



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 293 900**

51 Int. Cl.:

C12N 15/12 (2006.01)

C07K 14/435 (2006.01)

A61K 38/17 (2006.01)

G01N 33/50 (2006.01)

C12Q 1/68 (2006.01)

C07K 16/18 (2006.01)

C12N 5/20 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Número de solicitud europea: **00934824 .4**

86 Fecha de presentación : **06.06.2000**

87 Número de publicación de la solicitud: **1187916**

87 Fecha de publicación de la solicitud: **20.03.2002**

54

Título: **Identificación y caracterización molecular de proteínas, expresadas en las glándulas salivales de la garrapata.**

30

Prioridad: **09.06.1999 GB 9913425**

73

Titular/es: **Henogen S.A.**
rue des Prof. Jeener et Brachet 12
6041 Gosselies, BE

45

Fecha de publicación de la mención BOPI:
01.04.2008

72

Inventor/es: **Godfroid, Edmond;**
Bollen, Alex y
Leboulle, Gérard

45

Fecha de la publicación del folleto de la patente:
01.04.2008

74

Agente: **Curell Suñol, Marcelino**

ES 2 293 900 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

Identificación y caracterización molecular de proteínas, expresadas en las glándulas salivales de la garrapata.

5 **Campo de la invención**

La invención se refiere a la caracterización molecular de secuencias de ADN que codifican proteínas expresadas en las glándulas salivales de garrapatas, más específicamente del artrópodo garrapata *Ixodes ricinus*. Estas proteínas están implicadas en el mecanismo complejo de interacción entre este artrópodo y su mamífero hospedador. La invención se refiere a polinucleótidos recién identificados, a los polipéptidos codificados por los mismos y a la utilización de dichos polinucleótidos y polipéptidos, y a su producción.

Antecedentes de la invención

Las garrapatas son ectoparásitos que infestan un gran número de animales tales como mamíferos, aves, reptiles y anfibios. Están presentes en prácticamente cada zona del mundo. El propio proceso de alimentación de la garrapata es con frecuencia perjudicial para el hospedador. Además, muchas de estas garrapatas son vectores de virus, bacterias y protozoos que originan la morbilidad del hospedador y, en algunos casos, mortalidad, particularmente en seres humanos y ganado. Existen tres familias de garrapatas: las *Ixodidae* o garrapatas duras, las *Argasidae* o garrapatas blandas y las *Nuttalliellidae*. El ciclo vital de todas las garrapatas implica cuatro etapas (huevo - larva - ninfa - adulto). En la mayoría de las especies, como *Ixodes ricinus*, las garrapatas permanecen en el animal hospedador después de cada alimentación de sangre. Las larvas incuban los huevos y trepan a la vegetación en la que se introducen con facilidad para alcanzar a los animales que pasan. Una vez en el hospedador, le atacan y se alimentan de sangre. Las ninfas y adultos se alimentan en otros hospedadores y utilizan los mismos métodos de búsqueda del hospedador. La copulación de los adultos tiene lugar con frecuencia en el hospedador mientras están unidos y se alimentan. La puesta de huevos se produce después de cada separación.

La glándula salival de la garrapata desempeña una función importante en la realización de la alimentación de sangre y en la transmisión de patógenos. Durante la alimentación de sangre, las garrapatas *Ixodidae* concentran la sangre utilizando mecanismos especiales que eliminan el exceso de agua y de iones a través de las glándulas salivales. Una modificación sorprendente de la morfología y de la fisiología de las glándulas salivales se produce durante esta alimentación de sangre. El citoplasma y el núcleo de varias células de las glándulas salivales aumentan de volumen lo que conduce a un aumento importante en el tamaño y el peso de las glándulas salivales. Se produce también la síntesis del ARN mensajero (ARNm), dando como resultado la expresión de nuevas proteínas. Al final de la alimentación, la degeneración de las glándulas salivales es producida probablemente por la 20-hidroxicdisona, denominada también "factor de degeneración salival".

La glándula salival es rica en factores bioactivos: cemento, enzimas, inhibidores enzimáticos, agonistas y antagonistas de histamina, factores anticoagulantes, factores de modulación de la respuesta inmunitaria del hospedador, prostaglandina. Algunos de estos factores interactivos están ya presentes en las glándulas salivales de las garrapatas no alimentadas; pero otros, principalmente las proteínas, se producen durante la fase de alimentación del alimento sanguíneo. Estos factores proteínicos inducidos parece que desempeñan una función importante en la modulación de la respuesta inmunitaria del hospedador. Uno de estos factores, una proteína de 65 kDa, fue aislado por Brossard y colaboradores (Ganapamo *et al.*, 1997). Esta proteína provoca, *in vitro*, una proliferación de linfocitos CD4⁺ T específicos de las células de los ganglios linfáticos de los ratones infestados con garrapatas *I. ricinus* (Ganapamo *et al.*, 1997). Estas células producen concentraciones elevadas de interleucina-4 (IL-4) y bajas concentraciones de interferón- γ (IFN- γ) cuando se estimulan con concanavalina-A (Con A). Esto sugiere una polarización T_H2 del modelo de citocina (Ganapamo *et al.*, 1995). La producción de IL-5 e IL-10 confirma este fenómeno (Ganapamo *et al.*, 1996). Esta respuesta polarizada podría constituir condiciones favorables para la transmisión de algunos patógenos.

La inhibición de la serie de reacciones alternativa de activación del complemento, la disminución de la síntesis de anticuerpos provocada por antígenos dependientes del timo en animales infestados, y la disminución de la actividad proliferante de los linfocitos T estimulados con mitógenos, contribuyen al establecimiento de estos procesos (Wikel *et al.*, 1996; Brossard y Wikel, 1997). Además, las prostaglandinas (PGE₂) y las proteínas salivales están implicadas en la supresión de la respuesta inmunitaria. Además, es sabido que algunos factores proteicos expresados por las glándulas salivales estimulan el crecimiento de *Borrelia burgdorferi*, agente etiológico de la enfermedad de Lyme, que es el principal patógeno humano transmitido por las garrapatas *Ixodidae* (De Silva *et al.*, 1995).

NEEDHAM G. *et al.* (Experimental and Applied Acarology nº 7, pág. 21 a 32, 1989) describen la caracterización de los productos génicos de las glándulas salivales de garrapatas *Ixode* obtenidas de un banco de ADNc.

BIOR *et al.* Abdel (Faseb Journal, vol. 13, nº 7, 1999) describen un proyecto para identificar y caracterizar genes de garrapata expresados en las glándulas salivales.

LUO C. *et al.* (Insect Molecular Biology, vol. 6 (3), pág. 267-271, 1997) describen la clonación y el secuenciado de un gen para el homólogo de una enzima desaturasa y obtenida de las glándulas salivales de una garrapata.

ES 2 293 900 T3

El documento WO 95/04750 describe el polipéptido inhibidor de la proteasa procedente de la garrapata y las secuencias polinucleotídicas, un procedimiento para producir estos polipéptidos y una célula hospedadora y una composición farmacéutica que comprende dicho polipéptido.

5 Sumario de la invención

La presente invención se refiere a un polinucleótido obtenido de una glándula salival de garrapata y que presenta más del 75%, 80%, 90%, 95%, 98 a 99% o 100% de identidad con la secuencia nucleotídica SEC. ID. n°: 24 o una secuencia complementaria a ésta así como la secuencia polipeptídica correspondiente.

10

La presente invención se refiere asimismo a un vector que comprende el polinucleótido o las secuencias polipeptídicas según la invención, a la estirpe celular transfectada por este vector, a los anticuerpos producidos contra estos polipéptidos, a las estirpes celulares de hibridoma que expresan estos anticuerpos y a una composición farmacéutica que comprende estos diversos elementos según la invención.

15

Otro aspecto de la presente invención se refiere también a la utilización de los polipéptidos según la invención para la preparación de un medicamento, destinados a producir anticuerpos o a la respuesta inmunitaria de linfocitos T para proteger un mamífero de las bacterias o a los virus transmitidos durante la alimentación de sangre de *Ixodes ricinus* y las especies relacionadas en el tratamiento y/o a la prevención de la enfermedad de Lyme o de las enfermedades por virus de encefalitis de garrapata.

20

La presente invención se refiere asimismo a un kit para diagnóstico destinado a detectar una enfermedad o a la propensión a una enfermedad producida o transmitida por garrapatas, especialmente *Ixodes ricinus* que comprende las secuencias polinucleotídicas o polipeptídicas de la invención o a un fago que presenta los anticuerpos según la invención.

25

Breve descripción de los dibujos

Figura 1

30

Análisis RACE (Frohman *et al.*, 1995) específico para las SEC. ID. n° 16 y 24. La etapa de transcripción inversa se llevó a cabo utilizando 10 ng de los ARNm extraídos de la glándula salival de garrapatas saciadas. Las bandas más brillantes representan los fragmentos de ADNc correspondientes al extremo 3' del ARNm diana. Los productos ampliados se sometieron a electroforesis en gel de agarosa seguido de tinción de los fragmentos de ADN con bromuro de etidio. Los marcadores de peso molecular (M) fueron el Smart ladder (Life technologies, Rockville, Maryland, UEA). Las flechas indican la posición de los productos ampliados esperados.

35

Figura 2

40

Análisis de la expresión diferencial de los 5 ADNc completos seleccionado y de 9 fragmentos de ADNc aislados en el banco sustractivo. Se llevaron a cabo análisis por PCR utilizando como plantilla de ADN los ADNc obtenidos a partir de un procedimiento de transcripción inversa en los ARNm extraídos de las glándulas salivales, ya sea de garrapatas saciadas (E) o de garrapatas sin alimentación (UF). Estos ARN mensajeros se utilizaron también como plantilla en análisis de transcriptasa inversa. Diez microlitros de una mezcla tanto de PCR como de RT-PCR se sometieron a electroforesis en gel de agarosa y tinción con bromuro de etidio para la detección de productos ampliados de ADN. [++] muy positivo; [+] positivo; [-] negativo.

45

Descripción de la invención

50

Los autores han caracterizado genes que se producen en las glándulas salivales de *Ixodes ricinus* durante la fase de alimentación lenta de la alimentación con sangre. La clonación de estos genes se llevó a cabo creando dos bancos de ADN complementario (ADNc). El primero es un banco sustractivo basado en la metodología descrita por Lisitsyn *et al.* (*Science* 259, 946-951, 1993) y mejorada por Diatchenko *et al.* (*Proc. Natl. Acad. Sci. USA* 93, 6025-6030, 1996). Este banco clonó el ARNm producido selectivamente durante la fase de alimentación de la garrapata. El segundo banco es un banco de ADNc completo que se ha construido utilizando la propiedad básica de los ARNm (presencia de una cola poliA en su extremo 3' y de la estructura capsular en su extremo 5'). El banco de ADNc permitió la clonación de los ADNc completos, correspondiente a algunas secuencias de ADNc incompletas estimadas de interés e identificadas en el banco sustractivo de ADNc.

55

60

El banco sustractivo se creó sustrayendo los ADNc no producidos (sintetizados a partir de los ARNm expresados en las glándulas salivales de las garrapatas no alimentadas) a partir de los ADNc producidos (sintetizados a partir de los ARNm expresados en las glándulas salivales al final de la fase de alimentación lenta. Los ADNc fueron digeridos por una enzima de restricción, se dividieron en dos alícuotas y se modificaron de manera diferenciada mediante la adición de adaptadores específicos. Por lo que respecta a los ADNc producidos, los ADNc no producidos fueron digeridos también por la misma enzima de restricción y a continuación se mezclaron en exceso con cada alícuota del ADNc producido modificado. Cada mezcla de los ADNc no producidos/producidos se sometió a una etapa de desnaturalización, seguida inmediatamente de una etapa de hibridación, lo que conduce a una captura de los ADNc homólogos producidos por el ADNc no producido. Cada mezcla se mezcló a continuación y se sometió de nuevo a un

65

nuevo ciclo de desnaturalización/hibridación. Entre las moléculas de ADNc hibridadas, esta mezcla final comprende los ADN producidos con diferentes adaptadores en sus extremos 5' y 3'. Estos ADNc relevantes fueron ampliados por la reacción en cadena de la polimerasa (PCR), utilizando cebadores específicos para cada adaptador situado en cada extremo de las moléculas de ADNc. Los productos de PCR se ligaron a continuación en un vector pCRII™ mediante clonación A-T y se clonaron en una cepa TOP-10 de *E. coli*. La heterogeneidad de este banco sustractivo se evaluó secuenciando los clones recombinantes. La propiedad “inducida” de estas secuencias de ADNc se comprobó por PCR de transcripción inversa (RT-PCR) en el ARNm extraído de las glándulas salivales de garrapatas saciadas y no alimentadas. Por último, el ADNc completo producido se obtuvo mediante el cribado del banco de expresión utilizado, como sonda, algunos ADNc incompletos producidos procedentes del banco sustractivo. Estas moléculas de ADN completas producidas se secuenciaron y se compararon con las secuencias polipeptídicas y polinucleotídicas conocidas.

El banco de ADNc completo se creó utilizando la estrategia desarrollada en el “CapFinder PCR cDNA Library Construction Kit” (Clontech). Este kit para la construcción del banco utiliza el oligonucleótido CapSwitch™ exclusivo (patente en trámite) en la síntesis de la primera cadena, seguido de una ampliación de PCR de larga distancia para generar altos rendimientos de los ADNc completos, de doble cadena. Todos los procedimientos de síntesis de ADNc normalmente utilizados se basan en la capacidad de la transcriptasa inversa para transcribir el ARNm en ADN monocatenario en la reacción de la primera cadena. Sin embargo, debido a que la transcriptasa inversa no puede transcribir siempre la secuencia completa de ARNm, los extremos 5' de los genes tienden a estar infrarrepresentados en la población de ADNc. Esto es particularmente cierto para los ARNm largos, especialmente si la síntesis de la primera está cebada con cebadores oligo(dT) solamente, o si el ARNm presenta una estructura secundaria persistente. Además, la utilización de la T4 ADN polimerasa para generar extremos truncados de ADNc después de la síntesis de la segunda cadena produce generalmente extremos 5' heterogéneos que son 5 a 30 nucleótidos más cortos que el ARNm original (D'Alessio, 1988). En el método de síntesis de ADNc CapFinder, se utiliza un cebador oligo(dT) modificado para cebar la reacción de la primera cadena, y el oligonucleótido CapSwitch actúa como una plantilla corta y ampliada en el extremo 5' para la transcriptasa inversa. Cuando la transcriptasa inversa alcanza el extremo 5' del ARNm, la enzima conecta las plantillas y continúa replicando el extremo del oligonucleótido CapSwitch. Esta conexión en la mayoría de los casos tiene lugar en la estructura capsular de 7-metilguanosa, que está presente en el extremo 5' de todos los ARNm eucarióticos (Furuichi & Miura, 1975). El ADNc monocatenario completo resultante contiene el extremo 5' completo del ARNm así como la secuencia complementaria al oligonucleótido CapSwitch, que sirve a continuación como secuencia universal de cebado de PCR (anclaje CapSwitch) en la ampliación ulterior. El ADNc monocatenario anclado por CapSwitch se utiliza directamente (sin intervención de la etapa de purificación) para PCR. Únicamente aquellos ADNc monocatenarios cebados con oligo(dT) que presentan una secuencia de anclaje CapSwitch en el extremo 5' pueden servir como plantillas y ampliarse exponencialmente utilizando los cebadores de PCR 3' y 5'. En la mayoría de los casos, los ADNc incompletos y el ADNc transcrito a partir de poliA-ARN no será reconocido por el anclaje CapSwitch y por consiguiente no se ampliará.

Al final de estas reacciones, los productos de PCR ADNc completo se ligaron en el vector de clonación pCRII (Invitrogen) y se utilizaron para la transformación de la cepa XL2 de *E. coli*. El banco de ADNc completo fue cribado a continuación utilizando, como sonda, los ADNc incompletos producidos aislados del banco sustractivo.

Ochenta y nueve clones del banco sustractivo se secuenciaron al azar, y su ADN y sus secuencias traducidas del ADN y de aminoácidos se compararon con las bases de datos del ADN y proteínas. Entre éstas, se identificaron 27 secuencias de familia distintas, y 3 de ellas se seleccionaron para la caracterización posterior de su correspondiente secuencia de ARNm completo. Estas 3 secuencias eran iguales a la secuencia de i) el inhibidor de la serie de reacciones del factor tisular humano (TFPI), ii) el gen inhibidor de la trombina humana, e iii) una proteína metalopeptidasa dependiente del cinc del veneno de serpiente. Estos genes codifican proteínas que podrían estar implicadas en la inhibición de la coagulación sanguínea. Las otras 24 secuencias familiares presentaban pocas o ninguna homología con las secuencias polinucleotídicas y polipeptídicas existentes en las bases de datos. El cribado del banco de ADNc completo que utiliza sondas oligonucleotídicas específicas para los 3 clones sustractivos seleccionados anteriormente conduce a la recuperación de los ADNc completos correspondientes. El cribado al azar de este banco conduce a la selección de otros dos clones. Uno es íntimamente homólogo a una proteína similar al interferón mientras que el otro presenta homologías con la proteína relacionada con el antígeno común del leucocito de *Rattus norvegicus*.

55 Definiciones

Polipéptidos “anticoagulantes, anticomplementarios e inmunomoduladores supuestos” se refieren a los polipéptidos que presentan la secuencia de aminoácidos codificada por los genes definidos en la tabla. Éstos presentan homología con polipéptidos anticoagulantes, anticomplementarios e inmunomoduladores ya existentes en las bases de datos. Estos polipéptidos pertenecen a las secuencias de Clase I y Clase II (véase la tabla).

ADN “anticoagulantes, anticomplementarios e inmunomoduladores supuestos” se refieren a los polinucleótidos que presentan la secuencia nucleotídica definida en la tabla o a las variantes alélicas de misma y/o a sus complementos. Éstos presentan homología con polinucleótidos anticoagulantes, anticomplementarios e inmunomoduladores ya existentes en las bases de datos. Estos ADNc pertenecen a las secuencias de Clase I y Clase II (véase la tabla).

Algunas secuencias polipeptídicas o polinucleotídicas presentan pocas o ninguna homología con los polipéptidos o polinucleótidos ya existentes en las bases de datos. Estos pertenecen a la Clase III (véase la tabla).

“Polipéptido” se refiere a cualquier péptido o proteína que comprende dos o más aminoácidos unidos entre sí por enlaces peptídicos o enlaces peptídicos modificados, es decir, isoésteres peptídicos. “Polipéptido” se refiere tanto a las cadenas cortas, normalmente denominadas péptidos, oligopéptidos u oligómeros, como a las cadenas largas, generalmente denominadas proteínas. Los polipéptidos pueden contener aminoácidos distintos de los 20 aminoácidos codificados por genes. “Polipéptidos” incluyen las secuencias de aminoácidos modificadas ya sea por procesos naturales, tales como el tratamiento tras la traducción o mediante técnicas de modificación química que son bien conocidas en la técnica. Dichas modificaciones están bien descritas en los textos básicos y en monografías más detalladas, así como en una voluminosa bibliografía de investigación. Pueden producirse modificaciones en cualquier parte en un polipéptido, incluyendo el eje central peptídico, las cadenas laterales de aminoácidos y los terminales amino o carboxilo. Se valorará que el mismo tipo de modificación pueda estar presente en los mismos o variables grados en diferentes puntos en un polipéptido dado. Asimismo, un polipéptido dado puede contener muchos tipos de modificaciones. Los polipéptidos pueden estar ramificados como resultado de la ubicuidad, y pueden ser cíclicos con o sin ramificación. Los polipéptidos cíclicos, ramificados y cíclicos ramificados pueden proceder de los procesos naturales tras la traducción o pueden realizarse por procedimientos de síntesis. Las modificaciones incluyen la acetilación, acilación, ribosilación de ADP, amidación, enlace covalente de flavina, enlace covalente de un resto hem, enlace covalente de un nucleótido o de un derivado nucleotídico, enlace covalente de un lípido o derivado lipídico, enlace covalente de fosfatidilinositol, reticulación, ciclación, formación del enlace disulfuro, desmetilación, formación de reticulaciones covalentes, formación de cistina, formación de piroglutamato, formilación, gamma-carboxilación, glucosilación, formación del anclaje GPI, hidroxilación, yodación, metilación, miristoilación, oxidación, tratamiento proteolítico, fosforilación, prenilación, racemización, selenoilación, sulfonación, adición mediada por ARN de transferencia de amino de aminoácidos a proteínas tal como arginilación, y ubiquitinación. Véase, por ejemplo, PROTEINS-STRUCTURE AND MOLECULAR PROPERTIES, 2ª Ed., T. E. Creighton, W. H. Freeman y Comany, Nueva York, 1993 y Wolf, F., Posttranslational Protein Modifications: Perspectives and Prospects, págs. 1-12 en POSTTRANSLATIONAL COVALENT MODIFICATION OF PROTEINS, B. C. Johnson, Ed., Academic Press, Nueva York, 1983; Seifter *et al.*, “Analysis for protein modifications and nonprotein cofactors”, *Meth. Enzymol.* (1990) 182: 626-646 y Rattan *et al.*, “Protein Synthesis: Posttranslational Modifications and Aging”, *Ann. NY. Acad. Sci.* (1992) 663: 48-62.

“Polinucleótido” se refiere generalmente a cualquier polirribonucleótido o polidesoxirribonucleótido, que puede ser ARN o ADN no modificados o ARN o ADN modificados. Los “polinucleótidos” incluyen, sin limitación ADN mono y bicatenario, ADN que es una mezcla de zonas mono y bicatenarias, ARN mono y bicatenario y ARN que es una mezcla de zonas mono y bicatenarias, moléculas híbridas que comprenden ADN y ARN que pueden ser monocatenarios o, más típicamente, bicatenarios o una mezcla de zonas mono y bicatenarias. Además, “polinucleótido” se refiere a zonas tricatenarias que comprenden ARN o ADN o tanto a ARN como a ADN. El término Polinucleótido también incluye los ADN o los ARN que contienen una o más bases modificadas y los ADN o los ARN con ejes centrales modificados por estabilidad o por otras razones. Las bases “modificadas” incluyen, por ejemplo, bases tritiladas y bases no habituales tal como inosina. Se ha introducido una variedad de modificaciones en el ADN y ARN; de este modo, “Polinucleótido” comprende formas modificadas química, enzimática o metabólicamente de polinucleótidos tal como se encuentran por regla general en la naturaleza, así como las formas químicas de ADN y ARN características de virus y células. “Polinucleótido” también comprende polinucleótidos relativamente cortos, con frecuencia denominados oligonucleótidos.

“Variante”, como se utiliza el término en la presente memoria, es un Polinucleótido o polipéptido que se diferencia de un polinucleótido o polipéptido de referencia respectivamente, pero que conserva propiedades esenciales. Una variante típica de un polinucleótido se diferencia en la secuencia nucleotídica de otro, polinucleótido de referencia. Los cambios en la secuencia nucleotídica de la variante pueden o no alterar la secuencia de aminoácidos de un polipéptido codificado por el polinucleótido de referencia. Los cambios nucleotídicos pueden producirse en sustituciones, adiciones, eliminaciones, fusiones y truncamientos de aminoácidos en el polipéptido codificado por la secuencia de referencia, como se expone a continuación. Una variante típica de un polipéptido se diferencia en la secuencia de aminoácidos de otro polipéptido de referencia. Generalmente, las diferencias están limitadas de modo que las secuencias del polipéptido de referencia y de la variante son muy similares en general y, en muchas zonas idénticas. Una variante y un polipéptido de referencia pueden diferenciarse en la secuencia de aminoácidos en una o más sustituciones (preferentemente conservadoras), adiciones, eliminaciones en cualquier combinación. Un resto de aminoácido sustituido o insertado puede o no ser codificado por el código genético. Una variante de un polinucleótido o polipéptido puede ser el que se produce en la naturaleza tal como una variante alélica, o puede ser una variante que no se conoce que se produzca en la naturaleza. Las variantes no naturales de polinucleótidos y polipéptidos pueden prepararse por técnicas de mutagénesis o por síntesis directa. Las variantes deberían conservar una o más de las actividades biológicas del polipéptido de referencia. Por ejemplo, deberían tener actividades antigénicas o inmunogénicas similares como el polipéptido de referencia. La antigenicidad puede probarse utilizando experimentos de inmunotransferencia habituales, preferentemente utilizando sueros policlonales contra el polipéptido de referencia. La inmunogenicidad puede probarse midiendo las respuestas del anticuerpo (utilizando sueros policlonales generados contra el polipéptido variante) contra el polipéptido de referencia purificado en un ensayo ELISA normalizado. Preferentemente, una variante conservaría todas las actividades biológicas anteriores.

“Identidad” es una medida de la identidad de las secuencias nucleotídicas o de las secuencias de aminoácidos. En general, las secuencias se alinean de modo que se obtiene la igualdad de orden superior. “Identidad” tiene un

significado por sí mismo reconocido en la técnica y puede calcularse utilizando las técnicas publicadas. Véase, p. ej.: (COMPUTATIONAL MOLECULAR BIOLOGY, Lesk, A.M., ed., Oxford University Press, Nueva York, 1988; BIOCMPUTING: INFORMATICS AND GENOME PROJECTS, Smith, D.W., ed., Academic Press, Nueva York, 1993; COMPUTER ANALYSIS OF SEQUENCE DATA, PART I, Griffin, A.M., y Griffin, H.G., eds., Humana Press, Nueva Jersey, 1994; SEQUENCE ANALYSIS IN MOLECULAR BIOLOGY, von Heijne, G., Academic Press, 1987; y SEQUENCE ANALYSIS PRIMER, Gribskov, M. y Devereux, J. eds., M. Stockton Press, Nueva York, 1991). Aunque existen numerosos procedimientos para medir la identidad entre dos secuencias polinucleotídicas o polipeptídicas, el término “identidad” es bien conocido por los expertos en lamateria (Carillo, H. y Lipton, D., SIAM J. *Applied Math.* (1998) 48: 1073). Los procedimientos utilizados normalmente para determinar la identidad o similitud entre dos secuencias incluyen, pero no se limitan a, los descritos en Guide to Huge Computers, Martin J. Bishop, ed., Academic Press, San Diego, 1994, y Carillo, H., y Lipton, D., SIAM J. *Applied Math.* (1988) 48: 1073. Los procedimientos para determinar la identidad y similitud están codificados en programas informáticos. Los procedimientos de programas informáticos preferidos para determinar la identidad y similitud entre dos secuencias incluyen, pero no se limitan al paquete del programa GCG (Devereux, J., *et al.*, *J. Molec. Biol.* (1990) 215: 403). Todavía más preferentemente, el programa utilizado para determinar los niveles de identidad fue el programa GAP, tal como se utilizó en los Ejemplos siguientes.

A título ilustrativo, mediante un polinucleótido que presenta una secuencia nucleotídica que tiene por lo menos, por ejemplo, el 95% de “identidad” con una secuencia nucleotídica de referencia se pretende que la secuencia nucleotídica del polinucleótido sea idéntica a la secuencia de referencia con la excepción de que la secuencia polinucleotídica puede incluir un promedio de hasta cinco mutaciones puntuales por cada 100 nucleótidos de la secuencia nucleotídica de referencia. Es decir, para obtener un polinucleótido que presenta una secuencia nucleotídica de por lo menos el 95% idéntica a una secuencia nucleotídica de referencia, hasta el 5% de los nucleótidos en la secuencia de referencia puede eliminarse o sustituirse por otro nucleótido, o un número de nucleótidos hasta el 5%, de los nucleótidos totales en la secuencia de referencia puede insertarse en la secuencia de referencia. Estas mutaciones de la secuencia de referencia pueden producirse en las posiciones terminales 5' o 3' de la secuencia nucleotídica de referencia o en cualquier parte entre las posiciones terminales, interdispersadas individualmente entre nucleótidos en la secuencia de referencia o en uno o más grupos contiguos dentro de la secuencia de referencia.

30 *Polipéptidos de la invención*

La presente invención se refiere a proteínas (o polipéptidos) segregadas por las glándulas salivales de *I. ricinus*. Estos polipéptidos incluyen los polipéptidos codificados por los ADNc definidos en la tabla; así como los polipéptidos que comprenden la secuencia de aminoácidos codificada por los ADNc definidos en la tabla; y a los polipéptidos que comprenden la secuencia de aminoácidos que presenta por lo menos el 75% de identidad con la codificada por los ADNc definidos en la tabla por encima de su longitud completa, y preferentemente por lo menos 80% de identidad, y más preferentemente por lo menos 90% de identidad. Los que presentan del 95 al 99% son muy preferidos.

Los polipéptidos de la glándulas salivales de *I. ricinus* pueden estar en forma de proteínas “madura” o pueden formar parte de una proteína mayor tal como una proteína de fusión. Puede presentar ventajas incluir una secuencia de aminoácidos adicional que contenga secuencias secretoras o líderes, prosecretoras, secuencias que ayudan a la purificación tales como restos múltiples de histidina o una secuencia adicional para la estabilidad durante la producción recombinante.

Los fragmentos de los polipéptidos de la glándula salival de *I. ricinus* están también comprendidos en la presente invención. Un fragmento es un polipéptido que presenta una secuencia de aminoácidos que es la misma como parte, pero no toda, de la secuencia de aminoácidos de los polipéptidos de las glándulas salivales de *I. ricinus* mencionados anteriormente. Como en los polipéptidos de las glándulas salivales de *I. ricinus*, el fragmento puede estar “libre-estático” o comprimido dentro de un polipéptido mayor del que forman una parte o zona, aún más preferentemente como una zona continua individual. Ejemplos representativos de fragmentos de polipéptido de la invención, incluyen, por ejemplo, fragmentos de aproximadamente el aminoácido número 1-20, 21-40, 41-60, 61-80, 81-100 y 101 hasta el final del polipéptido. En este contexto “aproximadamente” comprende los intervalos descritos particularmente o más pequeños por varios aminoácidos 5, 4, 3, 2 ó 1 en uno de los extremos o en ambos extremos.

Los fragmentos preferidos incluyen, por ejemplo, polipéptidos de truncamiento que presentan la secuencia de aminoácidos de los polipéptidos de las glándulas salivales de *I. ricinus*, excepto para la eliminación de una serie continua de restos que incluye el terminal amino, o una serie continua de restos que incluye el terminal carboxilo y/o la zona transmembranaria o la eliminación de dos series continuas de restos, una que incluye el terminal amino y la otra que incluye el terminal carboxilo. Asimismo se prefieren los fragmentos caracterizados por atributos estructurales o funcionales tales como los fragmentos que comprenden la hélice alfa y las zonas que forman la hélice alfa, lámina beta y las zonas que forman la lámina beta, las zonas curvas y que forman curva, las zonas de enrollamiento y que forman enrollamiento, zonas hidrófilas, zonas hidrófobas, zonas alfa anfipáticas, zonas beta anfipáticas, zonas flexibles, zonas que forman superficie, zona de enlace al sustrato y zonas de índice antigénico elevado. Otros fragmentos preferidos son los fragmentos biológicamente activos. Los fragmentos biológicamente activos son los que median la actividad de la proteína de la glándula salival de *I. ricinus*, incluyendo los que tienen una actividad similar o una actividad mejorada, o con una actividad disminuida indeseable. También están incluidos los que son antigénicos o inmunógenos en el animal o en una persona.

ES 2 293 900 T3

Preferentemente, todos estos fragmentos de polipéptido conservan partes de la actividad biológica (por ejemplo antigénica o inmunógena) de los polipéptidos de las glándulas salivales de *I. ricinus*, incluyendo la actividad antigénica. Las variantes de la secuencia definida y de los fragmentos también forman parte de la presente invención. Las variantes preferidas son las que varían procedentes de las referencias mediante sustituciones de aminoácido conservadoras, es decir, las que sustituyen un resto por otro de características similares. Dichas sustituciones típicas están entre Ala, Val, Leu e Ile, entre Ser y Thr; entre los restos ácidos Asp y Glu; entre Asn y Gln, y entre los restos básicos Lys y Arg; o los restos aromáticos Phe y Tyr. Son particularmente preferidas las variantes en las que se sustituyen varios aminoácidos 5-10, 1-5 o 1-2, se eliminan o se añaden en cualquier combinación. La mayoría de las variantes preferidas son las variantes alélicas naturales del polipéptido de las glándulas salivales de *I. ricinus* presentes en las glándulas salivales de *I. ricinus*.

Los polinucleótidos de las glándulas salivales de *I. ricinus* de la invención pueden prepararse de cualquier manera adecuada. Dichos polipéptidos incluyen los polipéptidos naturales aislados, los polipéptidos producidos de manera recombinante, los polipéptidos producidos por síntesis o los polipéptidos producidos por una combinación de estos procedimientos. Los medios para preparar dichos polipéptidos son bien comprendidos en la técnica.

Polinucleótidos de la invención

Otro aspecto de la invención se refiere a los ADNc (polinucleótidos) de las glándulas salivales de *I. ricinus*. Éstos incluyen los polinucleótidos aislados que codifican los polipéptidos y fragmentos de las glándulas salivales de *I. ricinus* respectivamente, y los polinucleótidos estrechamente relacionados con éstos. Más específicamente, los ADNc de las glándulas salivales de *I. ricinus* de la invención incluyen un polinucleótido que comprende la secuencia nucleotídica de los ADNc definida en la tabla, que codifica un polipéptido de las glándulas salivales de *I. ricinus*. Los ADNc de las glándulas salivales de *I. ricinus* incluyen además una secuencia polinucleotídica que presenta por lo menos el 75% de identidad sobre su longitud completa con una secuencia nucleotídica que codifica el polipéptido de las glándulas salivales de *I. ricinus* codificado por los ADNc definidos en la tabla, y un polinucleótido que comprende una secuencia nucleotídica que es por lo menos 75% idéntica a la de los ADNc definidos en la tabla. A este respecto, los polinucleótidos por lo menos 80% idénticos son particularmente preferidos, y los que presenta por lo menos el 90% son especialmente preferidos. Además, los que presentan por lo menos el 95% son muy preferidos y los que presentan por lo menos del 98 al 99% son los más preferidos, siendo por lo menos el 99% los más preferidos. Asimismo están incluidos en los ADNc de las glándulas salivales de *I. ricinus* una secuencia nucleotídica que tiene suficiente identidad con una secuencia nucleotídica del ADNc definido en la tabla para hibridarse en condiciones útiles para la ampliación o para su utilización como sonda o marcador. La invención proporciona también polinucleótidos que son complementarios con dichos ADNc de las glándulas salivales de *I. ricinus*.

La secuencia nucleotídica que codifica el polipéptido de la glándula salival de *I. ricinus* codificado por los ADNc definidos en tabla puede ser idéntica a la secuencia que codifica el polipéptido contenida en los genes definidos en la tabla, o puede ser una secuencia, que como resultado de la redundancia (degeneración) del código genético, codifica asimismo el polipéptido codificado por los genes definidos en la tabla respectivamente.

Cuando se utilizan los polinucleótidos de la invención para la producción recombinante de un polipéptido de las glándulas salivales de *I. ricinus*, el polinucleótido puede incluir la secuencia de codificación del polipéptido maduro o de un fragmento de la misma, por sí mismo; la secuencia de codificación del polipéptido maduro o fragmento en el marco de lectura con otras secuencias de codificación, tales como las que codifican una secuencia principal o secretora, una secuencia de pre-, pro- o de preproteína, u otras partes del péptido de fusión. Por ejemplo, puede codificarse una secuencia marcadora que facilita la purificación de este polipéptido fusionado. En determinadas formas de realización preferidas de este aspecto de la invención, la secuencia marcadora es un péptido de hexa-histidina como el proporcionado en el vector pQE (Qiagen, Inc.) y descrita en Gentz *et al.*, *Proc. Natl. Acad. Sci. USA* (1989) 86:821-824, o en una etiqueta HA, o es la glutatión-s-transferasa. El polinucleótido puede contener también secuencias no codificadoras 5' y 3', tales como las secuencias transcritas y no traducidas, señales de corte y empalme y de poliadenilación, secuencias de fijación del ribosoma y secuencias que estabilizan el ARNm.

Las formas de realización más preferidas son las variantes de la proteína de glándulas salivales de *I. ricinus* que codifican polinucleótidos que comprenden las secuencias de aminoácidos del polipéptido de las glándulas salivales de *I. ricinus* codificado por los ADNc definidos por la tabla respectivamente en las que varios restos de aminoácidos 10-25, 5-10, 1-5, 1-3, 1-2 ó 1 están sustituidos, eliminados o añadidos, en cualquier combinación. Los polinucleótidos variantes más preferidos son las secuencias de *I. ricinus* naturales que codifican variantes alélicas de las proteínas de las glándulas salivales de *I. ricinus* en *I. ricinus*.

La presente invención se refiere además a polinucleótidos que se hibridan a las secuencias descritas anteriormente en la presente memoria. A este respecto, la presente invención se refiere especialmente a polinucleótidos que se hibridan en condiciones severas con los polinucleótidos descritos anteriormente en la presente memoria. Tal como se utiliza en la presente memoria, la expresión "condiciones severas" significa que la hibridación se producirá solamente si existe por lo menos el 80%, y preferentemente por lo menos el 90%, y más preferentemente por lo menos el 95%, incluso aún más preferentemente del 97 al 99% de identidad entre las secuencias.

Los polinucleótidos de la invención que son idénticos o suficientemente idénticos a una secuencia nucleotídica de cualquier gen definida en la tabla o a un fragmento de la misma, pueden utilizarse como sondas de hibridación

para los clones de ADNc que codifican los polipéptidos de la glándula salival de *I. ricinus* respectivamente y para aislar los clones del ADNc de otros genes (incluyendo los homólogos y ortólogos que codifican los ADNc de otras especies aparte de *I. ricinus*) que presentan una gran similitud de secuencia con los ADNc de la glándula salival de *I. ricinus*. Dichas técnicas de hibridación son conocidas por los expertos en la materia. Por regla general estas secuencias nucleotídicas son 80% idénticas, preferentemente 90% idénticas, más preferentemente 95% idénticas a las de la referencia. Las sondas generalmente comprenderán por lo menos 15 nucleótidos preferentemente, dichas sondas tendrán por lo menos 30 nucleótidos y podrán tener por lo menos 50 nucleótidos. Particularmente las sondas preferidas estarán comprendidas entre 30 y 50 nucleótidos. En una forma de realización, para obtener un polinucleótido que codifica el polipéptido de las glándulas salivales de *I. ricinus*, incluyendo los homólogos y ortólogos de otras especies aparte de *I. ricinus*, comprende las etapas de cribado de un banco apropiado en condiciones de hibridación severas con una sonda marcada que presenta una secuencia nucleotídica contenida en una de las secuencias génicas definidas por la tabla, o un fragmento de la misma; y aislar los clones del ADNc completo que contienen dicha secuencia polinucleotídica. Por lo tanto en otro aspecto, los polinucleótidos de las glándulas salivales de *I. ricinus* de la presente invención incluyen además una secuencia nucleotídica que comprende una secuencia nucleotídica que se hibrida en condiciones severas a una secuencia nucleotídica que presenta una secuencia nucleotídica contenida en los ADNc definidos en la tabla, o un fragmento de la misma. Asimismo están incluidos con los polipéptidos de las glándulas salivales de *I. ricinus* los polipéptidos que comprenden secuencias de aminoácidos codificadas por las secuencias nucleotídicas obtenidas en las condiciones de hibridación anteriores. Dichas técnicas de hibridación son bien conocidas por los expertos en la materia. Las condiciones de hibridación severas son tales como las definidas anteriormente o, alternativamente, en condiciones de incubación durante la noche a 42°C en una solución que comprende: formamida al 50%, 5× SSC (NaCl 150 mM, citrato trisódico 15 mM), fosfato sódico 50 mM (pH 7,6), 5× solución de Denhardt, sulfato de dextrano al 10% y 20 microgramos/ml desnaturalizado, ADN de esperma de salmón despojado, seguido del lavado de los filtros en 0,1×SSC a aproximadamente 65°C.

Los polinucleótidos y polipéptidos de la presente invención pueden utilizarse como reactivos de investigación y materiales para el descubrimiento de los tratamientos y diagnósticos para enfermedades animales y humanas.

Ensayos de diagnóstico

La presente invención se refiere también a la utilización de los polipéptidos de las glándulas salivales de *I. ricinus* o a los polinucleótidos de las glándulas salivales de *I. ricinus*, para su utilización como reactivos de diagnóstico.

Los materiales para diagnóstico pueden obtenerse a partir de las células del paciente, tales como de la sangre, orina, saliva, biopsia tisular.

Por lo tanto en otro aspecto, la presente invención se refiere a un kit para diagnóstico para una enfermedad o propensión a una enfermedad que comprende:

(a) un polinucleótido de las glándulas salivales de *I. ricinus*, preferentemente la secuencia nucleotídica de una de las secuencias génicas definidas en la tabla o un fragmento de las mismas;

(b) una secuencia nucleotídica complementaria con la de (a);

(c) un polipéptido de las glándulas salivales de *I. ricinus*, preferentemente el polipéptido codificado por una de las secuencias génicas definidas en la tabla o un fragmento de las mismas;

(d) un anticuerpo contra un polipéptido de las glándulas salivales de *I. ricinus*, preferentemente el polipéptido codificado por una de las secuencias génicas definidas en la tabla; o

(e) un fago que presenta un anticuerpo contra un polipéptido de las glándulas salivales de *I. ricinus*, preferentemente el polipéptido codificado por una de las secuencias de los ADNc definidas en la tabla.

Se valorará que en cualquiera de dichos kits, (a), (b), (c), (d) o (e) puede comprender un componente básico.

Los anticuerpos del polipéptido de las glándulas salivales anti-*I. ricinus* pueden comprender anticuerpos policlonales. Los métodos de preparación de anticuerpos policlonales son conocidos por los expertos en la materia. Los anticuerpos policlonales pueden producirse en un mamífero por ejemplo, mediante una o más inyecciones de un agente inmunizante y, si se desea, un adyuvante. Por regla general, el agente inmunizante y/o el adyuvante se inyectarán en el mamífero mediante múltiples inyecciones subcutáneas o intraperitoneales. El agente inmunizante puede incluir el polipéptido de las glándulas salivales de *I. ricinus* o una proteína de fusión del mismo. Puede ser útil conjugar el agente inmunizante con una proteína conocida que es inmunógena en el mamífero que se está inmunizando. Ejemplos de dichas proteínas inmunógenas incluyen pero no se limitan a la hemocianina de la lapa californiana, la albúmina del suero, la tiroglobulina y el inhibidor de tripsina de la soja. Ejemplos de adyuvantes que pueden emplearse incluyen el adyuvante completo de Freund y el adyuvante MPL TDM. El protocolo de inmunización puede ser seleccionado por un experto en la materia sin excesiva experimentación.

Los anticuerpos contra el polipéptido de las glándulas salivales anti-*I. ricinus* pueden ser, alternativamente, anticuerpos monoclonales. Los anticuerpos monoclonales pueden prepararse utilizando procedimientos de hibridoma,

ES 2 293 900 T3

tales como los descritos por Kohler y Milstein, *Nature*, 256:495 (1975). En un procedimiento de hibridoma, un ratón, hámster u otro animal hospedador apropiado, se inmuniza por regla general con un agente inmunizante para producir linfocitos que producen o son capaces de producir anticuerpos que se unirán específicamente al agente inmunizante. Alternativamente, los linfocitos pueden ser inmunizados *in vitro*.

5

El agente inmunizante incluirá por regla general el polipéptido de las glándulas salivales de *I. ricinus* o una proteína de fusión del mismo. Generalmente, se utilizan linfocitos de la sangre periférica ("PBL") si se desean células de origen humano o se utilizan células de bazo o células de ganglio linfático si se desean fuentes de mamífero no humano. Los linfocitos se fusionan a continuación con una estirpe celular inmortalizada utilizando un agente de fusión adecuado, tal como polietilenglicol, para formar una célula de hibridoma (Goding, *Monoclonal Antibodies: Principles and Practice* Academia Press, (1986) págs. 59-103). Las estirpes celulares inmortalizadas son habitualmente células de mamífero transformadas, específicamente células de mieloma de roedor, bovino y de origen humano. Normalmente, se emplean estirpes celulares de mieloma de rata o ratón. Las células de hibridoma pueden cultivarse en un medio de cultivo adecuado que contiene preferentemente una o más sustancias que inhiben el crecimiento o la supervivencia de las células inmortalizadas, no fusionadas. Por ejemplo, si las células precursoras carecen de la enzima hipoxantina guanina fosforribosil transferasa (HGPRT o HPRT), el medio de cultivo para los hibridomas por regla general incluirá hipoxantina, aminopterina y timidina ("medio HAT"), cuyas sustancias impiden el crecimiento de células insuficientes en HGPRT.

10

15

20

25

Las estirpes celulares inmortalizadas preferidas son las que se fusionan de manera eficaz, soportan la expresión de alto nivel estable del anticuerpo por las células productoras de anticuerpo seleccionado y son sensibles a un medio tal como el medio HAT. Las estirpes celulares inmortalizadas más preferidas son las estirpes de mieloma murino, que pueden obtenerse, por ejemplo, en el Salk Institute Cell Distribution Center, San Diego, California y la American Type Culture Collection, Rockville, Maryland. Las estirpes celulares de mieloma humano y de heteromieloma de ratón-humano han sido descritas también para la producción de anticuerpos monoclonales humanos (Kozbor, *J. Immunol.*, 133:3001 (1984); Brodeur *et al.*, *Monoclonal Antibody Production Techniques and Applications* Marcel Dekker, Inc., Nueva York, (1987) págs. 51-63).

30

35

En el medio de cultivo en el que se cultivan las células de hibridoma puede analizarse a continuación la presencia de anticuerpos monoclonales dirigidos contra el polipéptido de las glándulas salivales de *I. ricinus*. Preferentemente, la especificidad de unión de los anticuerpos monoclonales producidos por las células de hibridoma se determina por inmunoprecipitación o por un ensayo de unión *in vitro*, tal como radioinmunoanálisis (RIA) o el ensayo inmunoabsorbente con enzima ligada (ELISA). Dichas técnicas y análisis son conocidos en la técnica. La afinidad de unión o del anticuerpo monoclonal puede determinarse, por ejemplo, por el análisis de Scatchard de Munson y Pollard, *Anal. Biochem.*, 107:220 (1980).

40

Una vez las células de hibridoma deseadas están identificadas, pueden subclonarse limitando los procedimientos de dilución y crecimiento por procedimientos normalizados (Goding, anteriormente). Los medios de cultivo adecuados para este fin incluyen, por ejemplo, Medio de Eagle Modificado por Dulbecco y medio RPMI-1640. Alternativamente las células de hibridoma pueden cultivarse *in vivo* como ascitis en un mamífero.

45

Los anticuerpos monoclonales segregados por los subclones pueden aislarse o purificarse del medio de cultivo o del fluido de ascitis por procedimientos convencionales de purificación de inmunoglobulina tal como, por ejemplo, proteína A-Sepharose, cromatografía con hidroxipatito, electroforesis en gel, diálisis o cromatografía por afinidad.

50

55

60

Los anticuerpos monoclonales pueden prepararse asimismo por procedimientos de ADN recombinante, tales como los descritos en la patente U.S. n° 4.816.567. El ADN que codifica los anticuerpos monoclonales de la invención puede aislarse fácilmente y secuenciarse utilizando procedimientos convencionales (p. ej., utilizando sondas de oligonucleótido que son capaces de unirse específicamente a genes que codifican las cadenas pesada y ligera de anticuerpos murinos). Las células de hibridoma de la invención sirven como fuente preferida de dicho ADN. Una vez aislado, el ADN puede colocarse en vectores de expresión, que se transfieren a continuación a células hospedadoras tales como las células COS de simio, las células de ovario de hámster chino (CHO) o células de mieloma que no producen de otro modo la proteína inmunoglobulina, para obtener la síntesis de anticuerpos monoclonales en las células hospedadoras recombinantes. El ADN puede modificarse también, por ejemplo, sustituyendo la secuencia de codificación por los dominios constantes de las cadenas pesada y ligera en lugar de las secuencias murinas homólogas (patente U.S. n° 4.816.567; Morrison *et al.*, anteriormente) o por enlace covalente con la secuencia de codificación de inmunoglobulina toda o parte de la secuencia de codificación para un polipéptido distinto de inmunoglobulina. Dicho polipéptido distinto de inmunoglobulina puede ser sustituido por los dominios constantes de un anticuerpo de la invención, o puede ser sustituido por los dominios variables de una secuencia de combinación con el antígeno de un anticuerpo de la invención para crear un anticuerpo bivalente híbrido.

65

Los anticuerpos pueden ser anticuerpos monovalentes. Los procedimientos para preparar anticuerpos monovalentes son bien conocidos en la técnica. Por ejemplo, un procedimiento implica la expresión recombinante de la cadena ligera de inmunoglobulina y de la cadena pesada modificada. La cadena pesada está truncada generalmente en cualquier punto en la zona Fc de modo que impide la reticulación de la cadena pesada. Alternativamente, los restos de cisteína pertinentes se sustituyen con otros restos de aminoácidos o se eliminan con el fin de impedir la reticulación.

ES 2 293 900 T3

Los procedimientos *in vitro* son también adecuados para preparar anticuerpos monovalentes. La digestión de anticuerpos para producir fragmentos de los mismos, particularmente, fragmentos Fab, puede llevarse a cabo utilizando técnicas de rutina conocidas en la técnica.

5 Como alternativa o suplemento para inmunizar un mamífero con un péptido, puede obtenerse un anticuerpo específico para una proteína obtenida a partir de un banco producido de manera recombinante de los dominios variables de la inmunoglobulina expresada, p. ej. utilizando el bacteriófago lambda o el bacteriófago filamentoso que presentan dominios de uniones de inmunoglobulina funcional en sus superficies; por ejemplo, véase el documento WO 92/01047. El banco puede ser natural, es decir construido de las secuencias obtenidas a partir de un organismo que no ha sido
10 inmunizado con ninguna de las proteínas (o fragmentos) o puede ser el construido utilizando secuencias obtenidas a partir de un organismo que ha sido expuesto al antígeno de interés.

Vacunas

15 Otro aspecto de la invención se refiere a un procedimiento para provocar una respuesta inmunológica en un mamífero que comprende inocular al mamífero con el polipéptido de las glándulas salivales de *I. ricinus* o fragmentos portadores del epítipo, análogos, vesículas de la membrana externa o células (atenuadas o de otro modo), adecuados para producir el anticuerpo y/o la respuesta inmunitaria a los linfocitos T para proteger dicho animal de las bacterias y virus que podrían transmitirse durante la alimentación de sangre de *I. ricinus* y especies relacionadas. En particular la invención se refiere a la utilización de los polipéptidos de las glándulas salivales de *I. ricinus* codificados por los ADNc definidos en la tabla. Aún otro aspecto de la invención se refiere a un procedimiento de producción de la respuesta inmunológica en un mamífero que comprende, administrar el polipéptido de las glándulas salivales de *I. ricinus* mediante un vector que dirige la expresión del polinucleótido de las glándulas salivales de *I. ricinus in vivo* con objeto de producir dicha respuesta inmunológica para producir el anticuerpo que proteja a dicho animal de las enfermedades (enfermedad de Lyme, enfermedad del virus de la encefalitis de la garrapata,...).

Otro aspecto de la invención se refiere a una composición inmunológica o formulación de vacuna que, cuando se introduce en un hospedador animal, produce una respuesta inmunológica en este mamífero contra un polipéptido de las glándulas salivales de *I. ricinus* en el que la composición comprende un ADNc de la glándula salival de *I. ricinus* o u polipéptido de la glándula salival de *I. ricinus* o fragmentos que llevan el epítipo, análogos, vesículas de la membrana externa o células (atenuadas o de otro modo), la formulación de la vacuna puede comprender además un vehículo adecuado. La composición de la vacuna con polipéptido de las glándulas salivales de *I. ricinus* se administra preferentemente por vía oral o parenteral (incluyendo la inyección subcutánea, intramuscular, intravenosa, intradérmica, etc.). Las formulaciones adecuadas para la administración parenteral incluyen las soluciones de inyección esterilizada acuosas y no acuosas que pueden contener antioxidantes, tampones, bacterioestáticos y solutos que hacen la formulación resulte isotónica con la sangre del receptor; y suspensiones esterilizadas acuosas y no acuosas que pueden incluir agentes de suspensión o agentes de espesamiento. Las formulaciones pueden presentarse en recipientes para dosis unitaria o multidosis, por ejemplo, ampollas y viales sellados y pueden almacenarse en un estado liofilizado que requiere solamente la adición del vehículo líquido esterilizado inmediatamente antes de su uso. La formulación de la vacuna puede incluir también sistemas adyuvantes para potenciar la inmunogenicidad para la formulación, tales como los sistemas aceite en agua y otros sistemas conocidos en la técnica. La dosis dependerá de la actividad específica de la vacuna y puede ser determinada fácilmente por experimentación de rutina.

45 Todavía otro aspecto se refiere a una formulación inmunológica de vacuna que comprende el polinucleótido de la invención. Dichas técnicas son conocidas en la técnica, véase por ejemplo Wolff *et al.*, *Sciences*, (1990) 247: 1465-8.

Agentes terapéuticos

50 Otro aspecto de la invención se refiere a la utilización de estos polipéptidos de las glándulas salivales de *I. ricinus* como agente terapéutico. Al considerar las áreas terapéuticas potenciales específicas para los probables productos, las disciplinas hospitalarias cubiertas por estos productos son: hematología (particularmente estudios clínicos de coagulación), trasplante (para control de la inmunodepresión), reumatología (para antiinflamatorios) y tratamiento general (para anestésicos específicos o mejorados).
55

60

65

ES 2 293 900 T3

TABLA 1

Secuencias identificadas en los bancos sustractivo y de ADNc completo

Gen	Motivos	Secuencias similares en las bases de datos	Puntuación	Clase
Sec. 1		Sin identidad significativa		III
Sec. 2		Sin identidad significativa		III
Sec. 3		Sin identidad significativa		III
Sec. 4		Sin identidad significativa		III
Sec. 5	Punto de unión de lípido de lipoproteína de membrana procariótica	Sin identidad significativa		III
Sec. 6		Fijación de nitrógeno por <i>R. melioli</i> (fi×F)	0,00089	III
		Apolipoproteína B-100 humana	0,0045	
		ARNm hu. para prot. de unión al elemento de respuesta AMPc (CRE-BP1)	0,057	III
				III
Sec. 7	Familia Kunitz del inhibidor de serina proteasa	Clon GS345D13 de BAC humano Inhibidor PI-2 de la serie de reacciones del factor tisular de <i>H. sapiens</i>	4,7 ¹³ 4 ¹²	I I
Sec. 8	Secuencia de unión del lípido de lipoproteína de membrana procariótica	Sin identidad significativa		III
Sec. 9		ARNm de guisante para la prot. de unión a GTP	0,48	III
Sec. 10		Sin identidad significativa		III
Sec. 11		Gen R-beta de IL-11	0,18	III
Sec. 12		Sin identidad significativa		III
Sec. 13		Gen de cutinasa de <i>C. gloeosporioides</i>	0,082	III
Sec. 14		Sin identidad significativa		III
Sec. 15		ARNm de ratón para prot. Secretora que cont. motivos de transpondina	0,014	III
Sec. 16	Familia de metalopeptidasa dependiente de cinc	ARNm de <i>B. jararaca</i> para jararhagina Precursor de la metaloproteínasa de <i>Agkistrodon contortrix</i>	1,1 ⁵ 3,9 ⁵	I I

ES 2 293 900 T3

TABLA 1 (continuación)

Gen	Motivos	Secuencias similares en las bases de datos	Puntuación	Clase
Sec. 17		Gen O. aries para INF-alfa ovina	0,7	II
		Interferón- omega 45	0,88	II
		Interferón-omega 20	0,89	II
		RCPT PGE2	0,85	III
		PGE Rcpt EP2	0,85	III
Sec. 18		Sin identidad significativa		III
Sec. 19		Cadena L de IgG1 dirigida contra IL2 humana recept. de prot. Tac	0,19	II
		Zona var. de la cadena ligera de MAK447/179	0,2	II
Sec. 20		Sin identidad significativa		III
Sec. 21		Sin identidad significativa		III
Sec. 22		Neuroactina Mus del músculo	0,42	III
Sec. 23		Sin identidad significativa		III
Sec. 24		Inhibidor de trombina de <i>H. sapiens</i>	2,1 ⁻¹²	I
		Prot. serina intracelular de 38 kDa antiproteinasa citoplásmica	2,3 ⁻¹²	I
Sec. 25		Sin identidad significativa		
Sec. 26		Sin identidad significativa		
Sec. 27		Factor ELF3 (fasta) de transcripción del músculo Mus	0,053	III
Sec. 28		ARNm de la proteína (SM15) relacionado con el supuesto interferón de <i>Homo sapiens</i>	1,70E-22	II
Sec. 29		ARNm de <i>R. norvegicus</i> para la proteína relacionada con el antígeno común de leucocitos	4,80E.09	II
Clase I: homólogos supuestos de anticoagulante; Clase II: homólogos supuestos de inmunomoduladores; Clase III: pocas o ninguna homología encontradas en las bases de datos.				

ES 2 293 900 T3

TABLA 2

Características biológicas de los clones seleccionados

Las secuencias completas de los clones S seleccionados se compararon con las bases de datos de EMBL/GenBank utilizando los algoritmos tFasta y Blastp. En estas secuencias se analizó la presencia de una secuencia de consenso Kozak indicada en la columna titulada "nucleótido en posición -3". Aplicando el algoritmo del programa GCG* se analizó la presencia de las unidades de proteína específica o motivos (motivo de algoritmo) y la presencia de una secuencia peptídica señal (análisis de von Heijne y McGeoch) en las secuencias amino deducidas.

Clon	Similitud de las secuencias completas con las bases de datos	Puntuaciones Fasta/Blastp) ^a	ORF (aa)	Motivos	Puntuaciones del péptido señal ^b	Longitud Sp/prob.	Nucleótido en la posición 3 ^c
Sec. 28	Gen (SKMc15) relacionado con el interferón supuesto de <i>Homo sapiens</i> [U09585]	1,8×10 ⁻³⁶ / 1,10 ⁻⁷¹	426		d 5,4/F ^b	48 aa / 8,4×10 ⁻¹	G
Sec. 29	ARNm (LAR) del antígeno común al leucocito de <i>R. norvegicus</i> [X83546]	7,8×10 ⁻¹¹ /N	274		10,2/S	18 aa/7,4×10 ⁻⁷	A
Sec. 16	ARNm de ratón para la proteína secretora que contiene motivos de trombospondina [D67076]	0,002/6×10 ⁻⁷	489	Metalopeptidasa	7,9/S	19 aa/7,4×10 ⁻⁴	G
Sec. 24	ARNm del inhibidor de elastasa de leucocitos de cerdo [P80229]	0/7×10 ⁻⁶⁷	378	Serpina	8,5/S	51 aa/3,28×10 ⁻³ A	
Sec. 7	Inhibidor de la serie de reacciones del factor tisular humano [P48307]	4,8×10 ⁻¹² / 2×10 ⁻³	87	Kunitz	6,5/S	19 aa/4,8×10 ⁻⁴	G
^a Sin puntuación (N) ^b Logrados (S) y fallidos (F) ^c Guanina (G) y Adenina (A) ^d análisis de von Heijne ^e análisis de McGeoch * GCG = Genetic Computer Group, Madison Wisconsin, versión 10.0 en Unix, enero de 1999.							

Ejemplos

Materiales biológicos utilizados en este estudio

Las glándulas salivales de garrapatas *Ixodes ricinus* adultas hembra patógenas de 5 días saciadas o no alimentadas exentas de patógenos se utilizaron en este trabajo. Al extirparlas, estas glándulas se congelaron inmediatamente en nitrógeno líquido y se almacenaron a -80°C. Para extraer los ARN mensajeros (ARNm), se machacaron las glándulas salivales en nitrógeno líquido utilizando un mortero y una mano de mortero. Los ARNm se purificaron utilizando una oligo-dT celulosa (kit Fast Track 2.0, Invitrogen, Groningen, Holanda). Se extrajeron dos microgramos de los ARNm procedentes de 200 glándulas salivales de garrapatas alimentadas; y 1,5 µg de los ARNm se extrajeron también de 1.000 glándulas salivales de garrapatas sin alimentar.

ES 2 293 900 T3

Ejemplo 1

Construcción de un banco sustractivo por análisis diferencial de la representación (RDA)

5 Se realizaron todos los procedimientos descritos por Hubank y Schatz (1994). Se sintetizaron los ADNc de doble cadena utilizando el Superscript Choice System (Life Technologies, Rockville, Maryland, EUA). Los ADNc se digirieron con enzima de restricción DpnII, se ligaron a enzimas R, se ampliaron con cebadores R-24 (Hubank y Schatz, 1994) y por último se digirieron de nuevo con la misma enzima para generar un grupo “probador” de los ADNc procedente de las glándulas salivales de las garrapatas alimentadas y un grupo “conductor” constituido por los
10 ADNc de las glándulas salivales de las garrapatas sin alimentar. La primera ronda del procedimiento de hibridación sustractivo utilizó una relación probador/conductor de 1:100. Las segunda y tercera rondas utilizaron unas relaciones de 1:400 y 1:200.000, respectivamente. Después de tres ciclos de sustracción y ampliación, se subdividieron los productos diferenciados digeridos con Dpn-II según el tamaño en 4 fracciones diferentes en un gel de agarosa de electroforesis al 1,7% y se subclonaron en la secuencia BamHI del vector de clonación pTZ19r. El producto ligado se
15 utilizó para transformar las células competentes TOP-10 de *E. coli* (Invitrogen, Groningen, Holanda). Nueve mil seiscientos clones de este banco sustractivo se seleccionaron al azar y se colocaron individualmente en 100 microplacas y se almacenaron a -80°C. Este banco sustractivo se analizó secuenciando 89 clones seleccionados al azar, utilizando cebadores M13 directos e inversos específicos para una zona situada en el vector de clonación pTI9r. Se compararon las secuencias de ADN de estos 89 clones y se identificaron 27 secuencias familiares distintas. La homología de estas secuencias con las secuencias existentes en las bases de datos se presenta en la Tabla 1. Las secuencias sustractivas 1 a 27 se presentan en el archivo de listado de secuencias (excepto para las secuencias 16 y 24 cuyas secuencias de ARNm completo se presentan; véase también el Ejemplo 2). Se seleccionaron estas secuencias (Sec. 7, 16 y 24) para la caracterización adicional de su correspondiente secuencia del ARNm completo. Estas 3 secuencias eran iguales a la secuencia de i) el inhibidor de la serie de reacciones del factor tisular humano (TFPI), ii) una proteína metalopeptidasa dependiente del cinc del veneno de serpiente y iii) la proteína del inhibidor de trombina humana, correspondiente a las
25 Sec. 7, 16 y 24 respectivamente. Estos genes codifican las proteínas que podrían estar implicadas en la inhibición de la coagulación sanguínea.

30 Ejemplo 2

Construcción del banco de ADNc completo y recuperación de las secuencias de los ADNc completos por cribado de este banco del ADNc completo

35 Este banco se creó utilizando los ARNm extraídos de las glándulas salivales de garrapatas saciadas. Se sometieron los ARNm (80 ng) a transcripción inversa utilizando un cebador oligo-dT degenerado (5'-A(T)30VN-3'), el oligonucleótido Smart™ (Clontech, Palo Alto, EUA) y la transcriptasa inversa Superscript II (Life Technologies, Rockville, Maryland, EUA). Se utilizó la mezcla de ADNc monocatenario como plantilla en un análisis de PCR con inicio en caliente incluyendo la LA Taq polimerasa (Takara, Shiga, Japón), el cebador oligo-dT modificado y un cebador 3'-
40 “Smart” específico para una zona situada en el extremo 5' del oligonucleótido Smart™. El protocolo de PCR aplicado fue: 1 min. a 95°C, seguido de 25 s. a 95°C/5 min. a 68°C, 25 veces y 10 min. a 72°C. La mezcla del ADNc ampliado de doble cadena se purificó con un concentrador Centricon 30 (Millipore, Bedford, EUA). Los ADNc se dividieron en 4 fracciones comprendidas entre 0,3 y 0,6 kb, 0,6 a 1 kb, 1 kb a 2 kb y 2 kb a 4 kb en un gel de electroforesis en agarosa de alta calidad al 0,8% y se recuperaron por separado utilizando el kit de extracción Qiaex II (Qiagen, Hilden, Alemania). Las 4 fracciones se ligaron individualmente en el vector de clonación pCRII incluido en el kit de clonación TOPO (Invitrogen, Groningen, Holanda). Las fracciones ligadas se utilizaron a continuación para transformar las células XL2-Blue ultracompetentes de *E. coli* (Stratagene, Heidelberg, Alemania). Los clones recombinantes resultantes se almacenaron individualmente en microplacas a -80°C. Se seleccionaron al azar diez clones para secuenciado parcial o completo. Como resultado de este procedimiento, se seleccionaron 2 secuencias de ADNc (Sec. 28 y Sec. 29, véase
50 la Tabla 1) por su homología con las bases de datos de secuencias. Una es íntimamente homóloga con una proteína de tipo interferón (Sec. 28), mientras que la otra presenta homología con la proteína relacionada con el antígeno común al leucocito de *Rattus norvegicus* (Sec. 29).

Las 4 fracciones diferentes del banco de ADNc completo se cribaron con sondas de oligonucleótido marcadas
55 con radio específicas para los clones seleccionados identificados en el banco de ADNc sustractivo. El marcado de estas oligosondas se llevó a cabo como se describe en Current Protocols in Molecular Biology (Ausubel *et al.*, 1995, J. Wiley and sons, eds.). Estas 4 fracciones diferentes se colocaron a continuación en membranas de nitrocelulosa y se cultivaron toda la noche a 37°C. Estas membranas se desnaturalizaron en NaOH 0,2 M/NaCl 1,5 M, se neutralizaron en Tris 0,5 M pH 7,5-NaCl 1,5 M y se fijaron en 2× SSC (NaCl 0,3 M/ácido cítrico trisódico dihidratado 0,03 M). Las membranas se calentaron durante 90 min. a 80°C, se incubaron en una solución de prehibridación (SSC 6×, 10× Denhardt, SDS al 0,1%) a 55°C durante 90 min. y finalmente se colocaron durante la noche en una solución de hibridación precalentada que contenía una sonda oligonucleotídica marcada con radio específica a 55°C. Las membranas hibridadas se lavaron 3 veces en solución SSC 6× a 55°C durante 10 min., se secaron y se expusieron a una película Kodak X-OMAT durante la noche a -80°C. El banco de ADNc completo se analizó también secuenciando
65 una serie de clones. Las secuencias de ADN resultantes se compararon con las bases de datos ENBL/Gene Bank y se utilizaron para crear sondas oligonucleotídicas destinadas a recuperar otros clones correspondientes. De esta manera, se confirmó la secuencia de ARNm de consenso completa de las Sec. 28 y 29 mediante la recuperación de otros dos clones correspondientes a estas secuencias. Únicamente se aisló un clon de ADNc completo correspondiente al clon

ES 2 293 900 T3

16 sustractivo. Por consiguiente, para identificar la secuencia completa de las Sec. 16 y 24, se aplicó el procedimiento de los extremos del ADNc (RACE).

5 Se llevó a cabo la metodología RACE como describe Frohman *et al.* (1995). La etapa de transcripción inversa se realizó utilizando 10 ng de los ARNm extraídos de las glándulas salivales de garrapatas saciadas y la transcriptasa inversa Thermoscript (Life technologies, Rockville, Maryland, EUA). Todos los cebadores específicos para el gen (GSP) presentaron una longitud de 18 bases con una relación G/C del 61%. Los productos ampliados se sometieron a una electroforesis en gel de agarosa y se recuperaron utilizando un procedimiento de isotacoforesis. Los ADNc se clonaron en un vector de clonación pCRII-TOPO (Invitrogen, Groningen, Holanda). Para identificar la secuencia de
10 ADNc de consenso, se secuenciaron diferentes clones y su secuencia se comparó con su correspondiente secuencia conocida. Por consiguiente, las secuencias del ADNc completo de los clones 16 y 24 aisladas en el banco sustractivo se obtuvieron por este procedimiento RACE (figura 1).

Ejemplo 3

15

Análisis de las secuencias completas de 5 clones seleccionados

Las secuencias de los clones seleccionados (Sec. 7, 16, 24, 28 y 29) permitieron la identificación de los marcos de lectura abierta, a partir de los cuales se dedujeron las secuencias amino. Estos productos de traducción potencial tienen un tamaño comprendido entre 87 y 489 aminoácidos (véase la tabla 2). Con objeto de evaluar, mediante simulación informática, sus propiedades respectivas, las secuencias de aminoácidos y las secuencias nucleotídicas de dichos 5 marcos abiertos se compararon con las bases de datos utilizando los algoritmos tFasta y Blastp. Con estas comparaciones se demuestra que la Sec. 7 es muy homóloga con el inhibidor de la serie de reacciones del factor tisular humano (TFPI). El TFPI es un inhibidor de las serina proteasas que presenta 3 dominios del tipo del inhibidor de proteasa de Kunitz (KPI). Cada una de estas unidades o motivos presenta una afinidad específica para diferentes tipos de proteasas. El primer y segundo dominios de KPI son sensibles a la inhibición respectiva de los factores de coagulación VIIa y Xa. El tercer dominio de KPI aparentemente no presenta actividad inhibidora. Debe observarse que la secuencia codificada por el clon de la Sec. 7 es homóloga con la región del primer dominio KPI de TFPI y que el KPI se mantiene perfectamente en éste. Esta similitud sugiere que la proteína de la Sec. 7 es un inhibidor potencial del factor VIIa.
30

La secuencia amino deducida del clon de la Sec. 28 presenta una gran homología con 3 secuencias de la base de datos, a saber: proteína TIS7 de ratón, proteína PC4 de rata y proteína SKMc 15 humana. Estas 3 proteínas se describen como factores supuestos de tipo interferón. Poseen zonas muy bien conservadas de la proteína B2 del interferón. Por consiguiente, se propone que la proteína de la Sec. 28 presenta actividad inmunomoduladora.
35

Se compararon las secuencias 16 y 24 con las bases de datos presentando de este modo su homología respectivamente con la metalopeptidasa M12b de *Gloydius halys* (suborden de ofidios) y el inhibidor de elastasa porcina que pertenece a la superfamilia de los inhibidores de serina proteasa (Serpin). Las secuencias amino de estos 2 clones también presentan unidades específicas de dichas familias. Se propone que dichas proteínas tengan propiedades anticoagulantes e inmunomoduladoras.
40

Por último, el clon de la Sec. 29 presenta una homología débil con el antígeno común al leucocito de *R. norvegicus* (LAR). Esta última es una molécula de adhesión. Por lo tanto es posible que la proteína de la Sec. 29 presente propiedades inmunomoduladoras relacionadas con las expresadas por la proteína LAR.
45

Debido a sus propiedades potenciales, es de esperar que la mayoría de las proteínas examinadas sean segregadas en la saliva de garrapata durante la alimentación de sangre. Por consiguiente, se hicieron pruebas para hallar la presencia de un péptido señal al comienzo de las secuencia amino deducidas. Todos los resultados por el procedimiento de análisis de von Heijne fueron positivos. Las secuencias del péptido señal se detectaron por el procedimiento de McGeoch para las secuencias amino deducidas de las Sec. 7, 16, 24 y 29. Parece que dichas proteínas se segregan en la glándula salival de la garrapata. Además, se observó la presencia de una secuencia de consenso Kozak corriente arriba de las secuencias de codificación de todos los clones examinados. Esto indica que sus ARNm podrían traducirse potencialmente a proteínas.
50
55

Ejemplo 4

Evaluación de la expresión diferencial de los clones de ADNc aislados en los bancos de ADNc sustractivo y completo

60 Se evaluó la expresión diferenciada de los ARNm correspondiente a los 5 clones completos seleccionados (Sec. 7, 16, 24, 28 y 29) y de 9 clones sustractivos utilizando análisis de PCR y de RT-PCR (figura 2).

Los análisis de PCR se realizaron utilizando como plantilla de ADN los ADNc obtenidos de un procedimiento de transcripción inversa en los ARNm extraídos de las glándulas salivales de garrapatas saciadas o no alimentadas. Cada uno de los análisis de PCR incluía un par de cebadores específicos para cada diana sustractiva o la secuencia completa de los ADNc. Se realizaron análisis de PCR en un volumen final de 50 μ l que contenía cebadores 1 μ M, desoxinucleótido 0,2 mM (ATPd, GTPd, GTPd y TTPd; Boehringer Mannheim GmbH, Mannheim, Alemania), tampón de PCR (Tris-HCl 10 mM, KCl 50 mM, MgCl₂ 2,5 mM, pH 8,3) y 2,5 U de Taq ADN polimerasa (Boehringer Mannheim
65

ES 2 293 900 T3

GmbH, Mannheim, Alemania). Se ampliaron las muestras de ADN durante 35 ciclos en las condiciones siguientes: 94°C durante 1 min., 72°C durante 1 min. y 64°C durante 1 min., seguido de una etapa de alargamiento final de 72°C durante 7 min.

5 Se realizó el análisis RT-PCR en los clones de ADNc completo seleccionados y en 5 clones sustractivos de ADNc. Los ARNm utilizados como plantilla en el análisis de transcripción inversa se extrajeron de las glándulas salivales de garrapatas *I. ricinus* saciadas y no alimentadas. Los análisis de transcripción inversa se realizaron utilizando un cebador específico (esta diana uno las secuencias seleccionadas) y la “Thermoscript Reverse transcriptase” (Life technologies, Rockville, Maryland, EUA) a 60°C durante 50 min. Cada análisis de PCR utilizó el cebador específico de transcripción
10 inversa y otro cebador específico. Se realizaron análisis de PCR en un volumen final de 50 µl que contenía cebadores 1 µM, desoxinucleótido 0,2 mM (ATPd, CTPd, GTPd y TTPd; Boehringer Mannheim GmbH, Mannheim, Alemania), tampón de PCR (Tris-HCl 10 mM, KCl 50 mM, MgCl₂ 2,5 mM, pH 8,3) y 2,5 U de Expand High Fidelity polymerase (Roche, Bruselas, Bélgica). Las muestras de ADN monocatenario se ampliaron durante 30 ciclos en las siguientes condiciones: 95°C durante 1 min., 72°C durante 30 s. y 60°C durante 1 min., seguido de una etapa de alargamiento
15 final de 72°C durante 7 min. La figura 2 demuestra que la expresión de las secuencias seleccionadas se produce en las glándulas salivales de garrapatas saciadas de 5 días, excepto para la secuencia 28 que se expresa a nivel similar en las glándulas salivales de garrapatas saciadas y no alimentadas. La expresión de los demás ARNm podría producirse específicamente o aumentarse durante la alimentación de sangre.

20 Ejemplo 5

Expresión de proteínas recombinantes en células de mamífero

El estudio de las propiedades de las secuencias aisladas implica la expresión de las mismas en sistemas de expresión
25 que permiten grandes cantidades de proteínas codificadas por estas secuencias para ser producidas y purificadas. Estos experimentos se describen a continuación así como en el ejemplo 6.

5.1. Subclonación de las secuencias en el vector p CDNA3.1-His/V5 (Invitrogen)

30 Las secuencias de ADN de los 5 clones seleccionados (Sec. 7, 16, 24, 28 y 29) se transfirieron al vector de expresión p CDNA3.1-His/V5. Dicho vector permite la expresión de las proteínas heterólogas fusionadas a una cola de 6 histidinas así como al epítipo V5 en células eucarióticas. Se produjeron diferentes ADN por RT-PCR utilizando cebadores específicos para los clones correspondientes. Estos cebadores se construyeron con el fin de eliminar el codón de terminación de cada marco o fase de lectura abierto con objeto de permitir que la proteína se fusione a la
35 cola 6×HIS/epítipo V5. Además, los cebadores contenían secuencias de restricción adaptadas a la clonación en el vector de expresión. Se tuvo cuidado en utilizar, al amplificar, una polimerasa Pfu de alta fidelidad (Promega).

Se midió la expresión transitoria de las proteínas recombinantes Sec. 16 y 24 después de la transfección de las construcciones Sec. 16 y Sec. 24-pCDNA3-1-His/V5 en células COSI, utilizando Fugen 6 (Boehringer). Se analizaron
40 los extractos de proteína del medio de cultivo correspondientes a los periodos de 24, 48 y 72 horas tras la transfección en gel de acrilamida tiñendo con azul Coomassie o por transferencia Western utilizando por una parte un anticuerpo anti-6× histidina o por otra parte lechos de quelato de níquel acoplados a fosfatasa alcalina. Estos análisis presentaban la expresión de dichas proteínas en el medio de cultivo celular.

45 Ejemplo 6

Expresión de las proteínas en E. coli

6.1. Inserción de secuencias de codificación en el vector de expresión pMAL-C2E

50 Se decidió también expresar varias proteínas en bacterias utilizando el vector pMAL-C2E (NEB). Dicho vector expresa varias proteínas fusionadas con la proteína de unión a la maltosa (MBP). La proteína de interés podría separarse de la MBP entonces gracias a una secuencia que separa la MBP de la proteína, siendo dicha secuencia específica para la proteasa enterocinasas.

55 Con objeto de expresar de manera óptima las 5 secuencias examinadas, que utilizan el vector pMAL-C2E, se construyeron los pares de cebador de PCR complementarios con las 20 bases situadas corriente arriba del codón de terminación y con 20 bases situadas corriente abajo del ATG del marco de lectura abierto o de la fase. De esta manera, los fragmentos de CDNA ampliados comprendían solamente la secuencia de codificación del ARNm diana con su
60 codón de terminación. La proteína de interés se fusionó a MBP por su extremo terminal N. Por otra parte, ya que estos cebadores contenían secuencias de restricción específicas para la expresión del vector fue posible efectuar la clonación directa de los ADNc. La utilización de Pfu ADN polimerasa (Promega) hizo posible ampliar los ADNc sin tener que temer por los errores introducidos en las secuencias ampliadas.

65 Las secuencias de codificación de los clones Sec. 7, 16, 24 y 29 se reconstruyeron de esta manera. Las células TG1 competentes de *E. coli* se transfirieron utilizando estas construcciones. Las digestiones enzimáticas de estas minipreparaciones del ADN plasmídico hizo posible comprobar que la mayoría de los clones Sec. 7, 16, 24 y 29-p-MALC2-E efectivamente fueron recombinantes.

ES 2 293 900 T3

6.2. Expresión de proteínas recombinantes

Partiendo de varias construcciones clonadas en TGI de *E. coli*, se inició el estudio de la expresión de proteínas recombinantes fusionadas con MBP para todas las secuencias de interés (es decir, las Sec. 7, 16, 24 y 29) excepto para la Sec. 28. Se emprendió el cultivo de clones representativos de las Sec. 7, 16, 24 y 29 así como de referencias negativa (plásmidos no recombinantes) para producir la expresión de proteínas recombinantes en éstas. Estos cultivos se centrifugaron y los sedimentos (fondos) se separaron del medio para ser puestos en suspensión en Tris 15 mM pH 7,5 y se pasaron a la prensa francesa. Los análisis de estas muestras por gel de acrilamida (10%) coloreado con azul de Coomassie o por transferencia Western utilizando anticuerpos anti-MBP de conejo, presentó la expresión de proteínas recombinantes Sec. 7 (~50 kDa), Sec. 16 (~92 kDa), Sec. 24 (~80 kDa) y Sec. 29 (~67 kDa).

Ejemplo 7

15 Producción de anticuerpos

Las proteínas de las Sec. 7, 16 y 24 se inyectaron en grupos de 4 ratones con objeto de producir anticuerpos dirigidos contra dichas proteínas. El primer antígeno inyectado se realizó con el adyuvante completo de Freund. Dos semanas después, se puso una inyección de recuerdo con adyuvante incompleto de Freund. Los sueros de los ratones inyectados con Sec. 16 proporcionaron pruebas positivas para los anticuerpos anti-MBP.

Bibliografía

25 **Ganapamo et al, 1997:** Identification of an *Ixodes ricinus* salivary gland fraction through its ability to stimulate CD4 T cells present in BALB/c mice lymph nodes draining the tick fixation site. Ganapamo-F; Rutti-B; Brossard-M. *Parasitology*; 1997 julio; 115 (Pt 1): 91-6.

30 **Ganapamo et al, 1995:** *In vitro* production of interleukin-4 and interferon-gamma by lymph node cells from BALB/c mice infested with nymphal *Ixodes ricinus* ticks. Ganapamo-F; Rutti-B; Brossard-M. *Immunology*; 1995 mayo; 85(1):120-4.

35 **Ganapamo et al, 1996:** Immunosuppression and cytokine production in mice infested with *Ixodes ricinus* ticks: a possible role of laminin and interleukin-10 on the *in vitro* responsiveness of lymphocytes to mitogens. Ganapamo-F; Rutti-B; Brossard-M. *Immunology*; 1996 febrero; 87 (2): 259-63.

Wikel et al. 1996: Host immunity to ticks. Wikel-SK. *Annu-Rev-Entomol.*; 1996; 41: 1-22

40 **Wikel y Brossard, 1997:** Immunology of interactions between ticks and hosts. Brossard-M; Wikel-SK. *Med-Vet-Entomol.*; 1997 julio; 11(3): 270-6.

De **Silva et al, 1995:** Growth and migration of *Borrelia burgdorferi* in *Ixodes* ticks during blood feeding. Aravinda M De Silva, Erol Fikrig. *Am. J. Trop. Med. Hyg.*; 53(4), 1995 páginas 397-4043.

45 **Hubank y Schatz, 1994:** Identifying differences in mRNA expression by representational difference analysis of cDNA. Hubank-M; Schatz-DG. *Nucleic-Acids-Res.*; 1994 diciembre 25; 22(25): 5640-8.

50 **Frohman, 1995:** Rapid amplification of cDNA Ends. In PCR Primer. A laboratory manual (Dieffenbach, C. W. y Dveksler, G. S., eds), páginas 381-409, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

55

60

65

ES 2 293 900 T3

REIVINDICACIONES

- 5 1. Polinucleótido obtenido de una glándula salival de garrapata y que presenta más del 75% de identidad con la secuencia nucleotídica SEC. ID. nº: 24 o una secuencia complementaria a la misma.
2. Polinucleótido según la reivindicación 1, que presenta por lo menos el 80% de identidad con la secuencia nucleotídica SEC. ID. nº: 24 o una secuencia complementaria a la misma.
- 10 3. Polinucleótido según la reivindicación 1 ó 2, que es por lo menos 90% idéntico a la secuencia nucleotídica SEC. ID. nº: 24 o una secuencia complementaria a la misma.
4. Polinucleótido según cualquiera de las reivindicaciones anteriores 1 a 3, que es por lo menos 95% idéntico a la secuencia nucleotídica SEC. ID. nº: 24 o una secuencia complementaria a la misma.
- 15 5. Polinucleótido según cualquiera de las reivindicaciones anteriores 1 a 4, que es por lo menos 98 a 99% idéntico a la secuencia nucleotídica SEC. ID. nº: 24 o una secuencia complementaria a la misma.
- 20 6. Polinucleótido según cualquiera de las reivindicaciones anteriores 1 a 5, que es por lo menos 99% idéntico a la secuencia nucleotídica SEC. ID. nº: 24 o una secuencia complementaria a la misma.
7. Polinucleótido que presenta la secuencia nucleotídica SEC. ID. nº: 24 o una secuencia complementaria a la misma.
- 25 8. Polipéptido de la glándula salival de garrapata codificado por el polinucleótido según cualquiera de las reivindicaciones anteriores 1 a 7.
9. Polipéptido según la reivindicación 8, modificado por o ligado a por lo menos un grupo de sustitución, seleccionado preferentemente de entre el grupo constituido por grupos amida, acetilo, fosforilo y/o glucosilo.
- 30 10. Polopéptido según la reivindicación 8 ó 9, en forma de proteína “madura”.
11. Polopéptido según la reivindicación 8 ó 9, como parte de una proteína mayor.
- 35 12. Polopéptido según la reivindicación 8 ó 9, como parte de una proteína de fusión.
13. Polinucleótido según cualquiera de las reivindicaciones 8 a 12 anteriores, que comprende además por lo menos una secuencia adicional de aminoácidos que contiene secuencias secretoras o líderes, prosequencias, secuencias que ayudan a la purificación tales como los restos de histidina múltiples, o secuencias adicionales para la estabilidad durante la protección de recombinación.
- 40 14. Vector que comprende por lo menos un elemento seleccionado de entre el grupo constituido por el polinucleótido según cualquiera de las reivindicaciones anteriores 1 a 7 o el polipéptido según cualquiera de las reivindicaciones anteriores 8 a 13.
- 45 15. Estirpe celular transfectada por el vector según la reivindicación 14.
16. Anticuerpo producido contra el polipéptido de *Ixodes ricinus* según cualquiera de las reivindicaciones anteriores 8 a 13.
- 50 17. Estirpe celular de hibridoma que expresa el anticuerpo según la reivindicación 16.
18. Composición farmacéutica que comprende un vehículo farmacéutico adecuado y un elemento seleccionado de entre el grupo constituido por el polinucleótido según cualquiera de las reivindicaciones anteriores 1 a 7 o el polipéptido según cualquiera de las reivindicaciones anteriores 8 a 13.
- 55 19. Composición farmacéutica según la reivindicación 18, que presenta propiedades inmunomoduladoras.
20. Composición inmrológica o vacuna para provocar una respuesta inmunológica en un mamífero hospedador a un polipéptido de la glándula salival de garrapata que comprende por lo menos un elemento de entre el grupo constituido por:
 - a) un ADN de glándula salival de garrapata según cualquiera de las reivindicaciones anteriores 1 a 7,
 - 65 b) un polipéptido de glándula salival de garrapata según cualquiera de las reivindicaciones anteriores 8 a 13,
 - c) fragmentos que llevan epítipo, vesículas de membrana externa o células (atenuadas o de otro modo) de los componentes a) o b); y

ES 2 293 900 T3

d) posiblemente un vehículo.

21. Utilización de la composición inmúnológica, una vacuna según la reivindicación 20 o el vector según la reivindicación 14, para la preparación de un medicamento para producir un anticuerpo y/o la respuesta inmunitaria por linfocitos T para proteger a un mamífero de las bacterias y virus transmitidos durante la alimentación con sangre de *Ixodes ricinus* y de especies relacionadas y en el tratamiento o la prevención de enfermedades de Lyme o de enfermedades por el virus de la encefalitis de la garrapata.

22. Kit de diagnóstico para detectar una enfermedad o propensión a una enfermedad producida o transmitida por garrapatas, especialmente *Ixodes ricinus*, que comprende:

a) un polinucleótido de la glándula salival de *Ixodes ricinus*, según cualquiera de las reivindicaciones anteriores 1 a 7,

b) una secuencia nucleotídica complementaria a la de (a),

c) un polipéptido de la glándula salival de *Ixodes ricinus*, según cualquiera de las reivindicaciones anteriores 8 a 13,

d) el anticuerpo según la reivindicación 16,

e) un fago que presenta el anticuerpo según la reivindicación 16.

23. Polipéptido que es idéntico a la secuencia nucleotídica según cualquiera de las reivindicaciones anteriores 1 a 7, para su utilización como sonda de hibridación para los clones de ADNc que codifican un polipéptido de la glándula salival de garrapatas, más específicamente de *Ixodes ricinus* o para aislar clones u otros genes similares a los ADNc de la glándula salival de garrapata.

30

35

40

45

50

55

60

65

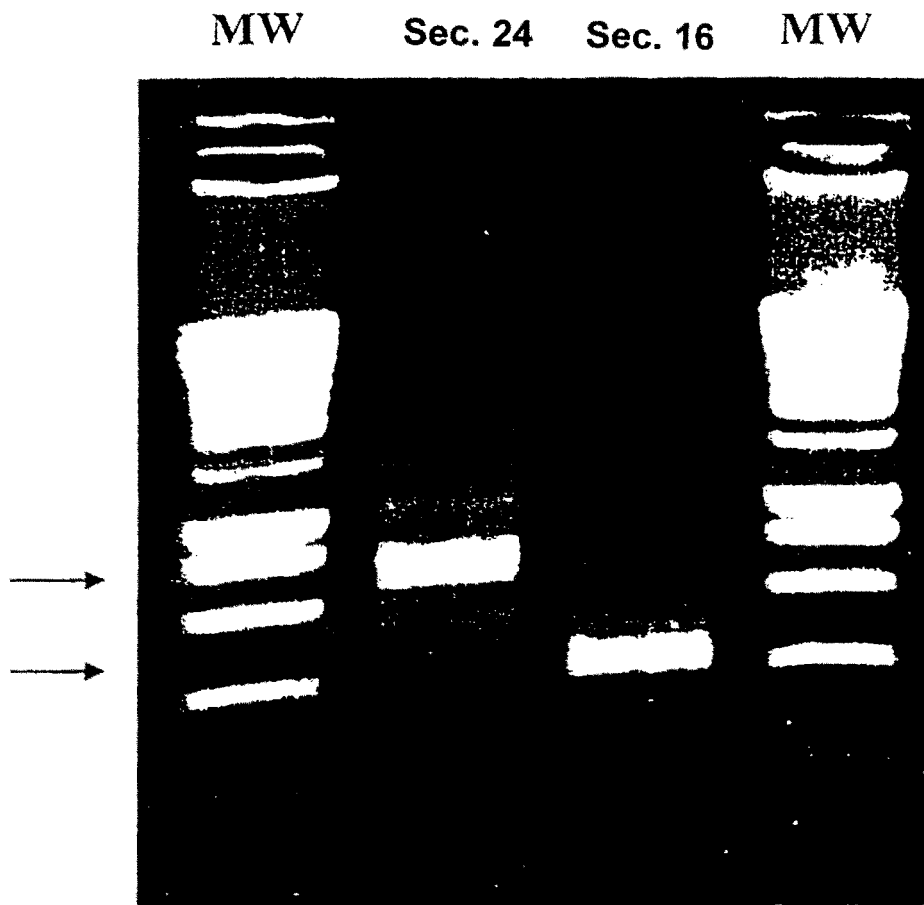


Figura 1.

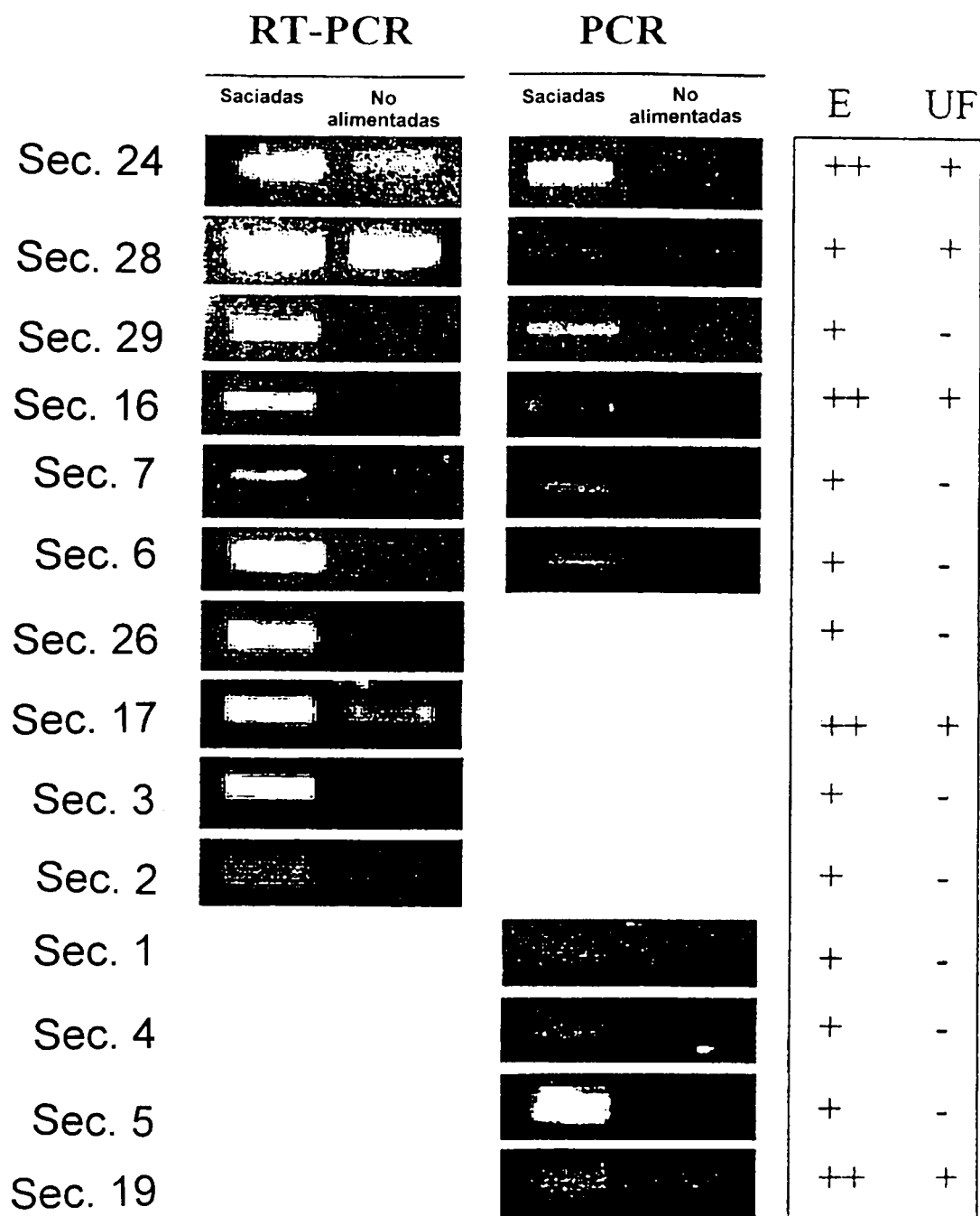


Figura 2.

ES 2 293 900 T3

LISTA DE SECUENCIAS

(1) INFORMACIONES PARA LA SEC. ID. nº: 1:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(a) LONGITUD: 194 pares de bases

(b) TIPO: ácido nucleico

10 (c) NÚMERO DE CADENAS: una

(d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

15 (iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 1:

1 ATACCTTCCA CTTGTAGCCC TTCCTCATCC GATATGGTGA CGGATGCCAT

51 TGCATCCTCG TCGTGGAAGA GGTCCCTCTC TAAATAAGAC CCATCCATAT

20 101 ATGTGTGTTT GCGAATGCCG TCGACGTAGC TCCTGACTAG AAACTCGTCG

151 GCTAGGACAG AACTTTTCTT CAGGTTTAGC GTAATGTCCT CGTT

25 (2) INFORMACIONES PARA LA SEC. ID. nº: 2:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(a) LONGITUD: 607 pares de bases

30 (b) TIPO: ácido nucleico

(c) NÚMERO DE CADENAS: una

(d) CONFIGURACIÓN: lineal

35 (ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 2:

1 TACCNGGGAA TCCAAAACCA ATTTTATTG GAACTTCCAC GTCTTCTTCA

40 51 AGGCGGTGGC ACCTCTGCAT TTATGAAGTT CGTCTTGGCA TTTTATTTTT

101 TGCTTCTTTC ATTGCRGAAC TCGCAAATGC ACTTCCCGTG CTTGTCCGAT

45 151 TTCGCCCAA AAGCGCATGG CATTCTTCC GGCAGATTAA CTTTTTCAAA

201 TTCACGGTTC TGAACCAATA ATAGATCGTG GCAATGTTG TGCTGTTTGC

50 251 GATTTGCAA CCAGCTGTAG CCACCATTGG ACTCAAAGGT GCGCACAACA

301 TGGCGCCGAA CTGTGAAAAA CAAATTAAGG CTNCTTTGTA ATAACGCTAG

55 351 TCTTGGTACG CCGTTAGAGG TCGATGTCCG GCCTCGCGAT TGCAAAGTCA

401 CTTGCACTTA TCAAGCTCCT GGAGAAAAAT GGGTGCAACG GGGGGATCAG

60 451 CGTITGTACT TGCAAACATT TGTGGAGACG GTAAACCWGT ATTTCCGCGGA

501 ACTCAGATGC TCCAGCGTGA AGCTCGTCTT AATAAAAAGTT GTAAATTCGA

65 551 GTATNGATGA AGAACTGAAA TTCGAGGCAT TTAGAAACAC CACGAGAAGC

601 AGCGGAA

ES 2 293 900 T3

(3) INFORMACIONES PARA LA SEC. ID. nº: 3:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 259 pares de bases
- (b) TIPO: ácido nucleico
- (c) NÚMERO DE CADENAS: una
- (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 3:

```
1   GATCCTACGC CTGAAAATGA GTGTCCATCG TCTTCACATA GTGCCACATT
5
10  51  GTAATTGGTA CAAGCTCCAT TTTCGTCAGC GCTGTTTGTG ATGCTGCCGC
15  101 CTACTTTTCC TTCGGCACTC CATAAGTTAA ACCCTGTCAT TATAAGTGIG
20  151 ATTGCCGTAT CTCGGCTGAA TGGGTTCAT TTTTCTCTTA AATAATCAGC
25  201 TGTCCATATT CCATGTATTG TGTTTCATGAG TATGTGATTC TCATCGTATA
30  251 TCTTCGCCT
```

(4) INFORMACIONES PARA LA SEC. ID. nº: 4:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 170 pares de bases
- (b) TIPO: ácido nucleico
- (c) NÚMERO DE CADENAS: una
- (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 4:

```
1   CCACTCGAAA ATGGAGGCTT TGAACATTT CAGTACCCCT GTGAACTCTG
5
10  51  GCTTTGCAAT GTAACAGCAA AAACACTTAC AGTTGAAGGG TGCAGTGTC A
15  101 GACGCTATGG AAGTTGCATC CACGAGCACR ACCCTGATTA CTACTGGCCA
20  151 CGTTGCTRTC CGGGTCGTCC
```

(5) INFORMACIONES PARA LA SEC. ID. nº: 5:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 168 pares de bases
- (b) TIPO: ácido nucleico
- (c) NÚMERO DE CADENAS: una
- (d) CONFIGURACIÓN: lineal

ES 2 293 900 T3

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 5:

5
1 GTATGTTACC ATGTCCAACC CGGTTATTAA ATACACCAAG TCGTAGGATT
51 TGTAGGCAGC TGCATTGCCC TTGACGTA CTCTCAACGT TGCCAAGGAC
10
101 TCAGGCCCAT AAATGTAGTG GGGTTGACCT TGAACCTCTC GTAAAAAGCG
151 TTCTTTCTCC GTCGTGAG
15

(6) INFORMACIONES PARA LA SEC. ID. nº: 6:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(a) LONGITUD: 247 pares de bases

(b) TIPO: ácido nucleico

(c) NÚMERO DE CADENAS: una

(d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 6:

20
1 CCGAAMATAA AACTTAGTCT CACCAATATA CGTTTGCTA ACGCGAAGGA
35
51 ACAGGCACAA ATATACTACG AGCAGGACAT TCTCAAGAAC ACGGTTACG
101 GAGTGTGGAC GAGAATTCAC TCAAAATATC CGTTCCCTGA AGATGAGGGA
40
151 ATTACACTGA TAATGACAGG GTTTGATTTA TGGAGTGCCG ATTTAACTGT
201 AGGCGGCACC ATAACAAACA GCGCTGAGAA AAGCGGAGCT TGTACGA
45

(7) INFORMACIONES PARA LA SEC. ID. nº: 7

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(a) LONGITUD: 261 pares de bases

(b) TIPO: ácido nucleico

(c) NÚMERO DE CADENAS: una

(d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) CARACTERÍSTICAS:

(a) DENOMINACIÓN/CLAVE: CDS

(b) POSICIÓN: 1...261

(iv) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 7:

65

ES 2 293 900 T3

229 TGC CAA GAG GAA TGC AGA GCA AAA AAA GTC TAG

Cys Gln Glu Glu Cys Arg Ala Lys Lys Val End

5

1 ATG CCT TTT ATT TTC GTG GTG AGC TTA GTC ATT GTG GCC TGC ATC GTG
GTA GAC ACA

10

Met Pro Phe Ile Phe Val Val Ser Leu Val Ile Val Ala Cys Ile Val
Val Asp Thr

15

58 GCC AAC CAC AAA GGT AGA GGG CGG CCT GCG AAG TGT AAA CTT CCT CCG
GAC GAC GGA

20

Ala Asn His Lys Gly Arg Gly Arg Pro Ala Lys Cys Lys Leu Pro Pro
Asp Asp Gly

25

115 CCA TGC AGA GCA CGA ATT CCG AGT TAC TAC TTT GAT AGA AAA ACC AAA
ACG TGC AAG

Pro Cys Arg Ala Arg Ile Pro Ser Tyr Tyr Phe Asp Arg Lys Thr Lys
Thr Cys Lys

30

172 GAG TTT ATG TAT GGC GGA TGC GAA GGA AAC GAA AAC AAT TTT GAA AAC
ATA ACT ACG

35

Glu Phe Met Tyr Gly Gly Cys Glu Gly Asn Glu Asn Asn Phe Glu Asn
Ile Thr Thr

40

229 TGC CAA GAG GAA TGC AGA GCA AAA AAA GTC TAG

Cys Gln Glu Glu Cys Arg Ala Lys Lys Val End

45 (8) INFORMACIONES PARA LA SEC. ID. n°: 8:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

50

(a) LONGITUD: 292 pares de bases

(b) TIPO: ácido nucleico

(c) NÚMERO DE CADENAS: una

(d) CONFIGURACIÓN: lineal

55

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 8:

60

65

ES 2 293 900 T3

1 CATCGMAGCC ATAGTATATT TTGCACTTGT CTTCCGTTTC GTCGTAGTAG
5 51 GACCGATTCC ACATTGTAGT ACACCGAGTCA CTTATATCCT GCGGGCGGTG
101 CTTGCATTG TCCTGAACAA ATCTTCCACA GCGCTTGTGG CACGCCTCCT
10 151 GGAATAGAA GCGCTTCTCT CCTCCGCATC TCCATTTGGA ATCATAGAAA
201 CATCTTTCAG TTTGAATATT GTAGCGATAA TAATCGGTAT CAGTTTCTTT
15 251 GCATGGTCCT GGGAGGGGTT TGGCGCAGGG GCCGTATTCA GG

(9) INFORMACIONES PARA LA SEC. ID. n°: 9:

- 20 (i) CARACTERÍSTICAS DE LA SECUENCIA:
(a) LONGITUD: 270 pares de bases
(b) TIPO: ácido nucleico
25 (c) NÚMERO DE CADENAS: una
(d) CONFIGURACIÓN: lineal
(ii) TIPO DE MOLÉCULA: ADNc
30 (iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 9:

1 GGTAATAGTT GTCAAATTCC ATTAATGTAT CCTGAAATGT GACCATATCT
35 51 TTGTTTCCCC TGTAATAATCT CATAAAAGGC TGTGTGTTTT CCTTAAGAAG
40 101 TGTAACAGCC ACGATGGTCA ATCTCACGGA TGGATGTGTG ACACTTTTAT
151 ATCTCAGGTT TGCCGACATT GCCATTACAG ATAAATAGTT GATAATTTCT
45 201 TTCTTGTTAT AGTTGTAAGC AGCGCATGTT GTTGCATCAA GCACCACATG
251 CACTTCAGGC AATATGGTTT

(10) INFORMACIONES PARA LA SEC. ID. n°: 10:

- 55 (i) CARACTERÍSTICAS DE LA SECUENCIA:
(a) LONGITUD: 316 pares de bases
(b) TIPO: ácido nucleico
(c) NÚMERO DE CADENAS: una
60 (d) CONFIGURACIÓN: lineal
(ii) TIPO DE MOLÉCULA: ADNc
(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 10:

65

ES 2 293 900 T3

1 AGAAAGCAGT CATATTGGCC ATCCACAGGT CACAATGGTT CTCTCCTTGA
5 51 CCTGGCATCG GGATTCTGAAG TATGGTGCAG TTCACGTAGT TGGAATACAA
101 CACGAAATGT GTTCGTTGGT ACGCCAATAG GGGTTCTCGC AAAGAACATA
10 151 TCATTGGAG GAAGGCGTAG TCCGTCGAGA TATCCCAAAA CTAGGGTTTC
201 ATTGCGTGCG AACCAACTGC CCCCACTTCT GTATGTGTAC TGTAAGGAGT
15 251 RGTGAACGG YGTCCTCTTT CCCATAACCT TGAAGTTTTT CACTGCAGA
301 GGATTACCTC TCAAAA

(11) INFORMACIONES PARA LA SEC. ID. n°: 11:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 241 pares de bases
- (b) TIPO: ácido nucleico
- (c) NÚMERO DE CADENAS: una
- (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 11:

35 1 AAGGTAGCAA GGGTGGTAGG CTTTCCTCAC AAAGAGTCTG GCTTCCGTGA
51 TAACCATATC CATTCTCAC CGTATACCCG TCATCCAACG TCAATTGTGT
40 101 TACAAGGCAG ATAATGTCAA AATGGCTCTG GTCCTATAA TAGTCGGATA
151 ATGTAGAAAT CGCTCCATGT GGCCAAATAG ATGTTCTCTT TTCATACTGT
45 201 TTTAACTTTA ATTGTAGGTC CGCCTCGTTC TCGAGGTATG T

(12) INFORMACIONES PARA LA SEC. ID. n°: 12:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 636 pares de bases
- (b) TIPO: ácido nucleico
- (c) NÚMERO DE CADENAS: una
- (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 12:

ES 2 293 900 T3

1 TTCCCCNAAT TGGCCTTGCG ANNCTTGCAA GTCGACNCTA GAGGCTCCGA
51 AGATGGACAG ATTGCCGATG AAATATTTGA AATCGAGCAG AATGGTGATT
5
101 TTAGGAGCGA TTATATTGTG CCACCCAGTT TAAAAGTGCA AGAACGCACA
151 GTGGTTTACC GTAACAAGTA CACCAGAGTT CCTGTAAATT TTACCGTCGA
10
201 AGTTGCCATG CTGATTGATA AGTATTTATA CWAGGAGTTC AAGAACGAGA
251 GCCACATCGT ACCGTACCTG GCTATGATAC TGACTTTGAT AAATCTGAGG
15
301 TATGCCGACA CACATGACCC GTACATCCAG TTTCTTCTCA CACAAGTGTT
351 CGTGGGGAAW WCTGGCGATC ATATGGGCCA CATGCCCTTC CGACGAGCGT
20
401 TCTTGTTTACG GCGCCGGCAT TATGCCGAGT TTAGGCCCAA TMACACCTTC
451 CACTTGTAAT TCTCCGTTGT TGGATAGTGT AAGTGAGGCC ATTGCATCAG
25
501 CATCGTGGAA GARGCCTTC TCCAAGTAGG AACCGCCAT TTAGGTTTGC
551 TTTCCCAATC CGCCAATTA ANTTTTAAAA AAAATTCCCC CCCCAAAAAT
30
601 TAATTTTTTTT TAAAGGTGGA TTGTGATTC TCCGTT

35 (13) INFORMACIONES PARA LA SEC. ID. n°: 13:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 40 (a) LONGITUD: 432 pares de bases
(b) TIPO: ácido nucleico
(c) NÚMERO DE CADENAS: una
(d) CONFIGURACIÓN: lineal

45 (ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 13:

1 GATCCCAAAA GTGCCCTGG ARCGACGGTT ACATCATGAG CTACGTCATA
50
51 AACTTCAAAA ACCACTTCAA ATTTTCTCCC TGCTGTGTAG AATCAATTCC
101 ATTTCGTCGA CGAGAGCGGG ACTGCCTCTA CAAAGTCAAT GCCAAGGATG
55
151 CTGTAAAAAG CCTAATATCT CTGCCCGGAT TTAGGATATC GCCAACGAGT
201 TTCTGTCAAT TTATGCATCC GCTTTACCGC GGTGTCCATA GCGATAAGAA
60
251 AGCAGGTCTG TCCGATTGCG TACAGACGTG TAGAACGGCC AAAAATCGAC
301 GAGGAGGCTA CCATTTCATGG ATTCACGCGG CACTTGACGG GGTTCTTTCG
65
351 GACAAGAGAA ACCCCAAGAA GGCCTGCATA AACGGGAAAT GCACCCTCCT
401 TAAGAGCATG CCCACAGAA CGTACCGGGA AT

ES 2 293 900 T3

(14) INFORMACIONES PARA LA SEC. ID. nº: 14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 466 pares de bases
- (b) TIPO: ácido nucleico
- (c) NÚMERO DE CADENAS: una
- (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 14:

```
1  AGGGGGTTCT TTGCITYACA GGGAAACRGCA TATGGGCCAC GTGACCTTCC
51  AATGACCGCT CCAAATCTGG CATAGGTTGA AYTCCGAAGT CGTGGCGCAG
101 CAGGCCTYCC ACATTCACCTC CATCCTCGTC TTTTAGGATG ACTGCCGCCA
151 TTTGTTTTGT ATCGTGGTAC AGGTGTTTGT TATGGTCCGA GCCGTGACAA
201 TAAGTATTGA CCAACGATCG GCCGAATGAT TACGGCTCAC CAAACACATC
251 AAATACCCCC GTCAAGTCAA GAGCTGGAAG CACAAAAGCAT AGTATGTACA
301 AGATACCCCTT GAAATCTTT CCCGAAGTTC ACCTTGTGCT GGACAGCACA
351 TTTGCCAAAG CTTTTAAATT TGACGTGTAC AAAGTAACGC GTTACTTCGC
401 AGTGCTTACA AATGCGGCTA ATCTTAGGTA TGCCAGCTTC GTATTTCCAA
451 AAGTACAGCT CAGGAT
```

(15) INFORMACIONES PARA LA SEC. ID. nº: 15:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 377 pares de bases
- (b) TIPO: ácido nucleico
- (c) NÚMERO DE CADENAS: una
- (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 15:

```
1  CTCGTCCACA CATTCTCCTA AAATGCAAGC CTTTTTTTTT CCACAAGGTG
51  TACCGTCGAC TACTGAGT CTCCAATAAA TATGTTTTCC GGTGCAATTT
101 ACCTTGAGT CTTTGACGCC GTATGTAGGG TCAGCGTGCA TGCCTTCGTC
151 GTACATATAC ACCCTCTGAC AGTAGTTGCT CAGTGTGTC ATCCTIACCAG
201 GAAGCTTAGA CGAACGTTTT ATTGTTTTTG TGGTGTATCG TTCTCTAAGG
251 CATTTGAATT CCGGACGGTT GTAGAGGTTT CTGACTTCTC GCTGGCAGCA
301 ATAAGAGAAC TGATACTGGC GCTCGTCTT CATCTTGTA CTCATGAGGT
351 ATCCGTCATC CCATGGGCAG TCCGCAG
```

ES 2 293 900 T3

(16) INFORMACIONES PARA LA SEC. ID. nº: 16:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 1670 pares de bases
- (b) TIPO: ácido nucleico
- (c) NÚMERO DE CADENAS: una
- (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) CARACTERÍSTICAS:

- (a) DENOMINACIÓN/CLAVE: CDS
- (b) POSICIÓN: 54...1520

(iv) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 16:

1 AAGGAAGAAG TTAGGCGTAG GCTTTGGGAA ACCGGTCATC CTCGAAACCA GAG

54 ATG TCG GGA CTC AGC CTG AAA TTG TGG ATT GTA GCG TTC TTT TCT
Met Ser Gly Leu Ser Leu Lys Leu Trp Ile Val Ala Phe Phe Ser

99 TTC TGC TTG GCC GAG AAA GAG CAT GGG ATC GTG TAC CCC AGG ATG CTT
Phe Cys Leu Ala Glu Lys Glu His Gly Ile Val Tyr Pro Arg Met Leu

147 GAA AGC AGA GCA GCA ACT GGA GAG AGA ATG CTT AAA ATC AAC GAT GAC
Glu Ser Arg Ala Ala Thr Gly Glu Arg Met Leu Lys Ile Asn Asp Asp

195 CTG ACG TTG ACG CTG CAG AAG AGT AAG GTC TTC GCT GAC GAC TTT CTC
Leu Thr Leu Thr Leu Gln Lys Ser Lys Val Phe Ala Asp Asp Phe Leu

243 TTC AGC ACG ACC GAC GGA ATT GAA CCT ATT GAT TAC TAC ATC AAA GCC

ES 2 293 900 T3

Phe Ser Thr Thr Asp Gly Ile Glu Pro Ile Asp Tyr Tyr Ile Lys Ala

5 291 GAA GAC GCT GAA CGT GAC ATC TAC CAC GAC GCA ACT CAC ATG GCA TCA
 Glu Asp Ala Glu Arg Asp Ile Tyr His Asp Ala Thr His Met Ala Ser

10 339 GTA AGG GTA ACG GAC GAT GAT GGC GTG GAA GTG GAA GGA ATT CTT GGA
 Val Arg Val Thr Asp Asp Asp Gly Val Glu Val Glu Gly Ile Leu Gly

15 387 GAG AGG CTT CGT GTT AAA CCT TTG CCG GCA ATG GCC CGC AGC AGC GAT
 Glu Arg Leu Arg Val Lys Pro Leu Pro Ala Met Ala Arg Ser Ser Asp

20 435 GGC CTC AGA CCG CAT ATG TTG TAC GAA GTC GAC GCA CAC GAA AAC GGC
 Gly Leu Arg Pro His Met Leu Tyr Glu Val Asp Ala His Glu Asn Gly

25 483 CGG CCA CAT GAT TAT GGT TCA CCG AAC ACA ACA AAT ACC CCC GTA GAG
 Arg Pro His Asp Tyr Gly Ser Pro Asn Thr Thr Asn Thr Pro Val Glu

30 531 AGA AGA GCT GGA GGC ACA GAA CCC CAG ATG TAC AAG ATA CCA GCG GAA
 Arg Arg Ala Gly Gly Thr Glu Pro Gln Met Tyr Lys Ile Pro Ala Glu

35 579 ATC TAT CCC GAA GTT TAC CTT GTG GCG GAT AGT GCC TTT GCC AAA GAA
 Ile Tyr Pro Glu Val Tyr Leu Val Ala Asp Ser Ala Phe Ala Lys Glu

40 627 TTT AAC TTT GAT GTG AAC GCC GTT AGG CGT TAC TTC GCA GTG CTT ACA
 Phe Asn Phe Asp Val Asn Ala Val Thr Arg Tyr Phe Ala Val Leu Thr

45 675 AAT GCG GCT AAT CTT AGG TAT GAA AGC TTC AAA TCT CCA AAG GTA CAG
 Asn Ala Ala Asn Leu Arg Tyr Glu Ser Phe Lys Ser Pro Lys Val Gln

50 723 CTC AGG ATC GTT GGC ATA ACG ATG AAC AAA AAC CCA GCA GAC GAG CCA
 Leu Arg Ile Val Gly Ile Thr Met Asn Lys Asn Pro Ala Asp Glu Pro

55 771 TAC ATT CAC AAT ATA CGG GGA TAT GAG CAG TAC CGG AAT ATT TTG TTT
 Tyr Ile His Asn Ile Arg Gly Tyr Glu Gln Tyr Arg Asn Ile Leu Phe

60 819 AAG GAA ACA CTG GAG GAT TTC AAC ACT CAG ATG AAG TCA AAA CAT TTT
 Lys Glu Thr Leu Glu Asp Phe Asn Thr Gln Met Lys Ser Lys His Phe

65 867 TAT CGT ACT GCC GAT ATC GTG TTT CTC GTG ACA GCA AAA AAT ATG TCC
 Tyr Arg Thr Ala Asp Ile Val Phe Leu Val Thr Ala Lys Asn Met Ser

 915 GAA TGG GTT GGT AGC ACA CTA CAA TCA TGG ACT GGC GGG TAC GCT TAC
 Glu Trp Val Gly Ser Thr Leu Gln Ser Trp Thr Gly Gly Tyr Ala Tyr

 963 GTA GGA ACA GCG TGT TCC GAA TGG AAA GTA GGA ATG TGT GAA GAC CGA
 Val Gly Thr Ala Cys Ser Glu Trp Lys Val Gly Met Cys Glu Asp Arg

ES 2 293 900 T3

5
10
15
20
25
30
35
40
45
50

1011 CCG ACA AGC TAT TAC GGA GCT TAC GTT TTC GCC CAT GAG CTG GCG CAT
Pro Thr Ser Tyr Tyr Gly Ala Tyr Val Phe Ala His Glu Leu Ala His

1059 AAT TTG GGT TGT CAA CAC GAT GGA GAT GGT GCC AAT AGC TGG GTC AAA
Asn Leu Gly Cys Gln His Asp Gly Asp Gly Ala Asn Ser Trp Val Lys

1107 GGC CAC ATC GGA TCT GCG GAC TGC CCA TGG GAT GAC GGA TAC CTT ATG
Gly His Ile Gly Ser Ala Asp Cys Pro Trp Asp Asp Gly Tyr Leu Met

1155 AGC TAC AAG ATG GAA GAC GAG CGC CAG TAT AAG TTT TCT CCC TAC TGC
Ser Tyr Lys Met Glu Asp Glu Arg Gln Tyr Lys Phe Ser Pro Tyr Cys

1203 CAG AGA GAA GTC AGG AAC CTC TAC AGG CGT CCG GAA TTC AAA TGC CTC
Gln Arg Glu Val Arg Asn Leu Tyr Arg Arg Pro Glu Phe Lys Cys Leu

1251 ACT GAA CGA AAA GCG AAA AAA ACA ATC CGC TCG TCT AAG CTA CCT GGT
Thr Glu Arg Lys Ala Lys Lys Thr Ile Arg Ser Ser Lys Leu Pro Gly

1299 GTG ATG ACA TCA TCG AGC AAC TAT TGC CGG AGG GTG TAC ATG TAC GAA
Val Met Thr Ser Ser Ser Asn Tyr Cys Arg Arg Val Tyr Met Tyr Glu

1347 AAA GGC ATG CAC GCC GAC GAG GCA TaT GGC GTC AAG GAC TGC AGG GTA
Lys Gly Met His Ala Asp Glu Ala Tyr Gly Val Lys Asp Cys Arg Val

1395 AAA TGC ACC ACC ACA TCA AGA ATG TAT TGG CTA CTC GGT GTA GTC GAC
Lys Cys Thr Thr Thr Ser Arg Met Tyr Trp Leu Leu Gly Val Val Asp

1443 GGT ACA CCT TGC GGA AAT GGA AAG GCT TGC ATT CTT GGG AAA TGC AGG
Gly Thr Pro Cys Gly Asn Gly Lys Ala Cys Ile Leu Gly Lys Cys Arg

1491 AAC AAA ATC AAA ATA AGC AAG AAG GAC TGA GAGGTIGATA ATATCAAATT
Asn Lys Ile Lys Ile Ser Lys Lys Asp End

1541 AATCATGATA TTTCAACCAC ATGACTTCGT GCTCAACTGG TAGCCCCAAA TAAATTTTAA

1601 AAAAAATCCC AATATGCGTG GTAGAAAAAG CAGCAAACAA TAAAATTCT AAAAAATGCT

1661 TGCAAAAATG

(17) INFORMACIONES PARA LA SEC. ID. n°: 17:

- 55
60
65
- (i) CARACTERÍSTICAS DE LA SECUENCIA:
 - (a) LONGITUD: 158 pares de bases
 - (b) TIPO: ácido nucleico
 - (c) NÚMERO DE CADENAS: una
 - (d) CONFIGURACIÓN: lineal
 - (ii) TIPO DE MOLÉCULA: ADNc
 - (iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 17:

ES 2 293 900 T3

1 CACCAGTGAT GCTTATTGTT GCACTGCACT TGTGATAAT ATCCGGTCGT

5

51 CGAATTGCAC TTGGGAACCT CCACTCCAAC TTGGCGAGCC GTGGATTTTG

101 ACTTCTCGTG ATGCTCCACC AGACAGTTGC AGGACTTCAG CTGCCTAGAT

10

151 GGAGCCTT

(18) INFORMACIONES PARA LA SEC. ID. n°: 18:

15

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(a) LONGITUD: 146 pares de bases

(b) TIPO: ácido nucleico

(c) NÚMERO DE CADENAS: una

20

(d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 18:

25

1 CTGTTGTTGA ACTGAAATAA ATAACAAAAA AATCATAAAG NTGGAGGAAA

30

51 GATGATCGAN TCCCGGCCCC TTGACAATCG TCCGATAAAA ACCAACTATA

101 TTCNGTCCTT TTTACAAACA ATTCCAANTG TCTGACCGAA CCGCGA

35

(19) INFORMACIONES PARA LA SEC. ID. n°: 19:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(a) LONGITUD: 140 pares de bases

40

(b) TIPO: ácido nucleico

(c) NÚMERO DE CADENAS: una

(d) CONFIGURACIÓN: lineal

45

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 19:

50

1 CTNGGACGAN GTCCTATGAC TTGGCCTTAN GTTTCTTAGT CTCCTTCGGT

51 TTCTTCTTTT TTIGCTTCGG TTTTTCGGTG GGCGCAGGTG TATAGTCATC

55

101 AGTGTCGGTG GGCCCATCCG AATGAGTTGT CAAATGACAT

(20) INFORMACIONES PARA LA SEC. ID. n°: 20:

60

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(a) LONGITUD: 143 pares de bases

(b) TIPO: ácido nucleico

(c) NÚMERO DE CADENAS: una

65

(d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

ES 2 293 900 T3

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 20:

5
10
1 **TGCCGAAAAA TAACGATGAT TTGACGTTGA CTCTGCAGAA GAGTAAGGTT**
51 **TTCACCGACA GTTTTCTGTT TAGCACGACG AAGGATAACG AGCCTATCGA**
101 **TTACTACGTG AGAGCCGAAG ATGCCGAACG AGACATATAT CAC**

(21) INFORMACIONES PARA LA SEC. ID. nº: 21:

15 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 140 pares de bases
(b) TIPO: ácido nucleico
(c) NÚMERO DE CADENAS: una
20 (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

25 (iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 21:

30
1 **TGTTGCTACA GACTCGACGT TTCGAGCTTG CTCGCCATTT MAAGACAÆCG**
51 **CACTCACAGA ATATTTAAGT GCGTTCGTGA WAGCTGTGGG CTTACGATTG**
101 **CAGGCGCTTC ANTCACCAGC TGTGATATTA MAGTTCCTAG**

35 (22) INFORMACIONES PARA LA SEC. ID. nº: 22:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 144 pares de bases
40 (b) TIPO: ácido nucleico
(c) NÚMERO DE CADENAS: una
(d) CONFIGURACIÓN: lineal

45 (ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 22:

50
55
1 **TCACGATAGT TGAACGTTG AAAC TTGAAA TACTCCCACA GTCGTTGGAT**
51 **GCTTCAGAAC TGCTAAGAAC TTCACACTTT GCAAGAAGTW CCAAAATGAA**
101 **AGCCGCGATG ACCGATGATT TAGCTTCCAT CTTCTATCAC TTGA**

(23) INFORMACIONES PARA LA SEC. ID. nº: 23:

60 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 95 pares de bases
(b) TIPO: ácido nucleico
65 (c) NÚMERO DE CADENAS: una
(d) CONFIGURACIÓN: lineal

ES 2 293 900 T3

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 23:

5 1 **GACCACCCCG TCCGAACTTG CTAAAKCAAG CAATGGAGTG AGGTGTTCTA**

 51 **TGCGGGTTGA TTACACCAAT GGCGCTGCGT GGTGCGTGTT GATTT**

10

(24) INFORMACIONES PARA LA SEC. ID. nº: 24:

15 (i) CARACTERÍSTICAS DE LA SECUENCIA:

 (a) LONGITUD: 1414 pares de bases

 (b) TIPO: ácido nucleico

 (c) NÚMERO DE CADENAS: una

20 (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) CARACTERÍSTICAS:

25 (a) DENOMINACIÓN/CLAVE: CDS

 (b) POSICIÓN: 143...1276

(iv) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 24:

30

 1 **GTAGGGCCGT GCAAGCGAAG GCAGCGAAGG CTGCGAGTGT ACGTGCAGTT CGGAAGTGCA**

35

 61 **ATATCCTGTT ATTAAGCTCT AATTAGCACA CTGTGAGTGG ATCAGAGGCC TCTCTTAACG**

121 **CCACATTGAA AAAGGATCCA AG ATG GAG GCA AGT CTG AGC AaC CAC ATC CTT**
 Met Glu Ala Ser Leu Ser Asn His Ile Leu

40

173 **AAC TTC TCC GTC GAC CTA TAC AAG CAG CTG AAA CCC TCC GGC AAA GAC**
 Asn Phe Ser Val Asp Leu Tyr Lys Gln Leu Lys Pro Ser Gly Lys Asp

45

221 **ACG GCA GGA AAC GTC TTC TGC TCA CCA TTC AGT ATT GCA GCT GCT CTG**
 Thr Ala Gly Asn Val Phe Cys Ser Pro Phe Ser Ile Ala Ala Ala Leu

50

269 **TCC ATG GCC CTC GCA GGA GCT AGA GGC AAC ACT GCC AAG CAA ATC GCT**
 Ser Met Ala Leu Ala Gly Ala Arg Gly Asn Thr Ala Lys Gln Ile Ala

55

365 **TTC CTT TGC AAG CTT CCC AGT TAC GCC CCA GAT GTG GCC CTG CAC ATC**
 Phe Leu Cys Lys Leu Pro Ser Tyr Ala Pro Asp Val Ala Leu His Ile

60

413 **GCC AAT GGC ATG TAC TCT GAG CAG ACC TTC CAT CCG AAA GCG GAG TAC**
 Ala Asn Arg Met Tyr Ser Glu Gln Thr Phe His Pro Lys Ala Glu Tyr

65

461 **ACA ACC CTG TTG CAA AAG TCC TAC GAC AGC ACC ATC AAG GCT GTT GAC**
 Thr Thr Leu Leu Gln Lys Ser Tyr Asp Ser Thr Ile Lys Ala Val Asp

ES 2 293 900 T3

509 TTT GCA GGA AAT GCC GAC AGG GTC CGT CTG GAG GTC AAT GCC TGG GTT
Phe Ala Gly Asn Ala Asp Arg Val Arg Leu Glu Val Asn Ala Trp Val

5

557 GAG GAA GTC ACC AGG TCA AAG ATC AGG GAC CTG CTC GCA CCT GGA ACT
Glu Glu Val Thr Arg Ser Lys Ile Arg Asp Leu Leu Ala Pro Gly Thr

10

605 GTT GAT TCA TCG ACA TCA CTT ATA TTA GTG AAT GCC ATT TAC TTC AAA
Val Asp Ser Ser Thr Ser Leu Ile Leu Val Asn Ala Ile Tyr Phe Lys

15

653 GGT CTG TGG GAT TCT CAG TTC AAG CCT AGT GGT ACG AAG CCG GGA GAT
Gly Leu Trp Asp Ser Gln Phe Lys Pro Ser Ala Thr Lys Pro Gly Asp

20

701 TTT CAC TTG ACA CCA CAG ACC TCA AAG AAA GTG GAC ATG ATG CAC CAG
Phe His Leu Thr Pro Gln Thr Ser Lys Lys Val Asp Met Met His Gln

25

749 GAA GGG GAC TTC AAG ATG GGT CAC TGC AGC GAC CTC AAG GTC ACT GCG
Glu Gly Asp Phe Lys Met Gly His Cys Ser Asp Leu Lys Val Thr Ala

30

797 CTT GAG ATA CCC TAC AAA GGC AAC AAG ACG TCG ATG GTC ATT CTC CTG
Leu Glu Ile Pro Tyr Lys Gly Asn Lys Thr Ser Met Val Ile Leu Leu

35

845 CCC GAA GAT GTA GAG GGA CTC TCA GTC CTG GAG GAA CAC TTG ACC GCT
Pro Glu Asp Val Glu Gly Leu Ser Val Leu Glu Glu His Leu Thr Ala

40

893 CCG AAA CTG TCG GCT CTG CTC GGC GGC ATG TAT GCG ACG TCC GAT GTC
Pro Lys Leu Ser Ala Leu Leu Gly Gly Met Tyr Ala Thr Ser Asp Val

45

941 AAC ITG CGC TTG CCG AAG TTC AAA CTA GAG CAG TCC ATA GGT TTG AAG
Asn Leu Arg Leu Pro Lys Phe Lys Leu Glu Gln Ser Ile Gly Leu Lys

50

989 GAT GTA CTG ATG GCG ATG GGA GTC AAG GAT TTC TTC ACG TCC CTT GCA
Asp Val Leu Met Ala Met Gly Val Lys Asp Phe Phe Thr Ser Leu Ala

55

1037 GAT CTT TCT GGC ATC AGC GCT GCG GGG AAT CTG TGC GCT TCG GAT GTC
Asp Leu Ser Gly Ile Ser Ala Ala Gly Asn Leu Cys Ala Ser Asp Val

60

1085 ATC CAC AAG GCT TTT GTG GAA GTT AAT GAG GAG GGC ACA GAG GCT GCA
Ile His Lys Ala Phe Val Glu Val Asn Glu Glu Gly Thr Glu Ala Ala

65

1133 GCT GCC ACT GCC ATA CCC ATT ATG TTG ATG TGT GCG AGA TTT CCA CAG
Ala Ala Thr Ala Ile Pro Ile Met Leu Met Cys Ala Arg Phe Pro Gln

70

1181 GTG GTG AAC TTT TTC GTT GAC CGC CCA TTC ATG TTC TTG ATC CAC AGC
Val Val Asn Phe Phe Val Asp Arg Pro Phe Met Phe Leu Ile His Ser

75

1229 CAT GAT CCA GAT GTT GTT CTC TTC ATG GGA TCC ATC CGT GAG CTC TAA

ES 2 293 900 T3

His Asp Pro Asp Val Val Leu Phe Met Gly Ser Ile Arg Glu Leu End

5
1277 AAAGCATATT CTTAACGGCG GCCAATCAGT CTGTGGAGTT ATCTCTTAGT CACTAATGTG
1337 TAACAATTCT GCAATATTCA GCTTGTGTAT TTCAGTAACT TGCTAGATCT TTGTGTTGTT
10
1397 GATGTTAGGC TTCTTGCG

(25) INFORMACIONES PARA LA SEC. ID. n°: 25:

- 15 (i) CARACTERÍSTICAS DE LA SECUENCIA:
(a) LONGITUD: 200 pares de bases
(b) TIPO: ácido nucleico
20 (c) NÚMERO DE CADENAS: una
(d) CONFIGURACIÓN: lineal
(ii) TIPO DE MOLÉCULA: ADNc
25 (iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 25:

1 ACCGTAACCA AAATTGTTTC TTTCCAGAAG AATGGTTCAA ACTTTTCAAA
30 51 CAGATTTTCGG AAACCTTTCT TGCACTTTTA AAATCCAATC TACAATCTTT
101 CCTCGCACTT CTGAATTGCA TTCCAGTTTA CTTTCCAAGC AAAGCTCTTT
35 151 TGGCAACTCC AGCCGTACTC CATTTCGGCA TACCACAGTG CATGCACTTG

(26) INFORMACIONES PARA LA SEC. ID. n°: 26:

- 40 (i) CARACTERÍSTICAS DE LA SECUENCIA:
(a) LONGITUD: 241 pares de bases
(b) TIPO: ácido nucleico
45 (c) NÚMERO DE CADENAS: una
(d) CONFIGURACIÓN: lineal
(ii) TIPO DE MOLÉCULA: ADNc
50 (iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 26:

55 1 CGTATTCTTT GAAGATTTGT ATACGAAACA TAAATTCGTC ATGCATACTT
51 TTGATCGTTA CACGACATGC GAAGCTGCCG ACAAAGAAGA CTGGGAAGAT
101 AAGAAGCACC TAGTTACGGT AGTCCGTGGA CCGGATAAAC GAAAGTACAC
60 151 GTTTCTACGC AACATTCTCA CTTACAACG GAGAGTGAGA GTTAGCAAAA
201 CAATGATTGA GCTCGTACGG AACATGTCCT GTAGGACATT T
65

ES 2 293 900 T3

(27) INFORMACIONES PARA LA SEC. ID. nº: 27:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(a) LONGITUD: 313 pares de bases

(b) TIPO: ácido nucleico

(c) NÚMERO DE CADENAS: una

(d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 27:

1 AAGCANCCGG ACTACCTGCT TGAAAACGTT GTACGGGCAA ACTTGGACGG

51 AAAACTCCCA GATGCTACTC CAGTTCCTCC CGGAAGCTAC ACGTACGCTG

101 AGAATGATAA CTTCACTGC TATTCCAGAA GTACACCGTT TCCGGATGGG

151 GTGAATGTTG TATAACGGCT GCTGGGTGCG GAAGACTATG ATGGATTACG

201 CAAAAAAGTT CTAACGAGT TGTTTCCCAT CCCGGAAAGT CTGCTGTATG

251 CTGACATGAT GCGACTTGTG GCTAAGAAAG ACAGAGTTGA TCACACTAGT

301 GGATGACCTG GGA

(28) INFORMACIONES PARA LA SEC. ID. nº: 28:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(a) LONGITUD: 2417 pares de bases

(b) TIPO: ácido nucleico

(c) NÚMERO DE CADENAS: una

(d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) CARACTERÍSTICAS:

(a) DENOMINACIÓN/CLAVE: CDS

(b) POSICIÓN: 218...1495

(iv) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. nº: 28:

ES 2 293 900 T3

1 GTCGTAGTCG TAGTCGTAGT CAGTTGCGCA TCGCGGGGC TTTCTGTCT TTCTTGCTT
5 61 TCTGCAGTCG TTCACCAACA TGTGGATACA GCTCCGGAGA TTTGTAAACA AATACTGCAC
121 ITTTAAGCAA GACTTGATAT TTAGATCGAT ATCCTCCTGT TGTCCGTCTT GATTAATCGG
10 181 CICTTTAGGG TTTTLAGAAT AGGCTTTTCG GTACGAG ATG CCC AAA GGA AAG AGG
Met Pro Lys Gly Lys Arg
236 GGA CCC AAA GCA GGT GGC GCC GCG CGC GGT GGC CGG TGC GAG GCC AGC
15 Gly Pro Lys Ala Gly Gly Ala Ala Arg Gly Gly Arg Cys Glu Ala Ser
284 CTG GCT CCG TCg TCC AGC GAC GAG GAG TCC AAC GCA GAC ACG GCG AGC
Leu Ala Pro Ser Ser Ser Asp Glu Glu Ser Asn Ala Asp Thr Ala Ser
20 332 GTG CTG AGC TGC GCC TCG GAG TCT CGC TGT GGC AGT GAC GCC ACC GTT
Val Leu Ser Cys Ala Ser Glu Ser Arg Cys Gly Ser Asp Gly Thr Val
380 GGA GAC CCA GAA GCG GAG GAG GCT GTG CTG CAT GAC GAC TTT GAA GAC
25 Gly Asp Pro Glu Ala Glu Glu Ala Val Leu His Asp Asp Phe Glu Asp
428 AAA CTC AAG GAG GCC ATC GAC GGA GCT TCG CAG AAG AGT GCC AAA GGA
30 Lys Leu Lys Glu Ala Ile Asp Gly Ala Ser Gln Lys Ser Ala Lys Gly
476 CGG CTG TCG TGC CTG GAG GCG ATT CGC AAG GCC TTT TcC ACC AAA TAC
Arg Leu Ser Cys Leu Glu Ala Ile Arg Lys Ala Phe Ser Thr Lys Tyr
35 524 CTG TAG GAC TTC CTC ATG GAC AGA CCG AGC ACG GTG TGC GAC CTG GTG
Leu Tyr Asp Phe Leu Met Asp Arg Pro Ser Thr Val Cys Asp Leu Val
40 572 GAG CGT GGG GTG CGC AAG GGC CGA GGG GAG GAG GCG GCC CTG TGC GCC
Glu Arg Gly Val Arg Lys Gly Arg Gly Glu Glu Ala Ala Leu Cys Ala
620 ACT CTC GGG GCC CTG GCC TGC GTC CAG CTC GGG GTC GGG GCC GAG GCG
45 Thr Leu Gly Ala Leu Ala Cys Val Gln Leu Gly Val Gly Ala Glu Ala
668 GAC GCC CTG TTC GAC GCC CTG CGC CAG CCG CTC TGC ACT TTG CTG CTT
50 Asp Ala Leu Phe Asp Ala Leu Arg Gln Pro Leu Cys Thr Leu Leu Leu
716 GAC GGG GCC CAG GGG CCC TCC CCC AGG GCC AGG TGT GCC ACT GCC CTC

55

60

65

ES 2 293 900 T3

Asp Gly Ala Gln Gly Pro Ser Pro Arg Ala Arg Cys Ala Thr Ala Leu

5 764 GGC CTC TGC TGc TTC GTG GTG GAC TCG GAC AAC CAG CTG GTG CTG CAG
 Gly Leu Cys Cys Phe Val Val Asp Ser Asp Asn Gln Leu Val Leu Gln

10 812 CCG TGC ATG GAG GTG CTC TGG CAG GTG GTG GGT GCC AAG GCG GGC CCC
 Pro Cys Met Glu Val Leu Trp Gln Val Val Gly Ala Lys Ala Gly Pro

15 860 GGC TCT CCG GTG CTC CAG GCA GCG GCC CTG CTC GCC TGG GGC CTC CTG
 Gly Ser Pro Val Leu Gln Ala Ala Ala Leu Leu Ala Trp Gly Leu Leu

20 908 CTC AGC GTG GCT CCC GTC GAC CGc CTG CTG GCG CTc ACG CGC ACG CAC
 Leu Ser Val Ala Pro Val Asp Arg Leu Leu Ala Leu Thr Arg Thr His

25 956 CTG CCC CGG CTG CAG GAG CTG CTG GAG AGC CCC GAC CTG GAC CTG CGC
 Leu Pro Arg Leu Gln Glu Leu Leu Glu Ser Pro Asp Leu Asp Leu Arg

30 1004 ATT GCG GCC GGG GAG GTG ATC GCC GTC ATG TAC GAG GGG GCC AGG GAC
 Ile Ala Ala Gly Glu Val Ile Ala Val Met Tyr Glu Gly Ala Arg Asp

35 1052 TAC GAC GAG GAC TTT GAG GAG CCC TCG GAG TCC CTG TGT GCC CAG CTG
 Tyr Asp Glu Asp Phe Glu Glu Pro Ser Glu Ser Leu Cys Ala Gln Leu

40 1100 CGC CAG CTG GCC ACG GAC AGC CAG AAG TTT CGG GCC AAG AAG GAG CGG
 Arg Gln Leu Ala Thr Asp Ser Gln Lys Phe Arg Ala Lys Lys Glu Arg

45 1148 CGC CAG CAG CGC TCC ACC TTC AGG GAC GTC TAC CGG GCC GTC AGG GAG
 Arg Gln Gln Arg Ser Thr Phe Arg Asp Val Tyr Arg Ala Val Arg Glu

50 1196 GGG GCC TCT CCC GAC GTG AGC GTC AAG TTT GGC CGG GAA GTC CTG GAA
 Gly Ala Ser Pro Asp Val Ser Val Lys Phe Gly Arg Glu Val Leu Glu

55 1244 CTG GAC ACC TGG AGT CGC AAG CTG CAG TAC GAC GCT TTC TGC CAG CTG
 Leu Asp Thr Trp Ser Arg Lys Leu Gln Tyr Asp Ala Phe Cys Gln Leu

60 1292 CTG GGC TCC GGC ATG AAC CTG CAC CTG GCC GTG AAC GAG CTG CTG AGG
 Leu Gly Ser Gly Met Asn Leu His Leu Ala Val Asn Glu Leu Leu Arg

65 1340 GAC ATC TTT GAA CTG GGG CAG GTG CTG GCA ACC GAG GAC CAC ATT ATC
 Asp Ile Phe Glu Leu Gly Gln Val Leu Ala Thr Glu Asp His Ile Ile

 1388 TCC AAG ATC ACC AAG TTC GAA AGG CAC ATG GTG AAC ATG GCC AGC TGC
 Ser Lys Ile Thr Lys Phe Glu Arg His Met Val Asn Met Ala Ser Cys

 1436 CGG GCC CGC ACC AAG ACA CGC AAC CGG CTG AGG GAC AAG CGC GCC GAC
 Arg Ala Arg Thr Lys Thr Arg Asn Arg Leu Arg Asp Lys Arg Ala Asp

ES 2 293 900 T3

1484 GTG GTC GCC TGA ACCTGCGGAG GGATGCTTAG CTATGCACTC GCCGGCCTAC
Val Val Ala End

1536 CCTGGCGGGA CTCGATGCCA CTCACGAGTC GCGGCTGGCA AATTCGCCGC CCATCGTTAC

1596 GCAATGGGAG ACAAAGCTGC TTTTGGCATT ACCGTTTGAG GTGGGCTCCA ACCCATAGAT

1656 GAATTTCTTT TTTGTGGCCG TTTCTGGGTT ACATGTTTTG GGGGAAGGGA GTGGAAGTGT

1716 CCGGTCTTTT GGCACACGTC AGGTTGCTCT TGATGCCGGA CGTGCTTGTA TTTGGGtACT

1776 GCCGACACCA AGCGTTTCGG CGATTCCCTGG AAAAGAGTGC CTCTCGCTCG ACGTTTGGTT

1836 GTTTTCTGCG TGGTCCGTCG TCGACCTTCG TTCGTCCAAA GACGCCGTCC GGTTTCAtAC

1896 TCCCCCGGC ACACATATCG AGGCCAATTA AATTGCTAAG GGTGCCGTTG TCGTGCATCT

1956 GGCAGGCTCA GAAGTGGCTT ATTTGTCTTT TAATTTTGGC GATGCACGCA AAAATTGTCA

2016 TTTCTTGAAA GTTTCTCTTT TATTGCGTAC ACAATTCAAC TTTTATGTAA TTTCTGATGG

2076 TCTGTTTTAC GTGTGCGTGT GTAAAACGTA ACTTTGGAAG AATTTTTATG CACACTGAAC

2136 AAACGCTCGG TCCTGGGGTT GAAAGTGCTC GGTGTGTGCA TGAGCTAAAG TGCAACTGCT

2196 TTGTTCCGAA GGTTTTCTAG TCGCCGAAAT GTACCATTGT GGACCTTGTI GCGAGAGACC

2256 TTGGTCTTCT GGGGGAGCTG CTGTAGCGTG GCAAGCCACT ATTTTGGGAG CGACATTGCA

2316 GAGAAAATCG GCTTTTAGAA AGGCACCTGC GCGGGGAGTG GACGTTTTTT CGTATATACT

2376 GCGAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AA

(29) INFORMACIONES PARA LA SEC. ID. n°: 29:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (a) LONGITUD: 933 pares de bases
- (b) TIPO: ácido nucleico
- (c) NÚMERO DE CADENAS: una
- (d) CONFIGURACIÓN: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(iii) CARACTERÍSTICAS:

- (a) DENOMINACIÓN/CLAVE: CDS
- (b) POSICIÓN: 32...853

(iv) DESCRIPCIÓN DE LA SECUENCIA: SEC. ID. n°: 29:

ES 2 293 900 T3

1 GATTGGGAAC CTCCTATTCC TCACTTGAAA C ATG GCT GGA CTC CGC TCC
Met Ala Gly Leu Arg Ser

5

50 TGC ATC CTC CTG GCT CTT GCC ACT AGT GCC TTC GCC GGC TAC CTT CAC
Cys Ile Leu Leu Ala Leu Ala Thr Ser Ala Phe Ala Gly Tyr Leu His

10

98 GGT GGC CTT ACC CAC GCC GCT GGG TAC GGT TAC GGT GTC GGC TAC GGT
Gly Gly Leu Thr His Gly Ala Gly Tyr Gly Tyr Gly Val Gly Tyr Gly

15

146 TCC GGC CTT GGC TAT GCC CTT GGC TAC GGT TCC GGC CTT GGC TAT GGA
Ser Gly Leu Gly Tyr Gly Leu Gly Tyr Gly Ser Gly Leu Gly Tyr Gly

20

194 CAT GCT GTT GGC CTT GGA CAC GCC TTT GCC TAT TCT GGT CTG ACC GGC
His Ala Val Gly Leu Gly His Gly Phe Gly Tyr Ser Gly Leu Thr Gly

25

242 TAC AGT GTG GCT GCC CCA GCT AGC TAC GCC GTT GCT GCT CCA GCC GTC
Tyr Ser Val Ala Ala Pro Ala Ser Tyr Ala Val Ala Ala Pro Ala Val

30

290 AGC CGC ACC GTT TCC ACT TAC CAC GCT GCT CCA GCT GTG GCC ACC TAC
Ser Arg Thr Val Ser Thr Tyr His Ala Ala Pro Ala Val Ala Thr Tyr

35

338 GCC GCT GCT CCT GTC GCC ACC TAT GCT GTT GCT CCA GCT GTC ACT AGG
Ala Ala Ala Pro Val Ala Thr Tyr Ala Val Ala Pro Ala Val-Thr Arg

40

386 GTT TCC CCC GTT CGC GCC GCC CCA GCT GTG GCC ACG TAC GCC GCC GCT
Val Ser Pro Val Arg Ala Ala Pro Ala Val Ala Thr Tyr Ala Ala Ala

45

434 CCA GTC GCC ACC TAC GCC GCT GCT CCA GCT GTG ACC AGG GTG TCC ACC
Pro Val Ala Thr Tyr Ala Ala Ala Pro Ala Val Thr Arg Val Ser Thr

50

482 ATT CAC GCT GCC CCG GCT GTG GCC AAT TAC GCC GTC GCT CCA GTC GCC
Ile His Ala Ala Pro Ala Val Ala Asn Tyr Ala Val Ala Pro Val Ala

55

530 ACC TAT GCC GCT GCT CCA GCT GTG ACC AGG GTC TCC ACC ATC CAC GCC
Thr Tyr Ala Ala Ala Pro Ala Val Thr Arg Val Ser Thr Ile His Ala

60

578 GCT CCA GCC GTG GCT AGC TAC CAG ACC TAC CAC GCT CCA GCT GTC GCC
Ala Pro Ala Val Ala Ser Tyr Gln Thr Tyr His Ala Pro Ala Val Ala

65

626 ACT GTG GCT CAT GCT CCA GCT GTG GCC AGC TAC CAG ACC TAC CAC GCT
Thr Val Ala His Ala Pro Ala Val Ala Ser Tyr Gln Thr Tyr His Ala

ES 2 293 900 T3

674 GCC CCA GCC GTG GCT ACC TAC GCC CAT GCC GCT GCC GTC TAC GGC TAT
Ala Pro Ala Val Ala Thr Tyr Ala His Ala Ala Pro Val Tyr Gly Tyr

5 722 GGT GTC GGT ACC CTC GGA TAT GGT GTC GGC CAC TAC GGC TAC GGA CAC
Gly Val Gly Thr Leu Gly Tyr Gly Val Gly His Tyr Gly Tyr Gly His

10 770 GGT CTT GGC AGC TAC GGC CTG AAC TAC GGT TAC GGC CTC GGC ACC TAC
Gly Leu Gly Ser Tyr Gly Leu Asn Tyr Gly Tyr Gly Leu Gly Thr Tyr

15 818 GGT GAC TAC ACC ACC CTT CTC CGC AAG AAG AAG TAA ATGGCA CATCTCAAGA
Gly Asp Tyr Thr Thr Leu Leu Arg Lys Lys Lys End

20 870 GAGCCCATG GACTGCCATC GACATTCTC TTCAATAAAA GAGCCCGAAG ATGGCATTAT

25

30

35

40

45

50

55

60

65 930 TTTT