



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 271 013**

51 Int. Cl.:

C12N 15/12 (2006.01)

C07K 14/47 (2006.01)

C12N 15/62 (2006.01)

C12Q 1/68 (2006.01)

C12N 5/10 (2006.01)

C07K 16/18 (2006.01)

G01N 33/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **01936104 .7**

86 Fecha de presentación : **20.03.2001**

87 Número de publicación de la solicitud: **1266008**

87 Fecha de publicación de la solicitud: **18.12.2002**

54

Título: **Ligando de proteína que se fija al calcio por inducción neuronal aguda de tipo 1.**

30

Prioridad: **21.03.2000 EP 00106110**

45

Fecha de publicación de la mención BOPI:
16.04.2007

45

Fecha de la publicación del folleto de la patente:
16.04.2007

73

Titular/es: **Merck Patent GmbH**
Frankfurter Strasse 250
64293 Darmstadt, DE

72

Inventor/es: **Den Daas, Izaak;**
Dücker, Klaus y
Hock, Björn

74

Agente: **Carvajal y Urquijo, Isabel**

ES 2 271 013 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

Ligando de proteína que se fija al calcio por inducción neuronal aguda de tipo 1.

5 **Ámbito de la invención**

La presente invención se refiere a polipéptidos recientemente descubiertos y a los polinucleótidos que codifican dichos polipéptidos, denominados en ocasiones “ligando ANIC-BP-1” en el presente documento, para su uso en el diagnóstico y en la identificación de compuestos que puedan ser agonistas, antagonistas y puedan ser útiles para la terapia, y para la producción de dichos polipéptidos y polinucleótidos.

Antecedentes de la invención

El proceso de descubrimiento de fármacos está pasando actualmente por una revolución fundamental al adoptar la “genómica funcional”, es decir, la biología basada en genomas o genes de alto rendimiento. Este enfoque como una forma para identificar los genes y los productos de éstos como objetivos terapéuticos está dejando atrás rápidamente los planteamientos basados en la “clonación posicional”. Consiste en identificar un fenotipo, es decir, una función biológica o enfermedad genética, la cual se rastrea hasta el gen responsable por medio de su posición en el mapa genético.

La genómica funcional se basa en gran parte en las tecnologías de secuenciación de ADN de alta capacidad y las diferentes herramientas de la bioinformática para identificar secuencias de genes que puedan ser de interés a partir de las numerosas bases de datos de biología molecular que existen actualmente. Hay una necesidad constante de identificar y caracterizar nuevos genes y los polipéptidos y/o proteínas relacionados con éstos como objetivo para el descubrimiento de fármacos.

Resumen de la invención

La presente invención se refiere al ligando ANIC-BP-1, en particular a los polipéptidos ligandos ANIC-BP-1 y a los polinucleótidos ligandos ANIC-BP-1, a los materiales recombinantes y a los métodos para su producción. Estos polipéptidos y polinucleótidos resultan de interés para los métodos de tratamiento de algunas enfermedades, como los derrames cerebrales, los traumatismos craneales, la esclerosis múltiple, la enfermedad de Parkinson, la enfermedad de Alzheimer, las lesiones de la médula espinal, denominadas en lo sucesivo “enfermedades de la invención”, aunque sin limitarse a ellas. En otro aspecto, la invención se refiere a los métodos para identificar agonistas y antagonistas (por ejemplo, inhibidores) que utilicen los materiales contemplados en la invención, y para el tratamiento de dolencias asociadas al desequilibrio del ligando ANIC-BP-1 con los compuestos identificados. En otro aspecto más, la invención se refiere a las pruebas diagnósticas destinadas a la detección de las enfermedades asociadas a actividades o niveles inadecuados de ligando ANIC-BP-1.

Descripción de la invención

En un primer aspecto, la presente invención se refiere a los polipéptidos ligandos ANIC-BP-1. Entre dichos polipéptidos se encuentran:

- (a) un polipéptido codificado por un polinucleótido que comprende la secuencia de SEQ ID NO:1;
- (b) un polipéptido que comprende una secuencia de polipéptidos que son idénticos al menos en un 95%, 96%, 97%, 98% o 99% a la secuencia de polipéptidos de SEQ ID NO:2;
- (c) un polipéptido que comprende la secuencia de polipéptidos de SEQ ID NO:2;
- (d) un polipéptido que es idéntico al menos en un 95%, 96%, 97%, 98% o 99% a la secuencia de polipéptidos de SEQ ID NO:2;
- (e) la secuencia de polipéptidos de SEQ ID NO:2, y
- (f) un polipéptido que posee o comprende una secuencia de polipéptidos que tiene un índice de identidad de 0,95, 0,96, 0,97, 0,98 ó 0,99 comparada con la secuencia de polipéptidos de SEQ ID NO:2.

Se cree que los polipéptidos de la presente invención son miembros del motivo estéril alfa que contienen polipéptidos de la familia de las proteínas. Éstos resultan de interés porque recientemente se han identificado tres miembros descritos de los genes de la familia ANIC-BP, denominados ANIC-BP-1, ANIC-BP-2 y ANIC-BP-1B. El ANIC-BP-1 se encontró regulado al alza en un modelo con ratas de traumatismo craneal descubierto con la técnica denominada presentación diferencial de ARNm.

Las propiedades biológicas del ligando ANIC-BP-1 se designan en lo sucesivo “actividad biológica del ligando ANIC-BP-1” o “actividad del ligando ANIC-BP-1”. Preferiblemente, los polipéptidos de la presente invención muestran al menos una actividad biológica del ligando ANIC-BP-1.

ES 2 271 013 T3

Pueden utilizarse fragmentos de los polipéptidos de la invención para producir los polipéptidos completos correspondientes mediante síntesis de péptidos; por lo tanto, estas variantes pueden utilizarse como intermedios para la producción de los polipéptidos completos de la invención. Los polipéptidos de la presente invención pueden adoptar la forma de proteínas “maduras” o bien pueden ser parte de una proteína mayor, como un precursor o una proteína de fusión. A menudo resulta conveniente incluir una secuencia adicional de aminoácidos que contenga secuencias segregantes o secuencias líderes, prosequencias, secuencias que ayuden a la purificación, por ejemplo diversos residuos de histidina o una secuencia adicional para la estabilidad durante la producción de recombinantes.

Los polipéptidos de la presente invención pueden prepararse de cualquier manera adecuada, por ejemplo, mediante aislamiento de fuentes naturales, a partir de células anfitrionas manipuladas genéticamente que comprendan sistemas de expresión (véase más abajo) o por síntesis química utilizando, por ejemplo, sintetizadores automatizados de péptidos, o por medio de una combinación de estos métodos. La técnica actual conoce muy bien los medios para preparar estos polipéptidos.

En otro aspecto, la presente invención se refiere a los polinucleótidos ligandos ANIC-BP-1. Entre dichos polinucleótidos se encuentran:

- (a) un polinucleótido que comprende una secuencia de polinucleótidos que son idénticos al menos en un 95%, 96%, 97%, 98% o 99% a la secuencia de polinucleótidos de SEQ ID NO:1;
- (b) un polinucleótido que comprende la secuencia de polinucleótidos de SEQ ID NO:1;
- (c) un polinucleótido que es idéntico al menos en un 95%, 96%, 97%, 98% o 99% al polinucleótido de SEQ ID NO:1;
- (d) el polinucleótido de la de SEQ ID NO:1;
- (e) un polinucleótido que comprende una secuencia de polinucleótidos que codifica una secuencia de polipéptidos que son idénticos al 95%, 96%, 97%, 98%, o 99% a la secuencia de polipéptidos de SEQ ID NO:2;
- (f) un polinucleótido que comprende una secuencia de polinucleótidos que codifica el polipéptido de SEQ ID NO:2;
- (g) un polinucleótido que tiene una secuencia de polinucleótidos que codifica una secuencia de polipéptidos que son idénticos al menos en un 95%, 96%, 97%, 98% o 99% a la secuencia de polipéptidos de SEQ ID NO:2;
- (h) un polinucleótido que codifica los polipéptidos de SEQ ID NO:2;
- (i) un polinucleótido que posee o comprende una secuencia de polinucleótidos que tiene un índice de identidad de 0,95, 0,96, 0,97, 0,98 ó 0,99 comparada con la secuencia de polinucleótidos de SEQ ID NO:1.
- (j) un polinucleótido que posee o comprende una secuencia de polinucleótidos que codifica una secuencia de polipéptidos que tiene un índice de identidad de 0,95, 0,96, 0,97, 0,98 ó 0,99 comparada con la secuencia de polipéptidos de SEQ ID NO:2; y polinucleótidos que son complementarios de los polinucleótidos antes mencionados en toda su longitud.

En otro aspecto, la presente invención presenta polinucleótidos que son transcripciones de ARN de las secuencias de ADN de la presente invención. Por consiguiente se presenta un polinucleótido de ARN que:

- (a) comprende una transcripción de ARN de la secuencia de ADN que codifica el polipéptido de SEQ ID NO:2;
- (b) es la transcripción de ARN de la secuencia de ADN que codifica el polipéptido de SEQ ID NO:2;
- (c) comprende una transcripción de ARN de la secuencia de ADN de SEQ ID NO:1; o bien
- (d) es la transcripción de ARN de la secuencia de ADN de SEQ ID NO:1;

y nucleótidos de ARN que son complementarios de estos últimos.

La secuencia de polinucleótidos de SEQ ID NO: 1 es homóloga a AB011096 (KIAA0524; Nagase T. *et al.*, DNA Research 5, 31 - 39; 1998). La secuencia de polinucleótidos de SEQ ID NO:1 es una secuencia de ADNc que codifica el polipéptido de SEQ ID NO:2. La secuencia de polinucleótidos que codifica el polipéptido de SEQ ID NO:2 puede ser idéntica al polipéptido que codifica la secuencia de SEQ ID NO:1 o bien puede ser una secuencia distinta a SEQ ID NO:1, la cuál, como resultado de la redundancia (degeneración) del código genético, codifica igualmente el polipéptido de SEQ ID NO:2. El polipéptido de SEQ ID NO:2 está relacionado con otras proteínas del motivo

ES 2 271 013 T3

estéril alfa que contiene una familia de familias de proteínas que presenta homología y/o similitud estructural con GI-3043572 (KIAA0524; Nagase T. *et al.*, DNA Research 5, 31 - 39; 1998).

5 Se espera que los polipéptidos y polinucleótidos preferidos de la presente invención tengan, entre otras cosas, funciones y/o propiedades biológicas similares a las de sus polipéptidos y polinucleótidos homólogos. Además, los polipéptidos y polinucleótidos preferidos de la presente invención tienen al menos una actividad ligando ANIC-BP.

10 Los polinucleótidos de la presente invención pueden obtenerse utilizando clonación ordinaria y técnicas de rastreo a partir de una biblioteca de ADNc derivada del ARNm en células humanas de hueso, cerebro, mama, colon, células germinales, corazón, ovarios, páncreas, paratiroides, placenta, bazo, amígdalas, útero y colon (véase por ejemplo, Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, 2nd Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989)). Los polinucleótidos de la invención pueden obtenerse igualmente de fuentes naturales, como bibliotecas genómicas de ADN o pueden sintetizarse mediante técnicas bien conocidas disponibles comercialmente.

15 Si se utilizan polinucleótidos de la presente invención para la producción recombinante de polipéptidos de la presente invención, el polinucleótido puede comprender la secuencia de codificación del polipéptido maduro, en sí misma, o bien la secuencia de codificación para el polipéptido maduro en marco de lectura con otras secuencias de codificación, como las que codifican una secuencia líder o secretante, una secuencia de preproteínas, proproteínas o preproproteínas u otras partes de péptidos de fusión. Por ejemplo, puede codificarse una secuencia de marcadores que facilite la purificación del polipéptido fusionado. En algunas realizaciones preferidas de este aspecto de la invención la secuencia de marcadores es un péptido de hexahistidina como el proporcionado en el vector pQE (Qiagen, Inc.) y descrito en Gentz *et al.*, Proc Natl Acad Sci USA (1989) 86:821-824, o bien es una marca HA. El polinucleótido también puede contener secuencias no codificadoras 5' y 3', como secuencias transcritas no traducidas, señales de empalme y poliadenilación, sitios de unión de ribosomas y secuencias estabilizadoras de ARNm.

20 Los polinucleótidos que son idénticos o que presentan una identidad suficiente a la secuencia de polinucleótidos de SEQ ID NO:1 pueden utilizarse como sondas de hibridación para ADNc y ADN genómico o como cebadores para una reacción de amplificación de ácidos nucleicos (por ejemplo, PCR). Dichas sondas y cebadores pueden utilizarse para aislar clones completos de ADNc y genómicos que codifiquen polipéptidos de la presente invención y para aislar clones de ADNc y genómicos de otros genes (incluyendo genes que codifiquen parálogos de fuentes humanas y ortólogos y parálogos de especies distintas a la humana) cuya secuencia tenga una gran similitud a SEQ ID NO:1, normalmente al menos un 95% de identidad. Las sondas y cebadores preferidos generalmente comprenden al menos 15 nucleótidos, preferiblemente al menos 30 nucleótidos, y pueden tener al menos 50, si no al menos 100 nucleótidos. Las sondas especialmente preferidas tendrán entre 30 y 50 nucleótidos. Los cebadores especialmente preferidos tendrán entre 20 y 25 nucleótidos.

25 Puede obtenerse un polinucleótido que codifique un polipéptido de la presente invención, incluyendo homólogos de especies distintas a la humana, mediante un proceso que comprende los pasos de rastreo de una biblioteca en estrictas condiciones de hibridación con una sonda marcada que tenga la secuencia de SEQ ID NO:1 o un fragmento de ésta, preferiblemente de al menos 15 nucleótidos; y aislamiento de clones completos de ADNc y genómicos que contengan dicha secuencia de polinucleótidos. Los especialistas en la materia conocen bien estas técnicas de hibridación. Las condiciones estrictas de hibridación preferidas comprenden la incubación nocturna a 42°C en una solución formada por: 50% de formamida, 5xSSC (NaCl 150 mM, citrato trisódico 15 mM), fosfato sódico 50 mM (pH 7.6), 5x de solución de Denhardt; 10% de sulfato de dextrano y 20 microgramos/ml de ADN de esperma de salmón desnaturalizado y cizallado; a continuación los filtros se lavan en 0,1x SSC a aproximadamente 65°C.

30 Los especialistas observarán que, en muchos casos, una secuencia aislada de ADNc estará incompleta, ya que la región de codificación correspondiente al polipéptido no se extiende todo a lo largo del extremo 5'. Esto se debe a la transcriptasa inversa, una enzima que tiene una "procesividad" (capacidad de la encima para mantenerse unida al modelo durante la reacción de polimerización) inherentemente baja, por lo que no completa una copia del ADN del modelo de ARNm durante la síntesis del ADNc de primera cadena.

35 Existen varios métodos bien conocidos por los especialistas en la materia para obtener ADNc completos o para extender ADNc cortos, por ejemplo, los basados en el método de Amplificación Rápida de extremos de ADNc (Rapid Amplification of cDNA ends - RACE) (véase, por ejemplo, Frohman *et al.*, Proc Nat Acad Sci USA 85, 8998-9002, 1988). Los cambios recientes introducidos en la técnica, ejemplificados por la tecnología Marathon (marca comercial, Clontech Laboratories Inc.), han simplificado considerablemente la búsqueda de ADNc más largos. En la tecnología Marathon (marca comercial), los ADNc se preparan a partir de ARNm extraídos de un tejido seleccionado y una secuencia de "adaptadores" ligada a cada extremo. A continuación se realiza una amplificación de ácido nucleico (PCR) para amplificar el extremo 5' "que falta" del ADNc utilizando una combinación de cebadores oligonucleótidos específicos para los genes y los adaptadores. A continuación se repite la reacción PCR utilizando cebadores "anidados", es decir, cebadores diseñados para templar dentro del producto amplificado (normalmente un cebador específico para el adaptador que temple otro 3' en la secuencia del adaptador y un cebador específico para el gen que temple otro 5' en la secuencia conocida del gen). Los productos de esta reacción pueden analizarse posteriormente por medio de una secuencia de ADN y puede construirse un ADNc completo ya sea uniendo el producto directamente al ADNc existente para obtener una secuencia completa, o bien llevando a cabo un PCR completo por separado utilizando la información de la nueva secuencia para diseñar el cebador 5'.

Los péptidos recombinantes de la presente invención pueden prepararse mediante procesos bien conocidos por la técnica a partir de células anfitrionas genéticamente manipuladas que contienen sistemas de expresión. Por consiguiente, en otro aspecto, la presente invención se refiere a sistemas de expresión que comprenden un polinucleótido o polinucleótidos de la presente invención, a células anfitrionas genéticamente manipuladas con dichos sistemas de expresión y a la producción de polipéptidos de la invención mediante técnicas de recombinación. Asimismo pueden emplearse sistemas de traducción sin células para producir dichas proteínas utilizando ARN derivados de las estructuras de ADN de la presente invención.

Para la producción de recombinantes, las células anfitrionas pueden modificarse genéticamente para incorporar sistemas de expresión o partes de éstos para los polinucleótidos de la presente invención. Los polinucleótidos pueden introducirse en las células anfitrionas mediante los métodos que se describen en numerosos manuales de laboratorio, como el que figura en Davis *et al.*, Basic Methods in Molecular Biology (1986) y Sambrook *et al. (ibid)*. Entre los métodos preferidos para introducir polinucleótidos en células anfitrionas se encuentran, por ejemplo, la transfección de fosfatos de calcio, la transfección mediada por DEAE-dextrano, la transvección, la microinyección, la transfección mediada por lípidos catiónicos, la electroporación, la transducción, la carga por raspado, la introducción balística o la infección.

Algunos ejemplos representativos de anfitriones adecuados comprenden células bacterianas, como *Streptococci*, *Staphylococci*, *E. coli*, *Streptomyces* y *Bacillus subtilis*; células fúngicas, como las células de levadura y de *Aspergillus*; células de insectos, como las células *Drosophila* S2 y *Spodoptera* Sf9; células de animales, como las células CHO, COS, HeLa, C127, 3T3, BHK, HEK 293 y de melanoma de Bowes; y células de plantas.

Se puede utilizar una gran variedad de sistemas de expresión, como por ejemplo, sistemas cromosómicos, episómicos y derivados de virus, por ejemplo, vectores derivados de plásmidos bacterianos, de bacteriófagos, de transposones, de episomas de levadura, de elementos de inserción, de elementos cromosómicos de levaduras, de virus, como los baculovirus, los virus de papova, como el SV40, los virus de vacunas, los adenovirus, virus de la viruela aviar, virus y retrovirus de pseudorabias, y vectores derivados de las combinaciones de éstos, como los derivados de elementos genéticos de plásmidos y bacteriófagos, como los cósmidos y fagémidos. Los sistemas de expresión pueden contener regiones de control que regulen y generen expresión. Por lo general se puede utilizar cualquier sistema o vector capaz de mantener, propagar o expresar un polinucleótido para producir un polipéptido en un anfitrión. La secuencia adecuada de polinucleótidos puede introducirse en un sistema de expresión mediante una gran variedad de técnicas bien conocidas y rutinarias, como por ejemplo, las expuestas en Sambrook *et al.*, (*ibid*). Las señales de secreción adecuadas pueden incorporarse en el polipéptido deseado para permitir que la secreción de la proteína traducida entre en el lumen del retículo endoplásmico, el espacio periplásmico o el entorno extracelular. Estas señales pueden ser endógenas al polipéptido o pueden ser señales heterólogas.

Si se desea expresar un polipéptido de la presente invención para su uso en ensayos de rastreo, generalmente es preferible que el polipéptido se produzca en la superficie de la célula. En este caso, las células pueden cosecharse antes de su uso en el ensayo de rastreo. Si el polipéptido se secreta en el medio, éste puede ser recuperado a fin de recuperar y purificar el polipéptido. Si se produce dentro de la célula, las células deben romperse antes de recuperar el polipéptido.

Los polipéptidos de la presente invención pueden recuperarse y purificarse a partir de cultivos de células recombinantes por medio de métodos bien conocidos, como precipitación con sulfato amónico o etanol, extracción ácida, cromatografía de intercambio de aniones o cationes, cromatografía de fosfocelulosa, cromatografía de interacción hidrofóbica, cromatografía de afinidad, cromatografía de hidroxilapatita y cromatografía de lectina. Preferiblemente se utiliza cromatografía líquida de alto rendimiento para la purificación. Se pueden utilizar técnicas bien conocidas para replegar proteínas a fin de regenerar la conformación activa cuando el polipéptido se desnaturaliza durante la síntesis intracelular, el aislamiento y/o la purificación.

Los polinucleótidos de la presente invención pueden utilizarse como reactivos de diagnóstico mediante la detección de mutaciones del gen correspondiente. La detección de una forma mutada del gen caracterizado por el polinucleótido de SEQ ID NO:1 en la secuencia de ADNc o secuencia genómica, y que está relacionada con una disfunción, ofrecerá una herramienta de diagnóstico que puede complementar o definir el diagnóstico de una enfermedad o propensión a una enfermedad, resultante de la infraexpresión, sobreexpresión o expresión alterada del gen en el espacio o en el tiempo. Los individuos portadores de mutaciones del gen pueden detectarse a nivel de ADN mediante diversas técnicas bien conocidas.

Los ácidos nucleicos para el diagnóstico pueden obtenerse de las células del sujeto, como sangre, orina, saliva, biopsias de tejidos o material de autopsia. El ADN genómico puede utilizarse directamente para fines de detección o puede amplificarse enzimáticamente utilizando PCR, preferiblemente RT-PCR, u otras técnicas de amplificación previas al análisis. Los ARN o ADNc pueden usarse de forma similar. Pueden detectarse supresiones e inserciones por medio de un cambio de tamaño del producto amplificado respecto al genotipo normal. Las mutaciones puntuales pueden identificarse mediante la hibridación de ADN amplificados a secuencias de nucleótidos ligados ANIC-BP marcados. Las secuencias perfectamente coincidentes pueden distinguirse de los duplicados no coincidentes mediante digestión de RNasa o por medio de las diferencias entre las temperaturas de fusión. La diferencia entre secuencias de ADN puede detectarse igualmente a través de las alteraciones de la movilidad electroforética de los fragmentos de ADN en geles, con o sin agentes desnaturalizantes, o mediante secuenciación directa de ADN (véase por ejemplo,

Myers *et al.*, Science (1985) 230:1242). Asimismo, los cambios de secuencia en lugares concretos pueden revelarse mediante ensayos de protección de la nucleasa, como la RNasa y la protección de S1, o por el método de escisión química (véase Cotton *et al.*, Proc Natl Acad Sci USA (1985) 85: 4397-4401).

5 Se puede construir una matriz de sondas de oligonucleótidos que comprendan una secuencia de polinucleótidos ligandos ANIC-BP o fragmentos de ésta para llevar a cabo un rastreo eficiente, por ejemplo, de mutaciones genéticas. Dichas matrices serán preferiblemente matrices o retículas de alta densidad. Los métodos basados en la tecnología de matrices son bien conocidos y pueden aplicarse de forma general, pudiendo ser utilizados para resolver una serie de cuestiones en genética molecular, como la expresión de genes, el vínculo entre genes y la variabilidad genética, véase, por ejemplo, M.Chee *et al.*, Science, 274, 610-613 (1996) y otras referencias citadas en dicho trabajo.

15 La detección de la disminución o aumento anormales de los niveles de expresión de los polipéptidos o de ARNm puede utilizarse también para diagnosticar o determinar la propensión de un sujeto a una de las enfermedades de la invención. La disminución o aumento de la expresión puede medirse a nivel de ARN mediante cualquiera de los métodos conocidos para cuantificar polinucleótidos, como por ejemplo, la amplificación de ácidos nucleicos, por ejemplo, PCR, RT-PCR, protección de RNasa, Northern blotting y otros métodos de hibridación. Las técnicas de análisis que pueden utilizarse para determinar los niveles de una proteína, como un polipéptido de la presente invención, en una muestra derivada de un anfitrión son bien conocidas entre los especialistas en la materia. Dichos métodos de análisis comprenden los ensayos de radioinmunología, los ensayos de unión competitiva, los análisis Western Blot y los ensayos ELISA,

De este modo, en otro aspecto, la presente invención ofrece un kit de diagnóstico formado por:

- 25 (a) un polinucleótido de la presente invención, preferiblemente la secuencia de nucleótidos de SEQ ID NO: 1 o una transcripción del ARN de ésta;
- (b) una secuencia de nucleótidos complementaria de la indicada en (a);
- 30 (c) un polipéptido de la presente invención, preferiblemente el polipéptido de SEQ ID NO:2.

Como puede observarse, en dicho kit, los elementos (a), (b) o (c) pueden comprender un componente sustancial. Dicho kit se utilizará para diagnosticar una enfermedad o propensión a una enfermedad, en particular las enfermedades de la invención, entre otras.

35 Las secuencias de polinucleótidos de la presente invención resultan útiles para los estudios de localización de cromosomas. La secuencia está dirigida concretamente a un lugar determinado de un cromosoma humano y puede hibridarse con él. El cartografiado de las secuencias correspondientes con los cromosomas que contempla la presente invención constituye un paso importante para establecer una correlación entre dichas secuencias y la enfermedad asociada al gen. Una vez que se ha cartografiado una secuencia en un lugar preciso del cromosoma, puede establecerse una correlación entre la posición física de la secuencia en el cromosoma y los datos del mapa genético. Dichos datos pueden encontrarse, por ejemplo, en V. McKusick, Mendelian Inheritance in Man (disponible en línea en la Welch Medical Library de la Universidad John Hopkins). A continuación se identifica la relación entre genes y enfermedades que se ha cartografiado en la misma región del cromosoma a través de un análisis de vínculos (co-herencia de genes físicamente adyacentes). Las ubicaciones precisas de cromosomas humanos para una secuencia genómica (fragmentos de genes, etc.) puede determinarse mediante Radiation Hybrid (RH) Mapping (Walter, M. Spillett, D., Thomas, P., Weissenbach, J., and Goodfellow, P., (1994) A method for constructing radiation hybrid maps of whole genomes, Nature Genetics 7, 22-28). ResearchGenetics (Huntsville, AL, USA) ofrece una serie de paneles de RH, por ejemplo, el panel GeneBridge4 RH (Hum Mol Genet 1996 Mar;5(3):339-46 A radiation hybrid map of the human genome. Gyapay G, Schmitt K, Fizames C, Jones H, Vega-Czarny N, Spillett D, Muselet D, Prud'Homme JF, Dib C, Auffray C, Morissette J, Weissenbach J, Goodfellow PN). Para determinar la ubicación de un gen en el cromosoma utilizando este panel, se realizan 93 PCR utilizando cebadores diseñados a partir del gen de interés sobre ADN de RH. Cada uno de estos ADN contiene fragmentos genómicos humanos aleatorios que se mantienen en un entorno de hámster (líneas celulares híbridas de humanos y hámsteres). Estos PCR dan por resultado 93 marcadores que indican la presencia o ausencia del producto PCR del gen correspondiente. Estos marcadores se comparan con marcadores obtenidos utilizando productos de PCR procedentes de secuencias genómicas cuya ubicación ya se conoce. Esta comparación puede verse en <http://www.genome.wi.mit.edu/>. El gen de la presente invención está situado en el cromosoma humano Chr. 17 (D17S922-D17S798).

60 Las secuencias de polinucleótidos de la presente invención constituyen igualmente herramientas útiles para los estudios de expresión de tejidos. Dichos estudios permiten determinar las pautas de expresión de los polinucleótidos de la presente invención que podrían ser una indicación de las pautas de expresión de los polipéptidos codificados en tejidos, al detectar los ARNm que los codifican. Las técnicas utilizadas son muy conocidas e incluyen las técnicas de hibridación *in situ* con clones que forman una retícula, como la hibridación de micromatrices de ADNc (Schena *et al.*, Science, 270, 467-470, 1995 y Shalon *et al.*, Genome Res, 6, 639-645, 1996), y las técnicas de amplificación de nucleótidos, como el PCR. Un método preferido utiliza la tecnología TAQMAN (marca comercial) de Perkin Elmer. Los resultados de estos estudios pueden indicar que el polipéptido funciona normalmente en el organismo. Además, los estudios comparativos entre la pauta de expresión normal de los ARNm y la pauta de expresión normal de los ARNm codificados por una forma alternativa del mismo gen (por ejemplo, que tenga una alteración en el potencial

de codificación del polipéptido o una mutación de regulación) pueden permitir conocer el papel que desempeñan los polipéptidos de la presente invención o la expresión anormal de éstos en enfermedades. Dicha expresión anormal puede ser de naturaleza temporal, espacial o simplemente cuantitativa.

5 Los polipéptidos de la presente invención se expresan en los huesos, el cerebro, las mamas, el colon, las células germinales, el corazón, los ovarios, el páncreas, la paratiroides, la placenta, el bazo, las amígdalas, el útero y el colon.

Otro aspecto de la presente invención sirve para revelar anticuerpos. Los polipéptidos de la invención o sus fragmentos, o las células que los expresan, pueden utilizarse como inmunógenos para producir anticuerpos inmunoespecíficos para los polipéptidos de la presente invención. Por "inmunoespecífico" se entiende que los anticuerpos tienen una afinidad considerablemente mayor con los polipéptidos de la invención que con otros polipéptidos anteriores.

10 Los anticuerpos generados contra los polipéptidos de la presente invención pueden obtenerse administrando los polipéptidos o fragmentos o células que contengan epítopos a un animal, preferiblemente un animal no humano, empleando para ello los protocolos rutinarios. Para la preparación de anticuerpos monoclonales se puede utilizar cualquier técnica que ofrezca anticuerpos producidos por cultivos de líneas celulares continuas. Por ejemplo, la técnica del hibridoma (Kohler, G. y Milstein, C., Nature (1975) 256: 495-497), la técnica del trioma, la técnica de hibridomas de células humanas B (Kozbor *et al.*, Immunology Today (1983) 4:72) y la técnica del hibridoma EBV (Cole *et al.*, Monoclonal Antibodies and Cancer Therapy, 77-96, Alan R. Liss, Inc., 1985).

20 Asimismo, las técnicas para la producción de anticuerpos de cadena simple, como los que se describen en la patente US nº 4,946,778, pueden adaptarse para producir anticuerpos de cadena simple a partir de polipéptidos de la presente invención. También pueden utilizarse ratones u otros organismos, incluso mamíferos, transgénicos para expresar anticuerpos humanizados.

25 Los anticuerpos antes descritos pueden utilizarse para aislar o identificar clones que expresen el polipéptido o para purificar polipéptidos mediante cromatografía de afinidad. Los anticuerpos contra los polipéptidos de la presente invención se pueden utilizar igualmente para el tratamiento de las enfermedades de la invención, entre otras.

30 Los polipéptidos y polinucleótidos de la presente invención pueden emplearse igualmente como vacunas. Por consiguiente, en otro aspecto, la presente invención se refiere a un método para inducir una respuesta inmunológica en un mamífero que comprende la inoculación en el animal de un polipéptido de la presente invención que sea capaz de producir anticuerpos y/o una respuesta inmune en las células T, incluyendo, por ejemplo, células T que produzcan citoquina o células T citotóxicas, para proteger a dicho animal contra la enfermedad, esté o no presente dicha enfermedad en el sujeto. Asimismo se puede provocar una respuesta inmunológica en un mamífero mediante un método consistente en administrar *in vivo* un polipéptido de la presente invención a través de un vector que dirija la expresión del polipéptido y codifique éste a fin de inducir la respuesta inmunológica para producir anticuerpos y proteger al animal contra las enfermedades de la invención. Una manera de administrar este vector es acelerándolo en las células deseadas en forma de recubrimiento de partículas. Dicho vector del ácido nucleico puede comprender ADN, ARN, un ácido nucleico modificado o un híbrido de ADN/ARN. Para el uso en vacunas, el polipéptido o el vector de ácido nucleico se administra normalmente en una formulación de vacuna (composición). Esta formulación puede incluir además un vehículo adecuado. Puesto que un polipéptido puede descomponerse en el estómago, es preferible su administración por vía parenteral (por ejemplo, inyección subcutánea, intramuscular, intravenosa o intradérmica). Entre las formulaciones adecuadas para su administración parenteral se encuentran las soluciones inyectables estériles acuosas y no acuosas que pueden contener antioxidantes, tampones, bacteriostatos y solutos que den a la fórmula propiedades instónicas en la sangre del receptor, así como suspensiones estériles acuosas y no acuosas que pueden incluir agentes de suspensión o espesantes. Las formulaciones pueden presentarse en recipientes de dosis unitarias o dosis múltiples, por ejemplo, ampollas y viales precintados, y pueden almacenarse en lugar fresco y seco, de modo que sólo requieran la adición del vehículo líquido estéril inmediatamente antes de su uso. Las formulaciones de vacunas pueden incluir igualmente sistemas adyuvantes para mejorar la inmunogenicidad de la formulación, como sistemas de aceite en agua y otros sistemas conocidos. La dosis dependerá de la actividad específica de la vacuna y puede determinarse fácilmente mediante experimentos rutinarios.

55 Los polipéptidos de la presente invención tienen una o varias funciones biológicas que revisten importancia en uno o varios estados patológicos, en particular en el caso de las enfermedades de la invención mencionadas en el presente documento. Por ello resulta conveniente identificar los compuestos que estimulan o inhiben la función o el nivel del polipéptido. Por consiguiente, en otro aspecto, la presente invención ofrece un método para rastrear compuestos para identificar aquellos que estimulan o inhiben la función o nivel del polipéptido. Dichos métodos identifican a los agonistas o antagonistas que pueden emplearse para fines terapéuticos o profilácticos con las enfermedades de la invención antes mencionadas. Los compuestos pueden identificarse a partir de diversas fuentes, por ejemplo, células, preparaciones sin células, bibliotecas químicas, colecciones de compuestos químicos y mezclas de productos naturales. Los agonistas o antagonistas así identificados pueden ser sustratos naturales o modificados, ligandos, receptores, enzimas, etc., según convenga, del polipéptido; una mimética estructural o funcional de éste (véase Coligan *et al.*, Current Protocols in Immunology 1 (2):Chapter 5 (1991)) o una molécula pequeña.

65 El método de rastreo puede simplemente medir la unión de un compuesto candidato al polipéptido o a células o membranas que contengan el polipéptido, o una proteína de fusión de éste, mediante una marca asociada directa o indirectamente con el compuesto candidato. De forma alternativa, el método de rastreo puede implicar la medición o

detección (cualitativa o cuantitativa) de la unión competitiva de un compuesto candidato al polipéptido respecto a un competidor marcado (por ejemplo, un agonista o antagonista). Además, estos métodos de rastreo pueden comprobar si el compuesto candidato emite una señal generada por la activación o inhibición del polipéptido, utilizando sistemas de detección adecuados para las células que contienen el polipéptido. Por lo general, los inhibidores de activación se comprueban en presencia de un agonista conocido, y se observa el efecto que la presencia del compuesto candidato tiene sobre la activación por parte del agonista. Además, los métodos de rastreo pueden consistir simplemente en mezclar un compuesto candidato con una solución que contenga un polipéptido de la presente invención para formar una mezcla, midiendo la actividad del ligando ANIC-BP y comparando la actividad del ligando ANIC-BP de la mezcla con una mezcla control que no contenga el compuesto candidato.

Los polipéptidos de la presente invención pueden utilizarse en métodos convencionales de rastreo de baja capacidad, así como en formatos de rastreo de alto rendimiento (HTS). Estos formatos HTS incluyen no sólo el uso de placas de microtitración de 96 pocillos, más recientemente de 384, sino también nuevos métodos, como el método de nanopozo que se describe en Schullek *et al.*, *Anal Biochem.*, 246, 20-29, (1997).

Las proteínas de fusión, como las construidas a partir de un fragmento Fc y un polipéptido ligando ANIC-BP, como las descritas en el presente documento, pueden utilizarse igualmente para ensayos de rastreo de alto rendimiento para identificar antagonistas para el polipéptido de la presente invención (véase D. Bennett *et al.*, *J Mol Recognition*, 8:52-58 (1995); y K. Johanson *et al.*, *J Biol Chem*, 270(16):

Técnicas de rastreo

Los polinucleótidos, polipéptidos y anticuerpos del polipéptido de la presente invención pueden utilizarse asimismo para configurar métodos de rastreo para detectar el efecto de los compuestos añadidos sobre la producción de ARNm y polipéptidos en células. Por ejemplo, se puede realizar un ensayo ELISA para medir los niveles de polipéptido secretados o asociados a las células, utilizando anticuerpos monoclonales y policlonales mediante métodos ordinarios bien conocidos. Este método puede utilizarse para descubrir agentes que puedan inhibir o incrementar la producción de polipéptido (denominados antagonistas y agonistas, respectivamente) a partir de células o tejidos debidamente manipulados.

Se puede emplear un polipéptido de la presente invención para identificar receptores unidos a la membrana o solubles, en caso de existir, a través de técnicas estándar de unión de receptores ya conocidas. Entre éstas se encuentran los ensayos de unión de ligandos y ensayos de reticulación en los que el polipéptido es marcado con un isótopo radiactivo (por ejemplo, ¹²⁵I), modificado químicamente (por ejemplo, biotilado) o fusionado con una secuencia de péptidos adecuada para la detección o purificación, e incubado con una fuente del receptor putativo (células, membranas celulares, sobrenadantes de células, extractos de tejidos, fluidos corporales). Otros métodos son las técnicas biofísicas, como la resonancia plasmónica de superficie y la espectroscopia. Estos métodos de rastreo pueden utilizarse igualmente para identificar agonistas y antagonistas del polipéptido que compitan con la unión del polipéptido a sus receptores, en caso de existir éstos. Actualmente ya se conocen métodos estándar para llevar a cabo dichos métodos.

Algunos ejemplos de antagonistas de los polipéptidos de la presente invención son los anticuerpos o, en algunos casos, los oligonucleótidos o proteínas que están estrechamente relacionados con los ligandos, sustratos, receptores, enzimas, etc. del polipéptido, por ejemplo, un fragmento de los ligandos, sustratos, receptores o enzimas; o bien una pequeña molécula que se una al polipéptido de la presente invención, pero que no provoque una respuesta, de forma que se evite la actividad del polipéptido.

Los métodos de rastreo pueden implicar igualmente el uso de tecnología transgénica y del gen ligando ANIC-BP-1. La técnica de construcción de animales transgénicos es ya bien conocida. Por ejemplo, puede introducirse el gen ligando ANIC-BP-1 mediante microinyección en el pronúcleo masculino de oocitos fertilizados, mediante transferencia retroviral en embriones antes o después de su implantación o bien mediante inyección de células madre genéticamente modificadas, por ejemplo, por electroporación, en blastocistos anfitriones. Resultan de especial utilidad los animales transgénicos denominados "knock-in", en los que se sustituye un gen animal por su equivalente humano dentro del genoma del animal. Los animales transgénicos *knock-in* son útiles para el proceso de descubrimiento de fármacos, para la validación de dianas, en la que el compuesto es específico para el objetivo humano. Otros animales transgénicos útiles son los denominados animales "knock-out", en los que se anula parcial o totalmente la expresión del ortólogo animal de un polipéptido de la presente invención y codificado por una secuencia endógena de ADN. El *knock-out* del gen puede ir dirigido a células o tejidos específicos, puede producirse únicamente en ciertas células o tejidos como consecuencia de las limitaciones de la tecnología o bien puede producirse en todas o casi todas las células del animal. La tecnología de animales transgénicos ofrece todo un sistema animal de clonación de expresiones en el que los genes introducidos se expresan para obtener grandes cantidades de polipéptidos de la presente invención.

Se presentan kits de rastreo para su uso con los métodos antes descritos. Estas unidades de rastreo comprenden:

- (a) un polipéptido de la presente invención;
- (b) una célula recombinante que expresa un polipéptido de la presente invención;
- (c) una membrana celular que expresa un polipéptido de la presente invención, o bien

(d) un anticuerpo para un polipéptido de la presente invención;

cuyo polipéptido es preferiblemente el de SEQ ID NO:2.

5 Cabe señalar que en dicho kit, los elementos (a), (b), (c) o (d) pueden comprender un componente sustancial.

Glosario

10 Las siguientes definiciones tiene por finalidad facilitar la comprensión de algunos de los términos usados con frecuencia en el presente documento.

A los efectos del presente documento, por “anticuerpos” se entienden los anticuerpos policlonales y monoclonales, quiméricos, de cadena simple y humanizados, así como los fragmentos Fab, incluidos los productos de un Fab u otra biblioteca de expresión de inmunoglobulina.

15 Por “aislado” se entiende alterado “por la mano del hombre” respecto a su estado natural, es decir, que si aparece en la naturaleza ha sido modificado o extraído de su entorno original o ambas cosas. Por ejemplo, un polipéptido o un polinucleótido que estén presentes de forma natural en un organismo vivo no están “aislados”, pero el mismo polipéptido o polinucleótido separados de los materiales que coexisten en su estado natural está “aislado”, en la
20 acepción utilizada en el presente documento. Además, un polinucleótido o polipéptido que ha sido introducido en un organismo mediante transformación, manipulación genética o por cualquier otro método recombinante está “aislado”, incluso si sigue presente en dicho organismo sea éste un organismo vivo o no.

25 En términos generales, el término “polinucleótido” se refiere a un poliribonucleótido (ARN) o a un polidesoxiribonucleótido (ADN), que puede ser ARN o ADN modificado o no. Entre los “polinucleótidos” se encuentran los ADN de cadena simple y cadena doble, los ADN que son una mezcla de regiones de cadena simple y cadena doble, los ARN de cadena simple y cadena doble y los ARN que son una mezcla de regiones de cadena simple y cadena
30 doble, moléculas híbridas que incluyen ADN y ARN que pueden tener cadena simple o, como es más común, cadena doble o una mezcla de regiones de cadena simple y cadena doble. Además, el término “polinucleótido” se refiere a las regiones de cadena triple que comprenden ARN o ADN, o ambos. El término “polinucleótido” incluye igualmente los ADN o ARN que contienen una o varias bases modificadas, y los ADN y ARN con estructuras modificadas por motivos de estabilidad o de otro tipo. Las bases “modificadas” incluyen, por ejemplo, las bases tritiladas y las bases poco habituales, como la inosina. Se pueden realizar diversas modificaciones a los ADN y ARN, por lo que el término
35 “polinucleótido” comprende formas de polinucleótidos modificadas química, enzimática y metabólicamente, como los que se encuentran en la naturaleza, así como las formas químicas de ADN y ARN que son características de virus y células. Asimismo, el término “polinucleótido” abarca los polinucleótidos relativamente cortos, a los que a menudo se denomina oligonucleótidos.

40 El término “polipéptido” se refiere a cualquier polipéptido que comprenda dos o más aminoácidos unidos entre sí por enlaces peptídicos o enlaces peptídicos modificados, es decir, isoésteres peptídicos. El término “polipéptido” se refiere tanto a los de cadena corta, conocidos comúnmente como péptidos, oligopéptidos u oligómeros, como a los de cadena larga, a los que generalmente se denomina proteínas. Los polipéptidos pueden contener aminoácidos distintos de los 20 aminoácidos codificados por genes. Los “polipéptidos” incluyen secuencias de aminoácidos modificados ya sea por procesos naturales, como los procesos postraslacionales, o bien por técnicas de modificación química que son
45 bien conocidas. Dichas modificaciones ya han sido descritas en los textos básicos y en monografías más detalladas, así como en la voluminosa bibliografía de investigación. Las modificaciones pueden tener lugar en cualquier parte de un polipéptido, incluyendo su estructura, las cadenas laterales de aminoácido o los términos amino o carboxilo. Cabe señalar que el mismo tipo de modificación puede estar presente al mismo grado o a diversos grados en diferentes lugares de un polipéptido determinado. Un polipéptido puede contener también muchos tipos de modificaciones. Los
50 polipéptidos pueden estar ramificados como resultado de la ubiquitinación, y pueden ser cíclicos, con o sin ramificaciones. Los polipéptidos cíclicos, ramificados y ramificados cíclicos pueden ser resultado de procesos naturales de postraducción o pueden producirse mediante métodos sintéticos. Entre estas modificaciones se encuentran la acetilación, la acilación, la ribosilación de ADP, la amidación, la biotilación, el enlace covalente de flavina, el enlace covalente de un nucleótido o un derivado de éste, el enlace covalente de un lípido o de un derivado de éste, el enlace
55 covalente de fosfatidilinositol, la reticulación, la ciclización, la formación de enlaces disulfuro, la desmetilación, la formación de entrecruzamientos covalentes, la formación de cistina, la formación de piroglutamato, la formilación, la carboxilación gamma, la glicosilación, la formación de anclajes GPI, la hidroxilación, la yodación, la metilación, la miristilación, la oxidación, el procesado proteolítico, la fosforilación, la prenilación, la racemización, la selenoilación, la sulfación, la adición mediada por ARN de transferencia de aminoácidos a proteínas, como la arginilación,
60 y la ubiquitinación (véase, por ejemplo, *Proteins - Structure and Molecular Properties*, 2nd Ed., T. E. Creighton, W. H. Freeman and Company, New York, 1993; Wold, F., *Post-translational Protein Modifications: Perspectives and Prospects*, 1-12, in *Post-translational Covalent Modification of Proteins*, B. C. Johnson, Ed., Academic Press, New York, 1983; Seifter *et al.*, “Analysis for protein modifications and nonprotein cofactors”, *Meth Enzymol*, 182, 626-646, 1990, y Rattan *et al.*, “Protein Synthesis: Post-translational Modifications and Aging”, *Ann NY Acad Sci*, 663,
65 48-62, 1992).

El término “fragmento” de una secuencia de polipéptidos se refiere a una secuencia de polipéptidos que es más corta que la secuencia de referencia, pero que mantiene las mismas funciones o actividades biológicas esenciales que

ES 2 271 013 T3

el polipéptido de referencia. El término “fragmento” de una secuencia de polinucleótido se refiere a una secuencia de polinucleótido que es más corta que la secuencia de referencia de SEQ ID NO:1.

5 El término “variante” se refiere a un polinucleótido o polipéptido que difiere del polinucleótido o polipéptido de referencia, pero mantiene las propiedades esenciales de éste. Una variante típica de un polinucleótido difiere del polinucleótido de referencia en la secuencia de nucleótidos. Los cambios en la secuencia de nucleótidos de la variante pueden modificar o no la secuencia de aminoácidos de un polipéptido codificado por el polinucleótido de referencia. Los cambios en los nucleótidos pueden tener como resultado sustituciones, adiciones, supresiones, fusiones y truncamientos de aminoácidos en el polipéptido codificado por la secuencia de referencia, como se expone más abajo. La
10 secuencia de aminoácidos de una variante típica de un polipéptido difiere del polipéptido de referencia. Por lo general, estas alteraciones son limitadas, por lo que las secuencias del polipéptido de referencia y de la variante son similares en general y, en muchas regiones, idénticas. Las secuencias de aminoácidos de una variante de polipéptido y de un polipéptido de referencia pueden diferir por una o más sustituciones, inserciones, supresiones en cualquier combinación. Un residuo de un aminoácido sustituido o insertado puede ser o no el codificado por el código genético. Entre las
15 sustituciones conservadoras encontramos Gly, Ala; Val, Ile, Leu; Asp, Glu; Asn, Gln; Ser, Thr; Lys, Arg; y Phe y Tyr. Una variante de un polinucleótido o polipéptido puede presentarse de forma natural, como en el caso de los alelos, o puede ser una variante que no se presente en la naturaleza. Las variantes de polinucleótidos y polipéptidos que no se presentan en la naturaleza pueden producirse mediante técnicas de mutagénesis o por síntesis directa. Otras variantes son los polipéptidos que tienen una o varias modificaciones postraducción, por ejemplo, la glicosilación, la fosforilación, la metilación, la ribosilación de ADP y modificaciones similares. Entre sus materializaciones encontramos la metilación del aminoácido N terminales, las fosforilaciones de serinas y treoninas, y la modificación de glicinas C-terminales.

25 El término “alelo” se refiere a una de las dos o más formas alternativas de un gen que tienen lugar en un lugar determinado del genoma.

El término “polimorfismo” se refiere a una variación de la secuencia de nucleótidos (y en su caso, de la secuencia de polipéptidos codificados) en un lugar determinado del genoma dentro de una población.

30 El término “polimorfismo de nucleótido simple” (Single Nucleotide Polymorphism - SNP) se refiere a la existencia de variabilidad del nucleótido en una única posición de éste en el genoma, dentro de una población. El SNP puede tener lugar dentro de un gen o dentro de regiones intergénicas del genoma. Los SNP pueden verificarse utilizando la amplificación específica de alelos (Allele Specific Amplification - ASA). Para este proceso se requiere un mínimo de tres cebadores. Un cebador común se utiliza en complemento inverso respecto al polimorfismo objeto del ensayo.
35 Este cebador común puede encontrarse entre 50 y 1500 bps de la base polimórfica. Los otros dos (o más) cebadores son idénticos entre sí, salvo que la base final 3' fluctúa para coincidir con uno de los dos (o más) alelos que forman el polimorfismo. A continuación se realizan dos (o más) reacciones PCR sobre la muestra de ADN, cada una de las cuales utiliza el cebador común y uno de los cebadores específicos del alelo.

40 El término “variante de empalme” en la acepción utilizada en el presente documento se refiere a moléculas de ADNc producidas a partir de moléculas de ARN transcritas inicialmente desde la misma secuencia genómica de ADN, pero que han sido objeto de un empalme alternativo de ARN. El empalme alternativo de ARN tiene lugar cuando una transcripción de ARN primario sufre un empalme, generalmente para la extracción de intrones, lo que tiene por resultado la producción de más de una molécula de ARNm, cada una de las cuales puede codificar diferentes
45 secuencias de aminoácidos. El término variante de empalme se refiere igualmente a las proteínas codificadas por las moléculas de ADNc antes mencionadas.

El término “identidad” refleja una relación entre dos o más secuencias de polipéptido o dos o más secuencias de polinucleótidos, que se determina mediante la comparación de dichas secuencias. Por lo general, la identidad se
50 refiere a una correspondencia exacta entre nucleótidos o entre aminoácidos de las dos secuencias de polinucleótidos o polipéptidos, respectivamente, a lo largo de las secuencias comparadas.

“Porcentaje de identidad” - En las secuencias en las que no existe una correspondencia exacta se puede determinar un “porcentaje de identidad”. Por lo general se alinean las dos secuencias que se desea comparar para obtener la
55 mayor correlación posible entre las secuencias. Esta operación puede incluir la inclusión de “huecos” en una o ambas secuencias con el fin de aumentar el grado de alineación. El porcentaje de identidad puede determinarse a lo largo de cada una de las secuencias comparadas (denominada alineación global), que resulta especialmente útil para secuencias de la misma longitud o de longitud muy similar, o bien a lo largo de tramos más cortos y definidos (denominada alineación local) que es más adecuada para secuencias de diferente longitud.

60 La “similitud” es otro parámetro, más sofisticado, de la relación entre dos secuencias de polipéptidos. Por lo general, la “similitud” implica una comparación entre los aminoácidos de dos cadenas de polipéptidos, residuo por residuo, en la que se tiene en cuenta no sólo la correspondencia exacta entre pares de residuos, uno de cada secuencia comparada (como en la identidad), sino también, cuando no existe una correspondencia exacta, la posibilidad de que un residuo pueda ser sustituido por el otro desde una perspectiva evolutiva. Esta posibilidad tiene un “marcador” asociado a ella, a partir del cual se puede determinar el “porcentaje de similitud” entre las dos
65 secuencias.

Los métodos para comparar la identidad y la similitud de dos o más secuencias son bien conocidos. Así por ejemplo, los programas BESTFIT y GAP de la versión 9.1 del Paquete de Análisis de Secuencias de Wisconsin [Wisconsin Sequence Analysis Package version 9.1] (Devereux J *et al*, Nucleic Acids Res, 12, 387-395, 1984) que comercializa Genetics Computer Group, Madison, Wisconsin, pueden emplearse para determinar el porcentaje de identidad entre dos polinucleótidos y el porcentaje de identidad y de similitud entre dos secuencias de polipéptidos. El programa BESTFIT utiliza el algoritmo de “homología local” de Smith y Waterman (J Mol Biol, 147,195-197, 1981, Advances in Applied Mathematics, 2, 482-489, 1981) y encuentra la región con mayor similitud entre las dos secuencias. BESTFIT es más apropiado para comparar dos secuencias de polinucleótidos o polipéptidos que tienen diferentes longitudes, ya que el programa supone que la secuencia más corta representa una parte de la más larga. En cambio, el programa GAP alinea las dos secuencias para encontrar la “similitud máxima” de acuerdo con el algoritmo de Needleman y Wunsch (J Mol Biol, 48, 443-453, 1970). El GAP es más adecuado para comparar secuencias que tienen aproximadamente la misma longitud y se espera una alineación a lo largo de toda su longitud. Los parámetros “ponderación de hueco” y “ponderación de longitud” que utiliza cada uno de estos programas serán preferiblemente 50 y 3 para las secuencias de polinucleótidos, y 12 y 14 para las secuencias de polipéptidos, respectivamente. Los porcentajes de identidad y similitud se determinan preferiblemente cuando las secuencias comparadas presentan una alineación óptima.

Se conocen otros programas para determinar la identidad y/o similitud entre secuencias, por ejemplo la familia BLAST de programas (Altschul S F *et al*, J Mol Biol, 215, 403-410, 1990, Altschul S F *et al*, Nucleic Acids Res., 25:389-3402, 1997, disponible en el Centro Nacional de Información de Biotecnología (National Center for Biotechnology Information - NCBI), Bethesda, Maryland, Estados Unidos, a la que se puede acceder a través de la página de NCBI: www.ncbi.nlm.nih.gov) y la familia de programas FASTA (Pearson W R, Methods in Enzymology, 183, 63-99, 1990; Pearson W R y Lipman D J, Proc Nat Acad Sci USA, 85, 2444-2448,1988, disponible como parte del Paquete de Análisis de Secuencias de Wisconsin).

Preferiblemente se utiliza la matriz de sustitución de aminoácidos BLOSUM62 (Henikoff S y Henikoff J G, Proc. Nat. Acad Sci. USA, 89, 10915-10919, 1992) para la comparación de secuencias de polipéptidos, incluso cuando las secuencias de nucleótidos se traducen primero en secuencias de aminoácidos antes de la comparación.

El programa BESTFIT se utiliza preferiblemente para determinar el porcentaje de identidad entre una secuencia de polinucleótidos o polipéptidos de consulta y una secuencia de polinucleótidos o polipéptidos de referencia, ya que la secuencia de consulta y la de referencia se alinean de forma óptima y los parámetros del programa se ajustan al valor predeterminado, como se describe más arriba.

El “índice de identidad” es un parámetro de la relación entre secuencias que puede utilizarse para comparar una secuencia seleccionada (de polinucleótidos o polipéptidos) y una secuencia de referencia. Así por ejemplo, una secuencia de polinucleótidos seleccionada que tenga, por ejemplo, un índice de identidad de 0,95 respecto a una secuencia de polinucleótidos de referencia, es idéntica a ésta última, excepto en que la secuencia de polinucleótidos seleccionada puede incluir en promedio hasta cinco diferencias por cada 100 nucleótidos de la secuencia de referencia. Dichas diferencias se seleccionan del grupo que haya sufrido la supresión, sustitución, incluyendo la transición o transversión, o la inserción de al menos un nucleótido. Estas diferencias pueden tener lugar en las posiciones terminales 5' o 3' de la secuencia de polinucleótidos de referencia o en cualquier lugar entre estas posiciones terminales, intercaladas ya sea de forma individual entre los nucleótidos de la secuencia de referencia o bien en uno o más grupos contiguos dentro de la secuencia de referencia. En otras palabras, para obtener una secuencia de polinucleótidos que tenga un índice de identidad de 0,95 respecto a una secuencia de polinucleótidos de referencia puede suprimirse, sustituirse, insertarse, o cualquier combinación de estas operaciones, una media de hasta 5 de cada 100 nucleótidos de la secuencia de referencia, de la forma antes descrita. Otro tanto se aplica por analogía a los demás valores del índice de identidad, por ejemplo, 0,96, 0,97, 0,98 y 0,99.

De forma similar en el caso de los polipéptidos, una secuencia candidata de polipéptidos que tenga, por ejemplo, un índice de identidad de 0,95 respecto a una secuencia de polipéptidos de referencia es idéntica a ésta última, excepto en que la secuencia de polipéptidos puede incluir en promedio hasta cinco diferencias por cada 100 aminoácidos de la secuencia de referencia. Dichas diferencias se seleccionan del grupo que haya sufrido la supresión, sustitución, ya sea conservadora o no, o la inserción de al menos un aminoácido. Estas diferencias pueden tener lugar en las posiciones terminales amino o carboxilo de la secuencia de polipéptidos de referencia o en cualquier lugar entre estas posiciones terminales, intercaladas ya sea de forma individual entre los aminoácidos de la secuencia de referencia o bien en uno o más grupos contiguos dentro de la secuencia de referencia. En otras palabras, para obtener una secuencia de polipéptidos que tenga un índice de identidad de 0,95 respecto a una secuencia de polipéptidos de referencia puede suprimirse, sustituirse, insertarse, o una combinación de estas operaciones, una media de hasta 5 de cada 100 aminoácidos de la secuencia de referencia, de la forma antes descrita. Otro tanto se aplica por analogía a los demás valores del índice de identidad, por ejemplo, 0,96, 0,97, 0,98 y 0,99.

La relación entre el número de diferencias de nucleótidos o aminoácidos y el índice de identidad puede expresarse con la siguiente ecuación:

$$n_a \leq x_a - (x_a \cdot I),$$

donde:

na es el número de diferencias de nucleótidos o aminoácidos,

5 xa es el número total de nucleótidos o aminoácidos en SEQ ID NO:1 o SEQ ID

NO:2, respectivamente,

I es el índice de identidad,

10

· es el símbolo del multiplicador, y

en la que los productos no enteros de xa e I se redondean al número entero inferior más cercano antes de restarlos de xa.

15

“Homólogo” es un término genérico que se utiliza en la técnica para indicar que una secuencia de polinucleótidos o polipéptidos posee un alto grado de relación con una secuencia de referencia. Dicha relación puede cuantificarse determinando el grado de identidad y/o similitud entre las dos secuencias, definidas anteriormente. Los términos “ortólogo” y “parálogo” forman parte de este término genérico. El término “ortólogo” se refiere a un polinucleótido o polipéptido que es el equivalente funcional del polinucleótido o polipéptido de otra especie. El término “parálogo” se refiere a un polinucleótido o polipéptido que tiene funciones similares dentro de la misma especie.

20

El término “proteína de fusión” se refiere a una proteína codificada por dos genes fusionados y no relacionados o fragmentos de los mismos. En las patentes US 5541087 y 5726044 se describen algunos ejemplos. En el caso del ligando Fc- ANIC-BP-1, el uso de una región Fc de una inmunoglobulina como parte de la proteína de fusión resulta conveniente para llevar a cabo la expresión funcional del ligando Fc- ANIC-BP-1 o de fragmentos de este ligando, para mejorar las propiedades farmacocinéticas de dicha proteína de fusión al utilizarla para fines terapéuticos y para generar un ligando ANIC-BP-1 dimérico. La estructura del ADN del ligando Fc- ANIC-BP-1 incluye, en dirección 5’ a 3’, un casete de secreción, es decir, una secuencia de señales que activa la exportación desde una célula de mamífero de un ADN que codifica un fragmento de la región Fc de una inmunoglobulina, como socio de fusión, y un ADN que codifica el ligando ANIC-BP-1 o fragmentos de éste. En algunos casos resulta conveniente poder modificar las propiedades funcionales intrínsecas (unión complementaria, unión del receptor Fc) mediante mutación de los aspectos funcionales Fc sin modificar el resto de la proteína de fusión o suprimir la parte Fc completamente después de la expresión.

35

Descripción de la figura

Figura 1

40 *Interacción del ligando ANIC-BP-1 con ANIC-BP-1 en S. cerevisiae EGY48/pSH18-32*

Se transformó *S. cerevisiae* EGY48/18-32 (Clontech) utilizando plásmidos que codifican proteínas de fusión del dominio de activación LexA y Gal4, como se indica en las ilustraciones. En la parte inferior de la figura. Se aislaron colonias individuales que se aplicaron a medios libres de histidina, triptófano y uracilo con (-WHU-Xgal) o sin (-WHU) 40 mg/l X-Gal y a un medio libre de histidina, triptófano, leucina y uracilo que contenía 40 mg/l X-Gal (-WHLU-XGal).

45

El crecimiento de todas las cepas en el medio -WHU indica la presencia de plásmidos. El color azul de las levaduras y el crecimiento en el medio libre de leucina indica la interacción de cebos y presas (carriles 5 y C), ya que la expresión de los genes indicadores depende de un esquema de selección de dos híbridos. Los carriles 1-5 y A-B se utilizan como carriles de control. El carril C indica interacción de ANIC y ANIC-BP1 en este sistema de dos híbridos de levadura.

50

Figura 2

55 Se transformó una cepa de *S. cerevisiae* EGY48-pSH18-34 mediante plásmidos que codifican las proteínas de fusión indicadas en la figura. Se comprobó la actividad del gen indicador de β -galactosidasa utilizando ONPG de la forma descrita (Yeast Protocols Handbook, Clontech).

55

Otros ejemplos

60

Estructuras de plásmidos

Clonación de pDBLeu - ANIC-BP-1

65 Se amplificó el ADNc que codifica ANIC-BP-1 mediante PCR utilizando los cebadores ANICBP-Y2H2-up (SEQ ID NO: 3, cebador 1) y ANICBP-Y2H-low (SEQ ID NO: 4, cebador 2), que se ligó al vector pCR2.1TOPO (Invitrogen). Posteriormente, el vector se partió utilizando NheI y NcoI, y el fragmento que codifica Mo25 se ligó en pDBLeu.

ES 2 271 013 T3

Clonación de pLexA-MCS - ANIC-BP-1

Se fijaron los cebadores de oligonucleótidos Y2H-MCS1 (SEQ ID NO: 7, cebador 3) y Y2H-MCS2 (SEQ ID NO: 8, cebador 4) y se ligaron a EcoRI y Sall situados en el vector pLexA (Clontech) para generar un vector pLexA-MCS que contenía lugares de restricción únicos EcoRI, Sall, XhoI y NotI.

Se aisló un fragmento que codifica ANIC-BP-1 del vector pDBLeu-Mo25 utilizando EcoRI y Sall, y se ligó al vector pLexA-MCS para generar pLexA-MCS-ANIC-BP-1. Todos los vectores se confirmaron por secuenciación.

En SEQ ID NO:5 se presenta la secuencia de péptidos de la proteína de fusión Gal4-ANIC-BP-1 (posición: 768-881) y una secuencia de proteínas ANIC-BP-1 completa unida de forma C terminal. En SEQ ID NO:6 se presenta la secuencia de péptidos correspondiente de la proteína de fusión LexA-ANIC-BP-1 que comprende la secuencia de la proteína LexA (posición: 1-202) y una secuencia de proteínas ANIC-BP-1 completa unida de forma C terminal.

15 *Rastreo de dos híbridos de levadura para seleccionar proteínas de interacción ANIC-BP-1*

Se realizó un rastreo de dos híbridos de levadura (Proquest, Life Technologies), utilizando como cebo pDBLeu-ANIC-BP-1, de la forma descrita (la selección se efectuó en un medio mínimo Trp, Leu y His que contenía 25 mM de 3-aminotriazol). La interacción se confirmó mediante el ensayo de filtros de β -galactosidasa y un medio libre de uracilo, triptófano y leucina, como se indica en el protocolo del fabricante.

Se aisló y secuenció un clon positivo. El ligando de la proteína de interacción correspondiente se denominó ligando ANIC-BP-1 y las secuencias se presentan en SEQ ID NO:1 y SEQ ID NO:2.

25 Confirmación de la interacción entre ligandos ANIC-BP-1 utilizando un esquema de selección de dos híbridos de levadura basado en LexA:

Para confirmar la interacción en un esquema de selección diferente se utilizó el sistema Matchmaker LexA Two-Hybrid (Clontech) en las condiciones indicadas por los fabricantes. El vector pLexA-MCS-ANICBP se utilizó como cebo. Se usó como presa el vector ligando pPC86-ANICBP, que fue aislado en el sistema de dos híbridos de levadura basado en Gal4, de la forma que se describe más arriba.

Se utilizaron las interacciones de EphrinB2 con PICK1 y un nuevo dominio PDZ que contenía proteína como controles positivos en la cepa de *S. cerevisiae* EGY48-pSH18-32. El crecimiento de la colonia en un medio libre de leucina y el color azul de la colonia indican la interacción de las proteínas cebo y presa.

40

45

50

55

60

65

ES 2 271 013 T3

REIVINDICACIONES

5 1. Proteína ligando ANICP-BP-1 aislada, que se une a una proteína de fusión que comprende la secuencia ANIC-BP-1 de SEQ ID 5 y/o 6 seleccionada del grupo formado por:

- (a) un polipéptido codificado por un polinucleótido que comprende la secuencia de SEQ ID NO:1;
- 10 (b) un polipéptido que comprende una secuencia de polipéptidos con una identidad de al menos el 95% con la secuencia de polipéptidos de SEQ ID NO:2;
- (c) un polipéptido que presenta una identidad de al menos el 95% con la secuencia de polipéptidos de SEQ ID NO:2;
- 15 (d) la secuencia de polipéptidos de SEQ ID NO:2.

2. Polinucleótido aislado seleccionado de uno de los grupos formados por:

- 20 (a) un polinucleótido aislado que comprende una secuencia de polinucleótidos con una identidad de al menos el 95% con la secuencia de polinucleótidos de SEQ ID NO:1;
- (b) un polinucleótido aislado que tiene una identidad de al menos un 95% con el polinucleótido de SEQ ID NO:1;
- 25 (c) un polinucleótido aislado que comprende una secuencia de polinucleótidos que codifica una secuencia de polipéptidos que tiene una identidad de al menos el 95% con la secuencia de polipéptidos SEQ ID NO:2;
- (d) un polinucleótido aislado que tiene una secuencia de polinucleótidos que codifica una secuencia de polipéptidos que tiene una identidad de al menos el 95% con la secuencia de polipéptidos SEQ ID NO:2;
- 30 (e) un polinucleótido aislado que es el ARN equivalente de un polinucleótido de (a) a (d);

o una secuencia de polinucleótidos complementaria a dicho polinucleótido aislado en toda la longitud de éste.

35 3. Sistema de expresión aislado que comprende un polinucleótido capaz de producir un polipéptido de los contemplados en las reivindicaciones 1 y 2, cuando dicho vector de expresión está presente en una célula anfitriona compatible.

40 4. Célula anfitriona recombinante no humana que comprende el vector de expresión contemplado en la reivindicación 3 o una membrana de éste que expresa el polipéptido contemplado en las reivindicaciones 1 y 2.

45

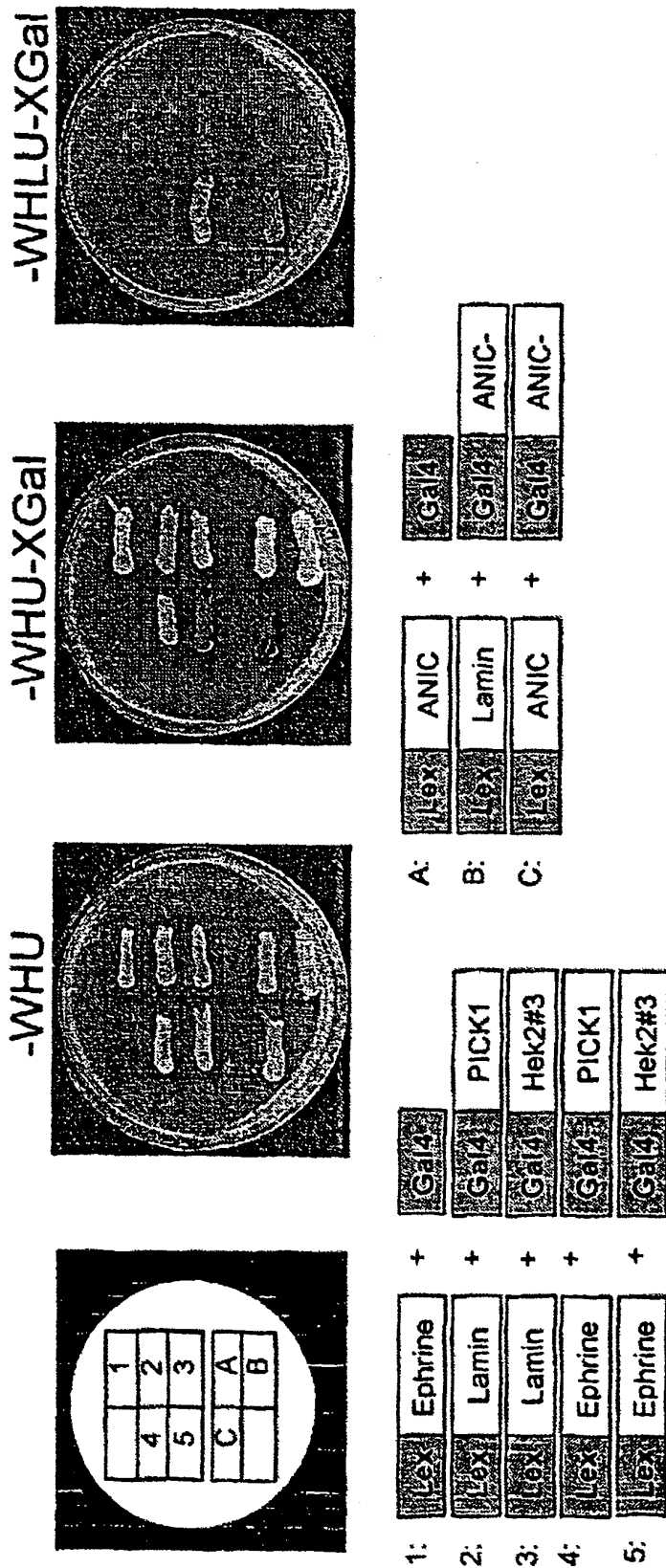
50

55

60

65

Figura 1 de 2



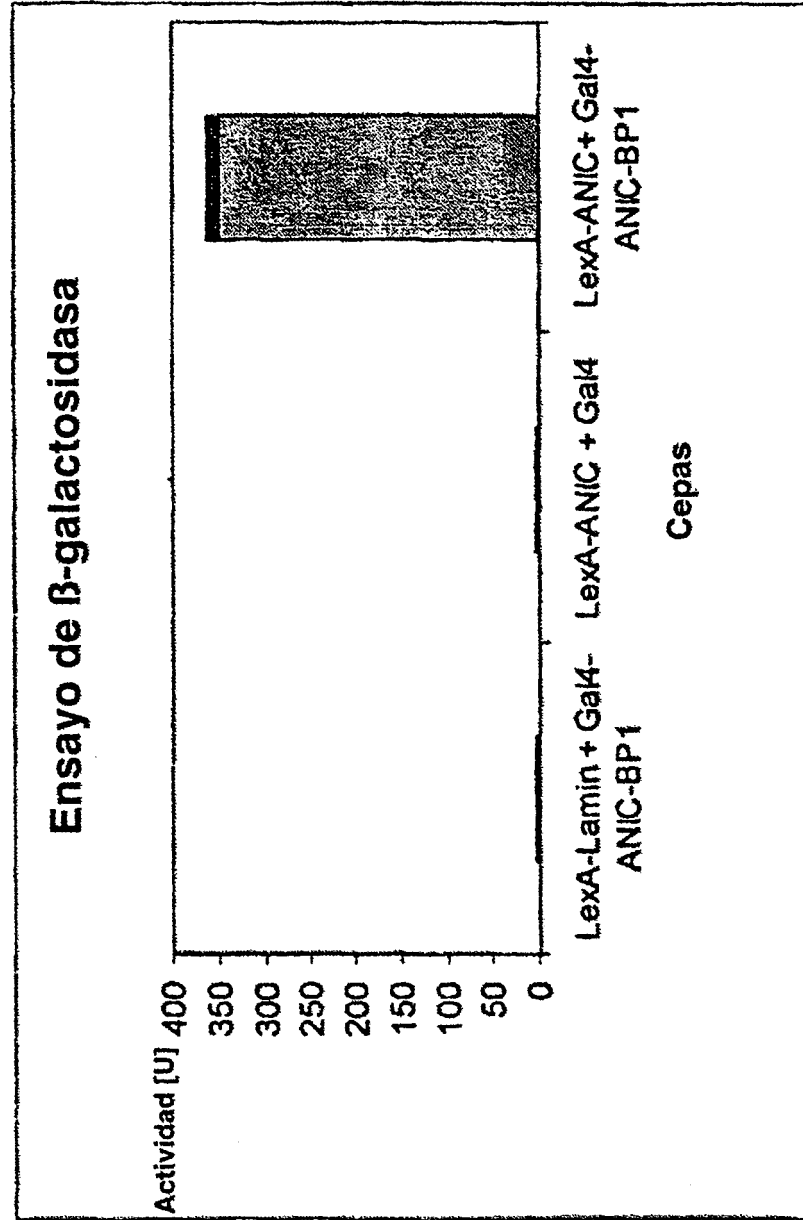


Figura 2 de 2

ES 2 271 013 T3

LISTA DE SECUENCIAS

5 <110> Merck Patent GmbH

<120> Ligando ANIC-BP1

10 <130> Ligando ANIC-BP-1

<140>

<141>

15 <160> 8

<170> PatentIn Ver. 2.1

20 <210> 1

<211> 2700

<212> ADN

25 <213> *Homo sapiens*

<220>

<221> CDS

30 <222> (363)..(2432)

<220>

<223> Descripción de proteína ligando ANIC-BP-1

35 <400> 1

```

agccatccct cgectgctcg ctctctcctt tcgccactc cctgcatctg ggectgcatc 60
acctttgcc aaccgctcccc cgatcctgcc gacactcctc ccccaaactt ctgaccggca 120
cccttgctg gtacccttct ctccattcct cccctccat cttcttctcc cgaccctct 180
cgggtccctc ttttccaaa acccgggtct ctccgcgtgg ccccgectcc aggcgggga 240
tgccccccgc ggccccgcgc ccatggctct gacgtgctt ctctccgect acaagctgtg 300
tcgcttcttc gccatgtcgg gccacggcg gggcgccgag cggctggcgg tgctgggcc 360

ag atg ggg gcg gtg gca cgg gcc cat ggt ggg ctg cgg gtg gcc cgg 407
Met Gly Ala Val Ala Arg Ala His Gly Gly Leu Arg Val Ala Arg
1 5 10 15

gcc cgc gaa agt gtc gcc ggg ggc agg cac cga ggt gca gga cgc cct 455
Ala Arg Glu Ser Val Ala Gly Gly Arg His Arg Gly Ala Gly Arg Pro
20 25 30

gga gcg cgc gct gcc gga gct gca gca ggc ctt gtc cgc gct gaa gca 503
Gly Ala Arg Ala Ala Gly Ala Ala Gly Leu Val Arg Ala Glu Ala
35 40 45

ggc ggg cgg cgc gcg ggc cgt ggg cgc cgg cct ggc cga ggt ctt cca 551
Gly Gly Arg Arg Ala Gly Arg Gly Arg Arg Pro Gly Arg Gly Leu Pro
50 55 60

act ggt gga gga ggc ctg gct gct gcg gcc gtg ggc cgc gag gta gcc 599
Thr Gly Gly Gly Gly Leu Ala Ala Ala Ala Val Gly Arg Glu Val Ala
65 70 75

cag ggt ctg tgc qac gcc atc cgc ctc gat ggc ggc ctc gac ctg ctg 647

```

ES 2 271 013 T3

	Gln	Gly	Leu	Cys	Asp	Ala	Ile	Arg	Leu	Asp	Gly	Gly	Leu	Asp	Leu	Leu	
	80					85					90					95	
5	ttg	cgg	ctg	ctg	cag	gcg	ccg	gag	ttg	gag	acg	cgt	gtg	cag	gcc	gcg	695
	Leu	Arg	Leu	Leu	Gln	Ala	Pro	Glu	Leu	Glu	Thr	Arg	Val	Gln	Ala	Ala	
					100					105					110		
10	cgc	ctg	ctg	gag	cag	atc	ctg	gtg	gct	gag	aac	cga	gac	cgc	gtg	gcg	743
	Arg	Leu	Leu	Glu	Gln	Ile	Leu	Val	Ala	Glu	Asn	Arg	Asp	Arg	Val	Ala	
				115					120					125			
15	cgc	att	ggg	ctg	ggc	gtg	atc	ctg	aac	ctg	gcg	aag	gaa	cgc	gaa	ccc	791
	Arg	Ile	Gly	Leu	Gly	Val	Ile	Leu	Asn	Leu	Ala	Lys	Glu	Arg	Glu	Pro	
			130					135					140				
20	gta	gag	ctg	gcg	cgg	agt	ggg	tca	ggc	atc	ttg	gag	cac	atg	ttc	aag	839
	Val	Glu	Leu	Ala	Arg	Ser	Gly	Ser	Gly	Ile	Leu	Glu	His	Met	Phe	Lys	
		145					150					155					
25	cat	tcg	gag	gag	aca	tgc	cag	agg	ctg	gtg	gcg	gcc	ggc	ggc	ctg	gac	887
	His	Ser	Glu	Glu	Thr	Cys	Gln	Arg	Leu	Val	Ala	Ala	Gly	Gly	Leu	Asp	
	160					165					170					175	
30	gcg	gtg	ctg	tat	tgg	tgc	cgc	cgc	acg	gac	ccc	gcg	ctg	ctg	cgc	cac	935
	Ala	Val	Leu	Tyr	Trp	Cys	Arg	Arg	Thr	Asp	Pro	Ala	Leu	Leu	Arg	His	
				180					185						190		
35	tgc	gcg	ctg	gcg	ctg	ggc	aac	tgc	gcg	ctg	cac	ggg	ggc	cag	gcg	gtg	983
	Cys	Ala	Leu	Ala	Leu	Gly	Asn	Cys	Ala	Leu	His	Gly	Gly	Gln	Ala	Val	
			195					200						205			
40	cag	cga	cgc	atg	gta	gag	aag	cgc	gca	gcc	gag	tgg	ctc	ttc	ccg	ctc	1031
	Gln	Arg	Arg	Met	Val	Glu	Lys	Arg	Ala	Ala	Glu	Trp	Leu	Phe	Pro	Leu	
			210					215					220				
45	gcc	ttc	tcc	aag	gag	gac	gag	ctg	ctt	tcg	ctg	cac	gcc	tgc	ctc	gca	1079
	Ala	Phe	Ser	Lys	Glu	Asp	Glu	Leu	Leu	Ser	Leu	His	Ala	Cys	Leu	Ala	
		225					230					235					
50	gta	gcg	gtg	ttg	gcg	act	aac	aag	gag	gtg	gag	cgc	gag	gtg	gag	cgc	1127
	Val	Ala	Val	Leu	Ala	Thr	Asn	Lys	Glu	Val	Glu	Arg	Glu	Val	Glu	Arg	
	240				245						250					255	
55	tcg	ggc	acg	ctg	gcg	ctc	gtg	gag	ccg	ctt	gtg	gcc	tcg	ctg	gac	cct	1175
	Ser	Gly	Thr	Leu	Ala	Leu	Val	Glu	Pro	Leu	Val	Ala	Ser	Leu	Asp	Pro	
				260					265						270		
60	ggc	cgc	ttc	gcc	cgc	tgt	ctg	gtg	gac	gcc	agc	gac	aca	agc	cag	ggc	1223
	Gly	Arg	Phe	Ala	Arg	Cys	Leu	Val	Asp	Ala	Ser	Asp	Thr	Ser	Gln	Gly	
			275						280					285			
65	cgc	ggg	ccc	gac	gac	ctg	cag	cgc	ctc	gtg	ccg	ttg	ctc	gac	tct	aac	1271
	Arg	Gly	Pro	Asp	Asp	Leu	Gln	Arg	Leu	Val	Pro	Leu	Leu	Asp	Ser	Asn	
			290					295					300				
70	cgc	ttg	gag	gcg	cag	tgc	atc	ggg	gct	ttc	tac	ctc	tgc	gcc	gag	gct	1319
	Arg	Leu	Glu	Ala	Gln	Cys	Ile	Gly	Ala	Phe	Tyr	Leu	Cys	Ala	Glu	Ala	
		305					310					315					
75	gcc	atc	aag	agc	ctg	caa	ggc	aag	acc	aag	gtg	ttc	agc	gac	atc	ggc	1367
	Ala	Ile	Lys	Ser	Leu	Gln	Gly	Lys	Thr	Lys	Val	Phe	Ser	Asp	Ile	Gly	

ES 2 271 013 T3

agt gtc atg ggt gcc cgc aac ttt gtg ttg gtg cta tca cct gga gca 2135
 Ser Val Met Gly Ala Arg Asn Phe Val Leu Val Leu Ser Pro Gly Ala
 580 585 590

ctg gac aag tgc atg caa gac cat gac tgc aag gat tgg gtg cat aag 2183
 Leu Asp Lys Cys Met Gln Asp His Asp Cys Lys Asp Trp Val His Lys
 595 600 605

gag att gtg act gct tta agc tgc ggc aag aac att gtg ccc atc att 2231
 Glu Ile Val Thr Ala Leu Ser Cys Gly Lys Asn Ile Val Pro Ile Ile
 610 615 620

gat ggc ttc gag tgg cct gag ccc cag gtc ctg cct gag gac atg cag 2279
 Asp Gly Phe Glu Trp Pro Glu Pro Gln Val Leu Pro Glu Asp Met Gln
 625 630 635

gct gtg ctt act ttc aac ggt atc aag tgg tcc cac gaa tac cag gag 2327
 Ala Val Leu Thr Phe Asn Gly Ile Lys Trp Ser His Glu Tyr Gln Glu
 640 645 650 655

gcc acc att gag aag atc atc cgc ttc ctg cag ggc cgc tcc tcc cgg 2375
 Ala Thr Ile Glu Lys Ile Ile Arg Phe Leu Gln Gly Arg Ser Ser Arg
 660 665 670

gac tca tct gca ggc tct gac acc agt ttg gag ggt gct gca ccc atg 2423
 Asp Ser Ser Ala Gly Ser Asp Thr Ser Leu Glu Gly Ala Ala Pro Met
 675 680 685

ggt cca acc taaccagtcc ccagttcccc agccctgctg tgacttccat 2472
 Gly Pro Thr
 690

ttccatcgtc ctttctgaag gaacagctcc tgaaaccagt ctccctgggc tgagacaacc 2532

tgggctcttc ttaggaaatg gctctcctc cccctgctcc ccaccctcat ggccccacctc 2592

caaccactt tcctcagtat ctggagaggg aagggaaagtc aggcttgggc acgggaggtt 2652

agaactcccc caggccctgc cattggggtg tctgtctccg tcatgggg 2700

45 <210> 2

<211> 690

<212> PRT

<213> *Homo sapiens*

50 <223> Descripción de proteína ligando ANIC-BP-1

<400> 2

Met Gly Ala Val Ala Arg Ala His Gly Gly Leu Arg Val Ala Arg Ala
 1 5 10 15

Arg Glu Ser Val Ala Gly Gly Arg His Arg Gly Ala Gly Arg Pro Gly
 20 25 30

Ala Arg Ala Ala Gly Ala Ala Ala Gly Leu Val Arg Ala Glu Ala Gly
 35 40 45

Gly Arg Arg Ala Gly Arg Gly Arg Arg Pro Gly Arg Gly Leu Pro Thr
 50 55 60

ES 2 271 013 T3

Gly Gly Gly Gly Leu Ala Ala Ala Ala Val Gly Arg Glu Val Ala Gln
 65 70 75 80
 5 Gly Leu Cys Asp Ala Ile Arg Leu Asp Gly Gly Leu Asp Leu Leu Leu
 85 90 95
 Arg Leu Leu Gln Ala Pro Glu Leu Glu Thr Arg Val Gln Ala Ala Arg
 100 105 110
 10 Leu Leu Glu Gln Ile Leu Val Ala Glu Asn Arg Asp Arg Val Ala Arg
 115 120 125
 15 Ile Gly Leu Gly Val Ile Leu Asn Leu Ala Lys Glu Arg Glu Pro Val
 130 135 140
 Glu Leu Ala Arg Ser Gly Ser Gly Ile Leu Glu His Met Phe Lys His
 145 150 155 160
 20 Ser Glu Glu Thr Cys Gln Arg Leu Val Ala Ala Gly Gly Leu Asp Ala
 165 170 175
 25 Val Leu Tyr Trp Cys Arg Arg Thr Asp Pro Ala Leu Leu Arg His Cys
 180 185 190
 Ala Leu Ala Leu Gly Asn Cys Ala Leu His Gly Gly Gln Ala Val Gln
 195 200 205
 30 Arg Arg Met Val Glu Lys Arg Ala Ala Glu Trp Leu Phe Pro Leu Ala
 210 215 220
 Phe Ser Lys Glu Asp Glu Leu Leu Ser Leu His Ala Cys Leu Ala Val
 225 230 235 240
 35 Ala Val Leu Ala Thr Asn Lys Glu Val Glu Arg Glu Val Glu Arg Ser
 245 250 255
 40 Gly Thr Leu Ala Leu Val Glu Pro Leu Val Ala Ser Leu Asp Pro Gly
 260 265 270
 Arg Phe Ala Arg Cys Leu Val Asp Ala Ser Asp Thr Ser Gln Gly Arg
 275 280 285
 45 Gly Pro Asp Asp Leu Gln Arg Leu Val Pro Leu Leu Asp Ser Asn Arg
 290 295 300
 50 Leu Glu Ala Gln Cys Ile Gly Ala Phe Tyr Leu Cys Ala Glu Ala Ala
 305 310 315 320
 Ile Lys Ser Leu Gln Gly Lys Thr Lys Val Phe Ser Asp Ile Gly Ala
 325 330 335
 55 Ile Gln Ser Leu Lys Arg Leu Val Ser Tyr Ser Thr Asn Gly Thr Lys
 340 345 350
 60 Ser Ala Leu Ala Lys Arg Ala Leu Arg Leu Leu Gly Glu Glu Val Pro
 355 360 365
 Arg Pro Ile Leu Pro Ser Val Pro Ser Trp Lys Glu Ala Glu Val Gln
 370 375 380
 65 Thr Trp Leu Gln Gln Ile Gly Phe Ser Lys Tyr Cys Glu Ser Phe Arg

ES 2 271 013 T3

<212> ADN

<213> Secuencia artificial

5 <220>

<223> Descripción de la secuencia artificial: cebador 1

<400> 3

10

cgatgctagc atgccgttcc cgtttggg

28

<210> 4

15 <211> 28

<212> ADN

<213> Secuencia artificial

20 <220>

<223> Descripción de la secuencia artificial: cebador 2

<400> 4

25

cgatccatgg ttaagcttct tgctgagc

28

<210> 5

<211> 498

30 <212> PRT

<213> Secuencia artificial

<220>

35

<223> Descripción de la secuencia artificial: proteína de fusión Gal4-ANIC-BP-1

<400> 5

40

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

Lys Lys Leu Lys Cys Ser Lys Glu Lys Pro Lys Cys Ala Lys Cys Leu
20 25 30

45

Lys Asn Asn Trp Glu Cys Arg Tyr Ser Pro Lys Thr Lys Arg Ser Pro
35 40 45

50

Leu Thr Arg Ala His Leu Thr Glu Val Glu Ser Arg Leu Glu Arg Leu
50 55 60

Glu Gln Leu Phe Leu Leu Ile Phe Pro Arg Glu Asp Leu Asp Met Ile
65 70 75 80

55

Leu Lys Met Asp Ser Leu Gln Asp Ile Lys Ala Leu Leu Thr Gly Leu
85 90 95

60

Phe Val Gln Asp Asn Val Asn Lys Asp Ala Val Thr Asp Arg Leu Ala
100 105 110

Ser Val Glu Thr Asp Met Pro Leu Thr Leu Arg Gln His Arg Ile Ser
115 120 125

65

Ala Thr Ser Ser Ser Glu Glu Ser Ser Asn Lys Gly Gln Arg Gln Leu
130 135 140

ES 2 271 013 T3

Thr Val Ser Ser Arg Ser Thr Pro Gly Ala Ser Met Pro Phe Pro Phe
 145 150 155 160
 5 Gly Lys Ser His Lys Ser Pro Ala Asp Ile Val Lys Asn Leu Lys Glu
 165 170 175
 Ser Met Ala Val Leu Glu Lys Gln Asp Ile Ser Asp Lys Lys Ala Glu
 180 185 190
 10 Lys Ala Thr Glu Glu Val Ser Lys Asn Leu Val Ala Met Lys Glu Ile
 195 200 205
 Leu Tyr Gly Thr Asn Glu Lys Glu Pro Gln Thr Glu Ala Val Ala Gln
 210 215 220
 Leu Ala Gln Glu Leu Tyr Asn Ser Gly Leu Leu Ser Thr Leu Val Ala
 225 230 235 240
 20 Asp Leu Gln Leu Ile Asp Phe Glu Gly Lys Lys Asp Val Ala Gln Ile
 245 250 255
 Phe Asn Asn Ile Leu Arg Arg Gln Ile Gly Thr Arg Thr Pro Thr Val
 260 265 270
 25 Glu Tyr Ile Cys Thr Gln Gln Asn Ile Leu Phe Met Leu Leu Lys Gly
 275 280 285
 Tyr Glu Ser Pro Glu Ile Ala Leu Asn Cys Gly Ile Met Leu Arg Glu
 290 295 300
 30 Cys Ile Arg His Glu Pro Leu Ala Lys Ile Ile Leu Trp Ser Glu Gln
 305 310 315 320
 35 Phe Tyr Asp Phe Phe Arg Tyr Val Glu Met Ser Thr Phe Asp Ile Ala
 325 330 335
 Ser Asp Ala Phe Ala Thr Phe Lys Asp Leu Leu Thr Arg His Lys Leu
 340 345 350
 40 Leu Ser Ala Glu Phe Leu Glu Gln His Tyr Asp Arg Phe Phe Ser Glu
 355 360 365
 Tyr Glu Lys Leu Leu His Ser Glu Asn Tyr Val Thr Lys Arg Gln Ser
 370 375 380
 45 Leu Lys Leu Leu Gly Glu Leu Leu Leu Asp Arg His Asn Phe Thr Ile
 385 390 395 400
 Met Thr Lys Tyr Ile Ser Lys Pro Glu Asn Leu Lys Leu Met Met Asn
 405 410 415
 50 Leu Leu Arg Asp Lys Ser Arg Asn Ile Gln Phe Glu Ala Phe His Val
 420 425 430
 Phe Lys Val Phe Val Ala Asn Pro Asn Lys Thr Gln Pro Ile Leu Asp
 435 440 445
 55 Ile Leu Leu Lys Asn Gln Ala Lys Leu Ile Glu Phe Leu Ser Lys Phe
 450 455 460
 60 Gln Asn Asp Arg Thr Glu Asp Glu Gln Phe Asn Asp Glu Lys Thr Tyr
 65

ES 2 271 013 T3

	465				470					475				480		
	Leu	Val	Lys	Gln	Ile	Arg	Asp	Leu	Lys	Arg	Pro	Ala	Gln	Gln	Glu	Ala
5					485					490					495	
	<210> 6															
	<211> 552															
	<212> PRT															
10	<213> Secuencia artificial															
	<220>															
15	<223> Descripción de la secuencia artificial: proteína de fusión LexA-ANIC-BP-1															
	<400> 6															
20	Met	Lys	Ala	Leu	Thr	Ala	Arg	Gln	Gln	Glu	Val	Phe	Asp	Leu	Ile	Arg
	1				5					10					15	
	Asp	His	Ile	Ser	Gln	Thr	Gly	Met	Pro	Pro	Thr	Arg	Ala	Glu	Ile	Ala
				20					25					30		
25	Gln	Arg	Leu	Gly	Phe	Arg	Ser	Pro	Asn	Ala	Ala	Glu	Glu	His	Leu	Lys
			35					40					45			
	Ala	Leu	Ala	Arg	Lys	Gly	Val	Ile	Glu	Ile	Val	Ser	Gly	Ala	Ser	Arg
30		50					55					60				
	Gly	Ile	Arg	Leu	Leu	Gln	Glu	Glu	Glu	Glu	Gly	Leu	Pro	Leu	Val	Gly
	65					70					75					80
35	Arg	Val	Ala	Ala	Gly	Glu	Pro	Leu	Leu	Ala	Gln	Gln	His	Ile	Glu	Gly
					85					90					95	
	His	Tyr	Gln	Val	Asp	Pro	Ser	Leu	Phe	Lys	Pro	Asn	Ala	Asp	Phe	Leu
40				100					105					110		
	Leu	Arg	Val	Ser	Gly	Met	Ser	Met	Lys	Asp	Ile	Gly	Ile	Met	Asp	Gly
45			115					120					125			
	Asp	Leu	Leu	Ala	Val	His	Lys	Thr	Gln	Asp	Val	Arg	Asn	Gly	Gln	Val
	130						135					140				
50	Val	Val	Ala	Arg	Ile	Asp	Asp	Glu	Val	Thr	Val	Lys	Arg	Leu	Lys	Lys
	145					150					155					160
	Gln	Gly	Asn	Lys	Val	Glu	Leu	Leu	Pro	Glu	Asn	Ser	Glu	Phe	Lys	Pro
55					165					170					175	
	Ile	Val	Val	Asp	Leu	Arg	Gln	Gln	Ser	Phe	Thr	Ile	Glu	Gly	Leu	Ala
				180					185					190		
60	Val	Gly	Val	Ile	Arg	Asn	Gly	Asp	Trp	Leu	Glu	Phe	Arg	Ser	Thr	Pro
			195					200					205			
	Gly	Ala	Ser	Met	Pro	Phe	Pro	Phe	Gly	Lys	Ser	His	Lys	Ser	Pro	Ala
65		210					215					220				

ES 2 271 013 T3

Asp Ile Val Lys Asn Leu Lys Glu Ser Met Ala Val Leu Glu Lys Gln
 225 230 235 240
 5 Asp Ile Ser Asp Lys Lys Ala Glu Lys Ala Thr Glu Glu Val Ser Lys
 245 250 255
 Asn Leu Val Ala Met Lys Glu Ile Leu Tyr Gly Thr Asn Glu Lys Glu
 260 265 270
 10 Pro Gln Thr Glu Ala Val Ala Gln Leu Ala Gln Glu Leu Tyr Asn Ser
 275 280 285
 Gly Leu Leu Ser Thr Leu Val Ala Asp Leu Gln Leu Ile Asp Phe Glu
 290 295 300
 15 Gly Lys Lys Asp Val Ala Gln Ile Phe Asn Asn Ile Leu Arg Arg Gln
 305 310 315 320
 20 Ile Gly Thr Arg Thr Pro Thr Val Glu Tyr Ile Cys Thr Gln Gln Asn
 325 330 335
 Ile Leu Phe Met Leu Leu Lys Gly Tyr Glu Ser Pro Glu Ile Ala Leu
 340 345 350
 25 Asn Cys Gly Ile Met Leu Arg Glu Cys Ile Arg His Glu Pro Leu Ala
 355 360 365
 30 Lys Ile Ile Leu Trp Ser Glu Gln Phe Tyr Asp Phe Phe Arg Tyr Val
 370 375 380
 Glu Met Ser Thr Phe Asp Ile Ala Ser Asp Ala Phe Ala Thr Phe Lys
 385 390 395 400
 35 Asp Leu Leu Thr Arg His Lys Leu Leu Ser Ala Glu Phe Leu Glu Gln
 405 410 415
 His Tyr Asp Arg Phe Phe Ser Glu Tyr Glu Lys Leu Leu His Ser Glu
 420 425 430
 40 Asn Tyr Val Thr Lys Arg Gln Ser Leu Lys Leu Leu Gly Glu Leu Leu
 435 440 445
 45 Leu Asp Arg His Asn Phe Thr Ile Met Thr Lys Tyr Ile Ser Lys Pro
 450 455 460
 Glu Asn Leu Lys Leu Met Met Asn Leu Leu Arg Asp Lys Ser Arg Asn
 465 470 475 480
 50 Ile Gln Phe Glu Ala Phe His Val Phe Lys Val Phe Val Ala Asn Pro
 485 490 495
 55 Asn Lys Thr Gln Pro Ile Leu Asp Ile Leu Leu Lys Asn Gln Ala Lys
 500 505 510
 Leu Ile Glu Phe Leu Ser Lys Phe Gln Asn Asp Arg Thr Glu Asp Glu
 515 520 525
 60 Gln Phe Asn Asp Glu Lys Thr Tyr Leu Val Lys Gln Ile Arg Asp Leu
 530 535 540
 65 Lys Arg Pro Ala Gln Gln Glu Ala
 545 550

ES 2 271 013 T3

<210> 7

<211> 29

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador 3

10

<400> 7

aattccaggt cgacctcgag gcggcgct

15

29

<210> 8

<211> 29

<212> ADN

20 <213> Secuencia artificial

<220>

<223> Descripción de la secuencia artificial: cebador 4

25

<400> 8

tcgaaccggc cgacctgagg tcgacctgg

30

29

35

40

45

50

55

60

65