescencia requiere el uso de un anticuerpo específico para la proteína diana para el análisis de la expresión. El ELISA está basado en el uso de antígenos o anticuerpos marcados con enzimas de modo que los conjugados formados entre el antígeno diana y el anticuerpo marcado dan como resultado la formación de complejos enzimáticamente activos. Puesto que uno de los componentes (el antígeno o el anticuerpo marcado) están inmovilizados sobre un soporte, los complejos antígeno-anticuerpo están inmovilizados sobre el soporte y de esta manera, se pueden detectar mediante la adición de un sustrato que es convertido por la enzima en un producto que es detectable mediante, por ejemplo, espectrofotometría o fluorimetría. Cuando se usa un método inmunológico, se puede usar cualquier anticuerpo o reactivo que se sabe se une a las proteínas diana con alta afinidad para detectar la cantidad de proteínas diana. Sin embargo, se prefiere el uso de un anticuerpo, por ejemplo sueros policlonales, sobrenadantes de hibridomas o anticuerpos monoclonales, fragmentos de anticuerpos, Fv,

Fab, Fab' y F(ab')<sub>2</sub>, scFv, díacuerpos, triacuerpos, tetracuerpos y anticuerpos humanizados. También es frecuente usar un control de fondo.

5 En una realización preferida de la invención, cuando la variante de CRM1 está fusionada a un fluoróforo (por ejemplo, YFP) la determinación de la correcta expresión de la variante de CRM1 se lleva a cabo mediante visualización directa de la fluorescencia tal y como muestra el Ejemplo 2 de la presente solicitud.

10 La etapa (iv) del primer método de la invención comprende detectar la señal del dominio reportero de dicha proteína de fusión que comprende la variante de survivina. El polipéptido reportero permite la detección y/o purificación de dicha proteína de fusión. Virtualmente, se puede usar cualquier método convencional para detectarla, por ejemplo mediante inmunotransferencia o inmunofluorescencia. La inmunofluorescencia (IF) es una técnica basada en microscopía de fluorescencia utilizada principalmente para testar muestras biológicas. Esta técnica usa la especificidad de los anticuerpos hacia su antígeno lo que permite la visualización de la molécula diana en la muestra. Se puede usar IF en líneas cultivadas o en células individuales, y se puede usar para analizar la distribución de proteínas, glicanos y pequeñas moléculas biológicas y no biológicas. Además, también se puede usar IF en combinación con otros métodos de tinción fluorescente no de anticuerpos, 15 tales como por ejemplo DAPI, para marcar ADN. Se puede usar más de un anticuerpo al mismo tiempo para detectar, por ejemplo más de una proteína. Se pueden usar varios diseños de microscopio para el análisis de muestras de IF, tal como el microscopio de epifluorescencia y el microscopio confocal. También se pueden usar varios diseños de microscopio de super-resolución cuya capacidad de resolución es mucho mayor.

25 En una realización preferida de la invención, la etapa (iv) del primer método de la invención comprende detectar el epítipo flag. La detección de dicho epítipo flag puede llevarse a cabo mediante la técnica de inmunofluorescencia indirecta mediante el empleo de un anticuerpo primario específico que reconozca dicho epítipo y el empleo de un anticuerpo secundario específico para dicho anticuerpo primario acoplado a un grupo detectable. Si se desea, la señal de fluorescencia puede ser cuantificada. Por ejemplo, dicha al menos una célula que expresa la proteína de fusión según se ha definido en la presente invención, se puede fijar con una solución de formaldehído (2-4%) y permeabilizar con Tx-100 al 0,1-0,3%. La detección de la señal del dominio reportero puede llevarse a cabo mediante la 30 incubación de la célula con un anticuerpo específico primario que reconozca el epítipo flag (por ejemplo, Flag M2, Sigma) seguido de la incubación con un anticuerpo específico

secundario que reconozca dicho anticuerpo primario marcado con un fluoróforo, por ejemplo AlexaFluor594 (Invitrogen). En una realización preferida de la invención, la detección de dicha señal del dominio polipeptídico reportero se realiza al mismo tiempo que se detecta la correcta expresión de la variante de CRM1.

5

Opcionalmente, esta etapa (iv) comprende detectar una estructura o compartimento celular (por ejemplo, el núcleo, retículo endoplasmático, aparato de Golgi, mitocondrias etc.) que puede tomarse como referencia y facilite la determinación subcelular de la proteína de fusión en la etapa (v) del presente método. Así en una realización todavía más preferida de la invención, esta etapa (iv) comprende la tinción del núcleo celular mediante el empleo de cualquier colorante apropiado para ello, por ejemplo DAPI o el colorante Hoescht. El experto en la materia entenderá que para la detección simultánea de dos o más proteínas en una célula es necesario emplear fluoróforos de colores distintos para que los espectros de emisión de los mismos no solapen. El Ejemplo 2 de la solicitud muestra la detección simultánea de la proteína de fusión definida en la primera etapa del primer método de la invención mediante el empleo de un anticuerpo secundario unido a un fluoróforo, cuyo espectro de emisión está localizado en la región roja del espectro visible; y de la variante de CRM1 expresada como una proteína de fusión que comprende dicha variante de CRM1 y la proteína fluorescente YFP cuyo espectro de emisión está localizado en la zona amarilla del espectro visible.

10  
15  
20

La etapa (v) del primer método de la invención comprende determinar la localización subcelular de la proteína de fusión definida en la etapa (i) de dicho primer método de la invención. El término "localización subcelular" según se usa en la presente descripción, se refiere a la ubicación de dicha proteína de fusión en un compartimento celular. El término "compartimento celular" tal y como se utiliza en la presente descripción, se refiere a las partes cerradas dentro del citosol de una célula eucariota normalmente rodeadas por una membrana lipídica de capa única o doble capa. La mayoría de los orgánulos celulares son compartimentos celulares: mitocondrias, cloroplastos (en células vegetales), peroxisomas, liposomas, retículo endoplasmático, aparato de Golgi y núcleo celular. La determinación de la localización subcelular de dicha proteína de fusión se lleva a cabo analizando la señal del dominio polipeptídico reportero detectada en la etapa (iv) del primer método de la invención mediante cualquier técnica apropiada para ello, preferiblemente mediante inmunofluorescencia. Por lo tanto, si la señal del dominio polipeptídico reportero se detecta en el interior del compartimento nuclear entonces la localización subcelular de la proteína de

25  
30  
35

fusión es nuclear; si la señal del dominio polipeptídico reportero se detecta en el citoplasma, entonces la localización subcelular de la proteína de fusión es citoplasmática.

5 La proteína de fusión definida en la etapa (i) del primer método de la invención comprende secuencias de localización, NLS y NES que determinan la localización subcelular de dicha proteína. En una realización particular de la invención, la proteína de fusión según se ha definido en la etapa (i) del primer método de la invención, comprende un dominio polipeptídico que comprende una señal NES, preferiblemente la señal NES de survivina, y dos dominios polipeptídicos en donde cada uno de dichos dos dominios comprende una  
10 señal NLS, preferiblemente la señal NLS del antígeno T del virus SV40. En esta forma de realización, la localización subcelular de la proteína de fusión es nuclear.

La etapa (vi) del primer método de la invención comprende comparar la localización subcelular de la proteína de fusión determinada en la etapa (v) de dicho primer método con  
15 dicha localización subcelular obtenida para dicha proteína de fusión en una muestra de referencia. El término "muestra de referencia" tal y como se utiliza en el presente documento, se refiere a una muestra que comprende al menos una célula que comprende:  
a) la proteína de fusión según se ha definido en la etapa (i) del primer método de la invención y b) CRM1 o una variante de CRM1 cuya capacidad de unión a secuencias NES  
20 no está alterada. Técnicas para determinar si CRM1 es capaz de unir secuencias NES son conocidas en el estado de la técnica. Ilustrativamente dichas técnicas incluyen, pero no están limitadas a, ensayos *in vitro* que permiten la reconstitución *in vitro* de un complejo de exportación formado por CRM1 y RanGTP y en los que se determina dicha interacción mediante ensayos basados en la cuantificación de la hidrólisis de GTP (Askjaer et al., Mol.  
25 Cell Biol., 19: 6276-6285). Otros métodos conocidos en el estado de la técnica apropiados para determinar si dicha variante de CRM1 no tiene alterada la capacidad de interactuar con secuencias NES incluyen pero no están limitados a ensayos de cromatografía de afinidad, pull down, inmunoprecipitación, colocalización etc.

30 En una realización preferida de la invención, dicha muestra de referencia comprende la proteína CRM1. En una realización todavía más preferida de la invención, dicha CRM1 es CRM1 humana.

En esta etapa (vi) dicha muestra de referencia es una muestra que comprende al menos una  
35 célula que comprende una proteína de fusión según se ha definido anteriormente y CRM1 o una variante de la misma con capacidad de unión a secuencias NES. Por lo tanto, dicha al

menos una célula se ha modificado genéticamente de forma que exprese dichas proteínas. Métodos adecuados para la introducción de material genético en una célula han sido detallados en el contexto de la etapa (i) del primer método de la invención. En una realización preferida de la invención, dicha muestra de referencia comprende al menos una  
5 célula que expresa de forma transitoria dichas proteínas. En una realización aún más preferida de la invención, la modificación de dicha célula para que exprese dichas proteínas se lleva a cabo mediante la técnica de lipofección.

Tal y como se ha explicado anteriormente, es necesario proporcionar condiciones  
10 adecuadas que permitan la expresión de dichas proteínas en la muestra de referencia. Condiciones adecuadas para la correcta expresión de la proteína de fusión y CRM1 o de una variante de CRM1 han sido detalladas en la etapa (ii) del primer método de la invención. En una forma de realización preferida, las células transformadas con dicha CRM1 son de la línea celular 293T. Con el fin de comprobar si dichas proteínas se expresan correctamente,  
15 el experto en la materia puede utilizar cualquier método conocido que permita determinar la expresión de una proteína. En una realización preferida de la invención, dicha CRM1 o variante de CRM1 con capacidad de unión a secuencias NES comprende, adicionalmente, una secuencia que codifica para un polipéptido reportero que permite la detección y/o purificación de dicha variante. En una realización todavía más preferida, dicho grupo  
20 detectable es un fluoróforo. Ejemplos ilustrativos no limitativos de fluoróforos han sido mencionados anteriormente en la descripción. En una realización preferida de la invención, dicha CRM1 o dicha variante de CRM1 con capacidad de unión a secuencias NES está fusionada al fluoróforo biológico YFP. Como el experto en la materia entenderá, la proteína de fusión, que comprende dicha CRM1 o dicha variante de CRM1 con capacidad de unión a  
25 secuencias NES y la proteína fluorescente YFP, puede comprender adicionalmente, un péptido enlazador entre ambas proteínas. En una realización todavía más preferida, dicha muestra de referencia comprende la proteína CRM1 fusionada a YFP que comprende la secuencia de aminoácidos mostrada en la SEQ ID NO: 60.

30 La etapa (vi) del primer método de la invención comprende detectar la señal del dominio polipeptídico reportero de la proteína de fusión en la muestra de referencia. Métodos para detectar dicha señal han sido detallados en el contexto de la etapa (iv) del primer método de la invención. En una realización preferida de la invención, el método empleado es inmunofluorescencia.

35

Preferiblemente el método empleado en la etapa (iv) de dicho primer método de la invención para detectar la señal del dominio polipeptídico reportero de la proteína de fusión en la muestra problema y el método empleado en la etapa (vi) para detectar la señal del dominio reportero de dicha proteína de fusión en la muestra de referencia es el mismo. En una realización preferida, el método empleado en las etapas (iv) y (vi) para detectar la señal del dominio polipeptídico reportero de la proteína de fusión en la muestra problema y en la muestra de referencia, de acuerdo al primer método de la invención es inmunofluorescencia.

La etapa (vi) del primer método de la invención comprende comparar la localización subcelular de la proteína de fusión definida anteriormente en la muestra que comprende al menos una célula en donde se desea determinar si CRM1 es funcional con la señal obtenida en una muestra de referencia que comprende CRM1 o una variante de CRM1 cuya capacidad de unión a secuencias NES no está alterada. Así, una alteración de la localización de dicha proteína de fusión es indicativa de una función alterada de CRM1.

Tal y como se ha detallado anteriormente en la etapa (i) del primer método de la invención, la proteína de fusión se localiza en el núcleo de la célula debido a que las señales NLS la dirigen hacia dicho compartimento celular. En presencia de la proteína CRM1 o de la variante de CRM1 en dicha muestra de referencia, la localización de la proteína de fusión será citoplasmática debido a que la capacidad de exportación de CRM1 o de dicha variante en la muestra de referencia está intacta. Por lo tanto, en la muestra de referencia la proteína de fusión se transporta desde el núcleo hacia el citoplasma.

La expresión “alteración de la localización subcelular de la proteína de fusión”, tal y como se utiliza en el presente documento, hace referencia a que la localización subcelular de dicha proteína de fusión es una localización subcelular distinta a la que ocupa en dicha muestra de referencia. En una realización particular de la invención, dicha alteración de la localización de la proteína de fusión se refiere a que, en presencia de la variante de CRM1 cuya función se desea determinar, la localización de la proteína de fusión es nuclear.

El primer método de la invención puede llevarse a cabo, ilustrativamente, tal y como se muestra en el Ejemplo 2 del presente documento.

Método para identificar un compuesto con capacidad de modular la función de CRM1

El diseño de la proteína de fusión que han desarrollado los inventores con capacidad de relocalizar al citoplasma en respuesta a una sobreexpresión de CRM1 funcional permite también la identificación de compuestos que modulen la función exportadora de CRM1.

5 Por lo tanto, en un segundo aspecto inventivo, la invención se refiere a un método, de aquí en adelante “segundo método de la invención”, para identificar un compuesto con capacidad de modular la función de CRM1 o de una variante funcionalmente equivalente de la misma en una muestra que comprende al menos una célula, comprendiendo dicho método:

10 (i) poner en contacto un compuesto candidato con una muestra que comprende al menos una célula en donde dicha al menos una célula comprende:

a) una proteína de fusión que comprende:

1) un primer dominio polipeptídico que comprende una variante de survivina seleccionada entre:

15 - una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional; y

20 - una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende una señal de exportación nuclear (NES) de survivina;

2) al menos, un segundo dominio polipeptídico que comprende una señal de localización nuclear (NLS); y

3) al menos, un tercer dominio polipeptídico reportero;

25 en donde si la variante de survivina carece de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además

4) un cuarto dominio polipeptídico que comprende una NES;

b) un casete de expresión o un plásmido que codifica para dicha CRM1 o dicha variante funcionalmente equivalente de la misma; y

30 (ii) detectar la señal del dominio reportero de dicha proteína de fusión;

(iii) determinar la localización subcelular de dicha proteína de fusión; y

(iv) comparar la localización subcelular de dicha proteína de fusión obtenida en la etapa (iii) con la localización subcelular de dicha proteína de fusión obtenida en una muestra que comprende al menos una célula que no ha sido tratada con dicho compuesto;

35 en donde una alteración de la localización subcelular de la proteína de fusión detectada en la etapa (iii) con respecto a dicha localización en dicha muestra no tratada es indicativa de

que dicho compuesto es capaz de modular la función de CRM1 o de una variante funcionalmente equivalente de la misma.

5 El término "variante funcionalmente equivalente de CRM1" tal y como se utiliza en el presente documento, se refiere a todos aquellos polipéptidos derivados de la secuencia de CRM1 mediante modificación, inserción y/o delección de uno o más aminoácidos siempre y cuando se mantenga sustancialmente la función de la proteína CRM1. En concreto, las variantes funcionalmente equivalentes mantendrán la capacidad de unión a secuencias NES. También serán variantes funcionalmente equivalentes de CRM1 aquellas con capacidad de exportar macromoléculas desde el núcleo al citoplasma. Si se desea, la capacidad de exportar macromoléculas (por ejemplo, proteínas) desde el núcleo al citoplasma para determinar si una variante de CRM1 es una variante funcionalmente equivalente de la misma puede ser determinada mediante el primer método de la invención. Métodos para determinar la unión de CRM1 a secuencias NES son conocidos en el estado de la técnica y se han detallado en el contexto del primer método de la invención.

15 Variantes funcionalmente equivalentes de CRM1 incluyen aquellas que muestran al menos un 25%, al menos 40%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, al menos 96%, al menos 97%, al menos 98% o al menos 99% de identidad de secuencia de aminoácidos con respecto a la secuencia de CRM1 indicada anteriormente en el contexto del primer método de la invención. El grado de identidad entre dos secuencias de aminoácidos puede determinarse por métodos convencionales mencionados en el contexto del primer método de la invención. El experto en la materia entenderá que las secuencias de aminoácidos a las que se hace referencia en esta descripción pueden estar modificadas químicamente, por ejemplo, mediante modificaciones químicas que son fisiológicamente relevantes, tales como, fosforilaciones, acetilaciones, etc.

20 La expresión "compuesto con capacidad de modular la función de CRM1 o de una variante funcionalmente equivalente de la misma" tal y como se usa en el presente documento, se refiere a cualquier compuesto que se une específicamente a CRM1 (o a la variante funcionalmente equivalente de la misma) y que al unirse es capaz de causar una alteración en la función de dicha proteína con respecto a un valor de referencia. En una realización preferida, dicha alteración en la función de CRM1 se refiere a la inhibición de la función de CRM1. Como el experto en la materia entenderá, la alteración en la función de dicha proteína puede llevar asociado una alteración en el nivel de ARNm o de proteína CRM1 con respecto a un valor de referencia.



La etapa (i) del segundo método de la invención comprende poner en contacto un compuesto candidato con una muestra que comprende una al menos una célula que comprende una proteína de fusión, en donde dicha proteína de fusión ha sido definida anteriormente en el contexto del primer método de la invención.

El término “poner en contacto” un compuesto candidato con dicha muestra, tal y como se utiliza en la presente descripción, se refiere a cualquier posible forma de llevar dicho compuesto candidato hasta el interior de la célula que expresa dicha proteína de fusión y dicha CRM1 o una variante funcionalmente equivalente de la misma. En caso de que el compuesto candidato sea una molécula de bajo peso molecular, es suficiente con añadir dicha molécula al medio de cultivo. El término “compuesto de bajo peso molecular”, tal y como se usa en el presente documento se refiere a compuestos cuyo peso molecular es de aproximadamente, igual o menor a 900 Daltons. Ejemplos ilustrativos no limitativos de dichos compuestos incluyen compuestos naturales como, por ejemplo, metabolitos secundarios tales como alcaloides, glicósidos, lípidos, péptidos no ribosomales, fezaninas, fenoles naturales (incluyendo flavonoides), poliquétidos, terpenos, tetrapirroles etc., y compuestos artificiales como fármacos sintetizados mediante síntesis química.

En caso de que el compuesto candidato sea una molécula de alto peso molecular (por ejemplo polímeros biológicos tales como un ácido nucleico o una proteína), es necesario aportar los medios para que dicha molécula pueda acceder al interior celular. El término “compuesto de alto peso molecular” tal y como se usa en la presente descripción, se refiere a compuestos cuyo peso molecular es, aproximadamente, mayor a 900 Daltons. En caso de que el compuesto candidato sea una proteína, la célula puede ponerse en contacto tanto con la proteína como con el ácido nucleico que la codifica acoplado a elementos que permitan su transcripción/traducción una vez que se encuentre en el interior celular. Para ello se puede emplear cualquiera de los métodos mencionados anteriormente en el contexto del primer método de la invención.

El compuesto a ensayar puede encontrarse aislado o formando parte de una mezcla más o menos compleja, bien derivada de una fuente natural, o bien formando parte de una biblioteca de compuestos. Ejemplos de bibliotecas de compuestos que pueden ser ensayadas según el método de la presente invención incluyen, sin limitación, bibliotecas de péptidos incluyendo tanto péptidos como análogos peptídicos que comprenden D-amino ácidos o péptidos que comprenden enlaces no peptídicos, bibliotecas de ácidos nucleicos

incluyendo ácidos nucleicos con enlaces no fosfodiéster del tipo de fosforotioato o ácidos nucleicos peptídicos, bibliotecas de anticuerpos, de carbohidratos, de compuestos de bajo peso molecular, preferiblemente moléculas orgánicas, de peptidomiméticos, y similares. En el caso de que se use una biblioteca de compuestos orgánicos de bajo peso molecular, la biblioteca puede haber sido preseleccionada para que contenga compuestos que puedan acceder al interior celular con mayor facilidad. Así, los compuestos se pueden seleccionar en base a determinados parámetros tales como tamaño, lipofilicidad, hidrofiliicidad, capacidad de formar puentes de hidrógeno, etc.

En caso de que el compuesto candidato se encuentre formando parte de una mezcla de mayor o menor complejidad, el método de la invención comprende adicionalmente una o varias etapas de fraccionamiento de la mezcla ensayada y la repetición de los pasos (i) a (iv) con cada una de las fracciones obtenidas hasta que se aísla el compuesto en la mezcla con capacidad de modular la función de CRM1 o una variante funcionalmente equivalente de la misma. Métodos para el fraccionamiento de compuestos presentes en una mezcla incluyen cromatografía (en capa fina, de gases o de exclusión molecular en gel, de afinidad), cristalización, destilación, filtración, precipitación, sublimación, extracción, evaporación, centrifugación, espectrometría de masas, adsorción y similares. Alternativamente, los compuestos a ensayar pueden estar formando parte de un extracto obtenido de una fuente natural. La fuente natural puede ser animal, vegetal obtenido de cualquier entorno, incluyendo, sin limitación, extractos de organismos terrestres, aéreos, marinos y similares.

El término “muestra” así como las realizaciones particulares de la muestra empleada en la presente invención, han sido definidos anteriormente en el primer método de la invención y se usan de la misma manera en este segundo aspecto.

El término “proteína de fusión” ha sido definido en el contexto del primer método de la invención y se usa de la misma manera en este segundo método de la invención. Las características de cada uno de los dominios polipeptídicos de la proteína de fusión definida en la etapa (i) del segundo método de la invención y las realizaciones particulares de dichos dominios polipeptídicos, han sido detalladas en la etapa (i) del primer método de la invención y se aplican aquí de igual manera .

En la etapa (i) del segundo método de la invención se pone en contacto el compuesto candidato con una muestra que comprende al menos una célula que comprende una proteína de fusión según se ha definido anteriormente y un casete de expresión o un

plásmido que codifica para CRM1 o una variante funcionalmente equivalente de la misma. Según esta etapa (i), dicha al menos una célula se ha modificado de forma que expresa dicha proteína de fusión y dicha CRM1 (o una variante funcionalmente equivalente de la misma). Métodos adecuados para modificar dicha al menos una célula para que exprese  
5 dicha proteína de fusión han sido detallados en el contexto del primer método de la invención. Dichos métodos son igualmente aplicables para la transfección de dicha al menos una célula con un casete de expresión o un plásmido que codifica para dicha CRM1 o una variante funcionalmente equivalente de la misma. En una realización preferida de la invención, dicha célula expresa de forma transitoria ambas proteínas. Técnicas apropiadas  
10 para la expresión transitoria de proteínas han sido detalladas en el contexto del primer método de la invención. El experto en la materia entenderá que en otras realizaciones de la invención, puede ser deseable que la expresión de dicha proteína de fusión y dicha CRM1 o una variante funcionalmente equivalente de la misma sea estable.

15 Tal y como se ha explicado en la etapa (iii) del primer método de la invención, dicho plásmido o dicho casete de expresión codifica para una proteína de fusión comprende dicha CRM1 o dicha variante funcionalmente equivalente de la misma, fusionada a un polipéptido reportero, preferiblemente un grupo detectable, aún más preferiblemente dicha CRM1 o una variante funcionalmente equivalente de la misma está fusionada al fluoróforo biológico YFP.

20 La etapa (ii) del segundo método de la invención comprende detectar la señal del dominio polipeptídico reportero de dicha proteína de fusión. Condiciones y métodos apropiados para la detección de dicho dominio polipeptídico reportero se han descrito en detalle en el primer método de la invención y se aplican aquí de igual manera. Opcionalmente, dicha etapa  
25 comprende la detección de una estructura celular o de un compartimento celular (por ejemplo, el núcleo celular) que puede ser utilizado como referencia para determinar la localización subcelular de la proteína de fusión en la etapa (iii) del segundo método de la invención.

30 La etapa (iii) del segundo método de la invención comprende determinar la localización subcelular de la proteína de fusión. El término "localización subcelular" ha sido definido en el primer método de la invención y se usa aquí de igual manera. La determinación de la localización subcelular de dicha proteína de fusión se lleva a cabo analizando la señal del dominio reportero detectada en la etapa (ii) del segundo método de la invención mediante  
35 cualquier técnica apropiada para ello, preferiblemente mediante inmunofluorescencia. Por lo tanto, si la señal del dominio polipeptídico reportero se detecta en el interior del

compartimento nuclear entonces la localización subcelular de la proteína de fusión es nuclear; Por el contrario, si la señal del dominio polipeptídico reportero se detecta en el citoplasma, entonces la localización subcelular de la proteína de fusión es citoplasmática.

5 La etapa (iv) del segundo método de la invención comprende comparar la localización subcelular de la proteína de fusión determinada en la etapa (iii) de dicho método con dicha localización subcelular obtenida para dicha proteína de fusión en una muestra que comprende al menos una célula que no ha sido tratada con dicho compuesto. El término “muestra que no ha sido tratada” tal y como se utiliza en el presente documento, se refiere a  
10 una muestra que comprende al menos una célula en donde dicha al menos una célula comprende dicha proteína de fusión y dicho casete o plásmido que codifica para CRM1 o una variante funcionalmente equivalente de la misma definidas en la primera etapa del segundo método de la invención, con la particularidad de que dicha muestra no se pone en contacto con el compuesto candidato que se desea analizar.

15 La modificación de dicha al menos una célula en la muestra no tratada para que exprese dichas proteína de fusión y CRM1, o una variante funcionalmente equivalente de la misma, puede llevarse a cabo empleando cualquier método del estado de la técnica apropiado para ello tal y como se ha detallado en el contexto del primer método de la invención.

20 Preferiblemente la muestra no tratada será de la misma naturaleza que la muestra que se pone en contacto con el compuesto candidato. Por ejemplo, si la muestra que se pone en contacto con el compuesto candidato es una población de células 293T entonces la muestra no tratada también es una población de células 293T, preferiblemente ambas muestras  
25 tendrán aproximadamente el mismo número de células y serán sometidas a las mismas condiciones de cultivo. Si se desea, la modificación de dicha al menos una célula de la muestra no tratada y de la muestra a analizar puede llevarse a cabo simultáneamente. Preferiblemente, se emplea la misma técnica para la transformación celular y bajo las mismas condiciones. Por ejemplo, si la muestra que va a ser tratada con el compuesto candidato se  
30 transfecta mediante lipofección con entre aproximadamente ente 0,1 y 3 µg de cada vector, preferiblemente 500 ng del vector de expresión que codifica para dicha proteína de fusión y 500 ng del vector que codifica para dicha CRM1, entonces preferiblemente, la muestra que no va a ser tratada se transfecta con dichas mismas cantidades de cada plásmido.

35 La detección del dominio reportero de la proteína de fusión en la muestra que no ha sido tratada puede llevarse a cabo por cualquier técnica apropiada para ello. Preferiblemente, se

lleva a cabo empleando la misma técnica empleada para la detección del dominio polipeptídico reportero en la muestra que se pone en contacto con el compuesto candidato. La detección del dominio polipeptídico reportero de la proteína de fusión puede llevarse a cabo en paralelo en la muestra tratada con el compuesto candidato y en la muestra no tratada. Tras la detección del dominio polipeptídico reportero en la muestra tratada con el compuesto candidato y en la muestra no tratada, la etapa (iv) del segundo método de la invención comprende comparar la localización subcelular de dicha proteína de fusión en ambos tipos de muestra.

Como se ha explicado anteriormente en el contexto del primer método de la invención, si dicha al menos una célula de la muestra que se está analizando expresa la proteína CRM1 funcional, entonces la localización subcelular de la proteína de fusión es citoplasmática puesto que dicha exportina reconoce a la proteína de fusión a través de la secuencia NES, preferiblemente la secuencia NES de survivina humana, y en asociación con una GTPasa transporta dicha proteína desde el núcleo hacia el citoplasma. Por lo tanto, de acuerdo con el segundo método de la invención, la localización subcelular de la proteína de fusión en la muestra no tratada es citoplasmática. Una alteración de la localización citoplasmática de la proteína de fusión en la muestra tratada es indicativa de que dicho compuesto podría impedir el correcto desarrollo de la función de CRM1 o de una variante equivalente de la misma, es decir de la podría modular la función de CRM1 o de dicha variante equivalente de la misma.

#### Proteína de fusión de la invención

Los autores de la presente invención han diseñado una proteína de fusión que actúa como un biosensor indicador de la actividad exportadora de la exportina CRM1.

Por lo tanto, en un tercer aspecto, la invención se relaciona con una proteína de fusión que comprende:

- (i) un primer dominio polipeptídico que comprende una variante de survivina seleccionada entre:
  - a) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional ; y

35

b) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende la señal de exportación nuclear (NES) de survivina

5 (ii) al menos, un segundo dominio polipeptídico que comprende una señal de localización nuclear (NLS); y

(iii) al menos, un tercer dominio polipeptídico reportero;

10 en donde si la variante de survivina carece de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además

(iv) un cuarto dominio polipeptídico que comprende una NES.

15 El término “proteína de fusión”, “dominio polipeptídico” y métodos para generar una proteína de fusión se han mencionado anteriormente y se usan de la misma manera en este tercer aspecto inventivo.

20 El primer, segundo, tercer y cuarto dominio polipeptídico de la proteína de fusión de la invención así como las realizaciones particulares de los mismos han sido detallados en el contexto del primer método de la invención y se aplican aquí de igual manera.

25 Tal y como se ha detallado en el primer y segundo aspecto inventivo, la proteína de fusión de la invención es útil para determinar si la función de una variante de CRM1 en una muestra está alterada y/o para identificar un compuesto con capacidad de modular la función de CRM1 o de una variante equivalente de la misma. Por lo tanto, en otro aspecto la invención se relaciona con el uso de la proteína de fusión de la invención para determinar si la función de una variante de CRM1 en una muestra está alterada y/o para identificar un compuesto con capacidad de modular la función de CRM1 o de una variante funcionalmente  
30 equivalente de la misma.

#### Ácidos nucleicos, casetes de expresión y vectores de la invención

35 En otro aspecto, la presente invención se relaciona con un ácido nucleico que codifica para la proteína de fusión de la invención y de cualquiera de sus realizaciones. El término “ácido nucleico”, según se usa en la presente descripción, se refiere a polímeros formados por la

repetición de monómeros llamados nucleótidos, unidos mediante enlaces fosfodiéster. En una realización particular, dicho ácido nucleico comprende la secuencia de nucleótidos mostrada en SEQ ID NO: 61.

5 Dicho ácido nucleico de la invención puede incorporar, operativamente unida, una secuencia reguladora de la expresión de las secuencias de nucleótidos que codifican para la proteína de fusión de la invención, constituyendo de este modo una construcción génica. Según se utiliza en el presente documento, la expresión "operativamente unida" significa que el polipéptido de la proteína de fusión codificado por la secuencia de ácido nucleico de la  
10 invención, es expresado en el marco de lectura correcto bajo el control de las secuencias de control o reguladoras de expresión. Por tanto, en otro aspecto, la invención proporciona un casete de expresión que comprende la construcción génica de la invención operativamente unida a una secuencia de control de expresión. La construcción génica de la invención puede obtenerse mediante el empleo de técnicas ampliamente conocidas en el estado de la  
15 técnica [Sambrook *et al.*, "Molecular cloning, a Laboratory Manual", 2<sup>nd</sup> ed., Cold Spring Harbor Laboratory Press, N.Y., 1989 Vol 1-3].

Las secuencias de control son secuencias que controlan y regulan la transcripción y, en su caso, la traducción de dicha proteína de fusión, e incluyen secuencias promotoras,  
20 secuencias codificantes para reguladores transcripcionales, secuencias de unión a ribosomas (RBS) y/o secuencias terminadoras de transcripción. El casete de expresión de la presente invención puede incluir además un potenciador, que puede estar adyacente o distante a la secuencia del promotor y puede funcionar aumentando la transcripción a partir del mismo. Dicha secuencia de control de expresión es funcional en células y organismos  
25 eucariotas, por ejemplo, células de insecto, células vegetales, células de mamífero, etc.

Se puede usar cualquier promotor disponible en la presente metodología. En una forma de realización preferida de la presente invención, el promotor utilizado para la construcción de ácido nucleico de la presente invención es activo en la población celular específica  
30 transformada. Ejemplos ilustrativos, no limitativos, de promotores ubicuos que pueden estar presentes en el casete de expresión proporcionado por esta invención incluyen el promotor de citomegalovirus humano (hCMV), el promotor de SV40, el promotor para EF1-alfa, y el promotor de ubiquitina C. Ejemplos ilustrativos, no limitativos, de promotores específicos de tipo celular y/o específicos de tejido incluyen promotores tales como albúmina que es  
35 específico de hígado [Pinkert *et al.*, (1987) Genes Dev. 1:268-277], promotores linfoides específicos [Calame *et al.*; (1988) Adv. Immunol. 43:235-275]; en particular promotores de

los receptores de células T [Winoto *et al.*, (1989) EMBO J. 8:729-733] e inmunoglobulinas; [Banerji *et al.* (1983) Cell 33729-740], promotores específicos de neuronas tal como el promotor de neurofilamentos [Byrne *et al.* (1989) Proc. Natl. Acad. Sci. USA 86:5473-5477], promotores específicos de páncreas [Edlunch *et al.* (1985) Science 230:912-916] o  
 5 promotores específicos de glándula mamaria tal como el promotor de suero de leche (patente de EE UU No. 4.873.316 y publicación de solicitud europea No. 264.166). El casete de expresión CAG-GS está compuesto de un elemento del enhancer CMV, el promotor de la  $\beta$ -actina de pollo y el elemento regulador postranscripcional (WPRE) del virus de la hepatitis de marmota (Woodchuck Hepatitis Virus, WHP) [Niwa *et al.* (1991) *Gene* 108: 193-9]. La  
 10 combinación del promotor y este elemento favorece unos niveles de expresión del transgen *in vivo* altos.

Ventajosamente, dicho casete de expresión comprende, además, un marcador o gen que codifica para un motivo o para un fenotipo que permita la selección de la célula hospedadora  
 15 transformada con dicho casete de expresión. Ejemplos ilustrativos de dichos marcadores que podrían estar presentes en el casete de expresión de la invención incluyen genes de resistencia a antibióticos, genes de resistencia a compuestos tóxicos, y, en general, todos aquellos que permitan seleccionar a las células manipuladas genéticamente.

20 La construcción génica de la invención, o el casete de expresión proporcionado por esta invención, puede ser insertada en un vector apropiado. Así, en otro aspecto, la invención se relaciona con un vector, tal como un vector de expresión, que comprende dicha construcción de génica de la invención o dicho casete de expresión. La elección del vector dependerá de la célula hospedadora en la que se va a introducir posteriormente. A modo ilustrativo, el  
 25 vector donde se introduce dicha secuencia de ácido nucleico puede ser un plásmido o un vector que, cuando se introduce en una célula hospedadora, se integra o no en el genoma de dicha célula. La obtención de dicho vector puede realizarse por métodos convencionales conocidos por los técnicos en la materia [Sambrook *et al.*, 1989, citado *supra*]. En una realización particular, dicho vector recombinante es un vector útil para transformar células  
 30 animales.

Dicho vector puede ser utilizado para transformar, transfectar o infectar células eucariotas susceptibles de ser transformadas, transfectadas o infectadas por dicho vector. Por tanto, en otro aspecto, la invención se relaciona con una célula hospedadora transformada,  
 35 transfectada o infectada con un vector proporcionado por esta invención. Dicha célula



transformada, transfectada o infectada comprende, por tanto, una construcción génica de la invención, o bien dicho casete de expresión o vector proporcionado por esta invención.

5 Células transformadas, transfectadas o infectadas pueden ser obtenidas por métodos convencionales conocidos por los técnicos en la materia (Sambrook et al., 1989, citado *supra*). Células adecuadas para llevar a cabo la invención, incluyen, sin limitación, células de mamíferos, plantas, insectos y de hongos. Ejemplos ilustrativos y no limitativos de dichas células se encuentran en el contexto del primer método de la invención y se usan aquí de igual manera.

10

En una realización preferida, dicha célula hospedadora es una célula animal transformada, transfectada o infectada con un vector apropiado, siendo dicha célula animal transformada, transfectada o infectada capaz de expresar la proteína de fusión proporcionada por esta invención, por lo que dichos vectores pueden utilizarse para la expresión en células animales de la proteína de fusión proporcionada por esta invención.

15

La construcción génica, o casete de expresión, o vector o célula hospedadora de la invención pueden ser utilizados para producir una proteína de fusión que comprende:

20

(i) un primer dominio polipeptídico que comprende una variante de survivina seleccionada entre:

25

a) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional ;  
y

30

b) una variante de survivina que carece de la totalidad o parte del dominio de homodimerización resultando en un dominio de homodimerización no funcional y comprende la señal de exportación nuclear (NES) de survivina

35

(i) al menos, un segundo dominio polipeptídico que comprende una señal de localización nuclear (NLS); y

(ii) al menos, un tercer dominio polipeptídico reportero;

en donde si la variante de survivina carece de la totalidad o parte de su señal NES, entonces dicha proteína de fusión comprende además

(iii) un cuarto dominio polipeptídico que comprende una NES.

5

\*\*\*

Los siguientes ejemplos son meramente ilustrativos y no se deben de considerar como limitativos de la presente invención.

10

## EJEMPLOS

### EJEMPLO 1

#### 15 Generación de la proteína de fusión

Para construir la proteína de fusión, se llevaron a cabo varias etapas de clonación utilizando métodos estándar de biología molecular (PCR, digestión con enzimas de restricción, ligación y transformación de bacterias competentes):

20

Etapa 1- En el vector pEGFP-N1, se reemplazó la secuencia de DNA que codifica GFP por otra que codifica el epítipo 3xFlag. Sitios de restricción utilizados: PinAI/NotI.

25

Etapa 2- Se añadió en el extremo 5' de 3xFlag una secuencia de DNA que codifica una NLS del antígeno T grande del virus SV40 (SV40NLS). Sitios de restricción utilizados: HindIII/BamHI.

30

Etapa 3.- Se añadió en el extremo 5' del cassette [SV40NLS-3xFlag] una secuencia de DNA que codifica el fragmento 1-100 de la proteína Survivin humana. Sitios de restricción utilizados: BglII/HindIII.

35

Etapa 4. Se reemplazó el cassette [SV40NLS-3xFlag] por otro cassette [SV40NLS-3xFlag-SV40NLS]. Sitios de restricción utilizados: HindIII/NotI

Tras cada etapa de clonación se confirmó la secuencia recombinante obtenida mediante secuenciación.

### EJEMPLO 2

Determinación de la funcionalidad de variantes de CRM1

Para determinar si una variante mutada de CRM1 es funcional mediante la proteína de fusión de la invención, se procedió del siguiente modo.

- 5           1. Se sembraron células 293T en tres pocillos de una placa de cultivo sobre cubreobjetos de cristal estériles, y se transfectaron el día siguiente utilizando el reactivo de transfección FugeneHD (Roche) de acuerdo con el protocolo del fabricante. En una muestra se transfectó sólo el ácido nucleico que codifica para la proteína de fusión (SEQ ID NO: 52), en otra muestra se transfectó dicha proteína de fusión junto con YFP-CRM1 wild type (normal) (SEQ ID NO: 60) y en la tercera dicha proteína de fusión junto con YFP-CRM1 mutada (mutación A541K).(SEQ ID NO: 59) Se utilizaron aproximadamente 500 nanogramos de cada plásmido en la transfección.
- 10
- 15           2. Aproximadamente 24 horas más tarde, las células se fijaron con formaldehído 3,7% en PBS (30 min). Las células fijadas se permeabilizaron con 0.2% Triton X-100 (10 min) y se bloquearon con 3% BSA en PBS (1 h). Para detectar dicha proteína de fusión se utilizó un anticuerpo monoclonal anti Flag M2 (Sigma) diluido 1:300 en la solución de bloqueo (3% BSA en PBS). Se aplicó a continuación un anticuerpo secundario marcado con AlexaFluor594 (Invitrogen) y el colorante nuclear Hoechst.
- 20
- 25           Utilizando un microscopio de fluorescencia o un microscopio confocal se determinó la localización de la proteína de fusión en células que no expresan CRM1 ectópica (las células del control negativo) y en células que sí expresan CRM1 ectópica.

El análisis de las imágenes de microscopía mostró que dicha proteína de fusión es nuclear en células 293T que expresan únicamente CRM1 endógena. Este resultado indica que los niveles de CRM1 endógena en estas células no son suficientes para contrarrestar la importación al núcleo mediada por las dos NLSs de dicha proteína de fusión.

30

Por el contrario, dicha proteína de fusión mostró localización citoplásmica en aquellas células que expresan YFP-CRM1. Este resultado indica que la proteína CRM1 ectópica transfectada es funcional como receptor de exportación.

35

Las imágenes de microscopía mostraron que dicha proteína de fusión es nuclear en las células que expresan YFP-CRM1 mutada (A541K). Estudios *in vitro* han demostrado que esta mutación reduce notablemente la afinidad de CRM1 por las secuencias NES. Este resultado indica que la proteína YFP-CRM1 (A541K) presenta una actividad reducida en un entorno celular como receptor de exportación.

**REIVINDICACIONES**

1. Proteína de fusión que comprende:

- 5 (i) un primer dominio polipeptídico que comprende una variante de  
survivina seleccionada entre:
- 10 a) una variante de survivina que carece de la totalidad o parte del  
dominio de homodimerización resultando en un dominio de  
homodimerización no funcional; y
- b) una variante de survivina que carece de la totalidad o parte del  
dominio de homodimerización resultando en un dominio de  
homodimerización no funcional y comprende una señal de  
exportación nuclear (NES) de survivina;
- 15 (ii) al menos, un segundo dominio polipeptídico que comprende una  
señal de localización nuclear (NLS);
- (iii) al menos, un tercer dominio polipeptídico reportero;
- 20 en donde si la variante de survivina carece de la totalidad o parte de su  
señal NES, entonces dicha proteína de fusión comprende además
- (iv) un cuarto dominio polipeptídico que comprende una secuencia  
25 NES de survivina humana,  
y, en donde dicha variante de survivina es de origen humano.

30 2. Proteína de fusión según la reivindicación 1, en donde dicha variante de  
survivina comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 2.

3. Proteína de fusión según la reivindicación 1, en donde dicha NES comprende  
la secuencia de aminoácidos mostrada en SEQ ID NO: 17.

35 4. Proteína de fusión según cualquiera de las reivindicaciones 1 a 3, en donde  
dicha NLS se selecciona del grupo formado por: secuencia NLS del antígeno T

grande del virus SV40, secuencia NLS de nucleoplasmina, secuencia NLS de c-myc, secuencia NLS del dominio IBB de la importina alfa, secuencia NLS de la proteína T de mioma, la secuencia NLS de la proteína p53 humana, secuencia NLS de la proteína c-abl IV murina, secuencia NLS de la proteína NS1 del virus de la gripe, la secuencia NLS del antígeno delta del virus de la hepatitis, la secuencia NLS de la proteína Mx1 murina, secuencia NLS de la poli(ADP-ribosa) polimerasa, la secuencia NLS de los receptores de hormonas esteroideas humanos y la secuencia NLS de la proteína p17 del reovirus aviar.

5

10

5. Proteína de fusión según la reivindicación 4, en donde dicha NLS comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 3.

15

6. Proteína de fusión según cualquiera de las reivindicaciones 1 a 5, en donde dicha proteína de fusión comprende dos segundos dominios polipeptídicos y en donde dichos dos segundos dominios polipeptídicos son idénticos, y en donde cada uno de dichos dos dominios polipeptídicos comprenden la secuencia mostrada en SEQ ID NO: 3.

20

7. Proteína de fusión según cualquiera de las reivindicaciones 1 a 6, en donde dicho al menos un tercer dominio polipeptídico comprende un polipéptido que se selecciona del grupo formado por: epítipo FLAG, epítipo HA, epítipo Myc, epítipos AU1 y AU5, epítipo Glu-Glu, epítipo KT3, epítipo IRS, epítipo Btag, epítipo de proteína quinasa-C, epítipo de VSV, polipéptido de hexahistidinas y péptido de unión a calmodulina.

25

8. Proteína de fusión según la reivindicación 7, en donde dicho tercer dominio polipeptídico comprende el epítipo flag.

30

9. Proteína de fusión según cualquiera de las reivindicaciones 1 a 8, en donde dicha proteína de fusión comprende tres terceros dominios polipeptídicos y en donde dichos tres terceros dominios polipeptídicos son idénticos, y en donde dichos tres terceros dominios polipeptídicos comprenden el epítipo flag.

35

10. Proteína de fusión según cualquiera de las reivindicaciones 1 a 9, en donde dicha proteína de fusión comprende la secuencia de aminoácidos mostrada en SEQ ID NO: 52.

5

11. Ácido nucleico que codifica para la proteína de fusión según cualquiera de las reivindicaciones 1 a 10.

12. Casete de expresión que comprende el ácido nucleico según la reivindicación 11 en donde dicho ácido nucleico se encuentra operativamente unido a un sistema de transcripción y/o traducción apropiado.

10

13. Plásmido que comprende el ácido nucleico según la reivindicación 11 o el casete de expresión según la reivindicación 12.

15

14. Célula hospedadora que comprende el ácido nucleico según la reivindicación 11, el casete de expresión según la reivindicación 12 o el plásmido según la reivindicación 13.

20

15. Uso de la proteína de fusión según cualquiera de las reivindicaciones 1 a 10 para determinar si la función de una variante de CRM1 está alterada en una muestra o para identificar un compuesto con capacidad de modular la función de CRM1 o de una variante funcionalmente equivalente de la misma en una muestra.

25

16. Método para determinar si la función de una variante de CRM1 está alterada en una muestra que comprende al menos una célula, en donde dicho método comprende:

30

(i) proporcionar una muestra que comprende al menos una célula en donde dicha célula comprende una proteína de fusión según cualquiera de las reivindicaciones 1 a 10;

(ii) introducir en dicha al menos una célula un casete de expresión o un plásmido que codifica para dicha variante de CRM1;

35

(iii) crecer las células bajo condiciones que permitan la correcta expresión de dicha variante de CRM1;

- (iv) detectar la señal del dominio reportero de dicha proteína de fusión;
- (v) determinar la localización subcelular de dicha proteína de fusión;
- (vi) comparar la localización subcelular determinada en la etapa (iv) con dicha localización subcelular obtenida para dicha proteína de fusión en una muestra de referencia;

5

en donde dicha variante de CRM1 tiene un grado de identidad de, al menos, el 60% con respecto a la secuencia de CRM1; y

en donde una alteración de la localización subcelular de dicha proteína de fusión es indicativa de una función alterada de CRM1.

10

17. Método para identificar un compuesto con capacidad de modular la función de CRM1 o de una variante funcionalmente equivalente de la misma en una muestra que comprende al menos una célula, comprendiendo dicho método:

15

(i) poner en contacto un compuesto candidato con una muestra que comprende al menos una célula en donde dicha célula comprende:

a) una proteína de fusión según cualquiera de las reivindicaciones 1 a 10; y

b) un casete de expresión o un plásmido que codifica para CRM1 o una variante funcionalmente equivalente de la misma;

20

(ii) detectar la señal del dominio reportero de dicha proteína de fusión;

(iii) detectar la localización subcelular de dicha proteína de fusión;

(iv) comparar la localización subcelular de dicha proteína de fusión obtenida en la etapa (iv) con la localización subcelular de dicha proteína de fusión obtenida en una muestra que comprende al menos una célula que no ha sido tratada con dicho compuesto;

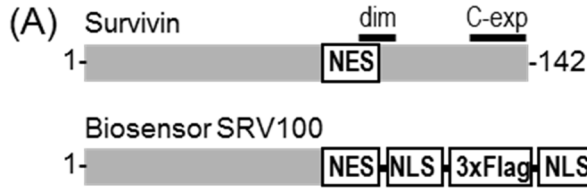
25

en donde dicha variante de CRM1 tiene un grado de identidad de, al menos, el 60% con respecto a la secuencia de CRM1; y

30

en donde una alteración de la localización subcelular de la proteína de fusión detectada en la etapa (iii) con respecto a dicha localización subcelular en dicha muestra no tratada es indicativa de que dicho compuesto es capaz de modular la función de dicha CRM1 o de dicha variante funcionalmente equivalente de la misma.





(B) MGAPTLPPAW QPFLKDHRIS TFKNWPFLG CACTPERMAE AGFIHCPTEN  
 EPDLAQCFFC FKELEGWEPD DDPIEEHKKH SSGCAFLSVK KQFEELTLGE  
 KLATMPKKKR KVEDPPVDYK DHDGDYKDHD IDYKDDDDKY RPMPKKRKV

(C)

ttg cct caA GAT CTt gcc acc atg ggt gcc ccg acg ttg ccc cct  
 gcc tgg cag ccc ttt ctc aag gac cac cgc atc tct aca ttc aag  
 aac tgg ccc ttc ttg gag ggc tgc gcc tgc acc ccg gag cgg atg  
 gcc gag gct ggc ttc atc cac tgc ccc act gag aac gag cca gac  
 ttg gcc cag tgt ttc ttc tgc ttc aag gag ctg gaa ggc tgg gag  
 cca gat gac gac ccc ata gag gaa cat aaa aag cat tcg tcc ggt  
tgc gct ttc ctt tct gtc aag aag cag ttt gaa gaa tta acc ctt  
 ggt gaa AAG CTT gcc acc atg cca aag aag aag agg aag gtg gag  
gat cca ccg gtc gac tac aaa gac cat gac ggt gat tat aaa gat  
cat gat atc gat tac aag gat gac gat gac aag tat cgg ccg atg  
 cca aag aag aag agg aag gtg **tag** CGG CCG CGa

# ES 2 549 488 A1

## LISTA DE SECUENCIAS

<110> UNIVERSIDAD DEL PAÍS VASCO  
<120> BIOSENSOR PARA DETERMINAR LA FUNCIÓN EXPORTADORA DE CRM1  
<130> P10178ES00  
<160> 61  
<170> PatentIn version 3.5  
<210> 1  
<211> 1071  
<212> PRT  
<213> Artificial Sequence  
  
<220>  
<223> CRM1A541K  
  
<400> 1  
Met Pro Ala Ile Met Thr Met Leu Ala Asp His Ala Ala Arg Gln Leu  
1                   5                   10                   15  
  
Leu Asp Phe Ser Gln Lys Leu Asp Ile Asn Leu Leu Asp Asn Val Val  
          20                   25                   30  
  
Asn Cys Leu Tyr His Gly Glu Gly Ala Gln Gln Arg Met Ala Gln Glu  
          35                   40                   45  
  
Val Leu Thr His Leu Lys Glu His Pro Asp Ala Trp Thr Arg Val Asp  
  50                   55                   60  
  
Thr Ile Leu Glu Phe Ser Gln Asn Met Asn Thr Lys Tyr Tyr Gly Leu  
65                   70                   75                   80  
  
Gln Ile Leu Glu Asn Val Ile Lys Thr Arg Trp Lys Ile Leu Pro Arg  
          85                   90                   95  
  
Asn Gln Cys Glu Gly Ile Lys Lys Tyr Val Val Gly Leu Ile Ile Lys  
          100                   105                   110  
  
Thr Ser Ser Asp Pro Thr Cys Val Glu Lys Glu Lys Val Tyr Ile Gly  
          115                   120                   125  
  
Lys Leu Asn Met Ile Leu Val Gln Ile Leu Lys Gln Glu Trp Pro Lys  
  130                   135                   140  
  
His Trp Pro Thr Phe Ile Ser Asp Ile Val Gly Ala Ser Arg Thr Ser  
145                   150                   155                   160

# ES 2 549 488 A1

Glu Ser Leu Cys Gln Asn Asn Met Val Ile Leu Lys Leu Leu Ser Glu  
 165 170 175

Glu Val Phe Asp Phe Ser Ser Gly Gln Ile Thr Gln Val Lys Ser Lys  
 180 185 190

His Leu Lys Asp Ser Met Cys Asn Glu Phe Ser Gln Ile Phe Gln Leu  
 195 200 205

Cys Gln Phe Val Met Glu Asn Ser Gln Asn Ala Pro Leu Val His Ala  
 210 215 220

Thr Leu Glu Thr Leu Leu Arg Phe Leu Asn Trp Ile Pro Leu Gly Tyr  
 225 230 235 240

Ile Phe Glu Thr Lys Leu Ile Ser Thr Leu Ile Tyr Lys Phe Leu Asn  
 245 250 255

Val Pro Met Phe Arg Asn Val Ser Leu Lys Cys Leu Thr Glu Ile Ala  
 260 265 270

Gly Val Ser Val Ser Gln Tyr Glu Glu Gln Phe Val Thr Leu Phe Thr  
 275 285

Leu Thr Met Met Gln Leu Lys Gln Met Leu Pro Leu Asn Thr Asn Ile  
 290 295 300

Arg Leu Ala Tyr Ser Asn Gly Lys Asp Asp Glu Gln Asn Phe Ile Gln  
 305 310 315 320

Asn Leu Ser Leu Phe Leu Cys Thr Phe Leu Lys Glu His Asp Gln Leu  
 325 330 335

Ile Glu Lys Arg Leu Asn Leu Arg Glu Thr Leu Met Glu Ala Leu His  
 340 345 350

Tyr Met Leu Leu Val Ser Glu Val Glu Glu Thr Glu Ile Phe Lys Ile  
 355 360 365

Cys Leu Glu Tyr Trp Asn His Leu Ala Ala Glu Leu Tyr Arg Glu Ser  
 370 375 380

Pro Phe Ser Thr Ser Ala Ser Pro Leu Leu Ser Gly Ser Gln His Phe  
 385 390 395 400

Asp Val Pro Pro Arg Arg Gln Leu Tyr Leu Pro Met Leu Phe Lys Val  
 405 410 415

# ES 2 549 488 A1

Arg Leu Leu Met Val Ser Arg Met Ala Lys Pro Glu Glu Val Leu Val  
 420 425 430

Val Glu Asn Asp Gln Gly Glu Val Val Arg Glu Phe Met Lys Asp Thr  
 435 440 445

Asp Ser Ile Asn Leu Tyr Lys Asn Met Arg Glu Thr Leu Val Tyr Leu  
 450 455 460

Thr His Leu Asp Tyr Val Asp Thr Glu Arg Ile Met Thr Glu Lys Leu  
 465 470 475 480

His Asn Gln Val Asn Gly Thr Glu Trp Ser Trp Lys Asn Leu Asn Thr  
 485 490 495

Leu Cys Trp Ala Ile Gly Ser Ile Ser Gly Ala Met His Glu Glu Asp  
 500 505 510

Glu Lys Arg Phe Leu Val Thr Val Ile Lys Asp Leu Leu Gly Leu Cys  
 515 520 525

Glu Gln Lys Arg Gly Lys Asp Asn Lys Ala Ile Ile Lys Ser Asn Ile  
 530 535 540

Met Tyr Ile Val Gly Gln Tyr Pro Arg Phe Leu Arg Ala His Trp Lys  
 545 550 555 560

Phe Leu Lys Thr Val Val Asn Lys Leu Phe Glu Phe Met His Glu Thr  
 565 570 575

His Asp Gly Val Gln Asp Met Ala Cys Asp Thr Phe Ile Lys Ile Ala  
 580 585 590

Gln Lys Cys Arg Arg His Phe Val Gln Val Gln Val Gly Glu Val Met  
 595 600 605

Pro Phe Ile Asp Glu Ile Leu Asn Asn Ile Asn Thr Ile Ile Cys Asp  
 610 615 620

Leu Gln Pro Gln Gln Val His Thr Phe Tyr Glu Ala Val Gly Tyr Met  
 625 630 635 640

Ile Gly Ala Gln Thr Asp Gln Thr Val Gln Glu His Leu Ile Glu Lys  
 645 650 655

# ES 2 549 488 A1

Tyr Met Leu Leu Pro Asn Gln Val Trp Asp Ser Ile Ile Gln Gln Ala  
 660 665 670

Thr Lys Asn Val Asp Ile Leu Lys Asp Pro Glu Thr Val Lys Gln Leu  
 675 680 685

Gly Ser Ile Leu Lys Thr Asn Val Arg Ala Cys Lys Ala Val Gly His  
 690 695 700

Pro Phe Val Ile Gln Leu Gly Arg Ile Tyr Leu Asp Met Leu Asn Val  
 705 710 715 720

Tyr Lys Cys Leu Ser Glu Asn Ile Ser Ala Ala Ile Gln Ala Asn Gly  
 725 730 735

Glu Met Val Thr Lys Gln Pro Leu Ile Arg Ser Met Arg Thr Val Lys  
 740 745 750

Arg Glu Thr Leu Lys Leu Ile Ser Gly Trp Val Ser Arg Ser Asn Asp  
 755 760 765

Pro Gln Met Val Ala Glu Asn Phe Val Pro Pro Leu Leu Asp Ala Val  
 770 775 780

Leu Ile Asp Tyr Gln Arg Asn Val Pro Ala Ala Arg Glu Pro Glu Val  
 785 790 795 800

Leu Ser Thr Met Ala Ile Ile Val Asn Lys Leu Gly Gly His Ile Thr  
 805 810 815

Ala Glu Ile Pro Gln Ile Phe Asp Ala Val Phe Glu Cys Thr Leu Asn  
 820 825 830

Met Ile Asn Lys Asp Phe Glu Glu Tyr Pro Glu His Arg Thr Asn Phe  
 835 840 845

Phe Leu Leu Leu Gln Ala Val Asn Ser His Cys Phe Pro Ala Phe Leu  
 850 855 860

Ala Ile Pro Pro Thr Gln Phe Lys Leu Val Leu Asp Ser Ile Ile Trp  
 865 870 875 880

Ala Phe Lys His Thr Met Arg Asn Val Ala Asp Thr Gly Leu Gln Ile  
 885 890 895

Leu Phe Thr Leu Leu Gln Asn Val Ala Gln Glu Glu Ala Ala Ala Gln  
 900 905 910

# ES 2 549 488 A1

Ser Phe Tyr Gln Thr Tyr Phe Cys Asp Ile Leu Gln His Ile Phe Ser  
915 920 925

Val Val Thr Asp Thr Ser His Thr Ala Gly Leu Thr Met His Ala Ser  
930 935 940

Ile Leu Ala Tyr Met Phe Asn Leu Val Glu Glu Gly Lys Ile Ser Thr  
945 950 955 960

Ser Leu Asn Pro Gly Asn Pro Val Asn Asn Gln Ile Phe Leu Gln Glu  
965 970 975

Tyr Val Ala Asn Leu Leu Lys Ser Ala Phe Pro His Leu Gln Asp Ala  
980 985 990

Gln Val Lys Leu Phe Val Thr Gly Leu Phe Ser Leu Asn Gln Asp Ile  
995 1000 1005

Pro Ala Phe Lys Glu His Leu Arg Asp Phe Leu Val Gln Ile Lys  
1010 1015 1020

Glu Phe Ala Gly Glu Asp Thr Ser Asp Leu Phe Leu Glu Glu Arg  
1025 1030 1035

Glu Ile Ala Leu Arg Gln Ala Asp Glu Glu Lys His Lys Arg Gln  
1040 1045 1050

Met Ser Val Pro Gly Ile Phe Asn Pro His Glu Ile Pro Glu Glu  
1055 1060 1065

Met Cys Asp  
1070

<210> 2  
<211> 83  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> variante de Survivina (1-83)

<400> 2

Met Gly Ala Pro Thr Leu Pro Pro Ala Trp Gln Pro Phe Leu Lys Asp  
1 5 10 15

His Arg Ile Ser Thr Phe Lys Asn Trp Pro Phe Leu Glu Gly Cys Ala  
20 25 30

# ES 2 549 488 A1

Cys Thr Pro Glu Arg Met Ala Glu Ala Gly Phe Ile His Cys Pro Thr  
35 40 45

Glu Asn Glu Pro Asp Leu Ala Gln Cys Phe Phe Cys Phe Lys Glu Leu  
50 55 60

Glu Gly Trp Glu Pro Asp Asp Asp Pro Ile Glu Glu His Lys Lys His  
65 70 75 80

Ser Ser Gly

<210> 3  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NLS de antígeno T grande del virus SV40

<400> 3

Pro Lys Lys Lys Arg Lys Val  
1 5

<210> 4  
<211> 16  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NLS de la nucleoplasmina

<400> 4

Lys Arg Pro Ala Ala Thr Lys Lys Ala Gly Gln Ala Lys Lys Lys Lys  
1 5 10 15

<210> 5  
<211> 11  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NLS similar al homeodominio de la proteína c-myc

<400> 5

Arg Gln Arg Arg Asn Glu Leu Lys Arg Ser Phe  
1 5 10

<210> 6  
<211> 41

# ES 2 549 488 A1

<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NLS del dominio IBB de la importina alfa

<400> 6

Arg Met Arg Lys Phe Lys Asn Lys Gly Lys Asp Thr Ala Glu Leu Arg  
1 5 10 15

Arg Arg Arg Val Glu Val Ser Val Glu Leu Arg Lys Ala Lys Lys Asp  
20 25 30

Glu Gln Ile Leu Lys Arg Arg Asn Val  
35 40

<210> 7  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NLS de la proteína T de mioma

<400> 7

Val Ser Arg Lys Arg Pro Arg Pro  
1 5

<210> 8  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NLS de la proteína T de mioma

<400> 8

Pro Pro Lys Lys Ala Arg Glu Asp  
1 5

<210> 9  
<211> 12  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NLS de la proteína c-abl IV de ratón

<400> 9

Ser Ala Leu Ile Lys Lys Lys Lys Lys Met Ala Pro  
1 5 10



## ES 2 549 488 A1

<210> 10  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NSL del virus de la gripe

<400> 10

Asp Arg Leu Arg Arg  
1 5

<210> 11  
<211> 7  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NSL del virus de la gripe

<400> 11

Pro Lys Gln Lys Lys Arg Lys  
1 5

<210> 12  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NLS del antígeno delta del virus de la hepatitis

<400> 12

Arg Lys Leu Lys Lys Lys Ile Lys Lys Leu  
1 5 10

<210> 13  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NLS de la proteína Mx1 de ratón

<400> 13

Arg Glu Lys Lys Lys Phe Leu Lys Arg Arg  
1 5 10

<210> 14  
<211> 20  
<212> PRT  
<213> Artificial Sequence

<220>

# ES 2 549 488 A1

<223> Secuencia NLS de poli(ADP-ribosa) polimerasa

<400> 14

Lys Arg Lys Gly Asp Glu Val Asp Gly Val Asp Glu Val Ala Lys Lys  
1                   5                   10                   15

Lys Ser Lys Lys  
          20

<210> 15

<211> 17

<212> PRT

<213> Artificial Sequence

<220>

<223> Secuencia NLS de receptores de hormonas esteroideas humanos

<400> 15

Arg Lys Cys Leu Gln Ala Gly Met Asn Leu Glu Ala Arg Lys Thr Lys  
1                   5                   10                   15

Lys

<210> 16

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Secuencia NLS de la proteína p17 del reovirus aviar

<400> 16

Ile Ala Ala Lys Arg Gly Arg Gln Leu Asp  
1                   5                   10

<210> 17

<211> 15

<212> PRT

<213> Artificial Sequence

<220>

<223> Secuencia NES de survivina humana

<400> 17

Cys Ala Phe Leu Ser Val Lys Lys Gln Phe Glu Glu Leu Thr Leu  
1                   5                   10                   15

<210> 18

<211> 10

<212> PRT

<213> Artificial Sequence

## ES 2 549 488 A1

<220>  
<223> Secuencia secuencia NES del inhibidor de la proteína quinasa dependiente de AMPc humano

<400> 18

Leu Ala Leu Lys Leu Ala Gly Leu Asp Ile  
1                           5                           10

<210> 19  
<211> 14  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NES de p53 humana

<400> 19

Glu Met Phe Arg Glu Leu Asn Glu Ala Leu Glu Leu Lys Asp  
1                           5                           10

<210> 20  
<211> 15  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NES de ubiquitin-proteín ligasa Mdm2 humana

<400> 20

Ser Ile Ser Leu Ser Phe Asp Glu Ser Leu Ala Leu Cys Val Ile  
1                           5                           10                           15

<210> 21  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NES de Xsmad4a de Xenopus

<400> 21

Gly Ile Asp Leu Ser Gly Leu Thr Leu Gln  
1                           5                           10

<210> 22  
<211> 17  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NES de catenina p120 humana (isoforma 3AB)

<400> 22

# ES 2 549 488 A1

Glu Ser Leu Glu Glu Glu Leu Asp Val Leu Val Leu Asp Asp Glu Gly  
1 5 10 15

Gly

<210> 23  
<211> 24  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NES de MAP quinasa 2 murina

<400> 23

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

Ala Thr Met Arg Val Asp Tyr Glu  
20

<210> 24  
<211> 19  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NES de p150 murina

<400> 24

Glu Ala Ile Asn Lys Leu Glu Ser Asn Leu Arg Glu Leu Gln Ile Cys  
1 5 10 15

Pro Ala Thr

<210> 25  
<211> 13  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Secuencia NES de ciclina B1 humana

<400> 25

Asp Leu Cys Gln Ala Phe Ser Asp Val Ile Leu Ala Val  
1 5 10

<210> 26  
<211> 12  
<212> PRT

## ES 2 549 488 A1

<213> Artificial Sequence

<220>

<223> Secuencia NES de actina de Saccharomyces

<400> 26

Ser Leu Pro His Ala Ile Leu Arg Ile Asp Leu Ala  
1                               5                               10

<210> 27

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Secuencia NES de la proteína Rev del virus HIV-1

<400> 27

Leu Pro Pro Leu Glu Arg Leu Thr Leu  
1                               5

<210> 28

<211> 5

<212> PRT

<213> Artificial Sequence

<220>

<223> Secuencia NES de la proteína Rex del virus T-linfotrópico tipo  
1

<400> 28

Leu Ser Leu Asp Ser  
1                               5

<210> 29

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Secuencia NES de MAP quinasa quinasa específica dual 1 de  
Xenopus

<400> 29

Leu Gln Lys Lys Leu Glu Glu Leu Glu Leu Asp Glu Gln  
1                               5                               10

<210> 30

<211> 12

<212> PRT

<213> Artificial Sequence

<220>

## ES 2 549 488 A1

<223> Secuencia NES de la proteína sustrato 15 del receptor del factor de crecimiento epidérmico humano

<400> 30

Leu Glu Leu Ala Ile Ala Leu Ser Lys Ser Glu Ile  
1 5 10

<210> 31

<211> 9

<212> PRT

<213> Artificial Sequence

<220>

<223> Secuencia NES de BRCA1 humana

<400> 31

Leu Glu Cys Pro Ile Cys Leu Glu Leu  
1 5

<210> 32

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Secuencia NES de p65 humana

<400> 32

Leu Ser Glu Ala Leu Leu Gln Leu Gln Phe  
1 5 10

<210> 33

<211> 19

<212> PRT

<213> Artificial Sequence

<220>

<223> Secuencia NES del factor de crecimiento similar a AP-1 de Schizosaccharomyces

<400> 33

Glu Ser Phe Asp Ile Asp Asp Leu Cys Ser Lys Leu Lys Asn Lys Ala  
1 5 10 15

Lys Cys Ser

<210> 34

<211> 9

<212> PRT

<213> Artificial Sequence

# ES 2 549 488 A1

<220>  
<223> Secuencia secuencia NES del factor de transcripción STAT-1 humano

<400> 34

Leu Leu Leu Lys Lys Met Tyr Leu Met  
1 5

<210> 35  
<211> 18  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Péptido enlazador

<400> 35

Gly Gly Ser Ser Arg Ser Ser Ser Ser Gly Gly Gly Gly Ser Gly Gly  
1 5 10 15

Gly Gly

<210> 36  
<211> 15  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Péptido enlazador

<400> 36

Gly Ser Gly Arg Ser Gly Gly Gly Gly Ser Gly Gly Gly Gly Ser  
1 5 10 15

<210> 37  
<211> 13  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Péptido enlazador

<400> 37

Glu Gly Ser Ser Gly Ser Gly Ser Glu Ser Lys Ser Thr  
1 5 10

<210> 38  
<211> 15  
<212> PRT  
<213> Artificial Sequence

<220>

# ES 2 549 488 A1

<223> Péptido enlazador

<400> 38

Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Ser Thr Gln  
1 5 10 15

<210> 39

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Péptido enlazador

<400> 39

Glu Gly Lys Ser Ser Gly Ser Gly Ser Glu Ser Lys Val Asp  
1 5 10

<210> 40

<211> 14

<212> PRT

<213> Artificial Sequence

<220>

<223> Péptido enlazador

<400> 40

Gly Ser Thr Ser Gly Ser Gly Lys Ser Ser Glu Gly Lys Gly  
1 5 10

<210> 41

<211> 18

<212> PRT

<213> Artificial Sequence

<220>

<223> Péptido enlazador

<400> 41

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

Leu Asp

<210> 42

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Péptido enlazador



# ES 2 549 488 A1

<400> 42

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

<210> 43

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Péptido enlazador

<400> 43

Thr Ala Thr Pro Ala Thr Thr Pro Thr Thr Ala Pro Thr Ala Gly Thr  
1 5 10 15

<210> 44

<211> 32

<212> PRT

<213> Artificial Sequence

<220>

<223> Péptido enlazador

<400> 44

Thr Ala Thr Pro Ala Thr Thr Pro Thr Thr Ala Pro Thr Ala Gly Thr  
1 5 10 15

Thr Ala Thr Pro Ala Thr Thr Pro Thr Thr Ala Pro Thr Ala Gly Thr  
20 25 30

<210> 45

<211> 7

<212> PRT

<213> Artificial Sequence

<220>

<223> Péptido enlazador

<400> 45

Gly Thr Lys Val His Met Lys  
1 5

<210> 46

<211> 13

<212> PRT

<213> Artificial Sequence

<220>

<223> Péptido enlazador

<400> 46

# ES 2 549 488 A1

Pro Gly Thr Ser Gly Gln Gln Pro Ser Val Gly Gln Gln  
1 5 10

<210> 47  
<211> 5  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Péptido enlazador

<400> 47

Gly Thr Ser Gly Gln  
1 5

<210> 48  
<211> 10  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Péptido enlazador

<400> 48

Pro Lys Pro Ser Thr Pro Pro Gly Ser Ser  
1 5 10

<210> 49  
<211> 11  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Péptido enlazador

<400> 49

Ala Pro Ala Glu Thr Lys Ala Glu Pro Met Thr  
1 5 10

<210> 50  
<211> 9  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Péptido enlazador

<400> 50

Ser Gly Gly Ser Gly Ser Gly Gly Gln  
1 5

<210> 51  
<211> 8

# ES 2 549 488 A1

<212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Péptido enlazador

<400> 51

Gly Gly Ser Ser Arg Ser Ser Ser  
 1 5

<210> 52  
 <211> 150  
 <212> PRT  
 <213> Artificial Sequence

<220>  
 <223> Proteína de fusión

<400> 52

Met Gly Ala Pro Thr Leu Pro Pro Ala Trp Gln Pro Phe Leu Lys Asp  
 1 5 10 15

His Arg Ile Ser Thr Phe Lys Asn Trp Pro Phe Leu Glu Gly Cys Ala  
 20 25 30

Cys Thr Pro Glu Arg Met Ala Glu Ala Gly Phe Ile His Cys Pro Thr  
 35 40 45

Glu Asn Glu Pro Asp Leu Ala Gln Cys Phe Phe Cys Phe Lys Glu Leu  
 50 55 60

Glu Gly Trp Glu Pro Asp Asp Asp Pro Ile Glu Glu His Lys Lys His  
 65 70 75 80

Ser Ser Gly Cys Ala Phe Leu Ser Val Lys Lys Gln Phe Glu Glu Leu  
 85 90 95

Thr Leu Gly Glu Lys Leu Ala Thr Met Pro Lys Lys Lys Arg Lys Val  
 100 105 110

Glu Asp Pro Pro Val Asp Tyr Lys Asp His Asp Gly Asp Tyr Lys Asp  
 115 120 125

His Asp Ile Asp Tyr Lys Asp Asp Asp Asp Lys Tyr Arg Pro Met Pro  
 130 135 140

Lys Lys Lys Arg Lys Val  
 145 150

## ES 2 549 488 A1

<210> 53  
<211> 8  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Dominio de transducción de proteína

<400> 53

Arg Lys Lys Arg Arg Gln Arg Arg  
1 5

<210> 54  
<211> 11  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Dominio de transducción de proteína

<400> 54

Tyr Ala Arg Lys Ala Arg Arg Gln Ala Arg Arg  
1 5 10

<210> 55  
<211> 11  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Dominio de transducción de proteína

<400> 55

Tyr Ala Arg Ala Ala Arg Arg Ala Ala Arg Arg  
1 5 10

<210> 56  
<211> 11  
<212> PRT  
<213> Artificial Sequence

<220>  
<223> Dominio de transducción de proteína

<400> 56

Tyr Ala Arg Ala Ala Arg Arg Ala Ala Arg Ala  
1 5 10

<210> 57  
<211> 11  
<212> PRT  
<213> Artificial Sequence

<220>

# ES 2 549 488 A1

<223> Dominio de transducción de proteína

<400> 57

Tyr Ala Arg Ala Ala Ala Arg Gln Ala Arg Ala  
1 5 10

<210> 58

<211> 16

<212> PRT

<213> Artificial Sequence

<220>

<223> Dominio de transducción de proteína

<400> 58

Arg Gln Ile Lys Ile Trp Phe Gln Asn Arg Arg Met Lys Trp Lys Lys  
1 5 10 15

<210> 59

<211> 1322

<212> PRT

<213> Artificial Sequence

<220>

<223> Proteína de fusión

<400> 59

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

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly  
20 25 30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile  
35 40 45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr  
50 55 60

Phe Gly Tyr Gly Leu Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys  
65 70 75 80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu  
85 90 95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu  
100 105 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly  
115 120 125

# ES 2 549 488 A1

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr  
 130 135 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn  
 145 150 155 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser  
 165 170 175

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

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu  
 195 200 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe  
 210 215 220

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser  
 225 230 235 240

Gly Leu Arg Ser Arg Glu Leu Val Pro Gln His Met Pro Ala Ile Met  
 245 250 255

Thr Met Leu Ala Asp His Ala Ala Arg Gln Leu Leu Asp Phe Ser Gln  
 260 265 270

Lys Leu Asp Ile Asn Leu Leu Asp Asn Val Val Asn Cys Leu Tyr His  
 275 280 285

Gly Glu Gly Ala Gln Gln Arg Met Ala Gln Glu Val Leu Thr His Leu  
 290 295 300

Lys Glu His Pro Asp Ala Trp Thr Arg Val Asp Thr Ile Leu Glu Phe  
 305 310 315 320

Ser Gln Asn Met Asn Thr Lys Tyr Tyr Gly Leu Gln Ile Leu Glu Asn  
 325 330 335

Val Ile Lys Thr Arg Trp Lys Ile Leu Pro Arg Asn Gln Cys Glu Gly  
 340 345 350

Ile Lys Lys Tyr Val Val Gly Leu Ile Ile Lys Thr Ser Ser Asp Pro  
 355 360 365

# ES 2 549 488 A1

Thr Cys Val Glu Lys Glu Lys Val Tyr Ile Gly Lys Leu Asn Met Ile  
 370 375 380

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

Ile Ser Asp Ile Val Gly Ala Ser Arg Thr Ser Glu Ser Leu Cys Gln  
 405 410 415

Asn Asn Met Val Ile Leu Lys Leu Leu Ser Glu Glu Val Phe Asp Phe  
 420 425 430

Ser Ser Gly Gln Ile Thr Gln Val Lys Ser Lys His Leu Lys Asp Ser  
 435 440 445

Met Cys Asn Glu Phe Ser Gln Ile Phe Gln Leu Cys Gln Phe Val Met  
 450 455 460

Glu Asn Ser Gln Asn Ala Pro Leu Val His Ala Thr Leu Glu Thr Leu  
 465 470 475 480

Leu Arg Phe Leu Asn Trp Ile Pro Leu Gly Tyr Ile Phe Glu Thr Lys  
 485 490 495

Leu Ile Ser Thr Leu Ile Tyr Lys Phe Leu Asn Val Pro Met Phe Arg  
 500 505 510

Asn Val Ser Leu Lys Cys Leu Thr Glu Ile Ala Gly Val Ser Val Ser  
 515 520 525

Gln Tyr Glu Glu Gln Phe Val Thr Leu Phe Thr Leu Thr Met Met Gln  
 530 535 540

Leu Lys Gln Met Leu Pro Leu Asn Thr Asn Ile Arg Leu Ala Tyr Ser  
 545 550 555 560

Asn Gly Lys Asp Asp Glu Gln Asn Phe Ile Gln Asn Leu Ser Leu Phe  
 565 570 575

Leu Cys Thr Phe Leu Lys Glu His Asp Gln Leu Ile Glu Lys Arg Leu  
 580 585 590

Asn Leu Arg Glu Thr Leu Met Glu Ala Leu His Tyr Met Leu Leu Val  
 595 600 605

Ser Glu Val Glu Glu Thr Glu Ile Phe Lys Ile Cys Leu Glu Tyr Trp  
 610 615 620

# ES 2 549 488 A1

Asn His Leu Ala Ala Glu Leu Tyr Arg Glu Ser Pro Phe Ser Thr Ser  
 625 630 635 640

Ala Ser Pro Leu Leu Ser Gly Ser Gln His Phe Asp Val Pro Pro Arg  
 645 650 655

Arg Gln Leu Tyr Leu Pro Met Leu Phe Lys Val Arg Leu Leu Met Val  
 660 665 670

Ser Arg Met Ala Lys Pro Glu Glu Val Leu Val Val Glu Asn Asp Gln  
 675 680 685

Gly Glu Val Val Arg Glu Phe Met Lys Asp Thr Asp Ser Ile Asn Leu  
 690 695 700

Tyr Lys Asn Met Arg Glu Thr Leu Val Tyr Leu Thr His Leu Asp Tyr  
 705 710 715 720

Val Asp Thr Glu Arg Ile Met Thr Glu Lys Leu His Asn Gln Val Asn  
 725 730 735

Gly Thr Glu Trp Ser Trp Lys Asn Leu Asn Thr Leu Cys Trp Ala Ile  
 740 745 750

Gly Ser Ile Ser Gly Ala Met His Glu Glu Asp Glu Lys Arg Phe Leu  
 755 760 765

Val Thr Val Ile Lys Asp Leu Leu Gly Leu Cys Glu Gln Lys Arg Gly  
 770 775 780

Lys Asp Asn Lys Ala Ile Ile Lys Ser Asn Ile Met Tyr Ile Val Gly  
 785 790 795 800

Gln Tyr Pro Arg Phe Leu Arg Ala His Trp Lys Phe Leu Lys Thr Val  
 805 810 815

Val Asn Lys Leu Phe Glu Phe Met His Glu Thr His Asp Gly Val Gln  
 820 825 830

Asp Met Ala Cys Asp Thr Phe Ile Lys Ile Ala Gln Lys Cys Arg Arg  
 835 840 845

His Phe Val Gln Val Gln Val Gly Glu Val Met Pro Phe Ile Asp Glu  
 850 855 860



# ES 2 549 488 A1

Ile Leu Asn Asn Ile Asn Thr Ile Ile Cys Asp Leu Gln Pro Gln Gln  
 865 870 875 880

Val His Thr Phe Tyr Glu Ala Val Gly Tyr Met Ile Gly Ala Gln Thr  
 885 890 895

Asp Gln Thr Val Gln Glu His Leu Ile Glu Lys Tyr Met Leu Leu Pro  
 900 905 910

Asn Gln Val Trp Asp Ser Ile Ile Gln Gln Ala Thr Lys Asn Val Asp  
 915 920 925

Ile Leu Lys Asp Pro Glu Thr Val Lys Gln Leu Gly Ser Ile Leu Lys  
 930 935 940

Thr Asn Val Arg Ala Cys Lys Ala Val Gly His Pro Phe Val Ile Gln  
 945 950 955 960

Leu Gly Arg Ile Tyr Leu Asp Met Leu Asn Val Tyr Lys Cys Leu Ser  
 965 970 975

Glu Asn Ile Ser Ala Ala Ile Gln Ala Asn Gly Glu Met Val Thr Lys  
 980 985 990

Gln Pro Leu Ile Arg Ser Met Arg Thr Val Lys Arg Glu Thr Leu Lys  
 995 1000 1005

Leu Ile Ser Gly Trp Val Ser Arg Ser Asn Asp Pro Gln Met Val  
 1010 1015 1020

Ala Glu Asn Phe Val Pro Pro Leu Leu Asp Ala Val Leu Ile Asp  
 1025 1030 1035

Tyr Gln Arg Asn Val Pro Ala Ala Arg Glu Pro Glu Val Leu Ser  
 1040 1045 1050

Thr Met Ala Ile Ile Val Asn Lys Leu Gly Gly His Ile Thr Ala  
 1055 1060 1065

Glu Ile Pro Gln Ile Phe Asp Ala Val Phe Glu Cys Thr Leu Asn  
 1070 1075 1080

Met Ile Asn Lys Asp Phe Glu Glu Tyr Pro Glu His Arg Thr Asn  
 1085 1090 1095

Phe Phe Leu Leu Leu Gln Ala Val Asn Ser His Cys Phe Pro Ala  
 1100 1105 1110

# ES 2 549 488 A1

Phe Leu Ala Ile Pro Pro Thr Gln Phe Lys Leu Val Leu Asp Ser  
 1115 1120 1125  
  
 Ile Ile Trp Ala Phe Lys His Thr Met Arg Asn Val Ala Asp Thr  
 1130 1135 1140  
  
 Gly Leu Gln Ile Leu Phe Thr Leu Leu Gln Asn Val Ala Gln Glu  
 1145 1150 1155  
  
 Glu Ala Ala Ala Gln Ser Phe Tyr Gln Thr Tyr Phe Cys Asp Ile  
 1160 1165 1170  
  
 Leu Gln His Ile Phe Ser Val Val Thr Asp Thr Ser His Thr Ala  
 1175 1180 1185  
  
 Gly Leu Thr Met His Ala Ser Ile Leu Ala Tyr Met Phe Asn Leu  
 1190 1195 1200  
  
 Val Glu Glu Gly Lys Ile Ser Thr Ser Leu Asn Pro Gly Asn Pro  
 1205 1210 1215  
  
 Val Asn Asn Gln Ile Phe Leu Gln Glu Tyr Val Ala Asn Leu Leu  
 1220 1225 1230  
  
 Lys Ser Ala Phe Pro His Leu Gln Asp Ala Gln Val Lys Leu Phe  
 1235 1240 1245  
  
 Val Thr Gly Leu Phe Ser Leu Asn Gln Asp Ile Pro Ala Phe Lys  
 1250 1255 1260  
  
 Glu His Leu Arg Asp Phe Leu Val Gln Ile Lys Glu Phe Ala Gly  
 1265 1270 1275  
  
 Glu Asp Thr Ser Asp Leu Phe Leu Glu Glu Arg Glu Ile Ala Leu  
 1280 1285 1290  
  
 Arg Gln Ala Asp Glu Glu Lys His Lys Arg Gln Met Ser Val Pro  
 1295 1300 1305  
  
 Gly Ile Phe Asn Pro His Glu Ile Pro Glu Glu Met Cys Asp  
 1310 1315 1320

<210> 60  
 <211> 1322  
 <212> PRT  
 <213> Artificial Sequence

# ES 2 549 488 A1

<220>

<223> Proteína de fusión

<400> 60

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

Val Glu Leu Asp Gly Asp Val Asn Gly His Lys Phe Ser Val Ser Gly  
20 25 30

Glu Gly Glu Gly Asp Ala Thr Tyr Gly Lys Leu Thr Leu Lys Phe Ile  
35 40 45

Cys Thr Thr Gly Lys Leu Pro Val Pro Trp Pro Thr Leu Val Thr Thr  
50 55 60

Phe Gly Tyr Gly Leu Gln Cys Phe Ala Arg Tyr Pro Asp His Met Lys  
65 70 75 80

Gln His Asp Phe Phe Lys Ser Ala Met Pro Glu Gly Tyr Val Gln Glu  
85 90 95

Arg Thr Ile Phe Phe Lys Asp Asp Gly Asn Tyr Lys Thr Arg Ala Glu  
100 105 110

Val Lys Phe Glu Gly Asp Thr Leu Val Asn Arg Ile Glu Leu Lys Gly  
115 120 125

Ile Asp Phe Lys Glu Asp Gly Asn Ile Leu Gly His Lys Leu Glu Tyr  
130 135 140

Asn Tyr Asn Ser His Asn Val Tyr Ile Met Ala Asp Lys Gln Lys Asn  
145 150 155 160

Gly Ile Lys Val Asn Phe Lys Ile Arg His Asn Ile Glu Asp Gly Ser  
165 170 175

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

Pro Val Leu Leu Pro Asp Asn His Tyr Leu Ser Tyr Gln Ser Ala Leu  
195 200 205

Ser Lys Asp Pro Asn Glu Lys Arg Asp His Met Val Leu Leu Glu Phe  
210 215 220

# ES 2 549 488 A1

Val Thr Ala Ala Gly Ile Thr Leu Gly Met Asp Glu Leu Tyr Lys Ser  
 225 230 235 240

Gly Leu Arg Ser Arg Glu Leu Val Pro Gln His Met Pro Ala Ile Met  
 245 250 255

Thr Met Leu Ala Asp His Ala Ala Arg Gln Leu Leu Asp Phe Ser Gln  
 260 265 270

Lys Leu Asp Ile Asn Leu Leu Asp Asn Val Val Asn Cys Leu Tyr His  
 275 280 285

Gly Glu Gly Ala Gln Gln Arg Met Ala Gln Glu Val Leu Thr His Leu  
 290 295 300

Lys Glu His Pro Asp Ala Trp Thr Arg Val Asp Thr Ile Leu Glu Phe  
 305 310 315 320

Ser Gln Asn Met Asn Thr Lys Tyr Tyr Gly Leu Gln Ile Leu Glu Asn  
 325 330 335

Val Ile Lys Thr Arg Trp Lys Ile Leu Pro Arg Asn Gln Cys Glu Gly  
 340 345 350

Ile Lys Lys Tyr Val Val Gly Leu Ile Ile Lys Thr Ser Ser Asp Pro  
 355 360 365

Thr Cys Val Glu Lys Glu Lys Val Tyr Ile Gly Lys Leu Asn Met Ile  
 370 375 380

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

Ile Ser Asp Ile Val Gly Ala Ser Arg Thr Ser Glu Ser Leu Cys Gln  
 405 410 415

Asn Asn Met Val Ile Leu Lys Leu Leu Ser Glu Glu Val Phe Asp Phe  
 420 425 430

Ser Ser Gly Gln Ile Thr Gln Val Lys Ser Lys His Leu Lys Asp Ser  
 435 440 445

Met Cys Asn Glu Phe Ser Gln Ile Phe Gln Leu Cys Gln Phe Val Met  
 450 455 460

Glu Asn Ser Gln Asn Ala Pro Leu Val His Ala Thr Leu Glu Thr Leu  
 465 470 475 480

# ES 2 549 488 A1

Leu Arg Phe Leu Asn Trp Ile Pro Leu Gly Tyr Ile Phe Glu Thr Lys  
485 490 495

Leu Ile Ser Thr Leu Ile Tyr Lys Phe Leu Asn Val Pro Met Phe Arg  
500 505 510

Asn Val Ser Leu Lys Cys Leu Thr Glu Ile Ala Gly Val Ser Val Ser  
515 520 525

Gln Tyr Glu Glu Gln Phe Val Thr Leu Phe Thr Leu Thr Met Met Gln  
530 535 540

Leu Lys Gln Met Leu Pro Leu Asn Thr Asn Ile Arg Leu Ala Tyr Ser  
545 550 555 560

Asn Gly Lys Asp Asp Glu Gln Asn Phe Ile Gln Asn Leu Ser Leu Phe  
565 570 575

Leu Cys Thr Phe Leu Lys Glu His Asp Gln Leu Ile Glu Lys Arg Leu  
580 585 590

Asn Leu Arg Glu Thr Leu Met Glu Ala Leu His Tyr Met Leu Leu Val  
595 600 605

Ser Glu Val Glu Glu Thr Glu Ile Phe Lys Ile Cys Leu Glu Tyr Trp  
610 615 620

Asn His Leu Ala Ala Glu Leu Tyr Arg Glu Ser Pro Phe Ser Thr Ser  
625 630 635 640

Ala Ser Pro Leu Leu Ser Gly Ser Gln His Phe Asp Val Pro Pro Arg  
645 650 655

Arg Gln Leu Tyr Leu Pro Met Leu Phe Lys Val Arg Leu Leu Met Val  
660 665 670

Ser Arg Met Ala Lys Pro Glu Glu Val Leu Val Val Glu Asn Asp Gln  
675 680 685

Gly Glu Val Val Arg Glu Phe Met Lys Asp Thr Asp Ser Ile Asn Leu  
690 695 700

Tyr Lys Asn Met Arg Glu Thr Leu Val Tyr Leu Thr His Leu Asp Tyr  
705 710 715 720

# ES 2 549 488 A1

Val Asp Thr Glu Arg Ile Met Thr Glu Lys Leu His Asn Gln Val Asn  
 725 730 735

Gly Thr Glu Trp Ser Trp Lys Asn Leu Asn Thr Leu Cys Trp Ala Ile  
 740 745 750

Gly Ser Ile Ser Gly Ala Met His Glu Glu Asp Glu Lys Arg Phe Leu  
 755 760 765

Val Thr Val Ile Lys Asp Leu Leu Gly Leu Cys Glu Gln Lys Arg Gly  
 770 775 780

Lys Asp Asn Lys Ala Ile Ile Ala Ser Asn Ile Met Tyr Ile Val Gly  
 785 790 795 800

Gln Tyr Pro Arg Phe Leu Arg Ala His Trp Lys Phe Leu Lys Thr Val  
 805 810 815

Val Asn Lys Leu Phe Glu Phe Met His Glu Thr His Asp Gly Val Gln  
 820 825 830

Asp Met Ala Cys Asp Thr Phe Ile Lys Ile Ala Gln Lys Cys Arg Arg  
 835 840 845

His Phe Val Gln Val Gln Val Gly Glu Val Met Pro Phe Ile Asp Glu  
 850 855 860

Ile Leu Asn Asn Ile Asn Thr Ile Ile Cys Asp Leu Gln Pro Gln Gln  
 865 870 875 880

Val His Thr Phe Tyr Glu Ala Val Gly Tyr Met Ile Gly Ala Gln Thr  
 885 890 895

Asp Gln Thr Val Gln Glu His Leu Ile Glu Lys Tyr Met Leu Leu Pro  
 900 905 910

Asn Gln Val Trp Asp Ser Ile Ile Gln Gln Ala Thr Lys Asn Val Asp  
 915 920 925

Ile Leu Lys Asp Pro Glu Thr Val Lys Gln Leu Gly Ser Ile Leu Lys  
 930 935 940

Thr Asn Val Arg Ala Cys Lys Ala Val Gly His Pro Phe Val Ile Gln  
 945 950 955 960

Leu Gly Arg Ile Tyr Leu Asp Met Leu Asn Val Tyr Lys Cys Leu Ser  
 965 970 975

# ES 2 549 488 A1

Glu Asn Ile Ser Ala Ala Ile Gln Ala Asn Gly Glu Met Val Thr Lys  
 980 985 990

Gln Pro Leu Ile Arg Ser Met Arg Thr Val Lys Arg Glu Thr Leu Lys  
 995 1000 1005

Leu Ile Ser Gly Trp Val Ser Arg Ser Asn Asp Pro Gln Met Val  
 1010 1015 1020

Ala Glu Asn Phe Val Pro Pro Leu Leu Asp Ala Val Leu Ile Asp  
 1025 1030 1035

Tyr Gln Arg Asn Val Pro Ala Ala Arg Glu Pro Glu Val Leu Ser  
 1040 1045 1050

Thr Met Ala Ile Ile Val Asn Lys Leu Gly Gly His Ile Thr Ala  
 1055 1060 1065

Glu Ile Pro Gln Ile Phe Asp Ala Val Phe Glu Cys Thr Leu Asn  
 1070 1075 1080

Met Ile Asn Lys Asp Phe Glu Glu Tyr Pro Glu His Arg Thr Asn  
 1085 1090 1095

Phe Phe Leu Leu Leu Gln Ala Val Asn Ser His Cys Phe Pro Ala  
 1100 1105 1110

Phe Leu Ala Ile Pro Pro Thr Gln Phe Lys Leu Val Leu Asp Ser  
 1115 1120 1125

Ile Ile Trp Ala Phe Lys His Thr Met Arg Asn Val Ala Asp Thr  
 1130 1135 1140

Gly Leu Gln Ile Leu Phe Thr Leu Leu Gln Asn Val Ala Gln Glu  
 1145 1150 1155

Glu Ala Ala Ala Gln Ser Phe Tyr Gln Thr Tyr Phe Cys Asp Ile  
 1160 1165 1170

Leu Gln His Ile Phe Ser Val Val Thr Asp Thr Ser His Thr Ala  
 1175 1180 1185

Gly Leu Thr Met His Ala Ser Ile Leu Ala Tyr Met Phe Asn Leu  
 1190 1195 1200

## ES 2 549 488 A1

Val Glu Glu Gly Lys Ile Ser Thr Ser Leu Asn Pro Gly Asn Pro  
 1205 1210 1215

Val Asn Asn Gln Ile Phe Leu Gln Glu Tyr Val Ala Asn Leu Leu  
 1220 1225 1230

Lys Ser Ala Phe Pro His Leu Gln Asp Ala Gln Val Lys Leu Phe  
 1235 1240 1245

Val Thr Gly Leu Phe Ser Leu Asn Gln Asp Ile Pro Ala Phe Lys  
 1250 1255 1260

Glu His Leu Arg Asp Phe Leu Val Gln Ile Lys Glu Phe Ala Gly  
 1265 1270 1275

Glu Asp Thr Ser Asp Leu Phe Leu Glu Glu Arg Glu Ile Ala Leu  
 1280 1285 1290

Arg Gln Ala Asp Glu Glu Lys His Lys Arg Gln Met Ser Val Pro  
 1295 1300 1305

Gly Ile Phe Asn Pro His Glu Ile Pro Glu Glu Met Cys Asp  
 1310 1315 1320

<210> 61

<211> 483

<212> DNA

<213> Artificial Sequence

<220>

<223> Secuencia de nucleótidos que codifica la proteína de fusión con  
 SEQ ID NO: 52

<400> 61

```

ttgcctcaag atcttgccac catgggtgcc cgcacgttgc ccctgcctg gcagcccttt      60
ctcaaggacc accgcatctc tacattcaag aactggcctt tcttgagggg ctgcgcctgc      120
accccgagac ggatggccga ggctggcttc atccactgcc cactgagaa cgagccagac      180
ttggcccagt gtttcttctg cttcaaggag ctggaaggct gggagccaga tgacgacccc      240
atagaggaac ataaaaagca ttcgtccggt tgcgctttcc tttctgtcaa gaagcagttt      300
gaagaattaa cccttggtga aaagcttgcc accatgcaa agaagaagag gaaggtggag      360
gatccaccgg tcgactacaa agacatgac ggtgattata aagatcatga tatcgattac      420
aaggatgacg atgacaagta tcggccgatg ccaaagaaga agaggaaggt gtagcggccg      480
cga                                                                                   483
  
```





②① N.º solicitud: 201431066

②② Fecha de presentación de la solicitud: 14.07.2014

③② Fecha de prioridad:

INFORME SOBRE EL ESTADO DE LA TECNICA

⑤① Int. Cl.: Ver Hoja Adicional

DOCUMENTOS RELEVANTES

Categoría	⑤⑥ Documentos citados	Reivindicaciones afectadas
X	Rodríguez J. A. et al. Subcellular localization and nucleocytoplasmic transport of the chromosomal passenger proteins before nuclear envelope breakdown. Oncogene. 2006, Vol. 25, páginas 4867–4879 (todo el documento)	1-17
X	Rodríguez J. A. et al. CRM1-mediated nuclear export determines the cytoplasmic localization of the antiapoptotic protein Survivin. Experimental Cell Research. 2002, Vol. 275, páginas 44–53 (todo el documento)	1-17
X	Engelsma D. et al. Homodimerization antagonizes nuclear export of Survivin. Traffic. 2007, Vol. 8, páginas 1495–1502 (todo el documento)	1-17
A	Stauber R. H. et al. Nucleocytoplasmic shuttling and the biological activity of mouse Survivin are regulated by an active nuclear export signal. Traffic. 2006, Vol. 7, páginas 1461–1472 (todo el documento)	1-17
A	Rodríguez J. A. et al. Cytoplasmic localization of Survivin is mediated by CRM1-dependent nuclear export. Proceedings of the American Association for Cancer Research. 03.2002, Vol. 43, página 530 (todo el documento)	1-17

Categoría de los documentos citados

X: de particular relevancia

Y: de particular relevancia combinado con otro/s de la misma categoría

A: refleja el estado de la técnica

O: referido a divulgación no escrita

P: publicado entre la fecha de prioridad y la de presentación de la solicitud

E: documento anterior, pero publicado después de la fecha de presentación de la solicitud

**El presente informe ha sido realizado**

para todas las reivindicaciones

para las reivindicaciones nº:

Fecha de realización del informe  
20.10.2015

Examinador  
M. Cumbreño Galindo

Página  
1/5

## CLASIFICACIÓN OBJETO DE LA SOLICITUD

**C07K19/00** (2006.01)

**C12N15/62** (2006.01)

**G01N33/533** (2006.01)

Documentación mínima buscada (sistema de clasificación seguido de los símbolos de clasificación)

C07K, C12N, G01N

Bases de datos electrónicas consultadas durante la búsqueda (nombre de la base de datos y, si es posible, términos de búsqueda utilizados)

INVENES, EPODOC, WPI, MEDLINE, NPL, EMBASE, BIOSIS

Fecha de Realización de la Opinión Escrita: 20.10.2015

**Declaración**

<b>Novedad (Art. 6.1 LP 11/1986)</b>	Reivindicaciones 2, 3, 6, 9, 10	<b>SI</b>
	Reivindicaciones 1, 4, 5, 7, 8, 11-17	<b>NO</b>
<b>Actividad inventiva (Art. 8.1 LP11/1986)</b>	Reivindicaciones	<b>SI</b>
	Reivindicaciones 1-17	<b>NO</b>

Se considera que la solicitud cumple con el requisito de aplicación industrial. Este requisito fue evaluado durante la fase de examen formal y técnico de la solicitud (Artículo 31.2 Ley 11/1986).

**Base de la Opinión.-**

La presente opinión se ha realizado sobre la base de la solicitud de patente tal y como se publica.

**1. Documentos considerados.-**

A continuación se relacionan los documentos pertenecientes al estado de la técnica tomados en consideración para la realización de esta opinión.

Documento	Número Publicación o Identificación	Fecha Publicación
D01	Rodríguez J. A. et al. Oncogene. Vol. 25, páginas 4867–4879	2006
D02	Rodríguez J. A. et al. Experimental Cell Research. Vol. 275, páginas 44–53	2002
D03	Engelsma D. et al. Traffic. Vol. 8, páginas 1495–1502	2007
D04	Stauber R. H. et al. Traffic. Vol. 7, páginas 1461–1472	2006
D05	Rodríguez J. A. et al. Proceedings of the American Association for Cancer Research. Vol. 43, página 530	03.2002

**2. Declaración motivada según los artículos 29.6 y 29.7 del Reglamento de ejecución de la Ley 11/1986, de 20 de marzo, de Patentes sobre la novedad y la actividad inventiva; citas y explicaciones en apoyo de esta declaración**

La presente invención tienen por objeto: una proteína de fusión (reivindicaciones 1 a 10), el ácido nucleico que la codifica (reivindicación 11), el casete de expresión que comprende el ácido nucleico (reivindicación 12), el plásmido que comprende el ácido nucleico o el casete de expresión (reivindicación 13), la célula hospedadora que comprende el ácido nucleico, el casete de expresión o el plásmido (reivindicación 14), el uso de la proteína de fusión para determinar si la función de una variante de CRM1 está alterada (reivindicación 15), un método para determinar si la función de una variante de CRM1 está alterada (reivindicación 16) y un método para identificar un compuesto con capacidad de modular la función de CRM1 (reivindicación 17).

D01 comprueba que survivina experimenta un transporte nucleocitoplasmático mediado por CRM1 utilizando polipéptidos reporteros.

D02 demuestra que survivina es una proteína cuya distribución subcelular está regulada por la importación al núcleo y por la exportación mediada por CRM1 del núcleo al citoplasma.

D03 demuestra que survivina contiene dos elementos de localización citoplasmáticos dependientes de CRM1.

D04 caracteriza el transporte nucleocitoplasmático de survivina<sub>140</sub> y de sus variantes survivina<sub>121</sub> y survivina<sub>40</sub>.

Demuestran que la localización intracelular de survivina<sub>140</sub> es mediada por una secuencia NES dependiente de CRM1, presente también en survivina<sub>121</sub> pero ausente en survivina<sub>40</sub>. Las variantes de survivina se fusionan con una proteína fluorescente por el extremo C-terminal, encontrándose que survivina<sub>140</sub>-GFP y survivina<sub>121</sub>-GFP son predominantemente citoplasmáticas. Además, las variantes de survivina con el residuo NES mutado no son exportadas por CRM1.

D05 investiga qué mecanismo determina la localización nuclear o citoplasmática de survivina, para lo cual se transfectan células con plásmidos que codifican diferentes versiones de la survivina humana. Encontraron que Flag-survivina y YFP-survivina se localizaban en el citoplasma, mientras que la proteína se re-localizaba en el núcleo cuando las células eran tratadas con leptomicina B, antifúngico que bloquea la exportación nuclear mediada por CRM1. Es decir, la localización nuclear o citoplasmática de survivina se debe a la actividad de CRM1, siendo necesaria la presencia de la región carboxi-terminal de survivina, comprendida entre los aminoácidos 89 y 142, para su transporte.

**NOVEDAD Y ACTIVIDAD INVENTIVA**

D01 comprueba que survivina experimenta un transporte nucleocitoplasmático mediado por CRM1 utilizando polipéptidos reporteros. Como control positivo utilizan Flag-survivina + NLS del antígeno T grande del virus SV40, así como YFP-survivina + NLS. Además, en el estudio se emplean los mutantes YFP-survivina(1–74) y YFP-survivina(74–142).

D02 demuestra que survivina es una proteína cuya distribución subcelular está regulada por la importación al núcleo y por la exportación mediada por CRM1 del núcleo al citoplasma. Utiliza variantes de survivina humana, que incluyen las secuencias codificantes de los epítomos Flag o c-Myc y la proteína fluorescente YFP. Entre las variantes ensayadas se encuentran: survivina(1-88), survivina(89-142), survivina(89-136) y survivina(119-142). También las isoformas survivina-2B y survivina-ΔEx3, así como los mutantes (Δ81-83), (Δ89-91) y (Δ81-83/89-91).

D03 demuestra que survivina contiene dos elementos de localización citoplasmáticos dependientes de CRM1, uno de los cuales es enmascarado en los dímeros de survivina ya que el dominio NES central de survivina sólo es activo en los monómeros. Se generan mutantes de survivina humana que contienen las secuencias NES pero sin el extremo C-terminal (1-109) o que contienen sólo dicho extremo (119-142), así como otros con mutaciones en la secuencia NES, observándose su localización nuclear o citoplasmática. Entre los mutantes generados: YFP-survivina(1-109), YFP-survivina(89-109), YFP-survivina (84-109), YFP-survivina(1-109)<sup>L96A</sup>, YFP-survivina(89-109)<sup>L96A</sup>, YFP-survivina (1-109)<sup>F101A/L102A</sup> y YFP-survivina(84-109)<sup>F101A/L102A</sup>.

Por tanto, D01 afecta a la novedad de las reivindicaciones 1, 4, 5, 7, 8 y 11-17 y a la actividad inventiva de las reivindicaciones 2, 3, 6, 9 y 10.

Respecto a D02 y D03, afectan a la novedad de las reivindicaciones 1, 4, 7, 8 y 11-17 y a la actividad inventiva de las reivindicaciones 2, 3, 5, 6, 9 y 10.