



(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 340 210**

(51) Int. Cl.:

**C12N 15/12** (2006.01)

**C07K 14/705** (2006.01)

**C12N 15/62** (2006.01)

**C07K 16/28** (2006.01)

**C12N 15/70** (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(96) Número de solicitud europea: **07123073 .4**

(96) Fecha de presentación : **07.05.1998**

(97) Número de publicación de la solicitud: **1925666**

(97) Fecha de publicación de la solicitud: **28.05.2008**

(54)

Título: **Proteínas humanas receptoras de tipo Toll, reactivos asociados y métodos.**

(30)

Prioridad: **07.05.1997 US 44293 P**  
**22.01.1998 US 72212 P**  
**05.03.1998 US 76947 P**

(73)

Titular/es: **SCHERING CORPORATION**  
**2000 Galloping Hill Road**  
**Kenilworth, New Jersey 07033-0530, US**

(45)

Fecha de publicación de la mención BOPI:  
**31.05.2010**

(72)

Inventor/es: **Hardiman, Gerard T.;**  
**Bazan, Fernando J.;**  
**Rock, Fernando L. y**  
**Kastelein, Robert A.**

(45)

Fecha de la publicación del folleto de la patente:  
**31.05.2010**

(74)

Agente: **Ungría López, Javier**

ES 2 340 210 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

Proteínas humanas receptoras de tipo Toll, reactivos asociados y métodos.

Esta solicitud reivindica prioridad sobre las Solicitudes de Patente de los Estados Unidos USSN 60/044,293, presentada el 7 de Mayo de 1997; USSN 60/072, 212, presentada el 22 de Enero de 1998; y USSN 60/076,947, presentada el 5 de Marzo de 1998, cada una de las cuales se incorpora a la presente memoria como referencia.

**Campo de la invención**

La presente invención se refiere a composiciones y métodos que causan efecto en la fisiología de mamíferos, incluyendo la morfogénesis o la función del sistema inmunitario. En particular, proporciona ácidos nucleicos, proteínas, y anticuerpos que regulan el desarrollo y/o el sistema inmunitario. También se describen los usos diagnósticos y terapéuticos de estas sustancias.

**Antecedentes de la invención**

La tecnología del ADN recombinante hace referencia generalmente a técnicas de integración de información genética de una fuente donadora en vectores para su posterior procesamiento, por ejemplo la introducción en un anfitrión, por medio del cual la información genética transferida es copiada y/o expresada en el nuevo entorno. Comúnmente, la información genética existe en forma de ADN complementario (ADNc) derivado de ARN mensajero (ARNm) que codifica una proteína producto deseada. El portador es frecuentemente un plásmido que tiene la capacidad de incorporar ADNc para su posterior replicación en un anfitrión y, en algunos casos, realmente para controlar la expresión del ADNc y de ese modo dirigir la síntesis del producto codificado en el anfitrión.

Durante algún tiempo, se ha sabido que la respuesta inmunitaria de los mamíferos está basada en una serie de interacciones celulares complejas, denominada "red inmunitaria". La investigación reciente ha proporcionado nuevas percepciones en los funcionamientos internos de esta red. Si bien queda claro que la respuesta inmunitaria, de hecho, gira en torno a interacciones de tipo red de los linfocitos, macrófagos, granulocitos, y otras células, los inmunólogos sostienen ahora en general la opinión de que proteínas solubles, conocidas como linfoquinas, citoquinas, o monoquinas, juegan un papel crítico en el control de estas interacciones celulares. De este modo, existe un considerable interés en el aislamiento, la caracterización, y los mecanismos de acción de los factores moduladores celulares, cuya comprensión conducirá a avances significativos en la diagnosis y terapia de numerosas anomalías médicas, p. ej., trastornos del sistema inmunitario.

Las linfoquinas median aparentemente las actividades celulares de diversas maneras. Se ha demostrado que apoyan la proliferación, el crecimiento, y/o la diferenciación de células madre hematopoyéticas pluripotenciales en un inmenso número de progenitores que comprenden diversos linajes celulares que pueden formar un sistema inmunitario complejo. Son necesarias interacciones apropiadas y equilibradas entre los componentes celulares para una respuesta inmunitaria saludable. Los diferentes linajes celulares a menudo responden de diferente manera cuando se administran junto con otros agentes.

Los linajes celulares especialmente importantes para la respuesta inmunitaria incluyen dos clases de linfocitos: las células B, que pueden producir y secretar inmunoglobulinas (proteínas con la capacidad de reconocer y unirse a una sustancia foránea para llevar a cabo su eliminación), y las células T de diferentes subgrupos que secretan linfoquinas e inducen o suprimen las células B y otras células diferentes (incluyendo otras células T) que forman la red inmunitaria. Estos linfocitos interactúan con muchos otros tipos de células.

Otro linaje celular importante son los mastocitos (que no han sido identificados positivamente en todas las especies de mamíferos), que son células del tejido conectivo que contiene gránulos localizadas cerca de los capilares por todo el organismo. Estas células se encuentran en concentraciones especialmente elevadas en los pulmones, la piel, y los tractos gastrointestinal y genitourinario. Los mastocitos juegan un papel central en los trastornos relacionados con la alergia, concretamente en la anafilaxis como sigue: cuando los antígenos seleccionados se entrecruzan con una clase de inmunoglobulinas unidas a los receptores de la superficie de los mastocitos, los mastocitos de desgranulan y liberan mediadores, p. ej., histamina, serotonina, heparina, y prostaglandinas, que ocasionan las reacciones alérgicas, p. ej., la anafilaxis.

La investigación para comprender y tratar mejor los diferentes trastornos inmunitarios ha estado impedida por la incapacidad general para mantener las células del sistema inmunitario *in vitro*. Los inmunólogos han descubierto que se puede llevar a cabo el cultivo de muchas de estas células a través del uso de sobrenadantes de células T y otras células, que contienen diferentes factores de crecimiento, incluyendo muchas de las linfoquinas.

La familia de proteínas de la interleuquina-1 incluye la IL-1 $\alpha$ , la IL-1 $\beta$ , la IL-1RA, y recientemente la IL-1 $\gamma$  (también denominada Factor Inductor de Interferón Gamma, IGIF). Esta familia relacionada de genes ha sido implicada en una amplia gama de funciones biológicas. Véanse Dinarello (1994) FASEB J. 8:1314-1325; Dinarello (1991) Blood 77:1627-1652; y Okamura, *et al.* (1995) Nature 378:88-91.

Además, existen diferentes factores de crecimiento y reguladores, que modulan el desarrollo morfogenético. Estos incluyen, p. ej., los ligandos Toll, que señalizan a través de la unión a receptores que comparten rasgos característicos

estructurales, y mecánicos de los receptores de IL-1. Véanse, p. ej., Lemaitre, *et al.* (1996) *Cell* 86:973-983; y Belvin y Anderson (1996) *Ann. Rev. Cell & Devel. Biol.* 12:393-416.

A partir de lo anterior, resulta evidente que el descubrimiento y desarrollo de nuevas proteínas solubles y sus receptores, incluyendo aquellas similares a las linfoquinas, deben contribuir a nuevas terapias para una amplia gama de afecciones degenerativas o anómalas que comprometen directamente o indirectamente el desarrollo, la diferenciación, o la función, p. ej., del sistema inmunitario y/o las células hematopoyéticas. En particular, el descubrimiento y la comprensión de receptores novedosos para las moléculas de tipo linfoquina que intensifican o potencian las actividades beneficiosas de otras linfoquinas resultarían muy ventajosos. La presente invención proporciona nuevos receptores para ligandos que muestran similitud con las composiciones de tipo interleuquina-1 y compuestos relacionados, y sus métodos de uso.

### Breve descripción de los dibujos

La Figura 1 muestra una comparación esquemática de las arquitecturas de las proteínas de los DTLR de *Drosophila* y humanos, y su relación con receptores de IL-1 de vertebrados y proteínas de resistencia a enfermedades de plantas. Se disponen tres DTLR de *Drosophila* (Dm) (Toll, 18w, y el fragmento Mst ORF) (Morisato y Anderson (1995) *Ann. Rev. Genet.* 29:371-399; Chiang y Beachy (1994) *Mech. Develop.* 47:225-239; Mitcham, *et al.* (1996) *J. Biol. Chem.* 271:5777-5783; y Eldon, *et al.* (1994) *Develop.* 120:885-899) al lado de cuatro receptores humanos (Hu) completos (DTLR 1-4) y uno parcial (DTLR5). Los LRR individuales de los ectodominios del receptor que están señalizados mediante PRINTS (Attwood, *et al.* (1997) *Nucleic Acids Res.* 25:212-217) se indican explícitamente por medio de recuadros; las agrupaciones ricas en Cys "superior" e "inferior" que flanquean los extremos C o N terminales de las matrices de LRR se dibujan respectivamente mediante semicírculos opuestos. La pérdida de la región rica en Cys interna de los DTLR 1-5 representa en gran parte sus ectodominios más pequeños (558, 570, 690, y 652 aa, respectivamente) cuando se comparan con las prolongaciones de 784 y 977 aa de Toll y 18w. Las cadenas incompletas de DmMst y HuDTLR5 (ectodominios de 519 y 153 aa, respectivamente) están representadas por líneas discontinuas. El módulo de señalización intracelular común a los DTLR, receptores de tipo IL-1 (IL-1R), la proteína intracelular Myd88, y el producto génico N de resistencia a las enfermedades del tabaco (DRgN) se indica debajo de la membrana. Véanse, p. ej., Hardiman, *et al.* (1996) *Oncogene* 13:2467-2475; y Rock, *et al.* (1998) *Proc. Nat'l Acad. Sci. USA* 95:588-. Los dominios adicionales incluyen el trío de módulos de tipo Ig de los IL-1R (bucles unidos a disulfuro); la proteína DRgN muestra un dominio NTPasa (recuadro) y Myd88 tiene un dominio de muerte (óvalo negro).

Las Figuras 2A-2B muestran patrones estructurales conservados en los dominios de señalización de los receptores de tipo Toll e IL-1, y dos proteínas modulares divergentes. La Figura 2A muestra un alineamiento de secuencia del dominio TH común. Los DTLR se marcan como en la Figura 1; los receptores de la familia de IL-1 humano (Hu) o ratón (Mo) (IL-1R1-6) se numeran sucesivamente como se ha propuesto antes (Hardiman, *et al.* (1996) *Oncogene* 13:2467-2475); Myd88 y las secuencias de tabaco (To) y lino, *L. usitatissimum* (Lu), representan dominios C y N terminales, respectivamente, de moléculas multidominio más grandes. Los bloques de secuencias sin espacios (numerados del 1-10) están recuadrados. Los triángulos indican mutaciones deletéreas, mientras los truncamientos N-terminales de la flecha eliminan la bioactividad en el IL-1R1 humano (Heguy, *et al.* (1992) *J. Biol. Chem.* 267:2605-2609). Se marcan las predicciones de la estructura secundaria de PHD (Rost y Sander (1994) *Proteins* 19:55-72) y DSC (King y Sternberg (1996) *Protein Sci.* 5:2298-2310) de una hélice  $\alpha$  (H), hebra  $\beta$  (E), o bobina (L). El esquema de aminoácidos con sombreado representa restos químicamente similares: hidrófobos, ácidos, alcalinos, Cys, aromáticos, de rotura de la estructura, y pequeños. Los patrones de las secuencias diagnóstico para los IL-1R, DTLR, y el alineamiento completo (ALL) se obtuvieron mediante Consensus con una restricción del 75%. Los símbolos para los subgrupos de aminoácidos son (véase el sitio de internet para más detalle): o, alcohol; l, alifático; ., cualquier aminoácido; a, aromático; c, cargado; h, hidrófobo; -, negativo; p, polar; +, positivo; s, pequeño; u, diminuto; t, de tipo giro. La Figura 2B muestra un diagrama de topología del pliegue del dominio  $\beta/\alpha$  de TH propuesto. La lámina  $\beta$  paralela (con hebras  $\beta$  A-E en forma de triángulos de color amarillo) se observa en su extremo C terminal; las hélices  $\alpha$  (círculos marcados como 1-5) conectan las hebras  $\beta$ ; las conexiones de la cadena están hacia delante (visibles) o hacia atrás (escondidas). Los restos cargados, conservados en el extremo C de la lámina  $\beta$  se indican en color gris (Asp) o como un único resto de color negro (Arg) (véase el texto).

La Figura 3 muestra la evolución de una superfamilia de dominios de señalización. Se utilizó el alineamiento del módulo TH múltiple de la Figura 2A para obtener un árbol filogenético mediante el Método del Vecino más Próximo (Thompson, *et al.* (1994) *Nucleic Acids Res.* 22:4673-4680). Las proteínas se marcaron como en el alineamiento; el árbol se presentó con TreeView.

Las Figuras 4A-4D muestran el mapeo cromosómico mediante FISH de los genes de DTLR humanos. Se hibridaron los cromosomas desnaturalizados de los cultivos sincronizados de linfocitos humanos con sondas de ADN de DTLR biotiniladas para su localización. La asignación de los datos del mapeo con FISH (izquierda, Figuras 4A, DTLR2; 4B, DTLR3; 4C, DTLR4; 4D, DTLR5) con bandas cromosómicas se logró superponiendo las señales FISH con los cromosomas que forman bandas en DAPI (paneles centrales). Heng y Tsui (1994) *Meth. Molec. Biol.* 33:109-122. Los análisis se resumen en forma de ideogramas de cromosomas humanos (paneles de la derecha).

Las Figuras 5A-5F muestran análisis de transferencia de ARNm de DTLR Humanos. Se sondearon transferencias de múltiples tejidos humanos (He, corazón; Br, cerebro; Pl, placenta; Lu, pulmón; Li, hígado; Mu, músculo; Ki, riñón; Pn, Páncreas; Sp, bazo; Th, timo; Pr, próstata; Te, testículo; Ov, ovario; SI, intestino delgado; Co, colon; PBL, linfocitos de sangre periférica) y línea de células cancerosas (leucemia promielocítica, HL60; cáncer cervical, HELAS3;

leucemia mielógena crónica, K562; leucemia linfoblástica, Molt4; adenocarcinoma colorrectal, SW480; melanoma, G361; Línea celular Raji de linfoma de Burkitt, Burkitt's; adenocarcinoma colorrectal, SW480; carcinoma de pulmón, A549) que contenían aproximadamente 2 µg de ARN poli(A)+ por calle con ADNc radiomarcados que codificaban DTLR1 (Figuras 5A-5C), DTLR2 (Figura 5D), DTLR3 (Figura 5E), y DTLR4 (Figura 5F) como se ha descrito. Las transferencias se expusieron a una película de rayos X durante 2 días (Figuras 5A-5C) o una semana (Figura 5D-5F) a -70°C con pantallas intensificadoras. En algunas calles aparece una especie de 0,3 kB anómala; los experimentos de hibridación excluyen un mensaje que codifica un fragmento citoplásmico de DTLR.

### Compendio de la invención

La presente invención está dirigida a una proteína o péptido esencialmente puros o recombinantes con actividad de receptor de tipo Toll, donde dicha proteína o péptido muestran una identidad de secuencia de al menos 70% con el SEQ ID NO: 6 en toda su longitud.

En la presente memoria se describen nueve receptores de mamífero relacionados, p. ej., estructuras moleculares de tipo receptor Toll, denominadas DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9, y DTLR10, y sus actividades biológicas. Asimismo se describen ácidos nucleicos que codifican los propios polipéptidos y los métodos para su producción y uso. Los ácidos nucleicos están caracterizados, en parte, por su homología con las secuencias de ADN complementario clonado (ADNc) incluidas en la presente memoria.

Asimismo se describe una composición de materia seleccionada del grupo de: una proteína o péptido de n DTLR2 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 4; un DTLR2 de secuencia natural de SEQ ID NO: 4; una proteína de fusión que comprende una secuencia de DTLR2; una proteína o péptido de DTLR3 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 6; un DTLR3 de secuencia natural de SEQ ID NO: 6; una proteína de fusión que comprende una secuencia de DTLR3; una proteína o péptido de DTLR4 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 26; un DTLR4 de secuencia natural de SEQ ID NO: 26; una proteína de fusión que comprende una secuencia de DTLR4; una proteína o péptido de DTLR5 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 10; un DTLR5 de secuencia natural de SEQ ID NO: 10; y una proteína de fusión que comprende la secuencia de DTLR5.

Asimismo se describe una composición de materia seleccionada del grupo de: una proteína o péptido de DTLR6 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 12; un DTLR6 de secuencia natural de SEQ ID NO: 12; una proteína de fusión que comprende la secuencia de DTLR6; una proteína o péptido de DTLR7 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 16 o 18 o; un DTLR7 de secuencia natural de SEQ ID NO: 16 o 18; una proteína de fusión que comprende la secuencia de DTLR7; una proteína o péptido de DTLR8 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 32; un DTLR8 de secuencia natural de SEQ ID NO: 32; una proteína de fusión que comprende la secuencia de DTLR8; una proteína o péptido de DTLR9 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 22; un DTLR9 de secuencia natural de SEQ ID NO: 22; y una proteína de fusión que comprende la secuencia de DTLR9; una proteína o péptido de DTLR10 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 34; un DTLR10 de secuencia natural de SEQ ID NO: 34; y una proteína de fusión que comprende la secuencia de DTLR10.

Preferiblemente, la proteína esencialmente pura o aislada comprende un segmento que muestra una identidad de secuencia con una porción correspondiente de un DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR 7, DTLR8, DTLR9, o DTLR10, donde: la homología es una identidad de al menos aproximadamente 90% y la porción es de al menos aproximadamente 9 aminoácidos; la homología es una identidad de al menos aproximadamente 80% y la porción es de al menos aproximadamente 17 aminoácidos; o la homología es una identidad de al menos aproximadamente 70% y la porción es de al menos aproximadamente 25 aminoácidos. En realizaciones específicas, la composición de materia: es DTLR2, que comprende una secuencia madura de SEQ ID NO: 4; o muestra un patrón de modificación post-traducciona distinto del de DTLR2 natural; es DTLR3, que comprende una secuencia madura de SEQ ID NO: 6; o muestra un patrón de modificación post-traducciona distinto del de DTLR3 natural; es DTLR4, que comprende una secuencia madura de SEQ ID NO: 26; o muestra un patrón de modificación post-traducciona distinto del de DTLR4 natural; o es DTLR5, que comprende la secuencia completa de SEQ ID NO: 10; o muestra un patrón de modificación post-traducciona distinto del de DTLR5 natural; o es DTLR6, que comprende una secuencia madura de SEQ ID NO: 12; o muestra un patrón de modificación post-traducciona distinto del de DTLR6 natural; es DTLR7, que comprende una secuencia madura de SEQ ID NO: 16 o 18; o muestra un patrón de modificación post-traducciona distinto del de DTLR7 natural; es DTLR8, que comprende una secuencia madura de SEQ ID NO: 32; o muestra un patrón de

modificación post-traducciona distinto del de DTLR8 natural; o es DTLR9, que comprende la secuencia completa de SEQ ID NO: 22; o muestra un patrón de modificación post-traducciona distinto del de DTLR9 natural; o es DTLR10, que comprende la secuencia completa de SEQ ID NO: 34; o muestra un patrón de modificación post-traducciona distinto del de DTLR10 natural; o la composición de materia puede ser una proteína o péptido que es de un animal de sangre caliente seleccionado entre un mamífero, incluyendo un primate, tal como un ser humano; comprende al menos un segmento polipeptídico de SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34; muestra una pluralidad de porciones que muestran dicha identidad; es una variante alélica natural de DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9, o DTLR10; tiene una longitud de al menos aproximadamente 30 aminoácidos; muestra al menos dos epítomos no solapantes que son específicos para un DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9, o DTLR10 de primate; muestra una identidad de secuencia de al menos aproximadamente 90% a lo largo de una longitud de al menos aproximadamente 20 aminoácidos con un DTLR2, DTLR3, DTLR4, DTLR5, DTLR6 de primate; muestra al menos dos epítomos no solapantes que son específicos para un DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9, o DTLR10 de primate; muestra una identidad de secuencia de al menos aproximadamente 20 aminoácidos con un DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9, o DTLR10 de primate; está glicosilado; tiene un peso molecular de al menos 100 kD con glicosilación natural; es un polipéptido sintético; está anclado a un sustrato sólido; está conjugado con otro radical químico; tiene 5 sustituciones o menos con respecto a la secuencia natural; o es una variante por delección o inserción a partir de una secuencia natural.

Asimismo se describe una composición que comprende: una proteína o péptido de DTLR2 estéril; o la proteína o péptido de DTLR2 y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral; una proteína o péptido de DTLR3 estéril; o la proteína o péptido de DTLR3 y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral; una proteína o péptido de DTLR4 estéril; o la proteína o péptido de DTLR4 y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral; una proteína o péptido de DTLR5 estéril; o la proteína o péptido de DTLR5 estéril y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral; una proteína o péptido de DTLR6 estéril; o la proteína o péptido de DTLR6 y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral; una proteína o péptido de DTLR7 estéril; o la proteína o péptido de DTLR7 y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral; una proteína o péptido de DTLR8 estéril; o la proteína o péptido de DTLR8 y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral; una proteína o péptido de DTLR9 estéril; o la proteína o péptido de DTLR9 y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral; una proteína o péptido de DTLR10 estéril; o la proteína o péptido de DTLR10 y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral.

Asimismo se describe una proteína de fusión que comprende la secuencia de la proteína madura de SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34; una etiqueta de detección o purificación, incluyendo una secuencia FLAG, His6, o Ig; o la secuencia de otra proteína receptora.

Asimismo se describe un kit que comprende una proteína o polipéptido de DTLR, y: un compartimento que comprende la proteína o polipéptido; y/o instrucciones para su uso o para la eliminación de los reactivos del kit.

Asimismo se describen compuestos de unión que comprenden un sitio de unión al antígeno de un anticuerpo, que se une específicamente a una proteína de DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9, o DTLR10 natural, donde: la proteína es una proteína de primate; el compuesto de unión es un fragmento Fv, Fab, o Fab2; el compuesto de unión se conjuga con otro radical químico; o el anticuerpo: se origina contra una secuencia peptídica de un polipéptido maduro de SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34; se origina contra un DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9 o DTLR10 maduro; se origina contra un DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9 o DTLR10 humano purificado; se inmunoselecciona; es un anticuerpo policlonal; se une a un DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9 o DTLR10 desnaturalizado; muestra una Kd para el antígeno de al menos 30  $\mu$ M; se ancla a un sustrato sólido, incluyendo una cuenta o membrana de plástico; está en una composición estéril; o está marcado detectablemente, incluyendo una marca radiactiva o fluorescente. Un kit de una composición de unión comprende a menudo el compuesto de unión, y: un compartimento que comprende dicho compuesto de unión; y/o instrucciones para el uso y eliminación de los reactivos del kit. A menudo el kit es capaz de realizar un análisis cualitativo o cuantitativo.

Otras composiciones incluyen una composición que comprende: un compuesto de unión estéril, o el compuesto de unión y un portador, donde el portador es: un compuesto acuoso, incluyendo agua, solución salina, y/o tampón; y/o se formula para la administración oral, rectal, nasal, tópica, o parenteral.

Los ácidos nucleicos descritos en la presente memoria incluyen un ácido nucleico aislado o recombinante que codifica una proteína o péptido de DTLR2-10 o una proteína de fusión, donde: el DTLR es de un mamífero; o el ácido

nucleico: codifica una secuencia de un péptido antigénico de SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34; codifica un pluralidad de secuencias de péptidos antigénicos de SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34; muestra una identidad de al menos aproximadamente 80% con un ADNc natural que codifica dicho segmento; es un vector de expresión; comprende adicionalmente un origen de replicación; es de una fuente natural; comprende una marca detectable; comprende una secuencia de nucleótidos sintética; tiene menos de 6 kb, preferiblemente menos de 3 kb; es de un mamífero, incluyendo un primate; comprende una secuencia codificante completa natural; es una sonda de hibridación para un gen que codifica dicho DTLR; o es un cebador de PCR, un producto de PCR, o un cebador de mutagénesis. Asimismo se proporciona una célula, tejido u órgano que comprende semejante ácido nucleico recombinante. Preferiblemente, la célula es: una célula procariótica; una célula eucariótica; una célula bacteriana; una célula de levadura; una célula de insecto; una célula de mamífero; una célula de ratón; una célula de primate; o una célula humana. Se proporcionan kits que comprenden tales ácidos nucleicos, y: un compartimento que comprende dicho ácido nucleico; un compartimento que comprende una proteína o polipéptido de DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9 o DTLR10 de primate; y/o instrucciones para el uso o eliminación de los reactivos del kit. A menudo, el kit es capaz de realizar un análisis cualitativo o cuantitativo.

Asimismo se describe un ácido nucleico que: hibrida en condiciones de lavado de 30°C y concentración de sal menor de 2M para el SEQ ID NO: 3; hibrida en condiciones de lavado de 30°C y concentración de sal menor de 2 M para el SEQ ID NO: 5; hibrida en condiciones de lavado de 30°C y concentración de sal menor de 2M para el SEQ ID NO: 25; hibrida en condiciones de lavado de 30°C y concentración de sal menor de 2 M para el SEQ ID NO: 9; hibrida en condiciones de lavado de 30°C y concentración de sal menor de 2M para el SEQ ID NO: 11; hibrida en condiciones de lavado de 30°C y concentración de sal menor de 2 M para el SEQ ID NO: 15 o 17; hibrida en condiciones de lavado de 30°C y concentración de sal menor de 2M para el SEQ ID NO: 31; hibrida en condiciones de lavado de 30°C y concentración de sal menor de 2 M para el SEQ ID NO: 21; hibrida en condiciones de lavado de 30°C y concentración de sal menor de 2 M para el SEQ ID NO: 33; muestra una identidad de al menos aproximadamente 85% a lo largo de un tramo de al menos aproximadamente 30 nucleótidos con un DTLR2, DTLR3, DTLR4, VPLR5, DTLR6, DTLR7, DTLR8, DTLR9 o DTLR10 de primate.

Preferiblemente, semejante ácido nucleico tendrá propiedades en las que: condiciones de lavado a 45°C y/o concentración de sal 500 mM; o la identidad es de al menos 90% y/o el tramo es de al menos 55 nucleótidos. Muy preferiblemente, las condiciones de lavado son a 55°C y/o concentración de sal 150 mM; o la identidad es de al menos 95% y/o el tramo es de al menos 75 nucleótidos.

Asimismo se describe un método de modulación de la fisiología o desarrollo de una célula o células de un cultivo de tejidos que comprende poner en contacto la célula con un agonista o antagonista de un DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9, o DTLR10 de mamífero.

## Descripción detallada de las realizaciones preferidas

### I. General

En la presente memoria se describen la secuencia de aminoácidos y la secuencia de ADN de mamífero, en la presente memoria las moléculas de receptor de tipo Toll de ADNc de primate (DTLR) tienen propiedades definidas concretas, tanto estructurales como biológicas. Estas han sido denominadas en la presente memoria DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9, y DTLR10, respectivamente, y aumentan el número de miembros de la familia de receptores de tipo Toll humanos de 1 a 10. Los diferentes ADNc que codifican estas moléculas fueron obtenidos de primate, p. ej., genotecas de secuencias de ADNc, humanas. También se desearían contrapartes de otros primates u otros mamíferos.

Algunos de los métodos convencionales aplicables son descritos o referidos, p. ej., por Maniatis, *et al.* (1982) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Laboratory, Cold Spring Harbor Press; Sambrook, *et al.* (1989) Molecular Cloning: A Laboratory Manual, (2ª ed.), vols 1-3, CSH Press, NY; Ausubel, *et al.*, Biology, Greene Publishing Associates, Brooklyn, NY; o Ausubel, *et al.* (1987 y suplementos periódicos) Current Protocols in Molecular Biology, Greene/Wiley, Nueva York.

En los SEQ ID NO: 1 y 2 se muestran una secuencia de nucleótidos completa y la correspondiente secuencia de aminoácidos de un segmento codificante de DTLR1 humano. Véase también Nomura, *et al.* (1994) DNA Res 1:27-35. En los SEQ ID NO: 3 y 4 se muestran una secuencia de nucleótidos completa y la correspondiente secuencia de aminoácidos de un segmento codificante de DTLR2 humano. En los SEQ ID NO: 5 y 6 se muestran una secuencia de nucleótidos completa y la secuencia de aminoácidos correspondiente de un DTLR3 humano. En los SEQ ID NO: 7 y 8 se muestran una secuencia de nucleótidos completa y la secuencia de aminoácidos correspondiente de un segmento codificante de DTLR4 humano. En los SEQ ID NO: 25 y 26 se proporcionan una secuencia de ácido nucleico y la secuencia de aminoácidos correspondiente de un segmento codificante de DTLR4 humano. En los SEQ ID NO: 9 y 10 se muestran una secuencia de nucleótidos parcial y la secuencia de aminoácidos correspondiente de un segmento codificante de DTLR5 humano. En los SEQ ID NO: 11 y 12 se muestran una secuencia de nucleótidos completa y la correspondiente secuencia de aminoácidos de un segmento codificante de DTLR6 humano y en los SEQ ID NO: 13 y 14 se proporciona una secuencia parcial de un DTLR6 de ratón. En los SEQ ID NO: 27 y 29 se proporciona una secuencia de DTLR6 de ratón adicional (secuencia de nucleótidos) y en los SEQ ID NO: 28 y 30 (secuencias de aminoácidos). También se proporciona una secuencia de nucleótidos parcial (SEQ ID NO: 15 y 17) y la correspondiente secuencia de aminoácidos (SEQ ID NO: 16 y 18) de un segmento codificante de DTLR7 humano. En los SEQ

## ES 2 340 210 T3

ID NO: 19 y 20 se muestran la secuencia de nucleótidos parcial y la correspondiente secuencia de aminoácidos de un segmento codificante de DTLR8 humano. En los SEQ ID NO: 31 y 32 se muestran una secuencia de nucleótidos más completa y la correspondiente secuencia de aminoácidos de un segmento codificante de DTLR humano. En el SEQ ID NO: 21 se muestran una secuencia de nucleótidos parcial y la correspondiente secuencia de aminoácidos de un segmento codificante de DTLR9 humano y en los SEQ ID NO: 23 y 24 se muestran una secuencia de nucleótidos parcial y la correspondiente secuencia de aminoácidos de un segmento codificante de DTLR10 humano. En los SEQ ID NO: 33 y 34 se muestra una secuencia de nucleótidos más completa y la correspondiente secuencia de aminoácidos de un segmento codificante de DTLR10 humano. En el SEQ ID NO: 35 se proporciona una secuencia de nucleótidos parcial para un segmento codificante de DTLR10 de ratón.

TABLA 1

*Comparación de dominios intracelulares de DTLR humanos. El DTLR1 es el SEQ ID NO: 2; el DTLR2 es el SEQ ID NO: 4; el DTLR3 es el SEQ ID NO: 6; el DTLR4 es el SEQ ID NO: 8; el DTLR5 es el SEQ ID NO: 10; y el DTLR6 es el SEQ ID NO: 12. Los restos particularmente importantes y conservados, p. ej., característicos, corresponden, a través de los DTLR, a los restos del SEQ ID NO: 18 tyr10-tyr13; trp26; cys46; trp52; pro54-gly55; ser69; lys71; trp134-pro135; y phe144-trp145*

DTLR1	QRNLQFHAFISYSGHD---SFVVKNELLPNLEKEG-----MQICLHERNF
DTLR9	KENLQFHAFISYSEHD---SAWVKSELVPYLEKED-----IQICLHERNF
DTLR8	-----NELIPNLEKEDGS---ILICLYESYF
DTLR2	SRNICYDAFVSYSERD---AYWVENLMVQELFNPP---FKLCLHKRDF
DTLR6	SPDCCYDAFIVYDTKDPVTEWVLAELVAKLEDPREK--HFNLCLLEERDW
DTLR7	TSQTFYDAYISYDTKDAVTDWVINELRYHLEESRDK--NVLLCLEERDW
DTLR10	EDALPYDAFVVDKTXSAVADWVYNELRGQLEECRGRW-ALRLCLEERDW
DTLR4	RGENIYDAFVIYSSQD---EDWVRNELVKNLEEGVPP---FQLCLHYRDF
DTLR5	PDMYKYDAYLCFSSKD---FTWVQNALLKHLDTQYSDQNRFNLCFEERDF
DTLR3	TEQFEYAAYIIHAYKD---KDWVWEHFSMEKEDQS---LKFCLEERDF
	: . . : * :
DTLR1	VPGKSIVENIITC-IEKSYKSIFVLSPNFVQSEWCH-YELYFAHHNLFHE
DTLR9	VPGKSIVENIINC-IEKSYKSIFVLSPNFVQSEWCH-YELYFAHHNLFHE
DTLR8	VPGKSISENVSF-IEKSYKSIFVLSPNFVQNEWCH-YEFYFAHHNLFHE
DTLR2	IPGKWIIDNIIDS-IEKSHKTVFVLSENFVKSEWCK-YELDFSHFRLFEE
DTLR6	LPGQPVLENLSQS-IQLSKKTVFVMTDKYAKTENFK-IAFYLSHQRLMDE
DTLR7	DPGLAIIDNLMQS-INQSKKTVFVLTKKYAKSWNFK-TAFYLLXQLRMGE
DTLR10	LPGKTLFENLWAS-VYGSRKTLFVLAHTDRVSGLLR-AIFLLAQQRLL-
DTLR4	IPGVAIAANIIEGHFKSRKVIVVVSQHFIQSRWCI-FEYETIAQTWQFLS
DTLR5	VPGENRIANIQDA-IWNSRKIVCLVSRHFLRDGWCL-EAFSYAQGRCLSD
DTLR3	EAGVFELEAIVNS-IKRSRKIFVITHLLKDPCKRKFVHHAVQQAIEQ
	. * : . * * : : :
DTLR1	GSNSLILILLEPIPQYSIPSSYHKLKSLMARRTYLEWPKEKSKRGLFWAN
DTLR9	GSNLIILILLEPIQNSIPNKYHKLKALMTQRTYLQWPKEKSKRGLFWA-
DTLR8	NSDHIILILLEPIPFYCIPTRYHKLKLEALLEKKAYLEWPKDRRKCGLFWAN
DTLR2	NNDAAIILILLEPIEKKAIPQRFCKLRKIMNTKTYLEWPMDEAOREGFVWN
DTLR6	KVDVILIFLEKPFQK---SKFLQLRKRLCGSSVLEWPTNPQAHFYFWQC
DTLR7	NMDVILIFILLEPVLQH---SPYLRLRQRICKSSILQWPDNPKAERLFWQT
DTLR10	-----
DTLR4	SRAGIIFIVLQKVEKT-LLRQQVELYRLLSRNTYLEWEDSVLGRHIFWRR
DTLR5	LNSALIMVVVGSLSQY-QLMKHQSIRGFVQKQOYLWPEDLQDVGWFLHK
DTLR3	NLDSIILVLFLEIIPDYKLNHALCLRRGMFKSHCILNWPVQKERIGAFRHK
DTLR1	LRAAINIKLTEQAKK-----
DTLR9	-----
DTLR8	LRAAVNVNVLATREMYELQTFTELNEESRGSTISLMRTDCL
DTLR2	LRAAIKS-----
DTLR6	LKNALATDNHVAYSQVFKETV-----
DTLR7	LXNVVLTENDSRYNMYVDSIKQY-----
DTLR10	-----
DTLR4	LRKALLDGKSWNPEGTVGTCGNWQEATSI-----
DTLR5	LSQQILKKEKEKKDNNIPLQTVATIS-----
DTLR3	LQVALGSKNSVH-----

Según se utiliza en la presente memoria, receptor 2 de tipo Toll de ADN (DTLR2) se utilizará para describir una proteína que comprende un segmento de proteína o péptido que tiene o comparte la secuencia de aminoácidos mostrada en el SEQ ID NO: 4, o uno de sus fragmentos esenciales. De un modo similar, para el DTLR3 y el SEQ ID NO: 6; el DTLR4 y el SEQ ID NO: 26; el DTLR5 y el SEQ ID NO: 10; el DTLR6 y el SEQ ID NO: 12; el DTLR7 y los SEQ ID NO: 16 y 18; el DTLR8 y el SEQ ID NO: 32; el DTLR9 y SEQ ID NO: 22; y el DTLR10 y el SEQ ID NO: 34.

También se describen variaciones de proteínas del respectivo alelo de DTLR cuya secuencia se proporciona, p. ej., un agonista o antagonista de muteína. Típicamente, tales agonistas o antagonistas mostrarán menos de aproximadamente 10% de diferencias en la secuencia, y de este modo a menudo tienen entre 1 y 11 sustituciones, p. ej., 2, 3, 5, 7, y otras. También abarca variantes alélicas y otras, p. ej., formas polimórficas naturales de la proteína descrita. Típicamente, se unirá a su correspondiente receptor biológico con una alta afinidad, p. ej., al menos aproximadamente 100 nM, normalmente más de aproximadamente 30 nM, preferiblemente más de aproximadamente 10 nM, y más preferiblemente más de aproximadamente 3 nM. El término también se utilizará en la presente memoria para referirse a formas naturales afines, p. ej., alelos, variantes polimórficas, y variantes metabólicas de la proteína de mamífero.

También se describen proteínas o péptidos que tienen una identidad de aminoácidos sustancial con la secuencia de aminoácidos del SEQ ID NO: 4. También incluirá variantes de la secuencia con relativamente pocas sustituciones, p. ej., preferiblemente menos de aproximadamente 3-5. Se aplican características similares a las otras secuencias de DTLR proporcionadas en los SEQ ID NO: 6, 26, 10, 12, 16, 18, 32, 22 y 34.

Un “fragmento” o “segmento” de polipéptido sustancial, es un tramo de restos aminoácido de al menos aproximadamente 8 aminoácidos, generalmente al menos 10 aminoácidos, más generalmente al menos 12 aminoácidos, a menudo al menos 14 aminoácidos, más a menudo al menos 16 aminoácidos, típicamente al menos 18 aminoácidos, más típicamente al menos 20 aminoácidos, normalmente al menos 22 aminoácidos, más normalmente al menos 24 aminoácidos, preferiblemente al menos 26 aminoácidos, más preferiblemente al menos 28 aminoácidos, y, en realizaciones particularmente preferidas, al menos aproximadamente 30 o más aminoácidos. Las secuencias de los segmentos de las diferentes proteínas se pueden comparar entre sí a lo largo de tramos de longitud apropiada.

La homología de la secuencia de aminoácidos, o la identidad de la secuencia, se determina optimizando los emparejamientos de restos, si fuera necesario, introduciendo espacios cuando se requiera. Véanse, p. ej., Needleham, *et al.*, (1970) J. Mol. Biol. 48:443-453; Sankoff, *et al.*, (1983) capítulo uno en Time Warps, String Edits and Macromolecules: The Theory and Practice of Sequence Comparison, Addison-Wesley, Reading, MA; y paquetes de soporte lógico de IntelliGenetics, Mountain View, CA; y the University of Wisconsin Genetics Computer Group (GCG), Madison, WI; cada uno de los cuales se incorpora en la presente memoria como referencia. Esto cambia cuando se consideran las sustituciones conservativas como emparejamientos. Las sustituciones conservativas incluyen típicamente sustituciones con los siguientes grupos: glicina, alanina; valina, isoleucina, leucina; ácido aspártico, ácido glutámico; asparagina, glutamina; serina, treonina; lisina, arginina; y fenilalanina, tirosina. Se pretende que las secuencias de aminoácidos homólogas incluyan variaciones alélicas naturales e interespecie en la secuencia de las citoquinas. Las proteínas o péptidos homólogos típicos tendrán una homología de 50-100% (si se pueden introducir espacios), a una homología de 60-100% (si se incluyen sustituciones conservativas) con un segmento de la secuencia de aminoácidos de los SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34. Las mediciones de la homología serán de al menos aproximadamente 70%, generalmente al menos 76%, más generalmente al menos 81%, a menudo al menos 85%, más a menudo al menos 88%, típicamente al menos 90%, más típicamente al menos 92%, normalmente al menos 94%, más normalmente al menos 95%, preferiblemente al menos 96%, y más preferiblemente al menos 97%, y en realizaciones particularmente preferidas, al menos 98% o más. El grado de homología variará con la longitud de los segmentos comparados. Las proteínas o péptidos homólogos, tales como las variantes alélicas, compartirán la mayoría de las actividades biológicas con las realizaciones descritas en los SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34. Las regiones de comparación particularmente interesantes, a nivel de aminoácidos o nucleótidos, corresponden a aquellas dentro de cada uno de los bloques 1-10, o regiones intrabloque, correspondientes a las indicadas en la Figura 2A.

Según se utiliza en la presente memoria, el término “actividad biológica” se utiliza para describir, sin limitación, los efectos de las respuestas inflamatorias, la inmunidad innata, y/o el desarrollo morfogénico por los respectivos ligandos. Por ejemplo, estos receptores deben, como los receptores de IL-1, mediar las actividades fosfatasa o fosforilasa, cuyas actividades son medidas fácilmente mediante procedimientos convencionales. Véanse, p. ej., Hardie, *et al.* (eds. 1995) The Protein Kinase FactBook vols. I y II, Academic Press, San Diego, CA; Hanks, *et al.* (1991) Meth. Enzymol. 200:38-62; Hunter, *et al.* (1992) Cell 70:375-388; Lewin (1990) Cell 61:743-752; Pines, *et al.* (1991) Cold Spring Harbor Symp. Quant. Biol. 56:449-463; y Parker, *et al.* (1993) Nature 363:736-738. Los receptores muestran actividades biológicas como las de las enzimas regulables, reguladas por la unión al ligando. No obstante, el número de recambio de la enzima está más próximo a una enzima que a un complejo receptor. Por otra parte, el número de receptores ocupados necesario para inducir semejante actividad enzimática es menor que en la mayoría de los sistemas receptores, y puede aproximarse a docenas por célula, en contraste con la mayoría de los receptores que se desencadenarán a un número de miles por célula. Los receptores, o sus porciones, pueden ser útiles como enzimas de marcaje con fosfato para marcar sustratos generales o específicos.

Los términos ligando, agonista, antagonista, y análogo de, p. ej., un DTLR, incluyen moléculas que modulan las respuestas celulares características a las proteínas de tipo ligando Toll, así como moléculas que poseen los rasgos de competición por la unión estructural más convencional de las interacciones ligando-receptor, p. ej., donde el receptor es



un receptor natural o un anticuerpo. Las respuestas celulares están mediadas probablemente por la unión de diferentes ligandos Toll a receptores celulares relacionados con, pero posiblemente distintos, de los receptores de IL-1 de tipo I o de tipo II. Véanse, p. ej., Belvin y Anderson (1996) *Ann. Rev. Cell Dev. Biol.* 12:393-416; Morisato y Anderson (1995) *Ann. Rev. Genetics* 29:371-3991 y Hultmark (1994) *Nature* 367:116-117.

Asimismo, un ligando es una molécula que sirve o bien como ligando natural al cual dicho receptor, o uno de sus análogos, se une, o una molécula que es un análogo funcional del ligando natural. El análogo funcional puede ser un ligando con modificaciones estructurales, o puede ser una molécula no relacionada en absoluto con los determinantes de unión del ligando apropiado. Los ligandos pueden servir como agonistas o antagonistas, véase, p. ej., Goodman, *et al.* (eds) (1990) *Goodman & Gilman's: The Pharmacological Bases of Therapeutics*, Pergamon Press, Nueva York.

El diseño racional de fármacos también puede estar basado en estudios estructurales de las formas moleculares de un receptor o anticuerpo y otros efectores o ligandos. Los efectores pueden ser otras proteínas que median otras funciones en respuesta a la unión al ligando, u otras proteínas que interaccionan normalmente con el receptor. Un método para determinar qué sitios interaccionan con otras proteínas específicas es una determinación de la estructura física, p. ej., cristalografía de rayos x o técnicas de RMN bidimensional. Estas proporcionarán las pautas en cuanto a qué restos aminoácido forman regiones de contacto molecular. Para una descripción detallada de la determinación estructural de proteínas, véase, p. ej., Blundell y Johnson (1976) *Protein Crystallography*, Academic Press, Nueva York.

## II. Actividades

Las proteínas receptoras de tipo Toll tendrán un número diferente de actividades biológicas, p. ej., en el metabolismo del fosfato, siendo añadidas o separadas de sustratos específicos, típicamente proteínas. Esto dará como resultado generalmente la modulación de una función inflamatoria, otra respuesta inmunitaria innata, o un efecto morfológico. Las proteínas de DTLR2, 3, 4, 5, 6, 7, 8, 9, o 10 son homólogas a otras proteínas receptoras de tipo Toll, pero cada una tiene diferencias estructurales. Por ejemplo, la secuencia codificante del gen de DTLR2 humano tiene probablemente una identidad de aproximadamente 70% con la secuencia de nucleótidos codificante del DTLR2 de ratón. A nivel de aminoácidos, también es probable que haya una identidad razonable.

Las actividades biológicas de los DTLR estarán relacionadas con la adición o eliminación de radicales fosfato de sustratos, típicamente de una manera específica, pero ocasionalmente de una manera no específica. Los sustratos pueden ser identificados, o las condiciones para la actividad enzimática pueden ser analizadas mediante métodos convencionales, p. ej., como describen Hardie, *et al.* (eds. 1995) en *The Protein Kinase FactBook* vols. I y II, Academic Press, San Diego, CA; Hanks, *et al.* (1991) *Meth. Enzymol.* 200:38-62; Hunter, *et al.* (1992) *Cell* 70:375-388; Lewin (1990) *Cell* 61:743-752; Pines, *et al.* (1991) *Cold Spring Harbor Symp. Quant. Biol.* 56:449-463; y Parker, *et al.* (1993) *Nature* 363:736-738.

## III. Ácidos Nucleicos

Esta invención contempla el uso de ácido nucleico aislado o fragmentos, p. ej., que codifican estas proteínas o proteínas íntimamente relacionadas, o sus fragmentos, p. ej., para codificar el correspondiente polipéptido, preferiblemente uno que sea biológicamente activo. Además, esta invención incluye ADN aislado o recombinante que codifica tales proteínas o polipéptidos que tienen secuencias características de los respectivos DTLR, individualmente o como grupo. Típicamente, el ácido nucleico es capaz de hibridar, en condiciones apropiadas, con un segmento de una secuencia de ácido nucleico mostrado en los SEQ ID NO: 3, 5, 25, 9, 11, 15, 17, 31, 21, o 33, pero preferiblemente no con un segmento correspondiente del SEQ ID NO: 1. Dicha proteína o polipéptido biológicamente activo puede ser una proteína completa, o un fragmento, y tendrá típicamente un segmento de una secuencia de aminoácidos altamente homóloga a la mostrada en los SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34. Adicionalmente, esta invención abarca el uso de un ácido nucleico aislado o recombinante, o sus fragmentos, que codifican proteínas que son equivalentes a las proteínas de DTLR2-10. Los ácidos nucleicos aislados pueden tener las respectivas secuencias reguladoras en los extremos 5' y 3', p. ej., promotores, intensificadores, señales de adición poli-A, y otras del gen natural.

Un ácido nucleico "aislado" es un ácido nucleico, p. ej., un ARN, ADN, o una mezcla de polímeros, que son esencialmente puros, p. ej., separados de otros componentes que acompañan naturalmente a la secuencia nativa, tales como ribosomas, polimerasas, y secuencias genómicas limítrofes de las especies de origen. El término abarca una secuencia de ácido nucleico que ha sido separada de su entorno natural, e incluye productos aislados de ADN recombinante o clonado, que son distinguibles de ese modo de las composiciones de origen natural, y análogos sintetizados químicamente o análogos sintetizados biológicamente por sistemas heterólogos. Una molécula esencialmente pura incluye las formas aisladas de la molécula, ya sean completamente o esencialmente puras.

Un ácido nucleico aislado será generalmente una composición homogénea de moléculas, pero, en algunas realizaciones, contendrá una heterogeneidad, preferiblemente minoritaria. Esta heterogeneidad se encuentra típicamente en los extremos del polímero o en porciones no críticas para la función o actividad biológica deseada.

Un ácido nucleico “recombinante” se define típicamente por su método de producción o su estructura. En referencia a su método de producción, p. ej., se elabora un producto por medio de un procedimiento, cuyo procedimiento consiste en el uso de técnicas de ácido nucleico recombinante, que implican la intervención humana en la secuencia de nucleótidos. Típicamente esta intervención implica la manipulación *in vitro*, aunque en ciertas circunstancias puede implicar técnicas de cría de animales más clásicas. Alternativamente, puede ser un ácido nucleico elaborado generando una secuencia que comprende la fusión de dos fragmentos que no son naturalmente contiguos entre sí, pero se pretende excluir productos de la naturaleza, p. ej., mutantes de origen natural encontrados en su estado natural. De este modo, por ejemplo, están incluidos los productos elaborados transformando células con cualquier vector de origen no natural, como los ácidos nucleicos que comprenden una secuencia derivada utilizando cualquier procedimiento de síntesis de oligonucleótidos. Semejante procedimiento se realiza a menudo para remplazar un codón con un codón redundante que codifica el mismo aminoácido o un aminoácido conservativo, a la vez que se introduce o se elimina típicamente un sitio de reconocimiento para una secuencia de una enzima de restricción. Alternativamente, el procedimiento se realiza para unir entre sí segmentos de ácido nucleico con funciones deseadas para generar una única entidad genética que comprende una combinación deseada de funciones no encontrada en las formas disponibles en la naturaleza, p. ej., una que codifica una proteína de fusión. Los sitios de reconocimiento para las enzimas de restricción son a menudo la diana de semejantes manipulaciones artificiales, pero mediante el diseño se pueden incorporar otras dianas específicas del sitio, p. ej., promotores, sitios de replicación de ADN, secuencias de regulación, secuencias de control, u otros rasgos útiles. Se pretende un concepto similar para un polipéptido recombinante, p. ej., de fusión. Esto incluirá una repetición dimérica. Se incluyen específicamente ácidos nucleicos sintéticos que, mediante redundancia del código genético, codifican polipéptidos equivalentes para fragmentos de DTLR2-10 y fusiones de secuencias de diversas moléculas relacionadas diferentes, p. ej., otros miembros de la familia de receptores de IL-1.

Un “fragmento” en un contexto de ácido nucleico es un segmento contiguo de al menos aproximadamente 17 nucleótidos, generalmente al menos 21 nucleótidos, más generalmente al menos 25 nucleótidos, normalmente al menos 30 nucleótidos, más normalmente al menos 35 nucleótidos, a menudo al menos 39 nucleótidos, más a menudo al menos 45 nucleótidos, típicamente al menos 50 nucleótidos, más típicamente al menos 55 nucleótidos, normalmente al menos 60 nucleótidos, más normalmente al menos 66 nucleótidos, preferiblemente al menos 72 nucleótidos, más preferiblemente al menos 79 nucleótidos, y en realizaciones particularmente preferidas será de al menos 85 o más nucleótidos. Típicamente, los fragmentos de secuencias genéticas diferentes se pueden comparar entre sí a lo largo de tramos de longitud apropiada, concretamente segmentos definidos tales como los dominios descritos más abajo.

Un ácido nucleico que codifica un DTLR2-10 será particularmente útil para identificar los genes, el ARNm, y las especies de ADNc que codifican el mismo o proteínas íntimamente relacionadas, así como ADN que codifican variantes polimórficas, alélicas, u otras variantes genéticas, p. ej., de diferentes individuos o de especies relacionadas. Las sondas preferidas para tales escrutinios son aquellas regiones de la interleuquina que están conservadas entre las diferentes variantes polimórficas o que contienen nucleótidos que carecen de especificidad, y serán preferiblemente completas o casi. En otras situaciones, las secuencias específicas de las variantes polimórficas serán más útiles.

Adicionalmente se describen moléculas de ácido nucleico recombinante y fragmentos que tienen una secuencia de ácido nucleico idéntica o muy homóloga al grupo de ADN aislados mostrados en la presente memoria. En particular, las secuencias estarán a menudo conectadas operablemente a segmentos de ADN que controlan la transcripción, traducción, y replicación del ADN. Estos segmentos adicionales ayudan típicamente a la expresión del segmento de ácido nucleico deseado.

Las secuencias de ácido nucleico homólogas, o altamente homólogas, cuando se comparan entre sí o las secuencias mostradas en los SEQ ID NO: 3, 5, 25, 9, 11, 15, 17, 31, 21, o 33 muestran una similitud significativa. Los patrones de homología en los ácidos nucleicos son medidas de la homología utilizadas generalmente en la técnica mediante comparación de secuencias o basándose en las condiciones de hibridación. Las condiciones de hibridación comparativas se muestran con mayor detalle más abajo.

La identidad sustancial en el contexto de la comparación de las secuencias de ácidos nucleicos significa que o bien los segmentos, o bien sus hebras complementarias, cuando se comparan son idénticas cuando se alinean óptimamente, con inserciones o deleciones de nucleótidos apropiadas, en al menos aproximadamente 60% de los nucleótidos, generalmente al menos 66%, normalmente al menos 71%, a menudo al menos 76%, más a menudo al menos 80%, normalmente al menos 84%, más normalmente al menos 88%, típicamente al menos 91%, más típicamente al menos aproximadamente 93%, preferiblemente al menos aproximadamente 95%, más preferiblemente al menos aproximadamente 96 a 98% o más, y en realizaciones concretas, tanto como aproximadamente 99% o más de los nucleótidos, incluyendo, p. ej., los segmentos que codifican los dominios estructurales tales como los segmentos descritos más abajo. Alternativamente, existirá una identidad sustancial cuando los segmentos hibriden en condiciones de hibridación selectivas, con una hebra o su complemento, típicamente utilizando una secuencia derivada de los SEQ ID NO: 3, 5, 25, 9, 11, 15, 17, 31, 21, o 33. Típicamente, la hibridación selectiva se producirá cuando haya una homología de al menos aproximadamente 55% a lo largo de un tramo de al menos aproximadamente 14 nucleótidos, más típicamente al menos aproximadamente 65%, preferiblemente al menos aproximadamente 75%, y más preferiblemente al menos aproximadamente 90%. Véase, Kanehisa (1984) Nuc. Acids Res. 12:203-213.

La longitud de la comparación por homología, como se ha descrito, puede ser a lo largo de tramos más largos, y en ciertas realizaciones será a lo largo de un tramo de al menos aproximadamente 17 nucleótidos, generalmente al menos aproximadamente 20 nucleótidos, normalmente al menos aproximadamente 24 nucleótidos, normalmente al menos

aproximadamente 28 nucleótidos, típicamente al menos aproximadamente 32 nucleótidos, más típicamente al menos aproximadamente 40 nucleótidos, preferiblemente al menos aproximadamente 50 nucleótidos, y más preferiblemente al menos aproximadamente 75 a 100 o más nucleótidos.

- 5 Las condiciones restrictivas, en referencia a la homología en el contexto de la hibridación, serán condiciones restrictivas combinadas de sal, temperatura, disolventes orgánicos, y otros parámetros controlados típicamente en las reacciones de hibridación. Las condiciones restrictivas de temperatura incluirán normalmente temperaturas de más de aproximadamente 30°C, más normalmente de más de aproximadamente 37°C, típicamente de más de aproximadamente 45°C, más típicamente de más de aproximadamente 55°C, preferiblemente de más de aproximadamente 65°C, y más  
10 preferiblemente de más de aproximadamente 70°C. Las condiciones restrictivas de sal serán normalmente de menos de aproximadamente 500 mM, normalmente menos de aproximadamente 400 mM, más normalmente menos de aproximadamente 300 mM, típicamente menos de aproximadamente 200 mM, preferiblemente menos de aproximadamente 100 mM, y más preferiblemente menos de aproximadamente 80 mM, incluso bajará a menos de aproximadamente 20 mM. No obstante, la combinación de parámetros es mucho más importante que la medida de cualquier parámetro  
15 individual. Véase, p. ej., Wetmur y Davidson (1968) J. Mol. Biol. 31:349-370.

- Alternativamente, para la comparación de secuencias, típicamente una secuencia actúa como secuencia de referencia, con la cual se comparan las secuencias de ensayo. Cuando se utiliza un algoritmo de comparación de la secuencia, se introducen en un ordenador las secuencias de ensayo y de referencia, se diseñan las coordenadas de la subsecuencia,  
20 si fuera necesario, y se diseñan los parámetros del programa del algoritmo de la secuencia. Después el algoritmo de comparación de la secuencia calcula el porcentaje de identidad de la secuencia para la secuencia o las secuencias de ensayo con respecto a la secuencia de referencia, basándose en los parámetros del programa designado.

- El alineamiento óptimo de secuencias para la comparación se puede realizar, p. ej., mediante el algoritmo de homología local de Smith y Waterman (1981) Adv. Appl. Math. 2:482, mediante el algoritmo de alineamiento por homología de Needleman y Wunsch (1970) J. Mol. Biol. 48:443, mediante el método de búsqueda de la similitud de Pearson y Lipman (1988) Proc. Nat'l Acad. Sci. USA 85:2444, mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA del Paquete de Soporte Lógico de Wisconsin Genetics Software, Genetics Computer Group, 575 Science Dr., Madison, WI), o mediante inspección visual (véase generalmente Ausubel  
30 *et al.*, *supra*).

- Un ejemplo de un algoritmo útil es PILEUP. PILEUP crea un alineamiento de secuencia múltiple a partir de un grupo de secuencias relacionadas utilizando alineamientos por pares, progresivos para mostrar la relación y el porcentaje de identidad de la secuencia. También traza un árbol o dendrograma que muestra las relaciones de agrupamiento utilizadas para crear el alineamiento. PILEUP utiliza una simplificación del método de alineamiento progresivo de Feng y Doolittle (1987) J. Mol. Evol. 35:351-360. El método utilizado es similar al método descrito por Higgins y Sharp (1989) CABIOS 5:151-153. El programa puede alinear hasta 300 secuencias, cada una de una longitud máxima de 5.000 nucleótidos o aminoácidos. El procedimiento de alineamiento múltiple comienza con el alineamiento por pares de las dos secuencias más similares, produciendo un agrupamiento de dos secuencias alineadas. Este agrupamiento  
40 es alineado después con la siguiente secuencia más relacionada o agrupación de secuencias alineadas. Se alinean dos agrupaciones de secuencias mediante una simple prolongación del alineamiento por pares de dos secuencias individuales. El alineamiento final se logra mediante una serie de alineamientos por pares, progresivo. El programa se lleva a cabo diseñando secuencias específicas y sus coordenadas de aminoácidos o nucleótidos para las regiones de comparación de la secuencia y diseñando los parámetros del programa. Por ejemplo, se puede comparar una secuencia de  
45 referencia con otras secuencias de ensayo para determinar la relación del porcentaje de identidad de la secuencia utilizando los siguientes parámetros: ponderación del espacio por defecto (3,00), ponderación de la longitud del espacio por defecto (0,10), y espacios finales ponderados.

- Otro ejemplo de algoritmo que es adecuado para determinar el porcentaje de identidad de la secuencia y de similitud de la secuencia es el algoritmo BLAST, que es descrito por Altschul, *et al.* (1990) J. Mol. Biol. 215:403-410. El soporte lógico para realizar los análisis BLAST es asequible públicamente a través del National Center for Biotechnology Information (<http://www.ncbi.nlm.nih.gov/>). Este algoritmo implica identificar primero pares de secuencias de puntuación elevada (HSP) identificando palabras cortas de longitud W en la secuencia problema, que se emparejan o satisfacen alguna puntuación umbral de valor positivo T cuando se alinean con una palabra de la misma longitud en una secuencia de una base de datos. T es referido como el umbral de puntuación de la palabra cercana (Altschul, *et al.*, *supra*). Estos éxitos de la palabra cercana inicial actúan como semillas para iniciar las búsquedas para encontrar HSP más largas que los contengan. Los éxitos de las palabras se extienden después en ambas direcciones a lo largo de cada secuencia tanto como se pueda incrementar la puntuación del alineamiento acumulativo. La extensión de los éxitos de las palabras en cada dirección se detiene cuando: la puntuación del alineamiento acumulativo cae fuera de la cantidad X de su valor máximo logrado; la puntuación acumulativa se va a cero o menos, debido a la acumulación de uno o más alineamientos de restos de puntuación negativa; o se alcanza el extremo de cada secuencia. Los parámetros del algoritmo BLAST W, T, y X determinan la sensibilidad y velocidad del alineamiento. El programa BLAST utiliza por defecto una longitud de palabra (W) de 11, la matriz de puntuación BLOSUM62 (véase Henikoff y Henikoff (1989) Proc. Nat'l Acad. Sci. USA 89:10915) alineamientos (B) de 50, una esperanza (E) de 10, M=5, N=4, y una  
65 comparación de ambas hebras.

Además de calcular el porcentaje de identidad de la secuencia, el algoritmo BLAST también realiza un análisis estadístico de la similitud entre dos secuencias (véase, p. ej., Karlin y Altschul (1993) Proc. Nat'l Acad. Sci. USA

90:5873-5787). Una medida de la similitud proporcionada por el algoritmo BLAST es la probabilidad de la suma más pequeña ( $P(N)$ ), que proporciona una indicación de la probabilidad por la cual se produciría un emparejamiento entre dos secuencias de nucleótidos o aminoácidos por casualidad. Por ejemplo, un ácido nucleico se considera similar a una secuencia de referencia si la probabilidad de la suma más pequeña en una comparación del ácido nucleico de ensayo con el ácido nucleico de referencia es menor de aproximadamente 0,1, más preferiblemente menor de aproximadamente 0,01, y muy preferiblemente menor de aproximadamente 0,001.

Una indicación adicional de que dos secuencias de ácido nucleico de polipéptidos son esencialmente idénticas es que el polipéptido codificado por el primer ácido nucleico experimenta reacción inmunológica cruzada con el polipéptido codificado por el segundo ácido nucleico, como se describe más abajo. De este modo, un polipéptido es típicamente esencialmente idéntico a un segundo polipéptido, p. ej., cuando los dos péptidos difieren solamente en las sustituciones conservativas. Otra indicación de que dos secuencias de ácido nucleico son esencialmente idénticas es que las dos moléculas hibridan entre sí en condiciones restrictivas, como se describe más abajo.

El ADN aislado puede ser fácilmente modificado mediante sustituciones de nucleótidos, deleciones de nucleótidos, inserciones de nucleótidos, e inversiones de tramos de nucleótidos. Estas modificaciones dan como resultado secuencias de ADN novedosas que codifican esta proteína o sus derivados. Estas secuencias modificadas se pueden utilizar para producir proteínas mutantes (muteínas) o para intensificar la expresión de especies variantes. La intensificación de la expresión puede implicar amplificación génica, aumento de la transcripción, aumento de la traducción, y otros mecanismos. Tales derivados de tipo DTLR mutantes incluyen mutaciones predeterminadas o específicas del sitio de la proteína o sus fragmentos, incluyendo mutaciones silenciosas utilizando la degeneración del código genético. "DTLR mutante" según se utiliza en la presente memoria incluye un polipéptido que de otro modo entra en la definición de homología del DTLR como se ha expuesto antes, pero que tiene una secuencia de aminoácidos que difiere de la de otras proteínas de tipo DTLR encontradas en la naturaleza, ya sea mediante delección, sustitución, o inserción. En particular, "el DTLR mutante específico del sitio" incluye una proteína que tiene una homología sustancial con una proteína del SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34, y típicamente comparte la mayoría de las actividades o efectos biológicos de las formas descritas en la presente memoria.

Aunque las mutaciones específicas del sitio son predeterminadas, no es necesario que los mutantes sean específicos del sitio. La mutagénesis de los DTLR de mamífero se puede lograr realizando inserciones o deleciones de aminoácidos en el gen, acopladas con la expresión. Se pueden generar sustituciones, deleciones, inserciones, o cualquier combinación para llegar a un constructo final. Las inserciones incluyen fusiones amino- o carboxi terminales. La mutagénesis al azar se puede llevar a cabo en un codón diana y los mutantes de DTLR de mamífero expresados pueden ser escrutados en busca de la actividad deseada. Los métodos para realizar mutaciones por sustitución en sitios predeterminados en el ADN que tiene una secuencia conocida son bien conocidos en la técnica, p. ej., mediante mutagénesis con cebadores de M13. Véase también Sambrook, *et al.* (1989) y Ausubel, *et al.* (1987 y Suplementos periódicos).

Las mutaciones en el ADN no deben situar normalmente secuencias codificantes fuera de los marcos de lectura y preferiblemente no crearán regiones complementarias que puedan hibridar para producir estructuras de ARNm secundario tales como bucles u horquillas.

El método de la fosforamidita descrito por Beaucage y Carruthers (1981) Tetra. Letts. 22:1859-1862, producirá fragmentos de ADN sintético adecuados. A menudo se obtendrá un fragmento de doble hebra sintetizando la hebra complementaria y recociendo las hebras juntas en condiciones apropiadas o añadiendo la hebra complementaria utilizando la ADN polimerasa con una secuencia cebadora apropiada.

Las técnicas de reacción en cadena de la polimerasa (PCR) a menudo pueden ser aplicadas en la mutagénesis. Alternativamente, los cebadores de mutagénesis son los métodos utilizados comúnmente para generar mutaciones definidas en sitios predeterminados. Véase, p. ej., Innis, *et al.* (eds. 1990) PCR Protocols: A Guide to Methods and Applications Academic Press, San Diego, CA; y Dieffenbach y Dveksler (1995; eds.) PCR Primer: A Laboratory Manual Cold Spring Harbor Press, CSH, NY.

#### IV. Proteínas, Péptidos

El DTLR-10 de primate, p. ej., cuyas secuencias se describen en los SEQ ID NOS: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34, se ha descrito antes. Asimismo se contemplan las variantes alélicas y otras, incluyendo, p. ej., proteínas de fusión que combinan porciones de tales secuencias con otras, que incluyen etiquetas epitópicas y dominios funcionales.

Asimismo se describen proteínas recombinantes, p. ej., proteínas de fusión heterólogas utilizando segmentos de estas proteínas de roedor. Una proteína de fusión heteróloga es una fusión de proteínas o segmentos que naturalmente no se fusionan normalmente de la misma manera. De este modo, el producto de fusión de un DTLR con un receptor de IL-1 es una molécula de proteína continua que tiene secuencias fusionadas en un enlace peptídico típico, típicamente elaborada como un único producto de la traducción y que muestra propiedades, p. ej., la secuencia o la antigenicidad, derivadas de cada péptido de origen. Se aplica un concepto similar a las secuencias de ácido nucleico heterólogas.

Además, se pueden elaborar nuevos constructos a partir de la combinación de dominios funcionales o estructurales similares de otras proteínas relacionadas, p. ej., receptores de IL-1 u otros DTLR, incluyendo variantes de especie. Por ejemplo, la unión al ligando u otros segmentos puede ser “intercambiada” entre diferentes polipéptidos o fragmentos de fusión nuevos. Véanse, p. ej., Cunningham, *et al.* (1989) *Science* 243:1330-1336; y O’Dowd, *et al.* (1988) *J. Biol. Chem.* 263:15985-15992.

De este modo, de la unión funcional de especificidades de unión al receptor, resultarán nuevos polipéptidos químicos que muestran nuevas combinaciones de especificidades. Por ejemplo, se pueden añadir los dominios de unión al ligando de otras moléculas receptoras relacionadas o sustituir por otros dominios de esta proteína o proteínas relacionadas. La proteína resultante a menudo tendrá funciones y propiedades híbridas. Por ejemplo, una proteína de fusión puede incluir un dominio de redireccionamiento que puede servir para proporcionar el secuestro de la proteína de fusión por un orgánulo subcelular concreto.

Los patrones y las secuencias de fusión candidato se pueden seleccionar de diferentes bases de datos de secuencias, p. ej., GenBank, c/o IntelliGenetics, Mountain View, CA; y BCG, University of Wisconsin Biotechnology Computing Group, Madison, WI.

Asimismo se describen muteínas que se unen a ligandos Toll, y/o que son afectadas en la transducción de la señal. El alineamiento estructural de los DTLR1-10 humanos con otros miembros de la familia de IL-1 muestra características/restos conservados. Véase, p. ej., la Figura 3A. El alineamiento de las secuencias de los DTLR humanos con otros miembros de la familia de IL-1 indica diferentes rasgos estructurales y funcionales compartidos. Véase también, Bazan, *et al.* (1996) *Nature* 379:591; Lodi, *et al.* (1994) *Science* 263:1762-1766; Sayle y Milner-White (1995) *TIBS* 20:374-376; y Gronenberg, *et al.* (1991) *Protein Engineering* 4:263-269.

Los ligandos de IL-1 $\alpha$  e IL-1 $\beta$  se unen a un receptor de IL-1 de tipo I como receptor primario y este complejo forma después un complejo receptor de alta afinidad con el receptor de IL-1 de tipo III. Tales subunidades del receptor son compartidas probablemente con los nuevos miembros de la familia de IL-1.

Las variaciones similares en otras contrapartes de especies de las secuencias de DTLR2-10, p. ej., en las regiones correspondientes, deben proporcionar interacciones similares con el ligando o el sustrato. Son particularmente preferidas las sustituciones con secuencias de ratón o secuencias humanas. En cambio, las sustituciones conservativas lejos de las regiones de interacción de unión al ligando conservarán probablemente la mayoría de las actividades de señalización.

Los “derivados” de los DTLR-10 de primate incluyen mutantes de la secuencia de aminoácidos, variantes de glicosilación, derivados metabólicos y productos conjugados covalentes o agregativos con otros radicales químicos. Los derivados covalentes se pueden preparar mediante conexión de funcionalidades a grupos que se encuentran en las cadenas laterales de los aminoácidos de DTLR o en los extremos N o C, p. ej., mediante métodos que son bien conocidos en la técnica. Estos derivados pueden incluir, sin limitación, ésteres o amidas alifáticos del extremo carboxilo, o de restos que contienen cadenas laterales carboxílicas, derivados O-acilados de restos que contienen grupos hidroxilo, y derivados N-acilados del aminoácido amino terminal o restos que contienen grupos amino, p. ej., lisina o arginina. Los grupos acilo se seleccionan del grupo de radicales alquilo incluyendo alquilo C3 a C18 normal, formando de ese modo especies alcanoil-arólicas.

En particular, se incluyen las alteraciones de la glicosilación, p. ej., realizadas modificando los patrones de glicosilación de un polipéptido durante su síntesis y procesamiento, o en etapas de procesamiento adicionales. Los métodos particularmente preferidos para lograrlo son mediante exposición del polipéptido a enzimas glicosilantes derivadas de células que normalmente proporcionan semejante procesamiento, p. ej., enzimas de glicosilación de mamífero. También se contemplan enzimas de desglicosilación. Asimismo se incluyen las versiones de la misma secuencia de aminoácidos primaria que tienen otras modificaciones menores, incluyendo restos aminoácido fosforilados, p. ej., fosfotirosina, fosfoserina, o fosfotreonina.

Un grupo principal de derivados son los conjugados covalentes de los receptores o sus fragmentos con otras proteínas o polipéptidos. Estos derivados se pueden sintetizar en un cultivo recombinante tal como fusiones N o C terminales o mediante el uso de agentes conocidos en la técnica por su utilidad en el entrecruzamiento de proteínas a través de grupos laterales reactivos. Los sitios de derivatización preferidos con agentes de entrecruzamiento están en los grupos amino libres, los radicales carbohidratados, y los restos cisteína.

También se proporcionan los polipéptidos de fusión entre los receptores y otras proteínas homólogas o heterólogas. Los polipéptidos homólogos pueden ser fusiones entre diferentes receptores, que dan como resultado, por ejemplo, una proteína híbrida que muestra especificidad de unión para múltiples ligandos Toll diferentes, o un receptor que puede tener una especificidad ampliada o debilitada del efecto del sustrato. Del mismo modo se pueden construir fusiones heterólogas que muestran una combinación de propiedades o actividades de las proteínas derivadas. Los ejemplos típicos son las fusiones de un polipéptido informador, p. ej., luciferasa, con un segmento o dominio de un receptor, p. ej., un segmento de unión al ligando, de manera que se pueda determinar fácilmente la presencia o localización de un ligando deseado. Véase, p. ej., Dull, *et al.*, Patente de los Estados Unidos Núm. 4.859.609, que se incorpora en la presente memoria como referencia. Otros patrones de fusión génica incluyen glutatión-S-transferasa (GST),  $\beta$ -galactosidasa bacteriana, trpE, Proteína A,  $\beta$ -lactamasa, alfa amilasa, alcohol deshidrogenasa, y factor de apareamiento alfa de levadura. Véase, p. ej., Godowski, *et al.* (1988) *Science* 241:812-816.

El método de la fosforamidita descrito por Beaucage y Carruthers (1981) Tetra. Letts. 22:1859-1862, producirá fragmentos de ADN sintético adecuados. A menudo se obtendrá un fragmento de doble hebra sintetizando la hebra complementaria y recociendo las hebras juntas en condiciones apropiadas o añadiendo la hebra complementaria utilizando ADN polimerasa con una secuencia cebadora apropiada.

Tales polipéptidos también pueden tener restos aminoácido que han sido modificados químicamente mediante fosforilación, sulfonación, biotinilación, o adición o eliminación de otros radicales, concretamente aquellos que tiene conformaciones moleculares similares a grupos fosfato. En algunas realizaciones, las modificaciones serán reactivos de marcaje útiles, o servirán como dianas de purificación, p. ej., ligandos de afinidad.

Las proteínas de fusión se elaborarán típicamente mediante métodos de ácidos nucleicos recombinantes o mediante métodos con polipéptidos sintéticos. Las técnicas para la manipulación y expresión de ácido nucleico son descritas generalmente, por ejemplo, por Sambrook, *et al.* (1989) Molecular Cloning: A Laboratory Manual (2ª ed.), Vols. 1-3, Cold Spring Harbor Laboratory, y Ausubel, *et al.* (eds. 1987 y suplementos periódicos) Current Protocols in Molecular Biology, Greene/Wiley, Nueva York, que se incorporan en la presente memoria como referencia. Las técnicas para la síntesis de polipéptidos son descritas, por ejemplo, por Merrifield (1963) J. Amer. Chem. Soc. 85:2149-2156; Merrifield (1986) Science 232: 341-347; y Atherton, *et al.* (1989) Solid Phase Peptide Synthesis: A Practical Approach, IRL Press, Oxford; cada una de las cuales se incorpora en la presente memoria como referencia. Véase también Dawson, *et al.* (1994) Science 266:776-779 para los métodos para elaborar polipéptidos más grandes.

Asimismo se contempla el uso de derivados de un DTLR-10 distintos de las variaciones en la secuencia de aminoácidos o la glicosilación. Tales derivados pueden implicar la asociación covalente o agregativa con radicales químicos. Estos derivados generalmente se dividen en tres clases: (1) sales, (2) modificaciones covalentes de la cadena lateral y el resto terminal, y (3) complejos de adsorción, por ejemplo con membranas celulares. Tales derivados covalentes o agregativos son útiles como inmunógenos, como reactivos en inmunoanálisis, o en métodos de purificación por ejemplo para la purificación por afinidad de un receptor u otra molécula de unión, p. ej., un anticuerpo. Por ejemplo, un ligando Toll puede ser inmovilizado mediante unión covalente en un soporte sólido tal como Sefarosa activada con bromuro de cianógeno, mediante métodos que son bien conocidos en la técnica, o adsorbido sobre superficies de poliolefinas, con o sin entrecruzamiento con glutaraldehído, para su uso en el análisis o purificación de un receptor DTLR, anticuerpos, u otras moléculas similares. El ligando también puede ser marcado con un grupo detectable, por ejemplo radioyodado mediante el procedimiento de la cloramina T, unido covalentemente a quelatos de tierras raras, o conjugado con otro radical fluorescente para su uso en análisis de diagnóstico.

Se puede utilizar un DTLR de esta invención como inmunógeno para la producción de antisueros o anticuerpos específicos, p. ej., capaces de distinguir entre otros miembros de la familia de receptores de IL-1, para los DTLR o sus diferentes fragmentos. El DTLR purificado se puede utilizar para escrutar anticuerpos monoclonales o fragmentos de unión al antígeno preparados mediante inmunización con diferentes formas de preparaciones impuras que contienen la proteína. En particular, el término "anticuerpos" también incluye fragmentos de unión al antígeno de anticuerpos naturales, p. ej., Fab, Fab2, Fv, etc. El DTLR purificado también se puede utilizar como reactivo para detectar anticuerpos generados en respuesta a la presencia de niveles elevados de expresión, o trastornos inmunológicos que conducen a la producción de anticuerpo para el receptor endógeno. Adicionalmente, los fragmentos de DTLR también pueden servir como inmunógenos para producir los anticuerpos de la presente invención, descritos inmediatamente más abajo. Por ejemplo, se contemplan anticuerpos que tienen afinidad de unión o que han sido originados contra las secuencias de aminoácidos mostradas en los SEQ ID NOS: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34, sus fragmentos, o diferentes homólogos peptídicos. En particular, se contemplan anticuerpos que tienen afinidad de unión para, o que han sido originados contra, fragmentos específicos que se pronostica que están, o realmente están, expuestos a la superficie exterior de la proteína del DTLR nativo.

El bloqueo de la respuesta fisiológica a los ligandos del receptor puede resultar de la inhibición de la unión del ligando al receptor, probablemente por medio de inhibición competitiva. De este modo, los análisis *in vitro* a menudo utilizarán anticuerpos o segmentos de unión a antígenos de estos anticuerpos, o fragmentos anclados a sustratos en fase sólida. Estos análisis también permitirán la determinación diagnóstica de los efectos de cualquiera de las mutaciones y modificaciones en la región de unión al ligando, u otras mutaciones y modificaciones, p. ej., que afecten a la señalización o a la función enzimática.

Asimismo se contempla el uso de análisis de escrutinio de fármacos competitivos, p. ej., en los que anticuerpos neutralizadores del receptor o sus fragmentos compiten con un compuesto de ensayo por la unión a un ligando u otro anticuerpo. De esta manera, se pueden utilizar los anticuerpos neutralizadores o fragmentos para detectar la presencia de un polipéptido que comparte uno o más sitios de unión a un receptor y también se pueden utilizar para ocupar sitios de unión sobre un receptor que podría de otro modo unirse a un ligando.

## V. Elaboración de Ácidos Nucleicos y Proteína

El ADN que codifica la proteína o sus fragmentos puede ser obtenido mediante síntesis química, escrutinio de genotecas de ADNc, o escrutinio de genotecas genómicas preparadas a partir de una amplia variedad de líneas celulares o muestras de tejidos. Las secuencias naturales pueden ser aisladas utilizando métodos convencionales y las secuencias

proporcionadas en la presente memoria. Se pueden identificar otras contrapartes de especies mediante técnicas de hibridación, o mediante diferentes técnicas de PCR, combinadas o mediante búsqueda en bases de datos de secuencias, p. ej., GenBank.

5 Este ADN puede ser expresado en una amplia variedad de células anfitrionas para la síntesis de un receptor completo o fragmentos que pueden a su vez ser utilizados, por ejemplo para generar anticuerpos policlonales o monoclonales; para estudios de unión; para la construcción y expresión de dominios de unión a ligandos o quinasas/fosfatasa modificados; y para estudios de estructura/función. Las variantes o fragmentos pueden ser expresados en células anfitrionas que son transformadas o transfectadas con vectores de expresión apropiados. Estas moléculas pueden estar esencialmente  
10 libres de contaminantes proteicos o celulares, distintos de los derivados del anfitrión recombinante, y por lo tanto son particularmente útiles en las composiciones farmacéuticas cuando se combinan con un portador y/o diluyente farmacéuticamente aceptable. La proteína, o sus porciones, pueden ser expresadas como fusiones con otras proteínas.

15 Los vectores de expresión son típicamente constructos de ADN o ARN auto-replicantes que contienen el gen del receptor deseado o sus fragmentos, normalmente conectado operablemente a elementos de control genético adecuados que son reconocidos en una célula anfitriona adecuada. Estos elementos de control son capaces de llevar a cabo la expresión en un anfitrión adecuado. El tipo específico de elementos de control necesarios para llevar a cabo la expresión dependerá de la célula anfitriona eventual utilizada. Generalmente, los elementos de control genético pueden incluir un sistema promotor procariótico o un sistema de control de la expresión de un promotor eucariótico, y típicamente incluyen un promotor transcripcional, un operador opcional para controlar el comienzo de la transcripción, intensificadores de la transcripción para elevar el nivel de expresión del ARNm, una secuencia que codifica un sitio de unión al ribosoma adecuado, y secuencias que terminan la transcripción y la traducción. Los vectores de expresión también contienen normalmente un origen de replicación que permite que el vector replique independientemente de la célula anfitriona.

25 Los vectores de esta invención incluyen aquellos que contienen ADN que codifica una proteína, como se ha descrito, o uno de sus fragmentos que codifican un polipéptido equivalente biológicamente activo. El ADN puede estar bajo el control de un promotor viral y puede codificar un marcador de selección. Esta invención contempla adicionalmente el uso de tales vectores de expresión que son capaces de expresar ADNc eucariótico que codifica semejante proteína en un anfitrión procariótico o eucariótico, donde el vector es compatible con el anfitrión y donde el ADNc eucariótico  
30 que codifica el receptor está insertado en el vector de manera que el crecimiento del anfitrión que contiene el vector expresa el ADNc en cuestión. Normalmente, se diseñan vectores de expresión para la replicación estable en sus células anfitrionas o para la amplificación para incrementar enormemente el número total de copias del gen deseable por célula. No siempre se requiere que un vector de expresión replique en una célula anfitriona, p. ej., es posible efectuar la expresión transitoria de la proteína o sus fragmentos en diferentes anfitriones utilizando vectores que no contengan un  
35 origen de replicación que sea reconocido por la célula anfitriona. También es posible utilizar vectores que ocasionen la integración de la proteína que codifica la porción o sus fragmentos en el ADN del anfitrión mediante recombinación.

Los vectores, utilizados en la presente memoria, comprenden plásmidos, virus, bacteriófagos, fragmentos de ADN integrables, y otros vehículos que permiten la integración de fragmentos de ADN en el genoma del anfitrión. Los  
40 vectores de expresión son vectores especializados que contienen elementos de control genético que llevan a cabo la expresión de genes conectados operablemente. Los plásmidos son la forma de vector más comúnmente utilizada pero todas las demás formas de vectores que sirven para una función equivalente y que son, o se vuelven, conocidos en la técnica son adecuados para su uso en la presente memoria. Véanse, p. ej., Pouwels, *et al.* (1985 y Suplementos) Cloning Vectors: A Laboratory Manual, Elsevier, N.Y., y Rodriguez, *et al.* (eds) Vectors: A Survey of Molecular Cloning  
45 Vectors and Their Uses, Butterworth, Boston, 1988, que se incorporan en la presente memoria como referencia.

Las células transformadas son células, preferiblemente de mamífero, que han sido transformadas o transfectadas con vectores de receptores contruidos utilizando técnicas de ADN recombinante. Las células anfitrionas transformadas expresan normalmente la proteína deseada o sus fragmentos, pero para la clonación, amplificación, y manipulación  
50 de su ADN, no se necesita que expresen la proteína sujeto. Esta invención contempla adicionalmente el cultivo de células transformadas en un medio nutriente, permitiendo de ese modo que el receptor se acumule en la membrana celular. La proteína puede ser recuperada, o bien del cultivo o bien, en ciertos casos, del medio de cultivo.

Para los fines de esta invención, las secuencias de ácido nucleico están conectadas operablemente cuando están  
55 funcionalmente relacionadas entre sí. Por ejemplo, el ADN para una presecuencia o un líder secretor está conectado operablemente a un polipéptido si éste es expresado en forma de preproteína o participa en el direccionamiento del polipéptido a la membrana celular o en la secreción del polipéptido. Un promotor está conectado operablemente a una secuencia codificante si controla la transcripción del polipéptido; un sitio de unión al ribosoma está conectado operablemente a una secuencia codificante si está situado para permitir la traducción. Normalmente, conectado  
60 operablemente significa contiguo y en el marco de lectura, no obstante, ciertos elementos genéticos tales como los genes represores no están conectados de manera contigua pero todavía se unen a secuencias operadoras que a su vez controlan la expresión.

Las células anfitrionas adecuadas incluyen procariotas, eucariotas inferiores, y eucariotas superiores. Los procariotas incluyen organismos tanto gram negativos como gram positivos, p. ej., *E. coli* y *B. subtilis*. Los eucariotas inferiores incluyen levaduras, p. ej., *S. cerevisiae* y *Pichia*, y especies del género *Dictyostelium*. Los eucariotas superiores incluyen líneas celulares de cultivos de tejido establecidas de células de animales, tanto de origen no mamífero, p. ej., células de insecto, y aves, como de origen mamífero, p. ej., seres humanos, primates, y roedores.

Los sistemas anfitrión procariótico-vector incluyen una amplia variedad de vectores para muchas especies diferentes. Según se utiliza en la presente memoria, se utilizarán genéricamente *E. coli* y sus vectores para incluir vectores equivalentes utilizados en otros procariotas. Un vector representativo para amplificar ADN es pBR322 o muchos de sus derivados. Los vectores que se pueden utilizar para expresar el receptor o sus fragmentos incluyen, pero no están limitados a, vectores tales como aquellos que contienen el promotor lac (serie pUC); el promotor trp (pBR322-trp); el promotor Ipp (la serie pIN); los promotores lambda-pP o pR (pOTS); o promotores híbridos tales como ptac (pDR540). Véase Brosius, *et al.* (1988) "Expression Vectors Employing Lambda-, trp-, lac-, and Ipp-derived Promoters", en *Vectors: A Survey of Molecular Cloning Vectors and Their Uses*, (eds. Rodríguez y Denhardt), Butterworth, Boston, Capítulo 10, págs. 205-236.

Los eucariotas inferiores, p. ej., levaduras y *Dictyostelium*, pueden ser transformados con vectores que contienen la secuencia de DTLR. Para los fines de esta invención, el anfitrión eucariótico inferior más común es la levadura panadera, *Saccharomyces cerevisiae*. Ésta se utilizará para representar genéricamente eucariotas inferiores aunque también se encuentran disponibles otras numerosas cepas y especies. Los vectores de levadura consisten típicamente en un origen de replicación (excepto el tipo integrante), un gen de selección, un promotor, un ADN que codifica el receptor o sus fragmentos, y secuencias para la terminación de la traducción, poliadenilación, y terminación de la transcripción. Los vectores de expresión adecuados para levadura incluyen promotores constitutivos tales como el de la 3-fosfoglicerato quinasa y otros promotores de genes de enzimas glicolíticas diferentes o promotores inducibles tales como el promotor de la alcohol deshidrogenasa 2 o el promotor de la metalotioneína. Los vectores adecuados incluyen derivados de los siguientes tipos: de bajo número de copias auto-replicante (tal como la serie YRp), de elevado número de copias auto-replicante (tal como la serie YE<sub>p</sub>); de tipo integrante (tal como la serie YIp), o mini-cromosomas (tal como la serie YC<sub>p</sub>).

Las células de cultivos de tejidos eucarióticos superiores son normalmente las células anfitrionas preferidas para la expresión de la proteína interleuquina funcionalmente activa. En principio, es factible cualquier línea celular de cultivo de tejido eucariótico superior, p. ej., sistemas de expresión de baculovirus en insectos, ya sean de origen invertebrado o vertebrado. No obstante, se prefieren células de mamífero. La transformación o transfección y propagación de tales células se ha convertido en un procedimiento rutinario. Los ejemplos de las líneas celulares útiles incluyen células HeLa, líneas celulares de ovario de hámster Chino (CHO), líneas celulares de riñón de cría rata (BRK), líneas celulares de insecto, líneas celulares de aves, y líneas celulares de mono (COS). Los vectores de expresión para tales líneas celulares incluyen normalmente un origen de replicación, un promotor, un sitio de inicio de la traducción, sitios de empalme del ARN (si se utiliza ADN genómico), un sitio de poliadenilación, un sitio de terminación de la transcripción. Estos vectores también contienen normalmente un gen de selección o un gen de amplificación. Los vectores de expresión adecuados pueden ser plásmidos, virus, o retrovirus que portan promotores derivados, p. ej., de fuentes tales como adenovirus, SV40, parvovirus, virus vaccinia, o citomegalovirus. Los ejemplos representativos de los vectores de expresión adecuados incluyen pADNC1; pCD, véase Okayama, *et al.* (1985) *Mol. Cell Biol.* 5:1136-1142; pMC1neo PolyA, véase Thomas, *et al.* (1987) *Cell* 51:503-512; y un vector de baculovirus tal como pAC 373 o pAC 610.

Para las proteínas secretadas, un marco de lectura abierto codifica normalmente un polipéptido que consiste en un producto maduro o secretado unido covalentemente en su extremo N a un péptido señal. El péptido señal es escindido antes de la secreción, del polipéptido maduro, o activo. El sitio de escisión puede ser pronosticado con un grado elevado de exactitud a partir de las reglas empíricas, p. ej., von-Heijne (1986) *Nucleic Acids Research* 14:4683-4690, y la composición de aminoácidos precisa del péptido señal no parece ser crítica para su función, p. ej., Randall, *et al.* (1989) *Science* 243:1156-1159; Kaiser *et al.* (1987) *Science* 235:312-317.

A menudo se deseará expresar estos polipéptidos en un sistema que proporcione un patrón de glicosilación específico o definido. En este caso, el patrón habitual será aquél proporcionado naturalmente por el sistema de expresión. No obstante, el patrón será modificable exponiendo el polipéptido, p. ej., una forma no glicosilada, a proteínas glicosilantes adecuadas introducidas en un sistema de expresión heterólogo. Por ejemplo, el gen receptor puede ser co-transformado con uno o más genes que codifiquen enzimas de glicosilación de mamífero u otras enzimas de glicosilación. Utilizando este enfoque, serán alcanzables ciertos patrones de glicosilación de mamíferos en células procariotas u otras células.

La fuente de DTLR puede ser un anfitrión eucariótico o procariótico que exprese un DTLR recombinante, tal como se ha descrito antes. La fuente también puede ser una línea celular tal como fibroblastos Swiss 3T3 de ratón, pero también se pueden contemplar otras líneas celulares de mamífero en esta invención, siendo la línea celular preferida de la especie humana.

Ahora que se conocen las secuencias, se pueden preparar DTLR de primate, sus fragmentos, o derivados mediante procedimientos convencionales para sintetizar péptidos. Estos incluyen procedimientos tales como los descritos por Stewart y Young (1984) *Solid Phase Peptide Synthesis*, Pierce Chemical Co., Rockford, IL; Bodanszky y Bodanszky (1984) *The Practice of Peptide Synthesis*, Springer-Verlag, Nueva York; y Bodanszky (1984) *The Principles of Peptide Synthesis*, Springer-Verlag, Nueva York; todas las cuales se incorporan en la presente memoria como referencia. Por ejemplo, se puede utilizar el procedimiento con azida, el procedimiento con cloruro de ácido, el procedimiento con anhídrido de ácido, a el procedimiento del anhídrido mixto, el procedimiento del éster activo (p. ej., éster p-nitrofenílico, éster N-hidroxisuccinimídico, o éster cianometílico), el procedimiento del carbodiimidazol, el procedimiento oxidativo-reductivo, o el procedimiento de la diciclohexilcarbodiimida (DCCD)/aditivo. Las síntesis en fase sólida y



en solución son aplicables ambas a los procedimientos anteriores. Se pueden utilizar técnicas similares con secuencias de DTLR parciales.

Las proteínas de DTLR, los fragmentos, o derivados se preparan adecuadamente de acuerdo con los procedimientos anteriores empleados típicamente en la síntesis de péptidos, generalmente mediante el denominado procedimiento por etapas que comprende condensar un aminoácido al aminoácido terminal, uno por uno en la secuencia, o acoplando los fragmentos peptídicos al aminoácido terminal. Los grupos amino que no están siendo utilizados en la reacción de acoplamiento deben ser protegidos típicamente para evitar el acoplamiento en una localización incorrecta.

Si se adopta una síntesis en fase sólida, el aminoácido C-terminal es unido a un portador o soporte insoluble a través de su grupo carboxilo. El portador insoluble no está particularmente limitado con tal que tenga capacidad de unión a un grupo carboxilo reactivo. Los ejemplos de tales portadores insolubles incluyen resinas halometiladas, tales como resina clorometilada o resina bromometilada, resinas hidroximetiladas, resinas fenólicas, resinas de t-alquiloxicarbonilhidrazida, y similares.

Se une por orden un aminoácido con el grupo amino protegido a través de la condensación de su grupo carboxilo activado y el grupo amino reactivo del péptido o cadena formados previamente, para sintetizar el péptido etapa por etapa. Después de sintetizar la secuencia completa, el péptido se escinde del portador insoluble para producir el péptido. Este enfoque en fase sólida es descrito generalmente por Merrifield, *et al.* (1963) en J. Am. Chem. Soc. 85:2149-2156, que se incorpora en la presente memoria como referencia.

La proteína preparada y sus fragmentos pueden ser aislados y purificados de la mezcla de reacción por medio de separación de péptidos, por ejemplo, mediante extracción, precipitación, electroforesis, diferentes formas de cromatografía, y similares. Los receptores de esta invención pueden ser obtenidos con grados variables de pureza dependiendo de los usos deseados. La purificación se puede completar mediante el uso de las técnicas de purificación de proteínas descritas en la presente memoria, véase más abajo, o mediante el uso de los anticuerpos de la presente memoria descritos en los métodos de cromatografía de afinidad por inmuoabsorción. Esta cromatografía de afinidad por inmuoabsorción se lleva a cabo conectando primero los anticuerpos a un soporte sólido y poniendo en contacto después los anticuerpos conectados con productos lisados solubilizados de las células apropiadas, productos lisados de otras células que expresan el receptor, o productos lisados o sobrenadantes de células que producen la proteína como resultado de las técnicas de ADN, véase más abajo.

Generalmente, la proteína purificada tendrá una pureza de al menos aproximadamente 40%, habitualmente una pureza de al menos aproximadamente 50%, normalmente una pureza de al menos aproximadamente 60%, típicamente una pureza de al menos aproximadamente 70%, más típicamente una pureza de al menos aproximadamente 80%, preferiblemente una pureza de al menos aproximadamente 90% y más preferiblemente una pureza de al menos aproximadamente 95%, y en realizaciones concretas, de 97%-99% o más. La pureza será normalmente en peso, pero también puede ser sobre una base molar. Se aplicarán diferentes análisis según sea apropiado.

## VI. Anticuerpos

Se pueden originar anticuerpos contra diferentes proteínas de DTLR de mamífero, p. ej., primate y sus fragmentos, tanto en forma nativa natural como en forma recombinante, estando la diferencia en que es más probable que los anticuerpos para el receptor activo reconozcan los epítomos que están presentes solamente en las conformaciones nativas. La detección del antígeno desnaturalizado también puede ser útil, p. ej., en el análisis Western. Asimismo se contemplan los anticuerpos anti-idiotípicos, que serían útiles como agonistas o antagonistas de un receptor natural o un anticuerpo.

Los anticuerpos, incluyendo los fragmentos de unión y las versiones de una sola cadena, contra fragmentos predefinidos de la proteína pueden ser originados mediante inmunización de animales con productos conjugados de los fragmentos con proteínas inmunogénicas. Los anticuerpos monoclonales se preparan a partir de células que secretan el anticuerpo deseado. Estos anticuerpos pueden ser escrutados en busca de su unión a la proteína normal o defectuosa, o escrutados en busca de actividad agonística o antagonística. Estos anticuerpos monoclonales se unirán normalmente con una  $K_D$  de al menos aproximadamente 1 mM, más normalmente al menos aproximadamente 300  $\mu$ M, típicamente al menos aproximadamente 100  $\mu$ M, más típicamente al menos aproximadamente 30  $\mu$ M, preferiblemente al menos aproximadamente 10  $\mu$ M, y más preferiblemente al menos aproximadamente 3  $\mu$ M o más.

Los anticuerpos, incluyendo los fragmentos de unión al antígeno, de esta invención pueden tener un valor diagnóstico o terapéutico significativo. Pueden ser potentes antagonistas que se unen al receptor e inhiben la unión al ligando o inhiben la capacidad del receptor para lograr una respuesta biológica, p. ej., actuar sobre su sustrato. También pueden ser útiles como anticuerpos no neutralizadores y pueden ser acoplados a toxinas o radionúclidos para unirse a las células productoras, o a células localizadas como fuente de la interleuquina. Adicionalmente, estos anticuerpos se pueden conjugar con fármacos u otros agentes terapéuticos, directamente o indirectamente por medio de un conector.

Los anticuerpos de esta invención también pueden ser útiles en aplicaciones de diagnóstico. Como anticuerpos de captura o no neutralizadores, se podrían unir al receptor sin inhibir la unión al ligando o sustrato. Como anticuerpos neutralizadores, pueden ser útiles en análisis de unión competitiva. También serán útiles en la detección o

cuantificación del ligando. Se pueden utilizar como reactivos para el análisis de transferencia Western, o para la inmunoprecipitación o inmunopurificación de la respectiva proteína.

Los fragmentos de proteína se pueden unir a otros materiales, concretamente polipéptidos, en forma de polipéptidos fusionados o unidos covalentemente para ser utilizados como inmunógenos. Los DTLR de mamífero y sus fragmentos pueden ser fusionados o unidos covalentemente a una variedad de inmunógenos, tales como hemocianina de lapa ojo de cerradura, albúmina de suero bovino, toxoide del tétanos, etc. Véanse Microbiology, Hoeber Medical Division, Harper y Row, 1969; Landsteiner (1962) *Specificity of Serological Reactions*, Dover Publications, Nueva York; y Williams, *et al.* (1967) *Methods in Immunology and Immunochemistry*, Vol. 1, Academic Press, Nueva York; cada una de las cuales se incorpora en la presente memoria como referencia, para las descripciones de los métodos de preparación de antisueros policlonales. Un método típico implica la hiperinmunización de un animal con un antígeno. Después se recoge la sangre del animal poco después de las inmunizaciones repetidas y se aísla la gamma-globulina.

En algunos casos, es deseable preparar anticuerpos monoclonales de diferentes anfitriones mamíferos, tales como ratones, roedores, primates, seres humanos, etc. La descripción de las técnicas para preparar tales anticuerpos monoclonales se puede encontrar, p. ej., en Stites, *et al.* (eds) *Basic and Clinical Immunology* (4ª ed.), Lange Medical Publications, Los Altos, CA, y referencias allí citadas; Harlow y Lane (1988) *Antibodies: A Laboratory Manual*, CSH Press; Goding (1986) *Monoclonal Antibodies: Principles and Practice* (2ª ed) Academic Press, Nueva York; y concretamente en Kohler y Milstein (1975) en *Nature* 256: 495-497, que comenta un método para generar anticuerpos monoclonales. Cada una de estas referencias se incorpora en la presente memoria como referencia. Brevemente resumido, este método implica inyectar al animal un inmunógeno. Después el animal es sacrificado y se recogen las células de su bazo, que después se fusionan con células de mieloma. El resultado es una célula híbrida o "hibridoma" que es capaz de reproducirse *in vitro*. La población de hibridomas es escrutada después para aislar clones individuales, cada uno de los cuales secreta una única especie de anticuerpo para el inmunógeno. De esta manera, las especies de anticuerpos individuales obtenidas son los productos de células B individuales inmortalizadas y clonadas a partir del animal inmune generadas en respuesta a un sitio específico reconocido sobre la sustancia inmunogénica.

Otras técnicas adecuadas implican la exposición de linfocitos *in vitro* a los polipéptidos antigénicos o alternatively a la selección de genotecas de anticuerpos en fagos o vectores similares. Véanse, Huse, *et al.* (1989) "Generation of a Large Combinatorial Library of the Immunoglobulin Repertoire in Phage Lambda", *Science* 24 6:1275-12 81; y Ward, *et al.* (1989) *Nature* 341:544-546, cada una de las cuales se incorpora en la presente memoria como referencia. Los polipéptidos y anticuerpos de la presente invención se pueden utilizar con o sin modificación, incluyendo anticuerpos quiméricos o humanizados. Frecuentemente los polipéptidos y anticuerpos se marcarán uniéndose, covalentemente o no covalentemente, una sustancia que proporcione una señal detectable. Se conocen una amplia variedad de marcas y técnicas de conjugación y son referidas extensamente en la literatura tanto científica como de patentes. Las marcas adecuadas incluyen radionúclidos, enzimas, sustratos, cofactores, inhibidores, radicales fluorescentes, radicales quimioluminescentes, partículas magnéticas, y similares. Las patentes, que ilustran el uso de tales marcas incluyen las Patentes de los Estados Unidos Núms. 3.817.837; 3.850.752; 3.939.350; 3.996.345; 4.277.437; 4.275.149; y 4.366.241. Asimismo, se pueden producir inmunoglobulinas recombinantes o quiméricas, véase Cabilly, Patente de los Estados Unidos Núm. 4.816.567; o elaborar en ratones transgénicos, véase Mendez, *et al.* (1997) *Nature Genetics* 15:146-156. Estas referencias se incorporan en la presente memoria como referencia.

Los anticuerpos de esta invención también se pueden utilizar para la cromatografía de afinidad en el aislamiento de los DTLR. Se pueden preparar columnas en las que los anticuerpos están unidos a un soporte sólido, p. ej., partículas, tales como agarosa, Sephadex, o similar, donde un producto lisado celular se puede hacer pasar a través de la columna, lavar la columna, seguido de concentraciones crecientes de un desnaturalizante suave, por medio de lo cual se liberará la proteína purificada. Se puede utilizar la proteína para purificar anticuerpos.

Los anticuerpos también se pueden utilizar para escrutar genotecas de expresión en busca de productos de expresión concretos. Normalmente los anticuerpos utilizados en semejante procedimiento estarán marcados con un radical que permita la fácil detección de la presencia del antígeno por la unión al anticuerpo.

Los anticuerpos originados contra un DTLR también se utilizarán para originar anticuerpos anti-idiotípicos. Estos serán útiles en la detección o diagnóstico de diferentes afecciones inmunológicas relacionadas con la expresión de la proteína o de células que expresan la proteína. También serán útiles como agonistas o antagonistas del ligando, que pueden ser inhibidores competitivos o sustitutos de los ligandos de origen natural.

Una proteína de DTLR que se une específicamente a, o que es específicamente inmunorreactiva con un anticuerpo generado contra un inmunógeno definido, tal como un inmunógeno que consiste en la secuencia de aminoácidos del SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34, es determinada típicamente en un inmunoanálisis. El inmunoanálisis utiliza típicamente un antisuero policlonal que es originado, p. ej., para una proteína de SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34. Este antisuero se selecciona para que tenga una reactividad cruzada baja frente a otros miembros de la familia de IL-1R, p. ej., DTLR1, preferiblemente de la misma especie, y se elimina cualquiera de tales reactividades cruzadas mediante inmunoabsorción antes de su uso en el inmunoanálisis.

Con el fin de producir antisueros para su uso en un inmunoanálisis, se aísla la proteína del SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34, o una de sus combinaciones, como se describe en la presente memoria. Por ejemplo, se puede producir la proteína recombinante en una línea celular de mamífero. En un anfitrión apropiado, p. ej., una

cepa endogámica de ratones tal como balb/c, es inmunizada con la proteína seleccionada, utilizando típicamente un coadyuvante convencional, tal como coadyuvante de Freund, y un protocolo de inmunización de ratón convencional (véase Harlow y Lane, *supra*). Alternativamente, se puede utilizar como inmunógeno un péptido sintético derivado de las secuencias descritas en la presente memoria y conjugado con una proteína portadora. Los sueros policlonales se recogen y se titulan frente a la proteína inmunogénica en un inmunoanálisis, p. ej., un inmunoanálisis en fase sólida con el inmunógeno inmovilizado sobre un soporte sólido. Se seleccionan los antisueros policlonales con un título de  $10^4$  o mayor y se someten a ensayo en busca de su reactividad cruzada frente a otros miembros de la familia de IL-1R, p. ej., DTLR de ratón o DTLR1 humano, utilizando un inmunoanálisis de unión competitiva tal como el descrito por Harlow y Lane, *supra*, en las páginas 570-573. Preferiblemente se utilizan al menos dos miembros de la familia de DTLR en esta determinación junto con cualquiera o alguno de los DTLR2-10 humanos. Estos miembros de la familia de IL-1R pueden ser producidos como proteínas recombinantes y aislados utilizando técnicas de biología molecular y de química de proteínas convencionales como las descritas en la presente memoria.

Se pueden utilizar inmunoanálisis en formato de unión competitiva para las determinaciones de la reactividad cruzada. Por ejemplo, las proteínas de SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34, o sus diferentes fragmentos, pueden ser inmovilizadas en un soporte sólido. Las proteínas añadidas al análisis compiten con la unión del antisuero al antígeno inmovilizado. La capacidad de las proteínas anteriores para competir con la unión del antisuero a la proteína inmovilizada se compara con la proteína de SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 y/o 34. Se calcula el porcentaje de reactividad cruzada para las proteínas anteriores utilizando cálculos convencionales. Se seleccionan y reúnen los antisueros con una reactividad cruzada menor del 10% con cada una de las proteínas enumeradas antes. Los anticuerpos que presentan reactividad cruzada se separan después de los antisueros reunidos mediante inmunoabsorción con las proteínas enumeradas antes.

Los antisueros inmunoabsorbidos y reunidos se utilizan después en un inmunoanálisis de unión competitiva como se ha descrito antes para comparar una segunda proteína con la proteína inmunogénica (p. ej., la proteína de tipo IL-1R del SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 y/o 34). Con el fin de realizar esta comparación, se analiza cada una de las dos proteínas a un intervalo de concentraciones muy amplio y se determina la cantidad de cada proteína requerida para inhibir el 50% de la unión del antisuero a la proteína inmovilizada. Si la cantidad de la segunda proteína requerida es menos de dos veces la cantidad de la proteína o las proteínas seleccionadas, se dice que la segunda proteína se une específicamente a un anticuerpo generado para el inmunógeno.

Se entiende que estas proteínas de DTLR son miembros de una familia de proteínas homólogas que comprenden al menos 10 genes identificados hasta ahora. Para un producto génico concreto, tal como el DTLR2-10, el término hace referencia no solamente a las secuencias de aminoácidos descritas en la presente memoria, si no también a otras proteínas que son variantes alélicas, no alélicas o de especie. También se entiende que los términos incluyen mutaciones no naturales introducidas mediante mutación deliberada utilizando la tecnología recombinante convencional tal como la mutación en un único sitio, o eliminando secciones cortas de ADN que codifica las respectivas proteínas, o sustituyendo nuevos aminoácidos, o añadiendo nuevos aminoácidos. Tales alteraciones minoritarias deben mantener esencialmente la inmunoidentidad de la molécula original y/o su actividad biológica. De este modo, estas alteraciones incluyen proteínas que son específicamente inmunorreactivas con una proteína relacionada con IL-1R de origen natural designada, por ejemplo, las proteínas de DTLR mostradas en los SEQ ID NO: 4, 6, 26, 10, 12, 16, 18, 32, 22 o 34. Las propiedades biológicas de las proteínas alteradas pueden ser determinadas expresando la proteína en una línea celular apropiada y midiendo el efecto apropiado sobre los linfocitos. Las modificaciones de las proteínas concretas consideradas minoritarias incluirían la sustitución conservativa de aminoácidos con propiedades químicas similares, como se ha descrito antes para la familia de IL-1R en su totalidad. Alineando una proteína óptimamente con la proteína de DTLR2-10 y utilizando los inmunoanálisis convencionales descritos en la presente memoria para determinar la inmunoidentidad, se pueden determinar las composiciones de las proteínas de la invención.

## VII. Kits y cuantificación

Las formas tanto naturales como recombinantes de las moléculas de tipo IL-1R de esta invención son particularmente útiles en kits y métodos de análisis. Por ejemplo, estos métodos también se aplicarían al escrutinio en busca de actividad de unión, p. ej., ligandos para estas proteínas. En los últimos años se han desarrollado varios métodos de análisis automático con el fin de permitir el escrutinio de decenas de miles de compuestos por año. Véase, p. ej., una estación de trabajo automatizada BIOMEK, Beckman Instruments, Palo Alto, California, y Fodor, *et al.* (1991) Science 251:767-773, que se incorpora en la presente memoria como referencia. La última describe métodos para someter a ensayo la unión de una pluralidad de polímeros definidos sintetizados sobre un sustrato sólido. El desarrollo de análisis adecuados para escrutar en busca de un ligando o proteínas homólogas agonísticas/antagónicas puede resultar muy facilitado por la disponibilidad de grandes cantidades de DTLR solubles, purificados en un estado activo como el proporcionado por esta invención.

El DTLR purificado puede ser aplicado como recubrimiento directamente sobre placas para su uso en las técnicas de escrutinio de ligandos anteriormente mencionadas. No obstante, se pueden utilizar anticuerpos no neutralizadores para estas proteínas como anticuerpos de captura para inmovilizar el respectivo receptor sobre la fase sólida, útiles, p. ej., en usos de diagnóstico.

Asimismo se contempla el uso de DTLR2-10, sus fragmentos, péptidos, y sus productos de fusión en una variedad de kits de diagnóstico y métodos para detectar la presencia de la proteína o su ligando. Alternativamente, o adicionalmente, se pueden incorporar anticuerpos contra las moléculas a los kits y métodos. Típicamente el kit tendrá un compartimento que contiene un péptido o segmento génico de DTLR definido, o un reactivo que reconoce el uno o el otro. Típicamente, el reactivo de reconocimiento, en el caso del péptido, sería un receptor o anticuerpo, o en el caso de un segmento génico, sería normalmente una sonda de hibridación.

Un kit preferido para determinar la concentración, p. ej., de DTLR4, de una muestra comprendería típicamente un compuesto marcado, p. ej., ligando o anticuerpo, con una afinidad de unión a DTLR4 conocida, una fuente de DTLR4 (de origen natural o recombinante) como control positivo, y un método para separar el compuesto marcado unido del libre, por ejemplo una fase sólida para inmovilizar el DTLR4 de la muestra de ensayo. Normalmente se proporcionarán los compartimentos que contengan los reactivos, y las instrucciones.

Los anticuerpos, incluyendo los fragmentos de unión al antígeno, específicos para el DTLR de mamífero o uno de sus fragmentos peptídicos, o fragmentos receptores son útiles en las aplicaciones de diagnóstico para detectar la presencia de niveles elevados de ligando y/o sus fragmentos. Los análisis de diagnóstico pueden ser homogéneos (sin una etapa de separación entre el reactivo libre y el complejo anticuerpo-antígeno) o heterogéneos (con una etapa de separación). Existen diferentes análisis comerciales, tales como el radioinmunoanálisis (RIA), el análisis de absorción con enzima ligada (ELISA), el inmunoanálisis enzimático (EIA), la técnica de inmunoanálisis multiplicado por enzimas (EMIT), el inmunoanálisis fluorescente con sustrato marcado (SLFIA) y similares. Por ejemplo, se pueden emplear anticuerpos no marcados utilizando un segundo anticuerpo que está marcado y que reconoce el anticuerpo para DTLR4 o uno de sus fragmentos concretos. Estos análisis también han sido extensamente comentados en la literatura. Véanse, p. ej., Harlow y Lane (1988) *Antibodies: A Laboratory Manual*, CSH., y Coligan (Ed.) (1991) y suplementos periódicos, *Current Protocols In Immunology* Greene/Wiley, Nueva York.

Los anticuerpos anti-idiotípicos pueden tener un uso similar para servir como agonistas o antagonistas de DTLR4. Estos deben ser útiles como reactivos terapéuticos en circunstancias apropiadas.

Frecuentemente, los reactivos para los análisis diagnósticos son proporcionados en kits, con el fin de optimizar la sensibilidad del análisis. Para la invención sujeto, dependiendo de la naturaleza del análisis, del protocolo, y de la marca, se proporciona un anticuerpo marcado o no marcado, o un ligando marcado. Esto se encuentra normalmente junto con otros aditivos, tales como tampones, estabilizadores, materiales necesarios para la producción de la señal tales como sustratos para enzimas, y similares. Preferiblemente, el kit también contendrá instrucciones para el uso y la eliminación apropiados de los contenidos después de su uso. Típicamente el kit tiene compartimentos para cada reactivo útil, y contendrá instrucciones para el uso y la eliminación apropiados de los reactivos. Deseablemente, los reactivos se proporcionan en forma de polvo liofilizado seco, donde los reactivos pueden ser reconstituidos en un medio acuoso que tenga concentraciones apropiadas para realizar el análisis.

Los constituyentes mencionados antes de los análisis diagnósticos pueden ser utilizados sin modificación o pueden ser modificados de diferentes maneras. Por ejemplo, se puede lograr el marcaje uniéndolo covalentemente o no covalentemente un radical que proporcione directamente o indirectamente una señal detectable. En cualquiera de estos análisis, el compuesto de ensayo, el DTLR, o los anticuerpos para éste pueden estar marcados directamente o indirectamente. Las posibilidades para el marcaje directo incluyen grupos marcadores: radiomarcas tales como  $I^{125}$ , enzimas (Patente de los Estados Unidos Núm. 3.645.090) tales como peroxidasa y fosfatasa alcalina, y marcas fluorescentes (Patente de los Estados Unidos Núm. 3.940.475) capaces de verificar el cambio en la intensidad de fluorescencia, el desplazamiento de la longitud de onda, o la polarización de la fluorescencia. Ambas patentes se incorporan en la presente memoria como referencia. Las posibilidades de marcaje indirecto incluyen biotinylación de un constituyente seguido de unión a avidina acoplada a uno de los grupos marcadores anteriores.

También existen numerosos métodos de separación del ligando unido del libre, o alternativamente el compuesto de ensayo unido del libre. El DTLR puede ser inmovilizado sobre diferentes matrices seguido de lavado. Las matrices adecuadas incluyen plásticos tal como una placa de ELISA, filtros, y cuentas. Los métodos de inmovilización del receptor a la matriz incluyen, sin limitación, la adherencia directa al plástico, el uso de un anticuerpo de captura, el acoplamiento químico, y biotina-avidina. La última etapa de enfoque implica la precipitación del complejo anticuerpo/antígeno mediante cualquiera de los diferentes métodos incluyendo aquellos que utilizan, p. ej., un disolvente orgánico tal como polietilenglicol o una sal tal como sulfato de amonio. Otras técnicas de separación adecuadas incluyen, sin limitación, el método de las partículas imantables de anticuerpo con fluoresceína descrito por Rattle, *et al.* (1984) *Clin. Chem.* 30(9):1457-14 61, y la separación de partículas magnéticas con doble anticuerpo como se describe en la Patente de los Estados Unidos Núm. 4.659.678.

Los métodos para conectar proteínas o fragmentos a diferentes marcas han sido referidos extensamente en la literatura y no requieren aquí un estudio detallado. Muchas de las técnicas implican el uso de grupos carboxilo activados por medio de la utilización de carbodiimida o ésteres activos para formar enlaces peptídicos, la formación de tioéteres mediante reacción de un grupo mercapto con un halógeno activado tal como cloroacetilo, o una olefina activada tal como maleimida, para la conexión, o similar. Las proteínas de fusión también encontrarán uso en estas aplicaciones.

Otro aspecto diagnóstico implica el uso de secuencias de oligonucleótidos o polinucleótidos tomadas de la secuencia de un DTLR. Estas secuencias pueden ser utilizadas como sondas para detectar niveles de los respectivos DTLR

en pacientes que se sospecha que tienen un trastorno inmunológico. La preparación de secuencias de nucleótidos de ARN y ADN, el marcaje de las secuencias, y el tamaño preferido de las secuencias han recibido una amplia descripción y discusión en la literatura. Normalmente una sonda oligonucleotídica debe tener al menos aproximadamente 14 nucleótidos, normalmente al menos aproximadamente 18 nucleótidos, y las sondas polinucleotídicas pueden tener hasta varias kilobases. Se pueden emplear diferentes marcas, muy comúnmente radionúclidos, concretamente  $P^{32}$ . Sin embargo, también se pueden emplear otras técnicas, por ejemplo el uso de nucleótidos modificados con biotina para su introducción en un polinucleótido. La biotina sirve después como sitio para la unión a avidina o anticuerpos, que pueden estar marcados con una amplia variedad de marcas, tales como radionúclidos, sustancias fluorescentes, enzimas, o similares. Alternativamente, se pueden emplear anticuerpos que pueden reconocer dúplex específicos, incluyendo dúplex de ADN, dúplex de ARN, dúplex híbridos de ADN-ARN, o dúplex de ADN-proteína. Los anticuerpos a su vez pueden estar marcados y se puede llevar a cabo un análisis en el que el dúplex está unido a una superficie, de manera que tras la formación del dúplex sobre la superficie, se puede detectar la presencia del anticuerpo unido al dúplex. El uso de sondas para el ARN anti-sentido novedoso se puede llevar a cabo mediante cualquier técnica convencional tal como la hibridación de ácido nucleico, el escrutinio más y menos, el sondeo recombinatorio, la traducción de híbridos liberados (HRT), y la traducción de híbridos detenidos (HART). Esto también incluye las técnicas de amplificación tales como la reacción en cadena de la polimerasa (PCR).

Asimismo se contemplan los kits diagnósticos que también someten a ensayo la presencia cualitativa o cuantitativa de otros marcadores. La diagnosis o la prognosis pueden depender de la combinación de múltiples indicaciones utilizadas como marcadores. De este modo, los kits pueden someter a ensayo las combinaciones de marcadores. Véase, p. ej., Viallet, *et al.* (1989) *Progress in Growth Factor Res.* 1:89-97.

### VIII. Utilidad terapéutica

Esta invención proporciona reactivos con un valor terapéutico significativo. Los DTLR (de origen natural o recombinantes), sus fragmentos, los receptores de muteína, y los anticuerpos, junto con los compuestos identificados por tener afinidad de unión para los receptores o anticuerpos, deben ser útiles en el tratamiento de afecciones que muestran una expresión anormal de los receptores de sus ligandos. Semejante anomalía se manifestará típicamente mediante trastornos inmunológicos. Adicionalmente, esta invención debe proporcionar valor terapéutico en diferentes enfermedades o trastornos asociados con la expresión anormal o el desencadenamiento anormal de la respuesta al ligando. Se ha sugerido que los ligandos Toll están implicados en el desarrollo morfológico, p. ej., la determinación de la polaridad dorso-ventral, y en las respuestas inmunitarias, concretamente las respuestas innatas primitivas. Véase, p. ej., Sun, *et al.* (1991) *Eur. J. Biochem.* 196:247-254; Hultmark (1994) *Nature* 367:116-117.

Las muteínas de DTLR recombinantes, los anticuerpos agonísticos o antagonísticos para estas, o los anticuerpos pueden ser purificados y después administrados a un paciente. Estos reactivos pueden ser combinados para su uso terapéutico con ingredientes activos adicionales, p. ej., en portadores o diluyentes farmacéuticamente aceptables convencionales, junto con estabilizadores y excipientes fisiológicamente inocuos. Estas combinaciones pueden ser estériles, p. ej., filtradas, y colocadas en formas de dosificación mediante liofilización en viales de dosificación o almacenadas en preparaciones acuosas estabilizadas. Esta invención también contempla el uso de anticuerpos o sus fragmentos de unión que no se unen al complemento.

Se puede realizar el escrutinio de ligandos utilizando DTLR o sus fragmentos para identificar las moléculas que tienen afinidad de unión con los receptores. Después se pueden utilizar análisis biológicos para determinar si un supuesto ligando puede proporcionar una unión competitiva, que pueda bloquear la actividad estimuladora intrínseca. Se pueden utilizar fragmentos receptores como bloqueador o antagonista ya que éste bloquea la actividad del ligando. Del mismo modo, un compuesto que tiene actividad estimuladora intrínseca puede activar el receptor y de este modo es un agonista ya que estimula la actividad del ligando, p. ej., induciendo la señalización.

Adicionalmente se contempla el uso terapéutico de anticuerpos para los DTLR como antagonistas.

Las cantidades de reactivos necesarias para la terapia eficaz dependerán de muchos factores diferentes, incluyendo los métodos de administración, el sitio diana, el estado fisiológico del paciente, y otros medicamentos administrados. De este modo, las dosis de tratamiento deben ser tituladas para optimizar la seguridad y la eficacia. Típicamente, las dosificaciones utilizadas *in vitro* pueden proporcionar unas pautas útiles en las cantidades útiles para la administración *in situ* de estos reactivos. El ensayo en animales de las dosis eficaces para el tratamiento de trastornos concretos proporcionará una indicación predictiva adicional de la dosificación en seres humanos. Se describen diferentes consideraciones, p. ej., Gilman, *et al.* (eds) (1990) *Goodman and Gilman's: The Pharmacological Bases of Therapeutics*, 8ª Ed., Pergamon Press; y Remington's *Pharmaceutical Sciences*, (edición actual), Mack Publishing Co., Easton, Penn.; cada una de las cuales se incorpora en la presente memoria como referencia. Los métodos para la administración son discutidos allí y más abajo, p. ej., para la administración oral, intravenosa, intraperitoneal, o intramuscular, la difusión transdérmica, y otros. Los portadores farmacéuticamente aceptables incluirán agua, solución salina, tampones, y otros compuestos descritos, p. ej., en el *Merck Index*, Merck & Co., Rahway, Nueva Jersey. Debido a la afinidad de unión probablemente elevada, o a los números de recambio, entre un supuesto ligando y sus receptores, se esperaría inicialmente que fueran eficaces dosis bajas de estos reactivos. Y la ruta de señalización sugiere que pueden tener efecto cantidades extremadamente bajas de ligando. De este modo, cabría esperar normalmente que los intervalos de dosificación estuvieran en cantidades menores que concentraciones 1 mM, típicamente concentraciones menores

de aproximadamente 10  $\mu$ M, normalmente menores de aproximadamente 100 nM, preferiblemente menores de aproximadamente 10 pM (picomolar), y muy preferiblemente menores de aproximadamente 1 fM (femtomolar), con un portador apropiado. A menudo se utilizarán formulaciones de liberación lenta, o aparatos de liberación lenta para la administración continua.

Los DTLR, sus fragmentos, y los anticuerpos o sus fragmentos, antagonistas, y agonistas, pueden ser administrados directamente al anfitrión que se va a tratar o, dependiendo del tamaño de los compuestos, puede ser deseable conjugarlos con proteínas portadoras tales como ovoalbúmina o albúmina de suero antes de su administración. Las formulaciones terapéuticas se pueden administrar en cualquier formulación de dosificación convencional. Si bien es posible administrar solo el ingrediente activo, es preferible presentarlo en forma de una formulación farmacéutica. Las formulaciones comprenden al menos un ingrediente activo, como se ha definido antes, junto con uno o más portadores aceptables del mismo. Cada portador debe ser farmacéuticamente y fisiológicamente aceptable en el sentido de ser compatible con los otros ingredientes y no perjudicial para el paciente. Las formulaciones incluyen aquellas adecuadas para la administración oral, rectal, nasal, o parenteral (incluyendo subcutánea, intramuscular, intravenosa e intradérmica). Las formulaciones pueden ser convenientemente presentadas en una forma de dosificación unitaria y se pueden preparar mediante cualquiera de los métodos bien conocidos en la técnica farmacéutica. Véase, p. ej., Gilman, *et al.* (eds) (1990) *Goodman and Gilman's: The Pharmacological Bases of Therapeutics*, 8ª Ed., Pergamon Press; y Remington's *Pharmaceutical Sciences* (edición actual), Mack Publishing Co., Easton, Penn.; Avis, *et al.* (eds. 1993) *Pharmaceutical Dosage Forms: Parenteral Medications* Dekker, NY; Lieberman, *et al.* (eds. 1990) *Pharmaceutical Dosage Forms: Tablets* Dekker, NY; y Lieberman, *et al.* (eds. 1990) *Pharmaceutical Dosage Forms: Disperse Systems* Dekker, NY. La terapia de esta invención puede ser combinada o utilizada asociada con otros agentes terapéuticos, concretamente agonistas o antagonistas de otros miembros de la familia de IL-1.

## IX. Ligandos

La descripción de los receptores Toll en la presente memoria proporciona medios para identificar ligandos, como se ha descrito antes. Semejante ligando se debería unir específicamente al respectivo receptor con una afinidad razonablemente elevada. Se encuentran disponibles diferentes constructos que permiten cualquier marcaje del receptor para detectar su ligando. Por ejemplo, el marcaje directo de DTLR, fusionando sobre él marcadores para el marcaje secundario, p. ej., FLAG u otras etiquetas epitópicas, etc., permitirá la detección del receptor. Este puede ser histológico, como un método de afinidad para la purificación bioquímica, o un marcaje o selección en un enfoque de clonación de la expresión. También se puede aplicar un sistema de selección de dos híbridos elaborando los constructos apropiados con las secuencias de DTLR disponibles. Véase, p. ej., Fields y Song (1989) *Nature* 340:245-246.

Generalmente, las descripciones de los DTLR serán aplicables análogamente a realizaciones específicas individuales dirigidas a reactivos y composiciones de DTLR2, DTLR3, DTLR4, DTLR5, DTLR6, DTLR7, DTLR8, DTLR9, y/o DTLR10.

El amplio alcance de esta invención se comprende mejor con referencia a los siguientes ejemplos, que no se pretende que limiten las invenciones a las realizaciones específicas.

## Ejemplos

### I. Métodos Generales

Algunos de los métodos convencionales son descritos o referidos, p. ej., por Maniatis, *et al.* (1982) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor Press; Sambrook, *et al.* (1989) *Molecular Cloning: A Laboratory Manual*, (2ª ed.), vols 1-3, CSH Press, NY; Ausubel, *et al.*, *Biology*, Greene Publishing Associates, Brooklyn, NY; o Ausubel, *et al.* (1987 y Suplementos) *Current Protocols in Molecular Biology*, Greene/Wiley, Nueva York. Los métodos para la purificación de proteína incluyen métodos tales como la precipitación con sulfato de amonio, la cromatografía en columna, la electroforesis, la centrifugación, la cristalización, y otros. Véanse, p. ej., Ausubel, *et al.* (1987 y suplementos periódicos); Coligan, *et al.* (ed. 1996) y suplementos periódicos, *Current Protocols In Protein Science* Greene/Wiley, Nueva York; Deutscher (1990) "Guide to Protein Purification" en *Methods in Enzymology*, vol. 182, y otros volúmenes de esta serie; y las publicaciones de los fabricantes sobre el uso de los productos de purificación de proteínas, p. ej., Pharmacia, Piscataway, N.J., o Bio-Rad, Richmond, CA. La combinación con técnicas recombinantes permite la fusión con segmentos apropiados, p. ej., a una secuencia FLAG o una equivalente que puede ser fusionada a través de una secuencia separable con proteasa. Véase, p. ej., Hochuli (1989) *Chemische Industrie* 12:69-70; Hochuli (1990) "Purification of Recombinant Proteins with Metal Chelate Absorbent" en Setlow (ed.) *Genetic Engineering, Principle and Methods* 12:87-98, Plenum Press, N.Y.; y Crowe, *et al.* (1992) *OIAexpress: The High Level Expression & Protein Purification System* QUIAGEN, Inc., Chatsworth, CA.

Las técnicas y análisis inmunológicos convencionales son descritos, p. ej., por Hertenberg, *et al.* (eds. 1996) *Weir's Handbook of Experimental Immunology* vols. 1-4, Blackwell Science; Coligan (1991) *Current Protocols in Immunology* Wiley/Greene, NY; y *Methods in Enzymology* volúmenes 70, 73, 74, 84, 92, 93, 108, 116, 121, 132, 150, 162, y 163.

Los análisis de las actividades biológicas vasculares son bien conocidos en la técnica. Abarcan tanto las actividades angiogénicas como angiostáticas en tumores, u otros tejidos, p. ej., proliferación de la musculatura lisa arterial (véase, p. ej., Koyoma, *et al.* (1996) *Cell* 87:1069-1078), la adherencia de monocitos al epitelio vascular (véase McEvoy, *et al.* (1997) *J. Exp. Med.* 185:2069-2077), etc. Véanse también Ross (1993) *Nature* 362:801-809; Rekhter y Gordon (1995) *Am. J. Pathol.* 147:668-677; Thyberg, *et al.* (1990) *Atherosclerosis* 10:966-990; y Gumbiner (1996) *Cell* 84:345-357.

Los análisis de las actividades biológicas de las células neurales son descritas, p. ej., por Wouterlood (ed. 1995) en *Neuroscience Protocols Modules 10*, Elsevier; *Methods in Neurosciences* Academic Press; y *Neuromethods Humana Press*, Totowa, NJ. La metodología de los sistemas evolutivos son descritos, p. ej., por Meisami (ed.) en *Handbook of Human Growth and Developmental Biology* CRC Press; y Chrispeels (ed.) *Molecular Techniques and Approaches in Developmental Biology* Interscience.

El análisis de secuencias por ordenador se realiza, p. ej., utilizando los programas de soporte lógico disponibles, incluyendo los de GCG (U. Wisconsin) y las fuentes de GenBank. Asimismo se utilizaron bases de datos de secuencias públicas, p. ej., de GenBank, NCBI, EMBO, y otras.

Muchas técnicas aplicables a los receptores de IL-10 pueden ser aplicadas a los DTLR, como se describe, p. ej., en el documento USSN 08/110.683 (receptor de IL-10), que se incorpora en la presente memoria como referencia para todos los fines.

## II. Familia Novedosa de Receptores Humanos

Abreviaturas: DTLR, receptor de tipo Toll; IL-1R, receptor de interleuquina-1; TH, homología Toll; LRR, repetición rica en leucina; EST, etiqueta de secuencia expresada; STS, sitio de secuencia etiquetada; FISH, hibridación fluorescente *in situ*.

El descubrimiento de la homología de secuencia entre los dominios citoplásmicos de los receptores Toll de *Drosophila* y de interleuquina-1 (IL-1) humana ha sembrado la convicción de que ambas moléculas desencadenan rutas de señalización relacionadas vinculadas a la translocación nuclear de los factores de transcripción de tipo Rel. Este esquema de señalización conservado determina una respuesta inmunitaria evolutivamente ancestral tanto en insectos como en vertebrados. Los autores de la presente invención informan sobre la clonación molecular de una clase novedosa de supuestos receptores humanos con una arquitectura proteica que es muy similar a Toll de *Drosophila* tanto en segmentos intra- como extra-celulares. Cinco receptores de tipo Toll humanos, denominados DTLR 1-5, son probablemente los homólogos directos de la molécula de la mosca, y como tales constituyen un componente importante y no reconocido de la inmunidad innata en seres humanos; fascinantemente, la conservación evolutiva de los DTLR en vertebrados puede indicar otro papel, semejante al de Toll en la dorso-ventralización del embrión de *Drosophila*, como reguladores de la formación de patrones morfogénicos tempranos. Las transferencias de ARNm de múltiples tejidos indican patrones de expresión marcadamente diferentes para los DTLR humanos. Utilizando los análisis las bases de datos de hibridación fluorescente *in situ* y del Sitio de Secuencia Etiquetada, los autores de la presente invención también demuestran que los genes de DTLR cognados residen en los cromosomas 4 (DTLR 1, 2, y 3), 9 (DTLR4), y 1 (DTLR5). La predicción de la estructura de los dominios de homología Toll (TH) alineados de DTLR de insectos variados y seres humanos, receptores de IL-1 de vertebrados, y factores MyD88, y proteínas de resistencia a enfermedades de plantas, reconoce un plegamiento  $\beta/\alpha$  paralelo con un sitio activo ácido; una estructura notablemente similar se repite en una clase de reguladores de la respuesta ampliamente implicados en la transducción de la información sensorial en bacterias.

Las semillas del abismo morfogénico que separa de manera espectacular las moscas de los seres humanos están plantadas en las formas y patrones embrionarios familiares, pero dan lugar a complejidades celulares muy diferentes. DeRobertis y Sasai (1996) *Nature* 380:37-40; y Arendt y Nübler-Jung (1997) *Mech. Develop.* 61:7-21. Esta divergencia de planes evolutivos entre insectos y vertebrados está coreografiada por rutas de señalización extraordinariamente similares, subrayando una mayor conservación de las redes de proteínas y de los mecanismos bioquímicos de repertorios génicos desiguales. Miklos y Rubin (1996) *Cell* 86:521-529; y Chothia (1994) *Develop.* 1994 Suppl., 27-33. Un modo potente de registrar gráficamente el diseño evolutivo de estas rutas reguladoras es deduciendo sus componentes moleculares probables (y funciones biológicas) por medio de la comparación interespecie de las secuencias y las estructuras de las proteínas. Miklos y Rubin (1996) *Cell* 86:521-529; Chothia (1994) *Develop.* 1994 Suppl., 27-33 (3-5); y Banfi, *et al.* (1996) *Nature Genet.* 13:167-174.

Una etapa universalmente crítica en el desarrollo embrionario es la especificación de los ejes corporales, ya sea procedentes de asimetrías innatas ya sea desencadenados por indicaciones externas. DeRobertis y Sasai (1996) *Nature* 380:37-40; y Arendt y Nübler-Jung (1997) *Mech. Develop.* 61:7-21. Como sistema modelo, se ha centrado una atención particular sobre la base filogenética y los mecanismos celulares de polarización dorsoventral. DeRobertis y Sasai (1996) *Nature* 380:37-40; y Arendt y Nübler-Jung (1997) *Mech. Develop.* 61:7-21. Ha surgido una estrategia molecular prototipo para esta transformación del embrión de *Drosophila*, donde la acción sucesiva de un pequeño número de genes da como resultado un gradiente de ventralización del factor de transcripción Dorsal. St. Johnston y Nüsslein-Volhard (1992) *Cell* 68:201-219; y Morisato y Anderson (1995) *Ann. Rev. Genet.* 29:371-399.

Esta ruta de señalización se centra en Toll, un receptor transmembrana que transduce la unión de un factor ventral secretado maternamente, Spätzle, al engranaje citoplásmico de Tube, una molécula accesoria, y la activación de Pelle, una Ser/Thr quinasa que cataliza la disociación de Dorsal desde el Cactus inhibidor y permite la migración de Dorsal a los núcleos ventrales (Morisato y Anderson (1995) *Ann. Rev. Genet.* 29:371-399; y Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416. La ruta de Toll también controla la inducción de factores antimicrobianos potentes en la mosca adulta (Lemaitre, *et al.* (1996) *Cell* 86:973-983); este papel en la defensa inmunitaria de *Drosophila* fortalece los paralelismos mecánicos con las rutas de la IL-1 que determinan una gran cantidad de respuestas inmunitarias e inflamatorias en vertebrados. Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416; y Wasserman (1993) *Molec. Biol. Cell* 4:767-771. Un dominio citoplásmico relacionado con Toll en los receptores de IL-1 dirige la unión de una quinasa de tipo Pelle, IRAK, y la activación de un complejo NF- $\kappa$ B/I- $\kappa$ B latente que refleja el abrazo de Dorsal y Cactus. Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416; y Wasserman (1993) *Molec. Biol. Cell* 4:767-771.

Los autores de la presente invención describen la clonación y la caracterización molecular de cuatro nuevas moléculas de tipo Toll en seres humanos, denominadas DTLR 2-5 (siguiendo Chiang & Beachy (1994) *Mech. Develop.* 47:225-239), que revelan una familia de receptores más íntimamente vinculada a homólogos Toll de *Drosophila* que a receptores de IL-1 de vertebrados. Las secuencias de DTLR derivan de EST humanas; estos ADNc parciales fueron utilizados para dibujar los perfiles de expresión completos en tejidos humanos para los cinco DTLR, mapear las localizaciones cromosómicas de genes cognados, y restringir la elección de genotecas de ADNc para recuperaciones de ADNc completos. Espoleados por otros esfuerzos (Banfi, *et al.* (1996) *Nature Genet.* 13:167-174; y Wang, *et al.* (1996) *J. Biol. Chem.* 271:4468-4476), los autores de la presente invención están recopilando, mediante conservación estructural y parquedad molecular, un sistema biológico en seres humanos que es la contraparte de un esquema regulador convincente en *Drosophila*. Además, se sugiere un mecanismo bioquímico que hace funcionar la señalización Toll por medio del plegamiento terciario propuesto del dominio de homología Toll (TH), un módulo central compartido por los DTLR, una amplia familia de receptores de IL-1, factores MyD88 de mamífero y proteínas de resistencia a enfermedades de plantas. Mitcham, *et al.* (1996) *J. Biol. Chem.* 271:5777-5783; y Hardiman, *et al.* (1996) *Oncogene* 13:2467-2475. Los autores de la presente invención proponen que una ruta de señalización que acopla la morfogénesis y la inmunidad primitiva en insectos, plantas, y animales (Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416; y Wilson, *et al.* (1997) *Curr. Biol.* 7:175-178) puede tener raíces en rutas bacterianas de dos componentes.

#### Análisis Computacional

Las secuencias humanas relacionadas con los DTLR de insecto fueron identificadas a partir de la base de datos de EST (dbEST) en el National Center for Biotechnology Information (NCBI) utilizando el servidor BLAST (Altschul, *et al.* (1994) *Nature Genet.* 6:119-129). Se utilizaron métodos basados en patrones y perfiles más sensibles (Bork y Gibson (1996) *Meth. Enzymol.* 266:162-184) para aislar los dominios de señalización de la familia de DTLR que son compartidos con proteínas de vertebrados y plantas en bases de datos no redundantes. El alineamiento progresivo de las secuencias de los dominios intra- o extracelulares de DTLR se llevó a cabo por medio de ClustalW (Thompson, *et al.* (1994) *Nucleic Acids Res.* 22:4673-4680); este programa también calculó el orden de ramificación de las secuencias alineadas mediante el algoritmo Neighbor-Joining algorithm (5000 replicaciones "bootstrap" proporcionaron valores de confianza para los tres agrupamientos).

Los patrones de alineamiento conservados, distinguidos a diferentes grados de restricción, fueron dibujados por el programa Consensus (internet URL <http://www.bork.embl-heidelberg.de/Alignment/consensus.html>). La genoteca PRINTS de huellas de proteínas (<http://www.biochem.ucl.ac.uk/bsm/dbbrowser/PRINTS/PRINTS.html>) (Attwood, *et al.* (1997) *Nucleic Acids Res.* 25:212-217) identificaron de forma fidedigna la miríada de repeticiones ricas en leucina (LRR) presentes en los segmentos extracelulares de los DTLR con un motivo compuesto (PRINTS codifica Leurichrpt) que empareja flexiblemente los rasgos N- y C-terminales de las LRR divergentes. Se utilizaron dos algoritmos de predicción cuya precisión en tres estados está por encima de 72% para obtener una estructura secundaria consenso para el alineamiento de dominios intracelulares, como puente para los esfuerzos de reconocimiento del plegamiento (Fischer, *et al.* (1996) *FASEB J.* 10:126-136). Tanto el programa de la red neural PHD (Rost y Sander (1994) *Proteins* 19:55-72) como el método de predicción estadística DSC (King y Sternberg (1996) *Protein Sci.* 5:2298-2310) tienen servidores de internet (URLs [http://www.embl-heidelberg.de/predictprotein/phd\\_pred.html](http://www.embl-heidelberg.de/predictprotein/phd_pred.html) y [http://bonsai.lif.icnet.uk/bmm/dsc/dsc\\_read\\_align.html](http://bonsai.lif.icnet.uk/bmm/dsc/dsc_read_align.html), respectivamente). La región intracelular codifica la región THD comentada, p. ej., por Hardiman, *et al.* (1996) en *Oncogene* 13:2467-2475; y Rock, *et al.* (1998) *Proc. Nat'l Acad. Sci. USA* 95:588-593, cada una de las cuales se incorpora en la presente memoria como referencia. Este dominio es muy importante en el mecanismo de señalización por los receptores, que transfiere un grupo fosfato a un sustrato.

#### Clonación de ADNc de DTLR humanos completos

Se utilizaron los cebadores de PCR obtenidos de la secuencia Humrsc786 de tipo Toll (código de acceso Genbank D13637) (Nomura, *et al.* (1994) *DNA Res* 1:27-35) para sondear una genoteca de ADNc derivado de la línea celular TF-1 eritroleucémica humana (Kitamura, *et al.* (1989) *Blood* 73:375-380) para proporcionar la secuencia de ADNc de DTLR1. Las secuencias de DTLR restantes se marcaron a partir de dbEST, y se obtuvieron los clones EST relevantes del consorcio I.M.A.G.E. (Lennon, *et al.* (1996) *Genomics* 33:151-152) por medio de Research Genetics (Huntsville,



AL): Núms. ID Clones 80633 y 117262 (DTLR2), 144675 (DTLR3), 202057 (DTLR4) y 277229 (DTLR5). Los ADNc completos para los DTLR 2-4 humanos se clonaron mediante escrutinio por hibridación de ADN de genotecas 5'-Stretch Plus (Clontech) del fago  $\lambda$ gt10, de pulmón adulto humano, de placenta, y de hígado fetal, respectivamente; la secuencia de DTLR5 se obtiene de una EST de una placa de escleriosis múltiple humana. Todos los clones positivos se secuenciaron y se alinearon para identificar los ORF de los DTLR individuales: DTLR1 (clon de 2366 pb, 786 aa ORF), DTLR2 (2600 pb, 784 aa), DTLR3 (3029 pb, 904 aa), DTLR4 (3811 pb, 879 aa) y DTLR5 (1275 pb, 370 aa). Las sondas para las hibridaciones de DTLR3 y DTLR4 fueron generadas mediante PCR utilizando genotecas de ADNc de placenta humana (Stratagene) e hígado adulto (Clontech) como moldes, respectivamente; los pares de cebadores se obtuvieron de las respectivas secuencias de EST. Las reacciones de PCR se llevaron a cabo utilizando ADN polimerasa Taqplus de *T. aquaticus* (Stratagene) en las siguientes condiciones: 1 x (94°C, 2 min) 30 x (55°C, 20 seg; 72°C 30 seg; 94°C 20 seg), 1 x (72°C, 8 min). Para el escrutinio de ADNc completo de DTLR2, se utilizó un fragmento de 900 pb generado mediante digestión con EcoRI/XbaI del primer clon EST (Núm. ID 80633) como sonda.

#### Transferencias de ARNm y localización cromosómica

Se adquirieron tejido múltiple humano (Núm. Cat. 1, 2) y transferencias de líneas celulares cancerosas (Núm. Cat. 7757-1), que contenían aproximadamente 2  $\mu$ g de ARN poli(A)<sup>+</sup> por calle, de Clontech (Palo Alto, CA). Para los DTLR 1-4, los ADNc completos sirvieron como sondas, para DTLR5 se utilizó el inserto del plásmido del clon EST (Núm. ID 277229). En resumen, las sondas fueron radiomarcadas con dATP [ $\alpha$ -P<sup>32</sup>] utilizando el kit de marcaje de cebadores al azar Amersham Rediprime (RPN1633). Se realizaron la prehibridación y las hibridaciones a 65°C en Na<sub>2</sub>HPO<sub>4</sub> 0,5M, SDS al 7%, EDTA 0,5 M (pH 8,0). Todos los lavados restrictivos se realizaron a 65°C con dos lavados iniciales en 2 x SSC, SDS al 0,1% durante 40 min seguido de un lavado posterior en 0,1 x SSC, SDS al 0,1% durante 20 min. Después las membranas se expusieron a -70°C a película de Rayos X (Kodak) en presencia de pantallas intensificadoras. Se realizaron estudios más detallados mediante método Southern de genotecas de ADNc (14) con clones de DTLR humanos seleccionados para examinar su expresión en subgrupos de células hematopoyéticas.

El mapeo cromosómico humano se llevó a cabo mediante el método de hibridación fluorescente *in situ* (FISH) como describen Heng y Tsui (1994) Meth. Molec. Biol. 33:109-122, utilizando diferentes clones de ADNc completos (DTLR 2-4) o parciales (DTLR5) como sondas. Estos análisis se realizaron como un servicio por SeeDNA Biotech Inc. (Ontario, Canadá). Se realizó una búsqueda de síndromes humanos (o defectos en ratones en loci sinténicos) asociados con los genes DTLR mapeados en la Dysmorphic Human-Mouse Homology Database mediante el servidor de internet ([http://www.hgmp.mrc.ac.uk/DHMHD/hum\\_chrome1.htm1](http://www.hgmp.mrc.ac.uk/DHMHD/hum_chrome1.htm1)).

#### Arquitectura conservada de ectodominios de DTLR de insectos y seres humanos

La familia Toll en *Drosophila* comprende al menos cuatro productos génicos distintos: Toll, el receptor prototipo implicado en la formación de patrones dorsoventrales del embrión de la mosca (Morisato y Anderson (1995) Ann. Rev. Genet. 29:371-399) y un segundo denominado "18 Wheeler" (18w) que también puede estar implicado en el desarrollo embrionario temprano (Chiang y Beachy (1994) Mech. Develop. 47:225-239; Eldon, *et al.* (1994) Develop. 120:885-899); dos receptores adicionales son pronosticados por los ORF de tipo Toll, incompletos aguas abajo del locus del transcrito específico del macho (Mst) (código Genbank X67703) o codificados por el "sitio de secuencia etiquetada" (STS) Dm2245 (código Genbank G01378) (Mitcham, *et al.* (1996) J. Biol. Chem. 271:5777-5783). Los segmentos extracelulares de Toll y 18w están compuestos inconfundiblemente por motivos LRR de ~24 aminoácidos, imperfectos (Chiang y Beachy (1994) Mech. Develop. 47:225-239; y Eldon, *et al.* (1994) Develop. 120:885-899). Disposiciones similares en tándem de las LRR forman comúnmente las antenas adherentes de moléculas de la superficie celular variadas y se presume que su estructura terciaria genérica imita un armazón en forma de herradura de un plegamiento inhibidor de la ribonucleasa, donde diecisiete LRR muestran un motivo de 28 restos, en horquilla  $\beta/\alpha$  repetitiva (Buchanan y Gay (1996) Prog. Biophys. Molec. Biol. 65:1-44). El reconocimiento específico de Spätzle por Toll puede seguir un modelo propuesto para la unión de hormonas de glicoproteína con un plegamiento en nudo de cistina por ectodominios multi-LRR de receptores de serpentina, utilizando el lado cóncavo de la lámina  $\beta$  curvada (Kajava, *et al.* (1995) Structure 3:867-877); fascinantemente, el patrón de cisteínas en Spätzle, y un ligando orfano de *Drosophila*, Trunk, pronostica una estructura terciaria en nudo de cistina similar (Belvin y Anderson (1996) Ann. Rev. Cell Develop. Biol. 12:393-416; y Casanova, *et al.* (1995) Genes Develop. 9:2539-2544).

Los ectodominios 22 y 31 LRR de Toll y 18w, respectivamente (el fragmento ORF de Mst presenta 16 LRR), están muy íntimamente relacionados con las disposiciones de 18, 19, 24, y 22 LRR comparables de los DTLR 1-4 (la cadena de DTLR5 incompleta incluye en este momento cuatro LRR próximos a la membrana) mediante análisis de la secuencia y del patrón (Altschul, *et al.* (1994) Nature Genet. 6:119-129; y Bork y Gibson (1996) Meth. Enzymol. 266:162-184) (Fig. 1). No obstante, una diferencia sorprendente en las cadenas de los DTLR humanos es la pérdida común de una región rica en cisteína de ~90 restos que está embebida variablemente en los ectodominios de Toll, 18w y el ORF de Mst (distanciado cuatro, seis y dos LRR, respectivamente, del límite de la membrana). Estas agrupaciones de cisteína son bipartitas, con mitades "superior" (terminando una LRR) e "inferior" (apiladas en lo alto de una LRR) distintas (Chiang y Beachy (1994) Mech. Develop. 47:225-239; Eldon, *et al.* (1994) Develop. 120:885-899; y Buchanan y Gay (1996) Prog. Biophys. Molec. Biol. 65:1-44); el módulo "superior" se repite en los DTLR de

*Drosophila* y de ser humano en forma de espaciador yuxtamembrana conservado (Fig. 1). Los autores de la presente invención sugieren que las agrupaciones de cisteína localizadas flexiblemente en los receptores de *Drosophila* (y otras proteínas con LRR), cuando se igualan la parte “superior” con la “inferior”, forman un módulo compacto con extremos emparejados que puede ser insertado entre cualquier par de LRR sin alterar el plegamiento global de los ectodominios de DTLR; dominios “extrudidos” análogos decoran las estructuras de otras proteínas (Russell (1994) *Protein Engin.* 7:1407-1410).

#### Diseño molecular del dominio de señalización TH

La comparación de la secuencia de los receptores Toll e IL-1 de tipo I (IL-1R1) ha descrito una semejanza distante de un dominio citoplásmico de ~200 aminoácidos que presumiblemente media la señalización por los factores de transcripción de tipo Rel similares. Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416; y (Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416; y Wasserman (1993) *Molec. Biol. Cell* 4:767-771). Las adiciones más recientes a este paradigma funcional incluyen un par de proteínas de resistencia a enfermedades de plantas de tabaco y lino que muestran un módulo TH N-terminal seguido de segmentos de unión a nucleótidos (NTPasa) y LRR (Wilson, *et al.* (1997) *Curr. Biol.* 7:175-178); en contraste, un “dominio de muerte” precede a la cadena TH de MyD88, un marcador de diferenciación mieloide intracelular (Mitcham, *et al.* (1996) *J. Biol. Chem.* 271:5777-5783; y Hardiman, *et al.* (1996) *Oncogene* 13:2467-2475) (Fig. 1). Los nuevos receptores de tipo IL-1 incluyen IL-1R3, una molécula de señalización accesoria, y receptores de orfano IL-1R4 (también denominados ST2/Fit-1/T1), IL-1R5 (proteína relacionada con IL-1R), e IL-1R6 (proteína 2 relacionada con IL-1R) (Mitcham, *et al.* (1996) *J. Biol. Chem.* 271:5777-5783; Hardiman, *et al.* (1996) *Oncogene* 13:2467-2475). Con las nuevas secuencias de los DTLR humanos, los autores de la presente invención han tratado de encontrar una definición estructural de hilo evolutivo analizando la conformación del módulo TH común: diez bloques de secuencia conservada que comprende 128 aminoácidos forman el pliegue del dominio TH mínimo; los espacios en el alineamiento marcan la localización probable de la secuencia y los bucles de longitud variable (Fig. 2a).

Dos algoritmos de predicción que sacan ventaja de los patrones de conservación y variación en secuencias de alineamiento múltiple, PHD (Rost y Sander (1994) *Proteins* 19:55-72) y DSC (King y Sternberg (1996) *Protein Sci.* 5:2298-2310), produjeron resultados fuertemente concordantes para el módulo de señalización TH (Fig. 2a). Cada bloque contiene un elemento estructural secundario discreto: la huella de hebras  $\beta$  alternantes (marcadas como A-E) y hélices  $\alpha$  (numeradas de 1-5) es el diagnóstico de un plegamiento de clase  $\beta/\alpha$  con hélices  $\alpha$  en ambas caras de una lámina  $\beta$  paralela. Se prevé que las hebras  $\beta$  hidrofóbicas A, C y D forman duelas “interiores” en la lámina  $\beta$ , mientras las hebras  $\beta$  anfipáticas B y E, más cortas se asemejan a unidades “borde” típicas (Fig. 2a). Esta asignación coincide con un orden de hebras B-A-C-D-E en la lámina  $\beta$  núcleo (Fig. 2b); los programas de comparación del pliegue (“mapeo”) y reconocimiento (“enhebrado”) (Fischer, *et al.* (1996) *FASEB J.* 10:126-136) devuelven fuertemente esta topología  $\beta/\alpha$  doblemente enrollada. Una predicción funcional, sorprendente de esta estructura esbozada para el dominio TH es que muchos de los restos cargados, conservados en el alineamiento múltiple se mapean en el extremo C-terminal de la lámina  $\beta$ : resto Asp16 (esquema de numeración de bloques - Fig. 2a) en el extremo de  $\beta$ A, Arg39 y Asp40 siguientes a  $\beta$ B, Glu75 en la primera vuelta de  $\alpha$ 3, y los restos Glu/Asp conservados más libremente en el bucle  $\beta$ D- $\alpha$ 4, o después de  $\beta$ E (Fig. 2a). La localización de otros cuatro restos conservados (Asp7, Glu28, y el par Arg57-Arg/Lys58) es compatible con una red de puentes salinos en el extremo N-terminal opuesto de la lámina  $\beta$  (Fig. 2a).

La función de señalización depende de la integridad estructural del dominio TH. Se han catalogado mutaciones o deleciones inactivantes en los límites del módulo (Fig. 2a) para IL-1R1 y Toll. Heguy, *et al.* (1992) *J. Biol. Chem.* 267:2605-2609; Croston, *et al.* (1995) *J. Biol. Chem.* 270:16514-16517; Schneider, *et al.* (1991) *Genes Develop.* 5:797-807; Norris y Manley. (1992) *Genes Develop.* 6:1654-1667; Norris y Manley (1995) *Genes Develop.* 9:358-369; y Norris y Manley (1996) *Genes Develop.* 10:862-872. Las cadenas de los DTLR 1-5 humanos que se extienden más allá del dominio TH mínimo (longitudes de 8, 0, 6, 22 y 18 restos, respectivamente) son muy similares a la “cola” de 4 aa, roma del ORF de Mst. Toll y 18w presentan colas de 102 y 207 restos no relacionadas (Fig. 2a) que pueden regular negativamente la señalización de los dominios TH fusionados. Norris y Manley (1995) *Genes Develop.* 9:358-369; y Norris y Manley (1996) *Genes Develop.* 10:862-872.

La relación evolutiva entre las proteínas disparejas que portan el dominio TH puede ser discernida mejor por medio del árbol filogenético derivado del alineamiento múltiple (Fig. 3). Cuatro ramas principales segregan las proteínas de plantas, los factores MyD88, los receptores de IL-1 y las moléculas de tipo Toll; la última rama agrupa los DTLR de *Drosophila* y humanos.

#### Dispersión cromosómica de genes de DTLR humanos

Con el fin de investigar la conexión genética de la familia de genes de los DTLR humanos naciente, los autores de la presente invención mapearon los loci cromosómicos de cuatro de los cinco genes mediante FISH (Fig. 4). El gen de DTLR1 ha sido previamente registrado gráficamente mediante el proyecto genoma humano: existe un locus de la base de datos STS (número de acceso dbSTS G06709, correspondiente a STS WI-7804 o SHGC-12827) para el ADNc Humrsc786 (Nomura, *et al.* (1994) *DNA Res* 1:27-35) y fija el gen en el intervalo del marcador del cromosoma 4 D4S1587-D42405 (50-56 cM) hacia 4p14. Esta asignación ha sido corroborada recientemente mediante análisis FISH. Taguchi, *et al.* (1996) *Genomics* 32:486-488. En el presente trabajo, los autores de la presente invención asignan los

genes de DTLR restantes a loci sobre el cromosoma 4q32 (DTLR2), 4q35 (DTLR3), 9q32-33 (DTLR4) y 1q33.3 (DTLR5). Durante el transcurso de este trabajo, se ha generado una STS para la EST de DTLR2 de origen (Núm. ID clon 80633) (número de acceso dbSTS T57791 para STS SHGC-33147) y se mapea en el intervalo del marcador del cromosoma 4 D4S424-D4S1548 (143-153 cM) en 4q32 - de acuerdo con los descubrimientos de los autores de la presente invención. Existe un espacio de ~50 cM entre los genes DTLR2 y DTLR3 sobre el brazo largo del cromosoma 4.

#### *Los genes DTLR son expresados diferencialmente*

Tanto Toll como 18w tienen patrones de expresión espaciales y temporales complejos en *Drosophila* que pueden apuntar a funciones más allá de la formación de patrones embrionarios. St. Johnston y Nusslein-Volhard (1992) *Cell* 68:201-219; Morisato y Anderson (1995) *Ann. Rev. Genet.* 29:371-399; Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416; Lemaitre, *et al.* (1996) *Cell* 86:973-983; Chiang y Beachy (1994) *Mech. Develop.* 47:225-239; y Eldon, *et al.* (1994) *Develop.* 120:885-899. Los autores de la presente invención han examinado la distribución espacial de los transcritos de DTLR mediante análisis de transferencia de ARNm que variaban las líneas celulares de tejidos humanos y cancerosas utilizando ADNc radiomarcados (Fig. 5). Se ha encontrado que DTLR1 es expresado ubicuamente, y a niveles más altos que los otros receptores. Reflejando presumiblemente el empalme alternativo, las formas del transcrito de DTLR1 "cortas" 3,0 kB y "largas" 8,0 kB están presentes en ovario y bazo, respectivamente (Fig. 5, paneles A y B). Un panel de ARNm de células cancerosas también muestra una expresión al alza destacada en una línea celular Raji de Linfoma de Burkitt (Fig. 5, panel C). El ARNm de DTLR2 es expresado menos ampliamente que el de DTLR1, con una especie de 4,0 kB detectada en pulmón y un transcrito de 4,4 kB evidente en corazón, cerebro y músculo. El patrón de distribución en el tejido de DTLR3 repite el de DTLR2 (Fig. 5, panel E). El DTLR3 también está presente en forma de dos transcritos principales de un tamaño de aproximadamente 4,0 y 6,0 kB, y los niveles más altos de expresión se observan en placenta y páncreas. En contraste, los mensajes de DTLR4 y DTLR5 parecen ser extremadamente específicos del tejido. Se detectó DTLR4 solamente en placenta como un único transcrito de un tamaño de ~7,0 kB. Se observó una señal débil de 4,0 kB para DTLR5 en ovario y monocitos de sangre periférica.

#### *Componentes de un sistema regulador evolutivamente ancestral*

Los planos moleculares originales y los destinos divergentes de las rutas de señalización pueden ser reconstruidos mediante enfoques genómicos comparativos. Miklos y Rubin (1996) *Cell* 86:521-529; Chothia (1994) *Develop.* 1994 Suppl., 27-33; Banfi, *et al.* (1996) *Nature Genet.* 13:167-174; y Wang, *et al.* (1996) *J. Biol. Chem.* 271:4468-4476. Los autores de la presente invención han utilizado esta lógica para identificar una familia de genes emergentes en seres humanos, que codifican, en este momento cinco parálogos de receptores, los DTLR 1-5, que son contrapartes evolutivas directas de una familia de genes de *Drosophila* encabezada por Toll (Figs. 1-3). La arquitectura conservada de los DTLR humanos y de mosca, los ectodominios LRR conservados y los módulos TH intracelulares (Fig. 1), da a entender que la ruta robusta acoplada a Toll en *Drosophila* (6, 7) sobrevive en vertebrados. La mejor evidencia se toma prestada de una ruta reiterada: el sistema de la IL-1 múltiple y su repertorio de dominios TH fusionados a receptores, IRAK, NF- $\kappa$ B y homólogos de I- $\kappa$ B (Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416; Wasserman (1993) *Molec. Biol. Cell* 4:767-771; Hardiman, *et al.* (1996) *Oncogene* 13:2467-2475; y Cao, *et al.* (1996) *Science* 271:1128-1131); también se ha caracterizado un factor de tipo Tube. No se sabe si los DTLR se pueden acoplar productivamente a la maquinaria de señalización de IL-1R, o en lugar de eso, se utiliza un grupo paralelo de proteínas. A diferencia de los receptores de IL-1, se prevé que el armazón de LRR de los DTLR humanos conserve una afinidad para los factores con nudo de cistina relacionados con Spätzle/Trunk; se han aislado ligandos de DTLR candidato (denominados PEN) que se ajustan a este molde.

Los mecanismos bioquímicos de transducción de la señal pueden ser evaluados mediante la conservación de los pliegues de la proteína interaccionante en una ruta. Miklos y Rubin (1996) *Cell* 86:521-529; Chothia (1994) *Develop.* 1994 Suppl., 27-33. En este momento, el paradigma de la señalización Toll implica algunas moléculas cuyos papeles están estrechamente definidos por sus estructuras, acciones o destinos: Pelle es una Ser/Thr quinasa (fosforilación), Dorsal es un factor de transcripción de tipo NF- $\kappa$ B (unión al ADN) y Cactus es un inhibidor de la repetición de ankirina (unión a Dorsal, degradación). Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416. En contraste, las funciones del dominio TH y Tube de Toll siguen siendo un enigma. Como otros receptores de citoquina (Heldin (1995) *Cell* 80:213-223), la dimerización mediada por ligando de Toll parece ser un evento desencadenante: las cisteínas libres en la región de la yuxtamembrana de Toll crean pares de receptores constitutivamente activos (Schneider, *et al.* (1991) *Genes Develop.* 5:797-807), y los receptores Torso-Toll quiméricos señalizan en forma de dímeros (Galindo, *et al.* (1995) *Develop.* 121:2209-2218); con todo, los truncamientos severos o la pérdida al por mayor del ectodominio Toll producen una señalización intracelular promiscua (Norris y Manley (1995) *Genes Develop.* 9:358-369; y Winans y Hashimoto (1995) *Molec. Biol. Cell* 6:587-596), reminiscente de los receptores oncogénicos con dominios catalíticos (Heldin (1995) *Cell* 80:213-223). Tube está localizado en la membrana, participa en el dominio N-terminal (muerte) de Pelle y está fosforilado, pero ninguna de las interacciones Toll-Tube o Toll-Pelle son registradas por el análisis de dos híbridos (Galindo, *et al.* (1995) *Develop.* 121:2209-2218; y Großhans, *et al.* (1994) *Nature* 372:563-566); esto último sugiere que el "estado" conformacional del dominio TH de Toll afecta de algún modo al reclutamiento del factor. Norris y Manley (1996) *Genes Develop.* 10:862-872; y Galindo, *et al.* (1995) *Develop.* 121:2209-2218.

En el corazón de estas desconcertantes cuestiones está la naturaleza estructural del módulo TH de Toll. Para estudiar esta cuestión, los autores de la presente invención se han aprovechado de la diversidad evolutiva de las secuencias de TH de insectos, plantas y vertebrados, que incorporan las cadenas de DTLR humanas, y han extraído el núcleo de la proteína conservada, mínima para la predicción de la estructura y el reconocimiento del plegamiento (Fig. 2). El plegamiento del dominio TH ( $\beta/\alpha$ )<sub>5</sub> fuertemente pronosticado con su agrupación asimétrica de restos ácidos es topológicamente idéntico a las estructuras de los reguladores de la respuesta en las rutas de señalización bacteriana de dos componentes (Volz (1993) *Biochemistry* 32:11741-11753; y Parkinson (1993) *Cell* 73:857-871) (Fig. 2). El regulador de la quimiotaxis prototipo CheY se une transitoriamente a un catión divalente en un “bolsillo de aspartato” en el extremo C de la lámina  $\beta$  núcleo; este catión proporciona estabilidad electrostática y facilita la fosforilación activadora de un Asp invariante. Volz (1993) *Biochemistry* 32:11741-11753. Del mismo modo, el dominio TH puede capturar cationes en su nido ácido, pero la activación, y la señalización aguas abajo, podría depender de la unión específica de un radical cargado negativamente: los ligandos aniónicos pueden superar potenciales en el sitio de unión intensamente negativos encerrándose en redes de enlaces de hidrógeno precisas. Ledvina, *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:6786-6791. Fascinantemente, el dominio DH puede no actuar simplemente como andamiaje pasivo para el ensamblaje de un complejo Tube/Pelle para Toll, o sistemas homólogos en plantas y vertebrados, en lugar de ello participa activamente como un verdadero disparador conformacional en la maquinaria de transducción de la señal. Quizás explicando la unión condicional de un complejo Tube/Pelle, la dimerización de Toll podría promover el desenmascaramiento, por medio de colas de receptores reguladores (Norris y Manley (1995) *Genes Develop.* 9:358-369; Norris y Manley (1996) *Genes Develop.* 10:862-872), o la unión por medio de activadores de molécula pequeña del bolsillo TH. No obstante, los módulos TH “libres” dentro de la célula (Norris y Manley (1995) *Genes Develop.* 9:358-369; Winans y Hashimoto (1995) *Molec. Biol. Cell* 6:587-596) podrían actuar como disparadores de tipo CheY catalíticos activando y acoplándose con complejos Tube/Pelle errantes.

## 25 *Receptores Morfogenéticos y defensa inmunitaria*

La conexión evolutiva entre los sistemas inmunitarios de insectos y vertebrados está impresa en el ADN: los genes que codifican los factores antimicrobianos en insectos presentan motivos aguas arriba similares a los elementos de respuesta en fase aguda que se sabe que se unen a los factores de transcripción NF- $\kappa$ B en mamíferos. Hultmark (1993) *Trends Genet.* 9:178-183. Dorsal, y dos factores relacionados con Dorsal, Dif y Relish, ayudan a inducir estas proteínas de defensa después de la sensibilización bacteriana (Reichhart, *et al.* (1993) *C. R. Acad. Sci. Paris* 316:1218-1224; Ip, *et al.* (1993) *Cell* 75:753-763; y Dushay, *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:10343-10347); Toll, u otros DTLR, modulan de un modo similar estas respuestas inmunitarias rápidas en *Drosophila* adulta (Lemaître, *et al.* (1996) *Cell* 86:973-983; y Rosetto, *et al.* (1995) *Biochem. Biophys. Res. Commun.* 209:111-116). Estos paralelismos mecánicos con la respuesta inflamatoria de IL-1 en vertebrados son la evidencia de la versatilidad funcional de la ruta de señalización de Toll, y sugieren una sinergia ancestral entre la formación del patrón embrionario y la inmunidad innata (Belvin y Anderson (1996) *Ann. Rev. Cell Develop. Biol.* 12:393-416; Lemaître, *et al.* (1996) *Cell* 86:973-983; Wasserman (1993) *Molec. Biol. Cell* 4:767-771; Wilson, *et al.* (1997) *Curr. Biol.* 7:175-178; Hultmark (1993) *Trends Genet.* 9:178-183; Reichhart, *et al.* (1993) *C. R. Acad. Sci. Paris* 316:1218-1224; Ip, *et al.* (1993) *Cell* 75:753-763; Dushay, *et al.* (1996) *Proc. Natl. Acad. Sci. USA* 93:10343-10347; Rosetto, *et al.* (1995) *Biochem. Biophys. Res. Commun.* 209:111-116; Medzhitov y Janeway (1997) *Curr. Opin. Immunol.* 9:4-9; y Medzhitov y Janeway (1997) *Curr. Opin. Immunol.* 9:4-9). La homología más próxima de las proteínas DTLR de insectos y humanas sugiere un solapamiento aún más fuerte de las funciones biológicas que reemplaza los paralelismos puramente inmunitarios para los sistemas de IL-1, y presta reguladores moleculares potenciales para las transformaciones dorso-ventrales y otras de los embriones de vertebrados. DeRobertis y Sasai (1996) *Nature* 380:37-40; y Arendt y Nübler-Jung (1997) *Mech. Develop.* 61:7-21.

La presente descripción de una familia de receptores robusta, emergente en seres humanos refleja el reciente descubrimiento de los receptores Frizzled de vertebrados para los factores de formación de patrones Wnt. Wang, *et al.* (1996) *J. Biol. Chem.* 271:4468-4476. Puesto que otros numerosos sistemas receptores de citoquina juegan papeles en el desarrollo temprano (Lemaire y Kodjabachian (1996) *Trends Genet.* 12:525-531), quizás los distintos contextos celulares de los embriones compactos y los adultos desgarrados simplemente dan como resultado rutas de señalización familiares y sus disparadores difundibles tienen diferentes resultados biológicos en diferentes momentos, p. ej., morfogénesis versus defensa inmunitaria para los DTLR. Para los sistemas relacionados con Toll de insectos, plantas, y seres humanos (Hardiman, *et al.* (1996) *Oncogene* 13:2467-2475; Wilson, *et al.* (1997) *Curr. Biol.* 7:175-178), estas señales cursan a través de un dominio TH regulador que se asemeja fascinantemente a una máquina de transducción bacteriana (Parkinson (1993) *Cell* 73:857-871).

En particular, el DTLR6 muestra rasgos estructurales que establecen su pertenencia a la familia. Por otra parte, los miembros de la familia han sido implicados en numerosas enfermedades evolutivas significativas y con función del sistema inmunitario innato. En particular, el DTLR6 ha sido mapeado en el cromosoma X en una localización que es un punto caliente para las principales anomalías evolutivas. Véase, p. ej., The Sanger Center: sitio en la red para el cromosoma X humano <http://www.sanger-ac.uk/HGP/ChrX/index.shtml>; y el sitio de la red Baylor College of Medicine Human Genome Sequencing <http://gc.bcm.tmc.edu:8088/cgi-bin/seq/home>.

El número de acceso para el PAC depositado es AC003046. Este número de acceso contiene la secuencia de dos PAC: RPC-164K3 y RPC-263P4. Estas dos secuencias PAC se mapearon en el cromosoma Xp22 humano en el sitio de la red entre los marcadores STS DXS704 y DXS7166. Esta región es un “punto caliente” para las anomalías evolutivas graves.

### III. Amplificación del fragmento DTLR mediante PCR

Se seleccionan dos secuencias de cebadores apropiados (véanse las Tablas 1 a 10). Se utiliza la RT-PCR sobre una muestra de ARNm apropiada seleccionada en busca de la presencia de mensaje para producir un ADNc parcial o completo, p. ej., una muestra que exprese el gen. Véase, p. ej., Innis, *et al.* (eds. 1990) PCR Protocols: A Guide to Methods and Applications Academic Press, San Diego, CA; y Dieffenbach y Dveksler (1995; eds.) PCR Primer: A Laboratory Manual Cold Spring Harbor Press, CSH, NY. Esto permitirá la determinación de una secuencia útil para sondear un gen completo en una genoteca de ADNc. El TLR6 es una secuencia contigua en el genoma, que puede sugerir que otros TLR también lo son. De este modo, la PCR en un ADN genómico puede producir una secuencia contigua completa, y en ese caso sería aplicable la metodología del paseo cromosómico. Alternativamente, las bases de datos de secuencias contendrán la secuencia correspondiente a las porciones de las realizaciones descritas, o formas íntimamente relacionadas, p. ej., empalme alternativo, etc. Las técnicas de clonación de la expresión también se pueden aplicar a genotecas de ADNc.

### IV. Distribución tisular de los DTLR

Ha sido detectado el mensaje para cada gen que codifica estos DTLR. Véanse las Figuras 5A-5F. Otras células y tejidos serán analizados mediante la tecnología apropiada, p. ej., PCR, inmunoanálisis, hibridación, o de otro modo. Las preparaciones de ADNc de tejidos y órganos se encuentran disponibles, p. ej., en Clontech, Mountain View, CA. La identificación de las fuentes de expresión natural es útil, como se ha descrito.

Análisis Southern: se digiere ADN (5  $\mu$ g) de una genoteca de ADNc amplificada primaria con las enzimas de restricción apropiadas para liberar los insertos, se hacen correr sobre un gel de agarosa al 1% y se transfieren a una membrana de nailon (Schleicher y Schuell, Keene, NH).

Las muestras para el aislamiento del ARNm humano incluirían típicamente, p. ej. : células mononucleares de sangre periférica (monocitos, células T, células NK, granulocitos, células B), en reposo (T100); células mononucleares de sangre periférica, activadas con anti-CD3 durante 2, 6, 12 h reunidas (T101); células T, clon TH0 Mot 72, en reposo (T102); células T, clon TH0 Mot 72, activadas con anti-CD28 y anti-CD3 durante 3, 6, 12 h reunidas (T103); células T, clon TH0 Mot 72, anérgicas tratadas con péptido específico durante 2, 7, 12 h reunidas (T104); células T, clon TH1 HY06, en reposo (T107); células T, clon TH1 HY06, activadas con anti-CD28 y anti-CD3 durante 3, 6, 12 h reunidas (T108); células T, clon TH1 HY06, anérgicas tratadas con péptido específico durante 2, 6, 12 h reunidas (T109); células T, clon TH2 HY935, en reposo (T110); células T, clon TH2 HY935, activadas con anti-CD28 y anti-CD3 durante 2, 7, 12 h reunidas (T111); células T CD4+ células T CD4 5RO-polarizadas 27 días en anti-CD28, IL-4, y anti IFN- $\gamma$ , TH2 polarizadas, activadas con anti-CD3 y anti-CD28 4 h (T116); líneas tumorales de células T Jurkat y Hut78, en reposo (T117); clones de células T, reunidas AD130.2, Tc783.12, Tc783.13, Tc783.58, Tc782.69, en reposo (T118); clones de células T y de células T al azar, en reposo (T119); Esplenocitos, en reposo (B100); Esplenocitos, activados con anti-CD40 y IL-4 (B101); líneas de células EBV de células B reunidas WT49, RSB, JY, CVIR, 721.221, RM3, HSY, en reposo (B102); línea de células B JY, activadas con PMA e ionomicina durante 1, 6 h reunidas (B103); clones NK 20 reunidos, en reposo (K100); clones NK 20 reunidos, activados con PMA e ionomicina durante 6 h (K101); clon NKL, derivado de sangre periférica de pacientes con leucemia LGL, tratado con IL-2 (K106); clon citotóxico NK 64 0-A3 0-1, en reposo (K107); línea precursora hematopoyética TF1, activada con PMA e ionomicina durante 1, 6 h reunidas (C100); línea premonocítica U937, en reposo (M100); línea premonocítica U937, activada con PMA e ionomicina durante 1, 6 h reunidas (M101); monocitos elutriados, activados con LPS, IFN $\gamma$ , anti-IL-10 durante 1, 2, 6, 12, 24 h reunidos (M102); monocitos elutriados, activados con LPS, IFN $\gamma$ , IL-10 durante 1, 2, 6, 12, 24 h reunidos (M103); monocitos elutriados, activados con LPS, IFN $\gamma$ , anti-IL-10 durante 4, 16 h reunidos (M106); monocitos elutriados, activados con LPS, IFN $\gamma$ , IL-10 durante 4, 16 h reunidos (M107); monocitos elutriados, activados con LPS durante 1 h (M108); monocitos elutriados, activados con LPS durante 6 h (M109); DC 70% CD1a+, de CD34+ GM-CSF, TNF $\alpha$  12 días, en reposo (D101); DC 70% CD1a+, de CD34+ GM-CSF, TNF $\alpha$  12 días, activadas con PMA e ionomicina durante 1 hr (D102); DC 70% CD1a+, de CD34+ GM-CSF, TNF $\alpha$  12 días, activadas con PMA e ionomicina durante 6 hr (D103); DC 95% CD1a+, de CD34+ GM-CSF, TNF $\alpha$  12 días clasificado mediante FACS, activadas con PMA e ionomicina durante 1, 6 h reunidas (D104); DC 95% CD14 +, ex CD34+ GM-CSF, TNF $\alpha$  12 días clasificadas mediante FACS, activadas con PMA e ionomicina 1, 6 hr reunidas (D105); DC CD1a+ CD86+, de CD34+ GM-CSF, TNF $\alpha$  12 días clasificadas mediante FACS, activadas con PMA e ionomicina durante 1, 6 h reunidas (D106); DC de monocitos GM-CSF, IL-4 5 días, en reposo (D107); DC de monocitos GM-CSF, IL-4 5 días, en reposo (D108); DC de monocitos GM-CSF, IL-4 5 días, activadas con LPS 4, 16 h reunidas (D109); DC de monocitos GM-CSF, IL-4 5 días, activadas con TNF $\alpha$ , monocito "supe" durante 4, 16 h reunidas (D110); tumor benigno de leiomioma L11 (X101); miometrio normal M5 (0115); leiomiomasarcoma GS1 maligno (X103); línea de sarcoma de fibroblastos de pulmón MRC5, activados con PMA e ionomicina durante 1, 6 h reunidos (C101); línea de células de carcinoma epitelial de riñón CHA, activada con PMA e ionomicina durante 1, 6 h reunida (C102); riñón fetal 28 wk masculino (0100); pulmón fetal 28 wk masculino (0101); hígado fetal 28 wk masculino (0102); corazón fetal 28 wk masculino (0103); cerebro fetal 28 wk masculino (0104); vesícula biliar fetal 28 wk masculina (0106); intestino delgado fetal 28 wk masculino (0107); tejido adiposo fetal 28 wk masculino (0108); ovario fetal 25 wk femenino (0109); útero fetal 25 wk femenino (0110); testículo fetal 28 wk masculino (0111); bazo fetal 28 wk masculino (0112); placenta adulta 28 wk (0113); y tonsila inflamada, de 12 años de edad (X100).

Las muestras para el aislamiento de ARNm de ratón pueden incluir, p. ej.: línea celular fibroblástica L de ratón en reposo (C200); células tranfectadas Braf:ER (fusión Braf a receptor de estrógeno), control (C201); células T, polarizadas con TH1 (Me114 brillante, células T CD4+ de bazo, polarizadas durante 7 días con IFN- $\gamma$  y anti IL-4; T200); células T, polarizadas con TH2 (Me114 brillante, células CD4+ de bazo, polarizadas durante 7 días con IL-4 y anti-IFN- $\gamma$ ; T201); células T, altamente polarizadas con TH1 (véase Openshaw, *et al.* (1995) J. Exp. Med. 182:1357-1367; activadas con anti-CD3 durante 2, 6, 16 h reunidas; T202); células T, altamente polarizadas con TH2 (véase Openshaw, *et al.* (1995) J. Exp. Med. 182:1357-1367; activadas con anti-CD3 durante 2, 6, 16 h reunidas; T203); células pre T CD44-CD25+, clasificadas de timo (T204); clon de células T TH1 D1.1, en reposo durante 3 después de la última estimulación con antígeno (T205); clon de células T TH1 D1.1, estimuladas con 10  $\mu$ g/ml ConA 15 h (T206); clon de células T TH2 CDC35, en reposo durante 3 semanas después de la última estimulación con antígeno (T207); clon de células T TH2 CDC35, estimuladas con 10  $\mu$ g/ml ConA 15 h (T208); células T no sometidas a tratamiento previo Me114+ de bazo, en reposo (T209); células T Me114+, polarizadas para Th1 con IFN- $\gamma$ /IL-12/anti-IL-4 durante 6, 12, 24 h reunidas (T210); células T Me114+, polarizadas para Th2 con IL-4/anti-IFN- $\gamma$  durante 6, 13, 24 h reunidas (T211); línea celular de leucemia de células B maduras no estimulada A20 (B200); línea de células B no estimuladas CH12 (B201); células B grandes no estimuladas de bazo (B202); células B de bazo total, activadas con LPS (B203); células dendríticas enriquecidas con metrizamida de bazo, en reposo (D200); células dendríticas de médula ósea, en reposo (D201); línea celular de monocitos RAW 264.7 activada con LPS 4 h (M200); macrófagos de médula ósea derivados con GM y M-CSF (M201); línea celular de macrófagos J774, en reposo (M202); línea celular de macrófagos J774 + LPS + anti-IL-10 a 0,5, 1, 3, 6, 12 h reunidos (M203); línea celular de macrófagos J774 + LPS + IL-10 a 0,5, 1, 3, 5, 12 h reunidos (M204); tejido de pulmón de ratón sensibilizado con aerosol, cebadores Th2, sensibilización con OVA en aerosol 7, 14, 23 h reunido (véase Garlisi, *et al.* (1995) Clinical Immunology and Immunopathology 75:75-83; X206); tejido de pulmón infectado con Nippostrongylus (véase Coffman, *et al.* (1989) Science 245:308-310; X200); pulmón adulto total, normal (0200); pulmón total, rag-1 (véase Schwarz, *et al.* (1993) Immunodeficiency 4:249-252; 0205); bazo IL-10 K.O. (véase Kuhn, *et al.* (1991) Cell 75:263-274; X201); bazo adulto total, normal (0201); bazo total, rag-1 (0207); placas de Peyer IL-10 K.O. (0202); placas de Peyer totales, normales (0210); nódulos linfáticos mesentéricos IL-10 K.O. (X203); nódulos linfáticos mesentéricos totales, normales (0211); colon IL-10 K.O. (X203); colon total, normal (0212); páncreas de ratón NOD (véase Makino, *et al.* (1980) Jikken Dobutsu 29:1-13; X205); timo total, rag-1 (0208); riñón total, rag-1 (0209); corazón total, rag-1 (0202); cerebro total, rag-1 (0203); testículo total, rag-1 (0204); hígado total, rag-1 (0206); tejido articular normal de rata (0300); y tejido articular artrítico de rata (X300).

30

#### V. Clonación de contrapartes de especies de DTLR

Se utilizan diferentes estrategias para obtener contrapartes de especies de estos DTLR, preferiblemente de otros primates. Un método consiste en la hibridación por cruzamiento utilizando sondas de ADN de especies íntimamente relacionadas. Puede resultar útil entrar en especies evolutivamente similares como etapas intermedias. Otro método consiste en utilizar sondas para PCR específicas basadas en la identificación de bloques de similitud o diferencia entre especies concretas, p. ej., genes, humanos, p. ej., zonas de secuencia polipeptídica o nucleotídica altamente conservadas o no conservadas. Alternativamente, se pueden utilizar anticuerpos para la clonación de la expresión.

40

#### VI. Producción de proteína DTLR de mamífero

Se diseña un constructo de fusión apropiado, p. ej., GST, para la expresión, p. ej., en *E. coli*. Por ejemplo, se construye un plásmido IGIF pGex de ratón y se transforma en *E. coli*. Las células recién transformadas se hacen crecer en medio LB que contiene 50  $\mu$ g/ml de ampicilina y se inducen con IPTG (Sigma, St. Louis, MO). Después de una inducción durante la noche, las bacterias se recogen y los sedimentos que contienen la proteína DTLR se aíslan. Los sedimentos se homogeneizan en tampón TE (Tris-base 50 mM pH 8,0, EDTA 10 mM y pefabloc 2 mM) en 2 litros. Este material se hace pasar a través de un microfluidificador (Microfluidics, Newton, MA) tres veces. El sobrenadante fluidificado se centrifuga en un rotor Sorvall GS-3 durante 1 h a 13.000 rpm. El sobrenadante resultante que contiene la proteína DTLR se filtra y se hace pasar por una columna de glutation-SEFAROSA equilibrada en Tris-base 50 mM pH 8,0. Las fracciones que contienen la proteína de fusión DTLR-GST se reúnen y se escinden con trombina (Enzyme Research Laboratories, Inc., South Bend, IN). La reserva escindida se hace pasar después a través de una columna Q-SEFAROSA equilibrada en Tris-base 50 mM. Las fracciones que contienen DTLR se reúnen y se diluyen en H<sub>2</sub>O destilada fría, para disminuir la conductividad, y se vuelven a pasar por una nueva columna Q-Sefarosa, sola o de manera sucesiva con una columna de anticuerpo de inmovinoafinidad. Las fracciones que contienen la proteína DTLR se reúnen, se toman alícuotas de las mismas, y se almacenan en el congelador a -70°C.

50

La comparación del espectro CD con la proteína DTLR1 puede sugerir que la proteína está correctamente plegada. Véase Hazuda, *et al.* (1969) J. Biol. Chem. 264:1689-1693.

60

#### VII. Análisis Biológicos con DTLR

Los análisis biológicos estarán dirigidos generalmente a la característica de unión al ligando de la proteína o a la actividad quinasa/fosfatasa del receptor. La actividad será típicamente reversible, como lo son muchas otras acciones enzimáticas, actividades mediadas por fosfatasa o fosforilasa, cuyas actividades son medidas fácilmente mediante procedimientos convencionales. Véanse, p. ej., Hardie, *et al.* (eds. 1995) The Protein Kinase FactBook vols. I y II,

65

Academic Press, San Diego, CA; Hanks, *et al.* (1991) *Meth. Enzymol.* 200:38-62; Hunter, *et al.* (1992) *Cell* 70:375-388; Lewin (1990) *Cell* 61:743-752; Pines, *et al.* (1991) *Cold Spring Harbor Symp. Quant. Biol.* 56:449-463; y Parker, *et al.* (1993) *Nature* 363:736-738.

- 5 La familia de interleuquinas 1 contiene moléculas, cada una de las cuales es un importante mediador de la enfermedad inflamatoria. Para una revisión comprensiva, véase Dinarello (1996) "Biologic basis for interleukin-1 in disease" *Blood* 87:2095-2147. Existen indicios de que los diferentes ligandos Toll pueden jugar papeles importantes en el inicio de la enfermedad, concretamente respuestas inflamatorias. El descubrimiento de proteínas novedosas relacionadas con la familia de IL-1 ofrece la identificación de moléculas que proporcionan la base molecular para el inicio de la enfermedad y permite el desarrollo de estrategias terapéuticas de una gama y una eficacia mejoradas.

#### VIII. Preparación de anticuerpos específicos, p. ej., para DTLR4

- 15 Se inmunizan ratones Balb/c endogámicos intraperitonealmente con formas recombinantes de la proteína, p. ej., DTLR4 purificado o células NIH-3T3 transfectadas estables. Los animales se refuerzan en momentos puntuales apropiados con proteína, con o sin coadyuvante adicional, para estimular adicionalmente la producción de anticuerpo. Se recoge el suero, o se producen hibridomas con los bazos recogidos.

- 20 Alternativamente, se inmunizan los ratones Balb/c con células transformadas con el gen o sus fragmentos, células endógenas o exógenas, o con membranas aisladas enriquecidas para la expresión del antígeno. Se recoge el suero en el momento apropiado, típicamente después de numerosas administraciones adicionales. Las diferentes técnicas de terapia génica pueden ser útiles, p. ej., en la producción de proteína *in situ*, para generar una respuesta inmunitaria.

- 25 Se pueden elaborar anticuerpos monoclonales. Por ejemplo, se fusionan esplenocitos con un compañero de fusión apropiado y se seleccionan los hibridomas en medio de crecimiento mediante procedimientos convencionales. Los sobrenadantes del hibridoma se escrutan en busca de la presencia de anticuerpos que se unen al DTLR deseado, p. ej., mediante ELISA u otro análisis. También se pueden seleccionar o preparar anticuerpos que reconocen específicamente las realizaciones de DTLR específicas.

- 30 En otro método, se presentan los péptidos sintéticos o la proteína purificada a un sistema inmunitario para generar anticuerpos monoclonales o policlonales. Véanse, p. ej., Coligan (1991) *Current Protocols in Immunology* Wiley/Greene; y Harlow y Lane (1989) *Antibodies: A Laboratory Manual* Cold Spring Harbor Press. En situaciones apropiadas, el reactivo de unión se marca como se ha descrito antes, p. ej., mediante fluorescencia o de otro modo, o se inmoviliza en un sustrato para métodos de selección repetitiva. Los ácidos nucleicos también pueden ser introducidos en células de un animal para producir el antígeno, lo que sirve para lograr una respuesta inmunitaria, véanse, p. ej., Wang, *et al.* (1993) *Proc. Nat'l. Acad. Sci.* 90:4156-4160; Barry, *et al.* (1994) *Bioethniques* 16:616-619; y Xiang, *et al.* (1995) *Immunity* 2: 129-135.

#### IX. Producción de proteínas de fusión, p. ej., con DTLR5

- 40 Se elaboran diferentes constructos de fusión con DTLR5. Esta porción del gen se fusiona con una etiqueta epitópica, p. ej., una etiqueta FLAG, o con un constructo de un sistema de dos híbridos. Véase, p. ej., Fields y Song (1989) *Nature* 340:245-246.

- La etiqueta epitópica se puede utilizar en un procedimiento de clonación de la expresión con detección con anticuerpos anti-FLAG para detectar un compañero de unión, p. ej., un ligando para el DTLR5 respectivo. El sistema de dos híbridos también se puede utilizar para aislar proteínas que se unen específicamente a DTLR5.

50

#### X. Mapeo cromosómico de los DTLR

- 55 Se preparan dispersiones cromosómicas. Se realiza la hibridación *in situ* sobre preparaciones de cromosoma obtenidas de linfocitos estimulados con fitohemaglutinina cultivados durante 72 h. Se añade 5-bromodesoxiuridina durante las siete horas finales de cultivo (60 µg/ml de medio), para asegurar un bandeo cromosómico post-hibridación de buena calidad.

- 60 Un fragmento apropiado, p. ej., un fragmento de PCR, amplificado con la ayuda de cebadores sobre un molde de ADNc de células B total, se clona en un vector apropiado. El vector se marca mediante traslado de cortes con H<sup>3</sup>. La sonda radiomarcada se hibrida con las dispersiones en metafase como describen Mattei, *et al.* (1985) *Hum. Genet.* 69:327-331.

- 65 Después de recubrir con emulsión Nuclear Track (KODAK NTB<sub>2</sub>), los portas se exponen, p. ej., durante 18 días a 4°C. Para evitar cualquier corrimiento de los granos de plata durante el procedimiento de bandeo, las dispersiones de cromosomas se tiñen primero con solución Giemsa tamponada y se fotografía la metafase. Después se realiza el bandeo R mediante el método del fluorocromo-fotolisis-Giemsa (FPG) y las metafases se vuelven a fotografiar antes del análisis.

## ES 2 340 210 T3

Alternativamente, se puede realizar FISH, como se ha descrito antes. Los genes de DTLR se localizan sobre diferentes cromosomas. DTLR2 y DTLR3 están localizados en el cromosoma 4 humano; DTLR4 está localizado en el cromosoma 9 humano, y DTLR5 está localizado en el cromosoma 1 humano. Véanse las Figuras 4A-4D.

### XI. Relación estructura actividad

La información sobre el carácter crítico de los restos concretos se determina utilizando procedimientos y análisis convencionales. El análisis de mutagénesis convencional se realiza, por ejemplo, generando muchas variantes diferentes en posiciones determinadas, p. ej., en las posiciones identificadas antes, y evaluando las actividades biológicas de las variantes. Esto se puede realizar hasta el punto de determinar las posiciones que modifican la actividad, o de centrarse en posiciones específicas para determinar los restos que pueden ser sustituidos para conservar, bloquear, o modular la actividad biológica.

Alternativamente, el análisis de las variantes naturales puede indicar qué posiciones toleran las mutaciones naturales. Esto puede resultar del análisis poblacional de la variación entre individuos, o a través de cepas o especies. Las muestras de los individuos seleccionados se analizan, p. ej., mediante análisis PCR y secuenciación. Esto permite la evaluación de los polimorfismos de la población.

### XI. Aislamiento de un ligando para un DTLR

Se puede utilizar un DTLR como reactivo de unión específico para identificar su compañero de unión, sacando ventaja de su especificidad de unión, muy probablemente se utilizaría un anticuerpo. Un reactivo de unión se marca como se ha descrito antes, p. ej., mediante fluorescencia o de otro modo, o se inmoviliza en un sustrato para los métodos de selección repetitiva.

La composición de unión se utiliza para escrutar una genoteca de expresión formada por una línea celular que expresa un compañero de unión, esto es, ligando, preferiblemente asociado a la membrana. Se utilizan técnicas de tinción convencionales para detectar o clasificar el ligando expresado en la superficie, o se escrutan mediante selección repetitiva las células transformadas que se expresan en la superficie. El escrutinio de la expresión intracelular se realiza mediante diferentes procedimientos de tinción o inmunofluorescencia. Véase también McMahan, *et al.* (1991) EMBO J. 10:2821-2832.

Por ejemplo, el día 0, se recubren previamente 2 portas Permanox de 2 cámaras con 1 ml por cámara de fibronectina, 10 ng/ml en PBS, durante 30 min a la temperatura ambiente. Se enjuaga una vez con PBS. Después se cultivan en placa células COS a 2-3 x 10<sup>5</sup> células por cámara en 1,5 ml de medio de crecimiento. Se incuban durante la noche a 37°C.

El día 1 para cada muestra, se preparan 0,5 ml de una solución de 66 µg/ml de DEAE-dextrano, 66 µM de cloroquina, y 4 µg de ADN en DME sin de suero. Para cada grupo, se prepara un control positivo, p. ej., de ADNc DTLR-FLAG a una dilución 1 y 1/200, y un supuesto negativo. Se enjuagan las células con DME sin de suero. Se añade la solución de ADN y se incuba 5 hr a 37°C. Se separa el medio y se añaden 0,5 ml de DMSO al 10% en DME durante 2,5 min. Se separa y se lava una vez con DME. Se añaden 1,5 ml de medio de crecimiento y se incuba durante la noche.

El día 2, se cambia el medio. Los días 3 o 4, las células se fijan y se tiñen. Se enjuagan las células dos veces con Solución Salina Tamponada de Hank (HBSS) y se fija en paraformaldehído al 4% (PFA)/glucosa durante 5 min. Se lava 3X con HBSS. Los portas se pueden almacenar a -80°C después de separar todo el líquido. Para cada cámara, se realizan incubaciones de 0,5 ml como sigue. Se añade HBSS/saponina (0,1%) con 32 µl/ml de NaN<sub>3</sub> 1 M durante 20 min. Después se lavan las células con HBSS/saponina 1X. Se añade el DTLR o el complejo DTLR/anticuerpo apropiado a las células y se incuban durante 30 min. Se lavan las células dos veces con HBSS/saponina. Si resulta apropiado, se añade primero el anticuerpo durante 30 minutos. Se añade el segundo anticuerpo, p. ej., anticuerpo anti-ratón Vector, a una dilución 1/200, y se incuba durante 30 min. Se prepara la solución de ELISA, p. ej., solución de peroxidasa de rábano picante Vector Elite ABC, y se preincuba durante 30 min. Se utilizan, p. ej., 1 gota de solución A (avidina) y 1 gota de solución B (biotina) por 2,5 ml de HBSS/saponina. Se lavan las células dos veces con HBSS/saponina. Se añade solución ABC HRP y se incuba durante 30 min. Se lavan las células dos veces con HBSS, el segundo lavado durante 2 min, que cierra las células. Después se añade ácido diaminobenzoico Vector (DAB) durante 5 a 10 min. Se utilizan 2 gotas de tampón más 4 gotas de DAB más 2 gotas de H<sub>2</sub>O<sub>2</sub> por 5 ml de agua destilada en vidrio. Se separa cuidadosamente la cámara y se enjuaga el porta en agua. Se seca al aire durante unos minutos, después se añade 1 gota de Crystal Mount y un cubre. Se cuece durante 5 min a 85-90°C.

Se evalúa la tinción positiva de las reservas y se subclonan progresivamente para el aislamiento de los genes individuales responsables de la unión.

Alternativamente, se utilizan los reactivos de DTLR para purificar por afinidad o clasificar las células que expresan un supuesto ligando. Véase, p. ej., Sambrook, *et al.* o Ausubel, *et al.*



## ES 2 340 210 T3

Otra estrategia consiste en escrutar un receptor unido a membrana mediante selección repetitiva. El ADNc del receptor se construye como se ha descrito antes. El ligando se puede inmovilizar y utilizar para inmovilizar las células de expresión. La inmovilización se puede lograr mediante el uso de anticuerpos apropiados que reconocen, p. ej., una secuencia FLAG de un constructo de fusión con DTLR, o mediante el uso de anticuerpos originados contra los primeros anticuerpos. Los ciclos recurrentes de selección y amplificación conducen al enriquecimiento de los clones apropiados y al aislamiento eventual de los clones que expresan el receptor.

Las genotecas de expresión de fagos pueden ser escrutadas en busca de DTLR de mamífero. Las técnicas de marcaje apropiadas, p. ej., anticuerpos anti-FLAG, permitirán el marcaje específico de los clones apropiados.

Las realizaciones específicas descritas en la presente memoria se ofrecen a modo de ejemplo solamente.

**REIVINDICACIONES**

- 5 1. Una proteína o péptido esencialmente puros o recombinantes con actividad de receptor de tipo Toll, donde dicha proteína o péptido muestra una identidad de secuencia de al menos 70% con el SEQ ID NO: 6 a lo largo de toda su longitud.
2. Un polipéptido esencialmente puro o recombinante que comprende la secuencia de aminoácidos del SEQ ID NO: 6.
- 10 3. Una proteína de fusión que comprende la proteína o péptido de la Reivindicación 1 o la Reivindicación 2.
4. Un anticuerpo o fragmento de anticuerpo que se une específicamente a la proteína o péptido de la Reivindicación 1 o la Reivindicación 2.
- 15 5. Un ácido nucleico que codifica la proteína o péptido de cualquiera de las Reivindicaciones 1 a 3.
6. Un ácido nucleico que muestra una identidad de al menos 80% con un ADNc que codifica el SEQ ID NO: 6 a lo largo de toda su longitud.
- 20 7. Un vector de expresión que comprende el ácido nucleico de la Reivindicación 5 o la Reivindicación 6.
8. Una célula anfitriona que comprende el vector de la Reivindicación 7.
- 25 9. Un procedimiento para producir recombinantemente un polipéptido de acuerdo con una cualquiera de las Reivindicaciones 1 a 3 que comprende cultivar la célula anfitriona de la Reivindicación 8 en condiciones en las que se expresa el polipéptido.

30

35

40

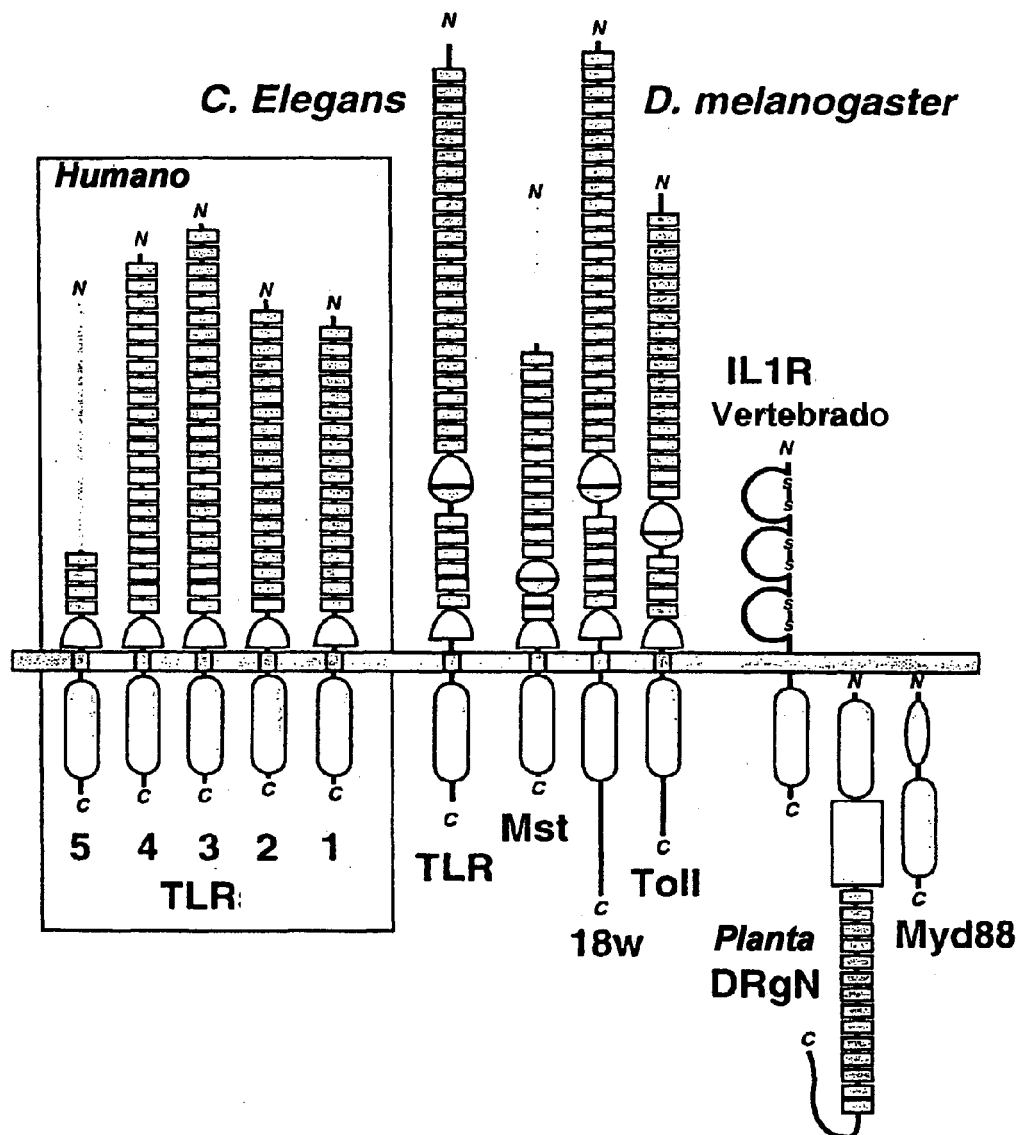
45

50

55

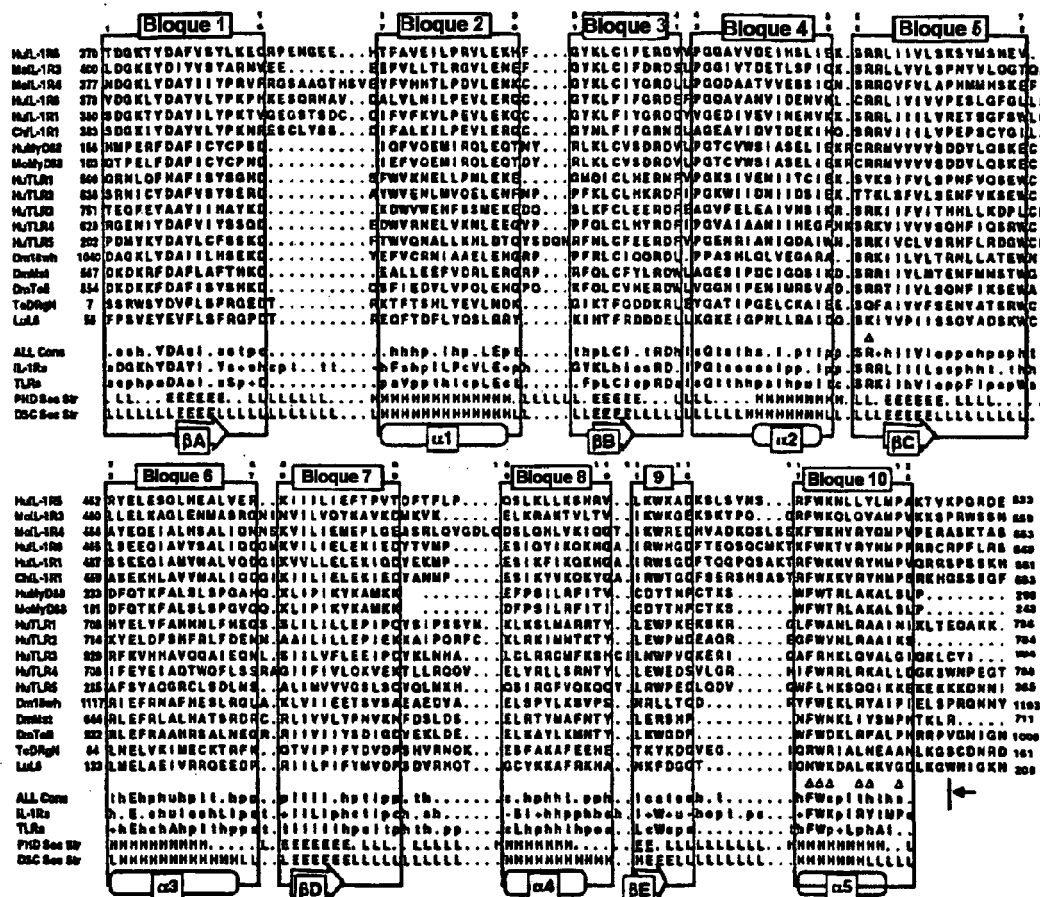
60

65

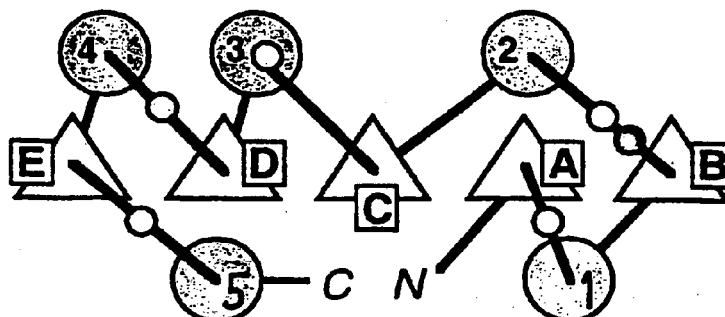


**FIG. 1**

**FIG. 2A**



**FIG. 2B**



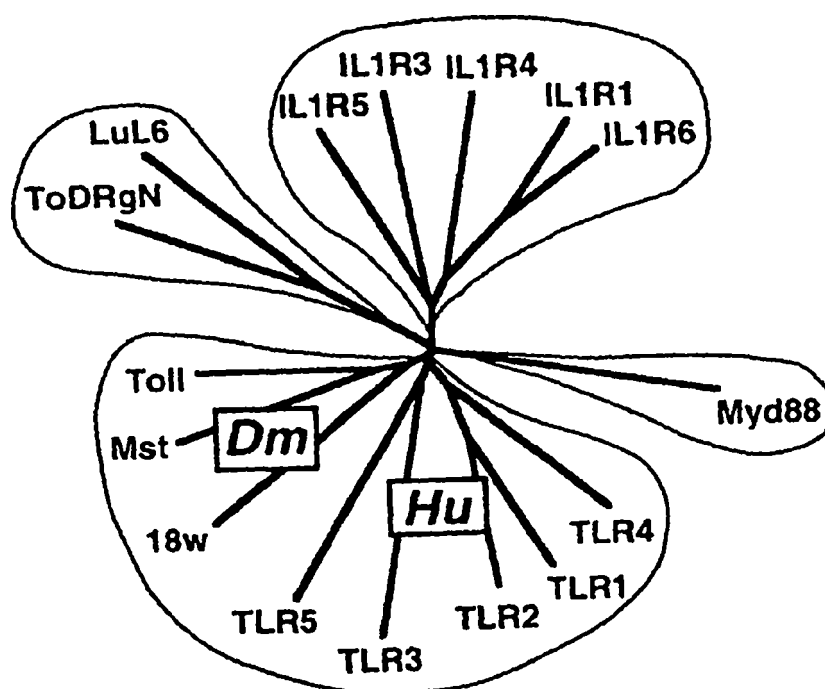


FIG. 3

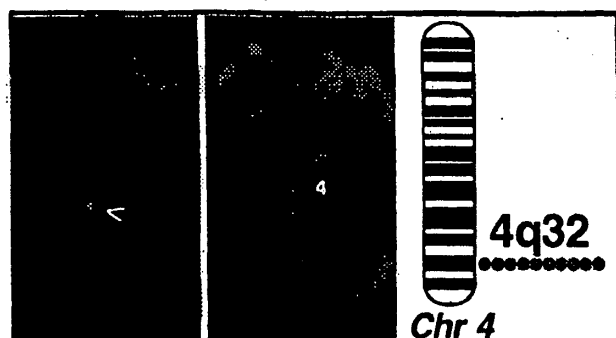


FIG. 4A

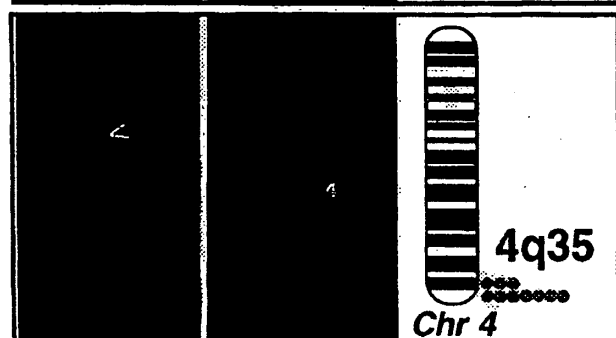


FIG. 4B

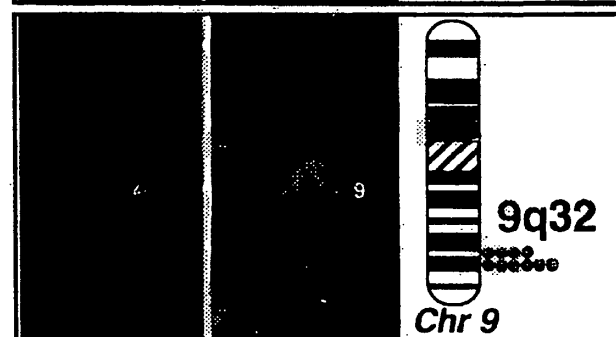


FIG. 4C



FIG. 4D

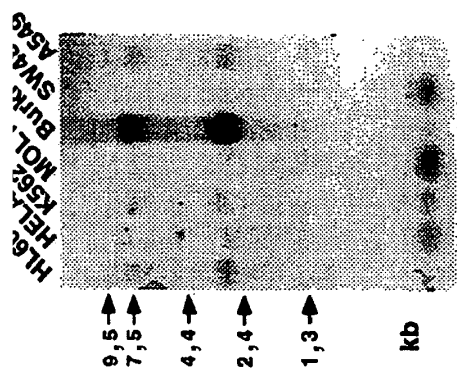


FIG. 5C

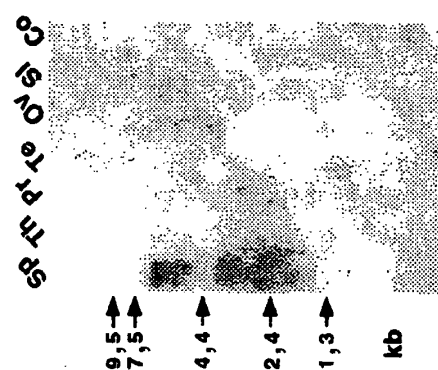


FIG. 5F

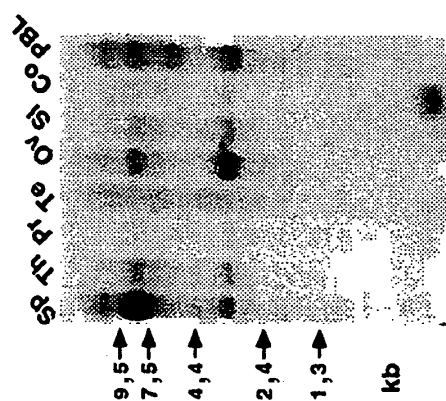


FIG. 5B

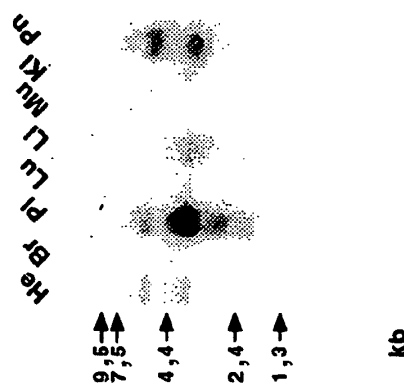


FIG. 5E

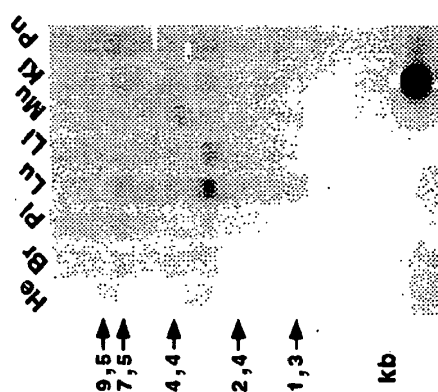


FIG. 5A

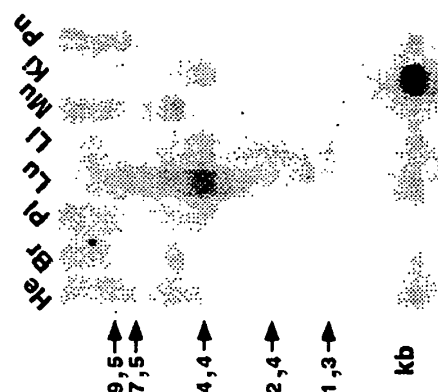


FIG. 5D

# ES 2 340 210 T3

## LISTA DE SECUENCIAS

### (1) INFORMACIÓN GENERAL:

#### (i) SOLICITANTE:

- (A) NOMBRE: Schering Corporation
- (B) CALLE: 2000 Galloping Hill Road
- (C) CIUDAD: Kenilworth
- (D) ESTADO: New Jersey
- (E) PAÍS: USA
- (F) CÓDIGO POSTAL: 07033
- (G) TELEFONO: (908) 298-4000
- (H) TELEFAX: (908) 298-5388

#### (ii) TÍTULO DE LA INVENCIÓN: PROTEÍNAS RECEPTORAS HUMANAS; REACTIVOS Y MÉTODOS RELACIONADOS

#### (iii) NÚMERO DE SECUENCIAS: 35

#### (iv) FORMA LEGIBLE CON ORDENADOR:

- (A) TIPO MEDIO: Disco Flexible
- (B) ORDENADOR: Macintosh Power PC
- (C) SISTEMA OPERATIVO: 8.0
- (D) SOPORTE LÓGICO: Microsoft Word 6.0

#### (v) DATOS DE LA SOLICITUD ACTUAL:

- (A) NÚMERO DE SOLICITUD:
- (B) FECHA DE PRESENTACIÓN:
- (C) CLASIFICACIÓN:

#### (vi) DATOS DE LA SOLICITUD ANTERIOR:

- (A) NÚM. DE SOLICITUD.: USSN 60/044,293
- (B) FECHA DE PRESENTACIÓN: 07-MAY-1997
- (A) NÚM. DE SOLICITUD.: USSN 60/072,212
- (B) FECHA DE PRESENTACIÓN: 22-ENERO-1998
- (A) NÚM. DE SOLICITUD.: USSN 60/076,947
- (B) FECHA DE PRESENTACIÓN: 05-MAR-1998

### (2) INFORMACIÓN PARA EL SEQ ID NO: 1:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 2367 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..2358



# ES 2 340 210 T3

(ix) RASGO:

(A) NOMBRE/CLAVE: mat\_péptido

(B) LOCALIZACIÓN: 67..2358

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:1:

10	ATG ACT AGC ATC TTC CAT TTT GCC ATT ATC TTC ATG TTA ATA CTT CAG Met Thr Ser Ile Phe His Phe Ala Ile Ile Phe Met Leu Ile Leu Gln -22 -20 -15 -10	48
15	ATC AGA ATA CAA TTA TCT GAA GAA AGT GAA TTT TTA GTT GAT AGG TCA Ile Arg Ile Gln Leu Ser Glu Glu Ser Glu Phe Leu Val Asp Arg Ser -5 1 5 10	96
20	AAA AAC GGT CTC ATC CAC GTT CCT AAA GAC CTA TCC CAG AAA ACA ACA Lys Asn Gly Leu Ile His Val Pro Lys Asp Leu Ser Gln Lys Thr Thr 15 20 25	144
25	ATC TTA AAT ATA TCG CAA AAT TAT ATA TCT GAG CTT TGG ACT TCT GAC Ile Leu Asn Ile Ser Gln Asn Tyr Ile Ser Glu Leu Trp Thr Ser Asp 30 35 40	192
30	ATC TTA TCA CTG TCA AAA CTG AGG ATT TTG ATA ATT TCT CAT AAT AGA Ile Leu Ser Leu Ser Lys Leu Arg Ile Leu Ile Ile Ser His Asn Arg 45 50 55	240
35	ATC CAG TAT CTT GAT ATC AGT GTT TTC AAA TTC AAC CAG GAA TTG GAA Ile Gln Tyr Leu Asp Ile Ser Val Phe Lys Phe Asn Gln Glu Leu Glu 60 65 70	288
40	TAC TTG GAT TTG TCC CAC AAC AAG TTG GTG AAG ATT TCT TGC CAC CCT Tyr Leu Asp Leu Ser His Asn Lys Leu Val Lys Ile Ser Cys His Pro 75 80 85 90	336
45	ACT GTG AAC CTC AAG CAC TTG GAC CTG TCA TTT AAT GCA TTT GAT GCC Thr Val Asn Leu Lys His Leu Asp Leu Ser Phe Asn Ala Phe Asp Ala 95 100 105	384
50	CTG CCT ATA TGC AAA GAG TTT GGC AAT ATG TCT CAA CTA AAA TTT CTG Leu Pro Ile Cys Lys Glu Phe Gly Asn Met Ser Gln Leu Lys Phe Leu 110 115 120	432
55	GGG TTG AGC ACC ACA CAC TTA GAA AAA TCT AGT GTG CTG CCA ATT GCT Gly Leu Ser Thr Thr His Leu Glu Lys Ser Ser Val Leu Pro Ile Ala 125 130 135	480
60	CAT TTG AAT ATC AGC AAG GTC TTG CTG GTC TTA GGA GAG ACT TAT GGG His Leu Asn Ile Ser Lys Val Leu Leu Val Leu Gly Glu Thr Tyr Gly 140 145 150	528
65	GAA AAA GAA GAC CCT GAG GGC CTT CAA GAC TTT AAC ACT GAG AGT CTG Glu Lys Glu Asp Pro Glu Gly Leu Gln Asp Phe Asn Thr Glu Ser Leu 155 160 165 170	576
70	CAC ATT GTG TTC CCC ACA AAC AAA GAA TTC CAT TTT ATT TTG GAT GTG His Ile Val Phe Pro Thr Asn Lys Glu Phe His Phe Ile Leu Asp Val 175 180 185	624
75	TCA GTC AAG ACT GTA GCA AAT CTG GAA CTA TCT AAT ATC AAA TGT GTG Ser Val Lys Thr Val Ala Asn Leu Glu Leu Ser Asn Ile Lys Cys Val 190 195 200	672
80	CTA GAA GAT AAC AAA TGT TCT TAC TTC CTA AGT ATT CTG GCG AAA CTT Leu Glu Asp Asn Lys Cys Ser Tyr Phe Leu Ser Ile Leu Ala Lys Leu	720

# ES 2 340 210 T3

	205	210	215	
5	CAA ACA AAT CCA AAG TTA TCA AGT CTT ACC TTA AAC AAC ATT GAA ACA Gln Thr Asn Pro Lys Leu Ser Ser Leu Thr Leu Asn Asn Ile Glu Thr 220 225 230			768
10	ACT TGG AAT TCT TTC ATT AGG ATC CTC CAA CTA GTT TGG CAT ACA ACT Thr Trp Asn Ser Phe Ile Arg Ile Leu Gln Leu Val Trp His Thr Thr 235 240 245 250			816
15	GTA TGG TAT TTC TCA ATT TCA AAC GTG AAG CTA CAG GGT CAG CTG GAC Val Trp Tyr Phe Ser Ile Ser Asn Val Lys Leu Gln Gly Gln Leu Asp 255 260 265			864
20	TTC AGA GAT TTT GAT TAT TCT GGC ACT TCC TTG AAG GCC TTG TCT ATA Phe Arg Asp Phe Asp Tyr Ser Gly Thr Ser Leu Lys Ala Leu Ser Ile 270 275 280			912
25	CAC CAA GTT GTC AGC GAT GTG TTC GGT TTT CCG CAA AGT TAT ATC TAT His Gln Val Val Ser Asp Val Phe Gly Phe Pro Gln Ser Tyr Ile Tyr 285 290 295			960
30	GAA ATC TTT TCG AAT ATG AAC ATC AAA AAT TTC ACA GTG TCT GGT ACA Glu Ile Phe Ser Asn Met Asn Ile Lys Asn Phe Thr Val Ser Gly Thr 300 305 310			1008
35	CGC ATG GTC CAC ATG CTT TGC CCA TCC AAA ATT AGC CCG TTC CTG CAT Arg Met Val His Met Leu Cys Pro Ser Lys Ile Ser Pro Phe Leu His 315 320 325 330			1056
40	TTG GAT TTT TCC AAT AAT CTC TTA ACA GAC ACG GTT TTT GAA AAT TGT Leu Asp Phe Ser Asn Asn Leu Leu Thr Asp Thr Val Phe Glu Asn Cys 335 340 345			1104
45	GGG CAC CTT ACT GAG TTG GAG ACA CTT ATT TTA CAA ATG AAT CAA TTA Gly His Leu Thr Glu Leu Glu Thr Leu Ile Leu Gln Met Asn Gln Leu 350 355 360			1152
50	AAA GAA CTT TCA AAA ATA GCT GAA ATG ACT ACA CAG ATG AAG TCT CTG Lys Glu Leu Ser Lys Ile Ala Glu Met Thr Thr Gln Met Lys Ser Leu 365 370 375			1200
55	CAA CAA TTG GAT ATT AGC CAG AAT TCT GTA AGC TAT GAT GAA AAG AAA Gln Gln Leu Asp Ile Ser Gln Asn Ser Val Ser Tyr Asp Glu Lys Lys 380 385 390			1248
60	GGA GAC TGT TCT TGG ACT AAA AGT TTA TTA AGT TTA AAT ATG TCT TCA Gly Asp Cys Ser Trp Thr Lys Ser Leu Leu Ser Leu Asn Met Ser Ser 395 400 405 410			1296
65	AAT ATA CTT ACT GAC ACT ATT TTC AGA TGT TTA CCT CCC AGG ATC AAG Asn Ile Leu Thr Asp Thr Ile Phe Arg Cys Leu Pro Pro Arg Ile Lys 415 420 425			1344
70	GTA CTT GAT CTT CAC AGC AAT AAA ATA AAG AGC ATT CCT AAA CAA GTC Val Leu Asp Leu His Ser Asn Lys Ile Lys Ser Ile Pro Lys Gln Val 430 435 440			1392
75	GTA AAA CTG GAA GCT TTG CAA GAA CTC AAT GTT GCT TTC AAT TCT TTA Val Lys Leu Glu Ala Leu Gln Glu Leu Asn Val Ala Phe Asn Ser Leu 445 450 455			1440

# ES 2 340 210 T3

	ACT GAC CTT CCT GGA TGT GGC AGC TTT AGC AGC CTT TCT GTA TTG ATC	1488
	Thr Asp Leu Pro Gly Cys Gly Ser Phe Ser Ser Leu Ser Val Leu Ile	
	460 465 470	
5	ATT GAT CAC AAT TCA GTT TCC CAC CCA TCA GCT GAT TTC TTC CAG AGC	1536
	Ile Asp His Asn Ser Val Ser His Pro Ser Ala Asp Phe Phe Gln Ser	
	475 480 485 490	
10	TGC CAG AAG ATG AGG TCA ATA AAA GCA GGG GAC AAT CCA TTC CAA TGT	1584
	Cys Gln Lys Met Arg Ser Ile Lys Ala Gly Asp Asn Pro Phe Gln Cys	
	495 500 505	
	ACC TGT GAG CTC GGA GAA TTT GTC AAA AAT ATA GAC CAA GTA TCA AGT	1632
	Thr Cys Glu Leu Gly Glu Phe Val Lys Asn Ile Asp Gln Val Ser Ser	
	510 515 520	
15	GAA GTG TTA GAG GGC TGG CCT GAT TCT TAT AAG TGT GAC TAC CCG GAA	1680
	Glu Val Leu Glu Gly Trp Pro Asp Ser Tyr Lys Cys Asp Tyr Pro Glu	
	525 530 535	
20	AGT TAT AGA GGA ACC CTA CTA AAG GAC TTT CAC ATG TCT GAA TTA TCC	1728
	Ser Tyr Arg Gly Thr Leu Leu Lys Asp Phe His Met Ser Glu Leu Ser	
	540 545 550	
25	TGC AAC ATA ACT CTG CTG ATC GTC ACC ATC GTT GCC ACC ATG CTG GTG	1776
	Cys Asn Ile Thr Leu Leu Ile Val Thr Ile Val Ala Thr Met Leu Val	
	555 560 565 570	
	TTG GCT GTG ACT GTG ACC TCC CTC TGC ATC TAC TTG GAT CTG CCC TGG	1824
	Leu Ala Val Thr Val Thr Ser Leu Cys Ile Tyr Leu Asp Leu Pro Trp	
	575 580 585	
30	TAT CTC AGG ATG GTG TGC CAG TGG ACC CAG ACC CGG CGC AGG GCC AGG	1872
	Tyr Leu Arg Met Val Cys Gln Trp Thr Gln Thr Arg Arg Arg Ala Arg	
	590 595 600	
35	AAC ATA CCC TTA GAA GAA CTC CAA AGA AAT CTC CAG TTT CAT GCA TTT	1920
	Asn Ile Pro Leu Glu Glu Leu Gln Arg Asn Leu Gln Phe His Ala Phe	
	605 610 615	
	ATT TCA TAT AGT GGG CAC GAT TCT TTC TGG GTG AAG AAT GAA TTA TTG	1968
	Ile Ser Tyr Ser Gly His Asp Ser Phe Trp Val Lys Asn Glu Leu Leu	
	620 625 630	
40	CCA AAC CTA GAG AAA GAA GGT ATG CAG ATT TGC CTT CAT GAG AGA AAC	2016
	Pro Asn Leu Glu Lys Glu Gly Met Gln Ile Cys Leu His Glu Arg Asn	
	635 640 645 650	
45	TTT GTT CCT GGC AAG AGC ATT GTG GAA AAT ATC ATC ACC TGC ATT GAG	2064
	Phe Val Pro Gly Lys Ser Ile Val Glu Asn Ile Ile Thr Cys Ile Glu	
	655 660 665	
	AAG AGT TAC AAG TCC ATC TTT GTT TTG TCT CCC AAC TTT GTC CAG AGT	2112
	Lys Ser Tyr Lys Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Ser	
	670 675 680	
50	GAA TGG TGC CAT TAT GAA CTC TAC TTT GCC CAT CAC AAT CTC TTT CAT	2160
	Glu Trp Cys His Tyr Glu Leu Tyr Phe Ala His His Asn Leu Phe His	
	685 690 695	

# ES 2 340 210 T3

	GAA GGA TCT AAT AGC TTA ATC CTG ATC TTG CTG GAA CCC ATT CCG CAG	2208
	Glu Gly Ser Asn Ser Leu Ile Leu Ile Leu Leu Glu Pro Ile Pro Gln	
	700 705 710	
5	TAC TCC ATT CCT AGC AGT TAT CAC AAG CTC AAA AGT CTC ATG GCC AGG	2256
	Tyr Ser Ile Pro Ser Ser Tyr His Lys Leu Lys Ser Leu Met Ala Arg	
	715 720 725 730	
10	AGG ACT TAT TTG GAA TGG CCC AAG GAA AAG AGC AAA CGT GGC CTT TTT	2304
	Arg Thr Tyr Leu Glu Trp Pro Lys Glu Lys Ser Lys Arg Gly Leu Phe	
	735 740 745	
15	TGG GCT AAC TTA AGG GCA GCC ATT AAT ATT AAG CTG ACA GAG CAA GCA	2352
	Trp Ala Asn Leu Arg Ala Ala Ile Asn Ile Lys Leu Thr Glu Gln Ala	
	750 755 760	
	AAG AAA TAGTCTAGA	2367
	Lys Lys	

## (2) INFORMACIÓN PARA EL SEQ ID NO:2:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 786 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

### (ii) TIPO DE MOLÉCULA: proteína

### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:2:

Met Thr Ser Ile Phe His Phe Ala Ile Ile Phe Met Leu Ile Leu Gln	
-22 -20 -15 -10	
Ile Arg Ile Gln Leu Ser Glu Glu Ser Glu Phe Leu Val Asp Arg Ser	
-5 1 5 10	
Lys Asn Gly Leu Ile His Val Pro Lys Asp Leu Ser Gln Lys Thr Thr	
15 20 25	
Ile Leu Asn Ile Ser Gln Asn Tyr Ile Ser Glu Leu Trp Thr Ser Asp	
30 35 40	
Ile Leu Ser Leu Ser Lys Leu Arg Ile Leu Ile Ile Ser His Asn Arg	
45 50 55	
Ile Gln Tyr Leu Asp Ile Ser Val Phe Lys Phe Asn Gln Glu Leu Glu	
60 65 70	
Tyr Leu Asp Leu Ser His Asn Lys Leu Val Lys Ile Ser Cys His Pro	
75 80 85 90	
Thr Val Asn Leu Lys His Leu Asp Leu Ser Phe Asn Ala Phe Asp Ala	
95 100 105	
Leu Pro Ile Cys Lys Glu Phe Gly Asn Met Ser Gln Leu Lys Phe Leu	
110 115 120	
Gly Leu Ser Thr Thr His Leu Glu Lys Ser Ser Val Leu Pro Ile Ala	
125 130 135	

# ES 2 340 210 T3

	His	Leu	Asn	Ile	Ser	Lys	Val	Leu	Leu	Val	Leu	Gly	Glu	Thr	Tyr	Gly	
	140						145					150					
5	Glu	Lys	Glu	Asp	Pro	Glu	Gly	Leu	Gln	Asp	Phe	Asn	Thr	Glu	Ser	Leu	
	155					160					165					170	
	His	Ile	Val	Phe	Pro	Thr	Asn	Lys	Glu	Phe	His	Phe	Ile	Leu	Asp	Val	
					175					180					185		
10	Ser	Val	Lys	Thr	Val	Ala	Asn	Leu	Glu	Leu	Ser	Asn	Ile	Lys	Cys	Val	
				190					195					200			
	Leu	Glu	Asp	Asn	Lys	Cys	Ser	Tyr	Phe	Leu	Ser	Ile	Leu	Ala	Lys	Leu	
			205					210					215				
15	Gln	Thr	Asn	Pro	Lys	Leu	Ser	Ser	Leu	Thr	Leu	Asn	Asn	Ile	Glu	Thr	
	220						225					230					
	Thr	Trp	Asn	Ser	Phe	Ile	Arg	Ile	Leu	Gln	Leu	Val	Trp	His	Thr	Thr	
	235					240					245					250	
20	Val	Trp	Tyr	Phe	Ser	Ile	Ser	Asn	Val	Lys	Leu	Gln	Gly	Gln	Leu	Asp	
					255					260						265	
	Phe	Arg	Asp	Phe	Asp	Tyr	Ser	Gly	Thr	Ser	Leu	Lys	Ala	Leu	Ser	Ile	
				270				275						280			
25	His	Gln	Val	Val	Ser	Asp	Val	Phe	Gly	Phe	Pro	Gln	Ser	Tyr	Ile	Tyr	
			285					290					295				
30	Glu	Ile	Phe	Ser	Asn	Met	Asn	Ile	Lys	Asn	Phe	Thr	Val	Ser	Gly	Thr	
	300						305					310					
	Arg	Met	Val	His	Met	Leu	Cys	Pro	Ser	Lys	Ile	Ser	Pro	Phe	Leu	His	
	315					320					325					330	
35	Leu	Asp	Phe	Ser	Asn	Asn	Leu	Leu	Thr	Asp	Thr	Val	Phe	Glu	Asn	Cys	
					335					340					345		
	Gly	His	Leu	Thr	Glu	Leu	Glu	Thr	Leu	Ile	Leu	Gln	Met	Asn	Gln	Leu	
				350					355					360			
40	Lys	Glu	Leu	Ser	Lys	Ile	Ala	Glu	Met	Thr	Thr	Gln	Met	Lys	Ser	Leu	
			365					370					375				
	Gln	Gln	Leu	Asp	Ile	Ser	Gln	Asn	Ser	Val	Ser	Tyr	Asp	Glu	Lys	Lys	
		380					385					390					
45	Gly	Asp	Cys	Ser	Trp	Thr	Lys	Ser	Leu	Leu	Ser	Leu	Asn	Met	Ser	Ser	
	395					400					405					410	
	Asn	Ile	Leu	Thr	Asp	Thr	Ile	Phe	Arg	Cys	Leu	Pro	Pro	Arg	Ile	Lys	
					415				420						425		
50	Val	Leu	Asp	Leu	His	Ser	Asn	Lys	Ile	Lys	Ser	Ile	Pro	Lys	Gln	Val	
				430					435					440			
	Val	Lys	Leu	Glu	Ala	Leu	Gln	Glu	Leu	Asn	Val	Ala	Phe	Asn	Ser	Leu	
			445					450					455				

# ES 2 340 210 T3

5 Thr Asp Leu Pro Gly Cys Gly Ser Phe Ser Ser Leu Ser Val Leu Ile  
 460 465 470  
 Ile Asp His Asn Ser Val Ser His Pro Ser Ala Asp Phe Phe Gln Ser  
 475 480 485 490  
 10 Cys Gln Lys Met Arg Ser Ile Lys Ala Gly Asp Asn Pro Phe Gln Cys  
 495 500 505  
 Thr Cys Glu Leu Gly Glu Phe Val Lys Asn Ile Asp Gln Val Ser Ser  
 510 515 520  
 15 Glu Val Leu Glu Gly Trp Pro Asp Ser Tyr Lys Cys Asp Tyr Pro Glu  
 525 530 535  
 Ser Tyr Arg Gly Thr Leu Leu Lys Asp Phe His Met Ser Glu Leu Ser  
 540 545 550  
 20 Cys Asn Ile Thr Leu Leu Ile Val Thr Ile Val Ala Thr Met Leu Val  
 555 560 565 570  
 Leu Ala Val Thr Val Thr Ser Leu Cys Ile Tyr Leu Asp Leu Pro Trp  
 575 580 585  
 25 Tyr Leu Arg Met Val Cys Gln Trp Thr Gln Thr Arg Arg Arg Ala Arg  
 590 595 600  
 Asn Ile Pro Leu Glu Glu Leu Gln Arg Asn Leu Gln Phe His Ala Phe  
 605 610 615  
 30 Ile Ser Tyr Ser Gly His Asp Ser Phe Trp Val Lys Asn Glu Leu Leu  
 620 625 630  
 Pro Asn Leu Glu Lys Glu Gly Met Gln Ile Cys Leu His Glu Arg Asn  
 635 640 645 650  
 35 Phe Val Pro Gly Lys Ser Ile Val Glu Asn Ile Ile Thr Cys Ile Glu  
 655 660 665  
 Lys Ser Tyr Lys Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Ser  
 670 675 680  
 40 Glu Trp Cys His Tyr Glu Leu Tyr Phe Ala His His Asn Leu Phe His  
 685 690 695  
 Glu Gly Ser Asn Ser Leu Ile Leu Ile Leu Leu Glu Pro Ile Pro Gln  
 700 705 710  
 45 Tyr Ser Ile Pro Ser Ser Tyr His Lys Leu Lys Ser Leu Met Ala Arg  
 715 720 725 730  
 Arg Thr Tyr Leu Glu Trp Pro Lys Glu Lys Ser Lys Arg Gly Leu Phe  
 735 740 745  
 50 Trp Ala Asn Leu Arg Ala Ala Ile Asn Ile Lys Leu Thr Glu Gln Ala  
 750 755 760  
 55 Lys Lys

## (2) INFORMACIÓN PARA EL SEQ ID NO:3:

- 60 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 2355 pares de bases  
 (B) TIPO: ácido nucleico  
 65 (C) CUALIDAD DE LA HEBRA: sencilla  
 (D) TOPOLOGÍA: lineal

## ES 2 340 210 T3

(ii) TIPO DE MOLÉCULA: ADNc

(ix) RASGO:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 1..2352

(ix) RASGO:

(A) NOMBRE/CLAVE: mat\_péptido

(B) LOCALIZACIÓN: 67..2352

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:3:

15	ATG CCA CAT ACT TTG TGG ATG GTG TGG GTC TTG GGG GTC ATC ATC AGC	48
	Met Pro His Thr Leu Trp Met Val Trp Val Leu Gly Val Ile Ile Ser	
	-22 -20 -15 -10	
20	CTC TCC AAG GAA GAA TCC TCC AAT CAG GCT TCT CTG TCT TGT GAC CGC	96
	Leu Ser Lys Glu Glu Ser Ser Asn Gln Ala Ser Leu Ser Cys Asp Arg	
	-5 1 5 10	
25	AAT GGT ATC TGC AAG GGC AGC TCA GGA TCT TTA AAC TCC ATT CCC TCA	144
	Asn Gly Ile Cys Lys Gly Ser Ser Gly Ser Leu Asn Ser Ile Pro Ser	
	15 20 25	
30	GGG CTC ACA GAA GCT GTA AAA AGC CTT GAC CTG TCC AAC AAC AGG ATC	192
	Gly Leu Thr Glu Ala Val Lys Ser Leu Asp Leu Ser Asn Asn Arg Ile	
	30 35 40	
35	ACC TAC ATT AGC AAC AGT GAC CTA CAG AGG TGT GTG AAC CTC CAG GCT	240
	Thr Tyr Ile Ser Asn Ser Asp Leu Gln Arg Cys Val Asn Leu Gln Ala	
	45 50 55	
40	CTG GTG CTG ACA TCC AAT GGA ATT AAC ACA ATA GAG GAA GAT TCT TTT	288
	Leu Val Leu Thr Ser Asn Gly Ile Asn Thr Ile Glu Glu Asp Ser Phe	
	60 65 70	
45	TCT TCC CTG GGC AGT CTT GAA CAT TTA GAC TTA TCC TAT AAT TAC TTA	336
	Ser Ser Leu Gly Ser Leu Glu His Leu Asp Leu Ser Tyr Asn Tyr Leu	
	75 80 85 90	
50	TCT AAT TTA TCG TCT TCC TGG TTC AAG CCC CTT TCT TCT TTA ACA TTC	384
	Ser Asn Leu Ser Ser Ser Trp Phe Lys Pro Leu Ser Ser Leu Thr Phe	
	95 100 105	
55	TTA AAC TTA CTG GGA AAT CCT TAC AAA ACC CTA GGG GAA ACA TCT CTT	432
	Leu Asn Leu Leu Gly Asn Pro Tyr Lys Thr Leu Gly Glu Thr Ser Leu	
	110 115 120	
60	TTT TCT CAT CTC ACA AAA TTG CAA ATC CTG AGA GTG GGA AAT ATG GAC	480
	Phe Ser His Leu Thr Lys Leu Gln Ile Leu Arg Val Gly Asn Met Asp	
	125 130 135	

# ES 2 340 210 T3

5	ACC TTC ACT AAG ATT CAA AGA AAA GAT TTT GCT GGA CTT ACC TTC CTT Thr Phe Thr Lys Ile Gln Arg Lys Asp Phe Ala Gly Leu Thr Phe Leu 140 145 150	528
10	GAG GAA CTT GAG ATT GAT GCT TCA GAT CTA CAG AGC TAT GAG CCA AAA Glu Glu Leu Glu Ile Asp Ala Ser Asp Leu Gln Ser Tyr Glu Pro Lys 155 160 165 170	576
15	AGT TTG AAG TCA ATT CAG AAC GTA AGT CAT CTG ATC CTT CAT ATG AAG Ser Leu Lys Ser Ile Gln Asn Val Ser His Leu Ile Leu His Met Lys 175 180 185	624
20	CAG CAT ATT TTA CTG CTG GAG ATT TTT GTA GAT GTT ACA AGT TCC GTG Gln His Ile Leu Leu Glu Ile Phe Val Asp Val Thr Ser Ser Val 190 195 200	672
25	GAA TGT TTG GAA CTG CGA GAT ACT GAT TTG GAC ACT TTC CAT TTT TCA Glu Cys Ser Glu Leu Arg Asp Thr Asp Leu Asp Thr Phe His Phe Ser 205 210 215	720
30	GAA CTA TCC ACT GGT GAA ACA AAT TCA TTG ATT AAA AAG TTT ACA TTT Glu Leu Ser Thr Gly Glu Thr Asn Ser Leu Ile Lys Lys Phe Thr Phe 220 225 230	768
35	AGA AAT GTG AAA ATC ACC GAT GAA AGT TTG TTT CAG GTT ATG AAA CTT Arg Asn Val Lys Ile Ser Asp Glu Ser Leu Phe Gln Val Met Lys Leu 235 240 245 250	816
40	TTG AAT CAG ATT TCT GGA TTG TTA GAA TTA GAG TTT GAT GAC TGT ACC Leu Asn Gln Ile Ser Gly Leu Leu Glu Leu Phe Asp Asp Cys Thr 255 260 265	864
45	CTT AAT GGA GTT GGT AAT TTT AGA GCA TCT GAT AAT GAC AGA GTT ATA Leu Asn Gly Val Gly Asn Phe Arg Ala Ser Asp Asn Asp Arg Val Ile 270 275 280	912
50	GAT CCA GGT AAA GTG GAA ACG TTA ACA ATC CGG AGG CTG CAT ATT CCA Asp Pro Gly Lys Val Glu Thr Leu Thr Ile Arg Arg Leu His Ile Pro 285 290 295	960
55	AGG TTT TAC TTA TTT TAT GAT CTG AGC ACT TTA TAT TCA CTT ACA GAA Arg Phe Tyr Leu Phe Tyr Asp Leu Ser Thr Leu Tyr Ser Leu Thr Glu 300 305 310	1008
60	AGA GTT AAA AGA ATC ACA GTA GAA AAC AGT AAA GTT TTT CTG GTT CCT Arg Val Lys Arg Ile Thr Val Glu Asn Ser Lys Val Phe Leu Val Pro 315 320 325 330	1056
65	TGT TTA CTT TCA CAA CAT TTA AAA TCA TTA GAA TAC TTG GAT CTC AGT Cys Leu Leu Ser Gln His Leu Lys Ser Leu Glu Tyr Leu Asp Leu Ser 335 340 345	1104
70	GAA AAT TTG ATG GTT GAA GAA TAC TTG AAA AAT TCA GCC TGT GAG GAT Glu Asn Leu Met Val Glu Glu Tyr Leu Lys Asn Ser Ala Cys Glu Asp 350 355 360	1152
75	GCC TGG CCC TCT CTA CAA ACT TTA ATT TTA AGG CAA AAT CAT TTG GCA Ala Trp Pro Ser Leu Gln Thr Leu Ile Leu Arg Gln Asn His Leu Ala 365 370 375	1200
80	TCA TTG GAA AAA ACC GGA GAG ACT TTG CTC ACT CTG AAA AAC TTG ACT	1248



# ES 2 340 210 T3

	Ser	Leu	Glu	Lys	Thr	Gly	Glu	Thr	Leu	Leu	Thr	Leu	Lys	Asn	Leu	Thr	
	380						385					390					
5	AAC	ATT	GAT	ATC	AGT	AAG	AAT	AGT	TTT	CAT	TCT	ATG	CCT	GAA	ACT	TGT	1296
	Asn	Ile	Asp	Ile	Ser	Lys	Asn	Ser	Phe	His	Ser	Met	Pro	Glu	Thr	Cys	
	395					400					405				410		
10	CAG	TGG	CCA	GAA	AAG	ATG	AAA	TAT	TTG	AAC	TTA	TCC	AGC	ACA	CGA	ATA	1344
	Gln	Trp	Pro	Glu	Lys	Met	Lys	Tyr	Leu	Asn	Leu	Ser	Ser	Thr	Arg	Ile	
					415					420					425		
15	CAC	AGT	GTA	ACA	GGC	TGC	ATT	CCC	AAG	ACA	CTG	GAA	ATT	TTA	GAT	GTT	1392
	His	Ser	Val	Thr	Gly	Cys	Ile	Pro	Lys	Thr	Leu	Glu	Ile	Leu	Asp	Val	
				430					435					440			
20	AGC	AAC	AAC	AAT	CTC	AAT	TTA	TTT	TCT	TTG	AAT	TTG	CCG	CAA	CTC	AAA	1440
	Ser	Asn	Asn	Asn	Leu	Asn	Leu	Phe	Ser	Leu	Asn	Leu	Pro	Gln	Leu	Lys	
				445				450					455				
25	GAA	CTT	TAT	ATT	TCC	AGA	AAT	AAG	TTG	ATG	ACT	CTA	CCA	GAT	GCC	TCC	1488
	Glu	Leu	Tyr	Ile	Ser	Arg	Asn	Lys	Leu	Met	Thr	Leu	Pro	Asp	Ala	Ser	
							465					470					
30	CTC	TTA	CCC	ATG	TTA	CTA	GTA	TTG	AAA	ATC	AGT	AGG	AAT	GCA	ATA	ACT	1536
	Leu	Leu	Pro	Met	Leu	Val	Leu	Lys	Ile	Ser	Arg	Asn	Ala	Ile	Thr		
						480					485				490		
35	ACG	TTT	TCT	AAG	GAG	CAA	CTT	GAC	TCA	TTT	CAC	ACA	CTG	AAG	ACT	TTG	1584
	Thr	Phe	Ser	Lys	Gln	Leu	Asp	Ser	His	Thr	Leu	Lys	Thr	Leu			
				495				500					505				
40	GAA	GCT	GGT	GGC	AAT	AAC	TTC	ATT	TGC	TCC	TGT	GAA	TTC	CTC	TCC	TTC	1632
	Glu	Ala	Gly	Gly	Asn	Asn	Phe	Ile	Cys	Ser	Cys	Glu	Phe	Leu	Ser	Phe	
				510				515					520				
45	ACT	CAG	GAG	CAG	CAA	GCA	CTG	GCC	AAA	GTC	TTG	ATT	GAT	TGG	CCA	GCA	1680
	Thr	Gln	Glu	Gln	Gln	Ala	Leu	Lys	Val	Leu	Ile	Asp	Trp	Pro	Ala		
				525				530					535				
50	AAT	TAC	CTG	TGT	GAC	TCT	CCA	TCC	CAT	GTG	CGT	GGC	CAG	CAG	GTT	CAG	1728
	Asn	Tyr	Leu	Cys	Asp	Ser	Pro	Ser	His	Val	Arg	Gly	Gln	Gln	Val	Gln	
				540			545					550					
55	GAT	GTC	CGC	CTC	TCG	GTG	TCG	GAA	TGT	CAC	AGG	ACA	GCA	CTG	GTG	TCT	1776
	Asp	Val	Arg	Leu	Ser	Val	Ser	Glu	Cys	His	Arg	Thr	Ala	Leu	Val	Ser	
						560					565				570		
60	GGC	ATG	TGC	TGT	GCT	CTG	TTC	CTG	CTG	ATC	CTG	CTC	ACG	GGG	GTC	CTG	1824
	Gly	Met	Cys	Cys	Ala	Leu	Phe	Leu	Leu	Ile	Leu	Leu	Thr	Gly	Val	Leu	
					575					580					585		
65	TGC	CAC	CGT	TTC	CAT	GGC	CTG	TGG	TAT	ATG	AAA	ATG	ATG	TGG	GCC	TGG	1872
	Cys	His	Arg	Phe	His	Gly	Leu	Trp	Tyr	Met	Lys	Met	Met	Trp	Ala	Trp	
				590				595						600			
70	CTC	CAG	GCC	AAA	AGG	AAG	CCC	AGG	AAA	GCT	CCC	AGC	AGG	AAC	ATC	TGC	1920
	Leu	Gln	Ala	Lys	Arg	Lys	Pro	Arg	Lys	Ala	Pro	Ser	Arg	Asn	Ile	Cys	
				605				610					615				
75	TAT	GAT	GCA	TTT	GTT	TCT	TAC	AGT	GAG	CGG	GAT	GCC	TAC	TGG	GTG	GAG	1968
	Tyr	Asp	Ala	Phe	Val	Ser	Tyr	Ser	Glu	Arg	Asp	Ala	Tyr	Trp	Val	Glu	

# ES 2 340 210 T3

	620	625	630	
5	AAC CTT ATG GTC CAG GAG CTG GAG AAC TTC AAT CCC CCC TTC AAG TTG Asn Leu Met Val Gln Glu Leu Glu Asn Phe Asn Pro Pro Phe Lys Leu 635 640 645 650	2016		
10	TGT CTT CAT AAG CGG GAC TTC ATT CCT GGC AAG TGG ATC ATT GAC AAT Cys Leu His Lys Arg Asp Phe Ile Pro Gly Lys Trp Ile Ile Asp Asn 655 660 665	2064		
15	ATC ATT GAC TCC ATT GAA AAG AGC CAC AAA ACT GTC TTT GTG CTT TCT Ile Ile Asp Ser Ile Glu Lys Ser His Lys Thr Val Phe Val Leu Ser 670 675 680	2112		
20	GAA AAC TTT GTG AAG AGT GAG TGG TGC AAG TAT GAA CTG GAC TTC TCC Glu Asn Phe Val Lys Ser Glu Trp Cys Lys Tyr Glu Leu Asp Phe Ser 685 690 695	2160		
25	CAT TTC CGT CTT TTT GAA GAG AAC AAT GAT GCT GCC ATT CTC ATT CTT His Phe Arg Leu Phe Glu Glu Asn Asn Asp Ala Ala Ile Leu Ile Leu 700 705 710	2208		
30	CTG GAG CCC ATT GAG AAA AAA GCC ATT CCC CAG CGC TTC TGC AAG CTG Leu Glu Pro Ile Glu Lys Lys Ala Ile Pro Gln Arg Phe Cys Lys Leu 715 720 725 730	2256		
35	CGG AAG ATA ATG AAC ACC AAG ACC TAC CTG GAG TGG CCC ATG GAC GAG Arg Lys Ile Met Asn Thr Lys Thr Tyr Leu Glu Trp Pro Met Asp Glu 735 740 745	2304		
40	GCT CAG CGG GAA GGA TTT TGG GTA AAT CTG AGA GCT GCG ATA AAG TCC Ala Gln Arg Glu Gly Phe Trp Val Asn Leu Arg Ala Ala Ile Lys Ser 750 755 760	2352		
	<b>TAG</b>	2355		

## (2) INFORMACIÓN PARA EL SEQ ID NO.:4:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 784 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

### (ii) TIPO DE MOLÉCULA: proteína

### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:4:

50	Met Pro His Thr Leu Trp Met Val Trp Val Leu Gly Val Ile Ile Ser -22 -20 -15 -10
55	Leu Ser Lys Glu Glu Ser Ser Asn Gln Ala Ser Leu Ser Cys Asp Arg -5 1 5 10
60	Asn Gly Ile Cys Lys Gly Ser Ser Gly Ser Leu Asn Ser Ile Pro Ser 15 20 25
65	Gly Leu Thr Glu Ala Val Lys Ser Leu Asp Leu Ser Asn Asn Arg Ile 30 35 40
	Thr Tyr Ile Ser Asn Ser Asp Leu Gln Arg Cys Val Asn Leu Gln Ala

# ES 2 340 210 T3

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

# ES 2 340 210 T3

	Ser	Leu	Glu	Lys	Thr	Gly	Glu	Thr	Leu	Leu	Thr	Leu	Lys	Asn	Leu	Thr	
	380					385						390					
5	Asn	Ile	Asp	Ile	Ser	Lys	Asn	Ser	Phe	His	Ser	Met	Pro	Glu	Thr	Cys	
	395					400					405					410	
	Gln	Trp	Pro	Glu	Lys	Met	Lys	Tyr	Leu	Asn	Leu	Ser	Ser	Thr	Arg	Ile	
					415					420					425		
10	His	Ser	Val	Thr	Gly	Cys	Ile	Pro	Lys	Thr	Leu	Glu	Ile	Leu	Asp	Val	
				430					435					440			
	Ser	Asn	Asn	Asn	Leu	Asn	Leu	Phe	Ser	Leu	Asn	Leu	Pro	Gln	Leu	Lys	
				445				450					455				
15	Glu	Leu	Tyr	Ile	Ser	Arg	Asn	Lys	Leu	Met	Thr	Leu	Pro	Asp	Ala	Ser	
	460					465						470					
	Leu	Leu	Pro	Met	Leu	Leu	Val	Leu	Lys	Ile	Ser	Arg	Asn	Ala	Ile	Thr	
	475					480					485					490	
20	Thr	Phe	Ser	Lys	Glu	Gln	Leu	Asp	Ser	Phe	His	Thr	Leu	Lys	Thr	Leu	
				495						500					505		
	Glu	Ala	Gly	Gly	Asn	Asn	Phe	Ile	Cys	Ser	Cys	Glu	Phe	Leu	Ser	Phe	
				510					515					520			
25	Thr	Gln	Glu	Gln	Gln	Ala	Leu	Ala	Lys	Val	Leu	Ile	Asp	Trp	Pro	Ala	
			525			530					535						
	Asn	Tyr	Leu	Cys	Asp	Ser	Pro	Ser	His	Val	Arg	Gly	Gln	Gln	Val	Gln	
	540					545					550						
30	Asp	Val	Arg	Leu	Ser	Val	Ser	Glu	Cys	His	Arg	Thr	Ala	Leu	Val	Ser	
	555					560					565					570	
	Gly	Met	Cys	Cys	Ala	Leu	Phe	Leu	Leu	Ile	Leu	Leu	Thr	Gly	Val	Leu	
				575						580					585		
35	Cys	His	Arg	Phe	His	Gly	Leu	Trp	Tyr	Met	Lys	Met	Met	Trp	Ala	Trp	
				590					595					600			
40	Leu	Gln	Ala	Lys	Arg	Lys	Pro	Arg	Lys	Ala	Pro	Ser	Arg	Asn	Ile	Cys	
			605			610					615						
	Tyr	Asp	Ala	Phe	Val	Ser	Tyr	Ser	Glu	Arg	Asp	Ala	Tyr	Trp	Val	Glu	
	620					625					630						
45	Asn	Leu	Met	Val	Gln	Glu	Leu	Glu	Asn	Phe	Asn	Pro	Pro	Phe	Lys	Leu	
	635					640				645					650		
	Cys	Leu	His	Lys	Arg	Asp	Phe	Ile	Pro	Gly	Lys	Trp	Ile	Ile	Asp	Asn	
				655					660					665			
50	Ile	Ile	Asp	Ser	Ile	Glu	Lys	Ser	His	Lys	Thr	Val	Phe	Val	Leu	Ser	
			670			675							680				
	Glu	Asn	Phe	Val	Lys	Ser	Glu	Trp	Cys	Lys	Tyr	Glu	Leu	Asp	Phe	Ser	
			685			690				695							

## ES 2 340 210 T3

	His	Phe	Arg	Leu	Phe	Glu	Glu	Asn	Asn	Asp	Ala	Ala	Ile	Leu	Ile	Leu
	700						705					710				
5	Leu	Glu	Pro	Ile	Glu	Lys	Lys	Ala	Ile	Pro	Gln	Arg	Phe	Cys	Lys	Leu
	715					720					725					730
	Arg	Lys	Ile	Met	Asn	Thr	Lys	Thr	Tyr	Leu	Glu	Trp	Pro	Met	Asp	Glu
					735					740					745	
10	Ala	Gln	Arg	Glu	Gly	Phe	Trp	Val	Asn	Leu	Arg	Ala	Ala	Ile	Lys	Ser
					750				755						760	

### (2) INFORMACIÓN PARA EL SEQ ID NO:5:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 2715 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..2712

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: mat\_péptido
- (B) LOCALIZACIÓN: 64..2712

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:5:

	ATG	AGA	CAG	ACT	TTG	CCT	TGT	ATC	TAC	TTT	TGG	GGG	GGC	CTT	TTG	CCC	
40	Met	Arg	Gln	Thr	Leu	Pro	Cys	Ile	Tyr	Phe	Trp	Gly	Gly	Leu	Leu	Pro	48
	-21	-20					-15					-10					
	TTT	GGG	ATG	CTG	TGT	GCA	TCC	TCC	ACC	ACC	AAG	TGC	ACT	GTT	AGC	CAT	96
45	Phe	Gly	Met	Leu	Cys	Ala	Ser	Ser	Thr	Thr	Lys	Cys	Thr	Val	Ser	His	
	-5					1				5					10		
	GAA	GTT	GCT	GAC	TGC	AGC	CAC	CTG	AAG	TTG	ACT	CAG	GTA	CCC	GAT	GAT	144
	Glu	Val	Ala	Asp	Cys	Ser	His	Leu	Lys	Leu	Thr	Gln	Val	Pro	Asp	Asp	
				15					20					25			
50	CTA	CCC	ACA	AAC	ATA	ACA	GTG	TTG	AAC	CTT	ACC	CAT	AAT	CAA	CTC	AGA	192
	Leu	Pro	Thr	Asn	Ile	Thr	Val	Leu	Asn	Leu	Thr	His	Asn	Gln	Leu	Arg	
			30					35					40				
	AGA	TTA	CCA	GCC	GCC	AAC	TTC	ACA	AGG	TAT	AGC	CAG	CTA	ACT	AGC	TTG	240
55	Arg	Leu	Pro	Ala	Ala	Asn	Phe	Thr	Arg	Tyr	Ser	Gln	Leu	Thr	Ser	Leu	
	45						50					55					
	GAT	GTA	GGA	TTT	AAC	ACC	ATC	TCA	AAA	CTG	GAG	CCA	GAA	TTG	TGC	CAG	288
	Asp	Val	Gly	Phe	Asn	Thr	Ile	Ser	Lys	Leu	Glu	Pro	Glu	Leu	Cys	Gln	
60	60					65					70					75	
	AAA	CTT	CCC	ATG	TTA	AAA	GTT	TTG	AAC	CTC	CAG	CAC	AAT	GAG	CTA	TCT	336
	Lys	Leu	Pro	Met	Leu	Lys	Val	Leu	Asn	Leu	Gln	His	Asn	Glu	Leu	Ser	

# ES 2 340 210 T3

	80										85										90										
5	CAA	CTT	TCT	GAT	AAA	ACC	TTT	GCC	TTC	TGC	ACG	AAT	TTG	ACT	GAA	CTC	384														
	Gln	Leu	Ser	Asp	Lys	Thr	Phe	Ala	Phe	Cys	Thr	Asn	Leu	Thr	Glu	Leu															
				95					100					105																	
10	CAT	CTC	ATG	TCC	AAC	TCA	ATC	CAG	AAA	ATT	AAA	AAT	AAT	CCC	TTT	GTC	432														
	His	Leu	Met	Ser	Asn	Ser	Ile	Gln	Lys	Ile	Lys	Asn	Asn	Pro	Phe	Val															
			110					115					120																		
15	AAG	CAG	AAG	AAT	TTA	ATC	ACA	TTA	GAT	CTG	TCT	CAT	AAT	GGC	TTG	TCA	480														
	Lys	Gln	Lys	Asn	Leu	Ile	Thr	Leu	Asp	Leu	Ser	His	Asn	Gly	Leu	Ser															
			125				130					135																			
20	TCT	ACA	AAA	TTA	GGA	ACT	CAG	GTT	CAG	CTG	GAA	AAT	CTC	CAA	GAG	CTT	528														
	Ser	Thr	Lys	Leu	Gly	Thr	Gln	Val	Gln	Leu	Glu	Asn	Leu	Gln	Glu	Leu															
			140			145					150					155															
25	CTA	TTA	TCA	AAC	AAT	AAA	ATT	CAA	GCG	CTA	AAA	AGT	GAA	GAA	CTG	GAT	576														
	Leu	Leu	Ser	Asn	Asn	Lys	Ile	Gln	Ala	Leu	Lys	Ser	Glu	Glu	Leu	Asp															
				160						165					170																
30	ATC	TTT	GCC	AAT	TCA	TCT	TTA	AAA	AAA	TTA	GAG	TTG	TCA	TCG	AAT	CAA	624														
	Ile	Phe	Ala	Asn	Ser	Ser	Leu	Lys	Lys	Leu	Glu	Leu	Ser	Ser	Asn	Gln															
				175					180					185																	
35	ATT	AAA	GAG	TTT	TCT	CCA	GGG	TGT	TTT	CAC	GCA	ATT	GGA	AGA	TTA	TTT	672														
	Ile	Lys	Glu	Phe	Ser	Pro	Gly	Cys	Phe	His	Ala	Ile	Gly	Arg	Leu	Phe															
			190				195						200																		
40	GGC	CTC	TTT	CTG	AAC	AAT	GTC	CAG	CTG	GGT	CCC	AGC	CTT	ACA	GAG	AAG	720														
	Gly	Leu	Phe	Leu	Asn	Asn	Val	Gln	Leu	Gly	Pro	Ser	Leu	Thr	Glu	Lys															
			205				210				215																				
45	CTA	TGT	TTG	GAA	TTA	GCA	AAC	ACA	AGC	ATT	CGG	AAT	CTG	TCT	CTG	AGT	768														
	Leu	Cys	Leu	Glu	Leu	Ala	Asn	Thr	Ser	Ile	Arg	Asn	Leu	Ser	Leu	Ser															
			220			225					230					235															
50	AAC	AGC	CAG	CTG	TCC	ACC	ACC	AGC	AAT	ACA	ACT	TTC	TTG	GGA	CTA	AAG	816														
	Asn	Ser	Gln	Leu	Ser	Thr	Thr	Ser	Asn	Thr	Thr	Phe	Leu	Gly	Leu	Lys															
				240					245					250																	
55	TGG	ACA	AAT	CTC	ACT	ATG	CTC	GAT	CTT	TCC	TAC	AAC	AAC	TTA	AAT	GTG	864														
	Trp	Thr	Asn	Leu	Thr	Met	Leu	Asp	Leu	Ser	Tyr	Asn	Asn	Leu	Asn	Val															
				255					260					265																	
60	GTT	GGT	AAC	GAT	TCC	TTT	GCT	TGG	CTT	CCA	CAA	CTA	GAA	TAT	TTC	TTC	912														
	Val	Gly	Asn	Asp	Ser	Phe	Ala	Trp	Leu	Pro	Gln	Leu	Glu	Tyr	Phe	Phe															
			270				275						280																		
65	CTA	GAG	TAT	AAT	AAT	ATA	CAG	CAT	TTG	TTT	TCT	CAC	TCT	TTG	CAC	GGG	960														
	Leu	Glu	Tyr	Asn	Asn	Ile	Gln	His	Leu	Phe	Ser	His	Ser	Leu	His	Gly															
			285				290					295																			
70	CTT	TTC	AAT	GTG	AGG	TAC	CTG	AAT	TTG	AAA	CGG	TCT	TTT	ACT	AAA	CAA	1008														
	Leu	Phe	Asn	Val	Arg	Tyr	Leu	Asn	Leu	Lys	Arg	Ser	Phe	Thr	Lys	Gln															
			300			305					310					315															
75	AGT	ATT	TCC	CTT	GCC	TCA	CTC	CCC	AAG	ATT	GAT	GAT	TTT	TCT	TTT	CAG	1056														
	Ser	Ile	Ser	Leu	Ala	Ser	Leu	Pro	Lys	Ile	Asp	Asp	Phe	Ser	Phe	Gln															
				320					325						330																

# ES 2 340 210 T3

	TGG CTA AAA TGT TTG GAG CAC CTT AAC ATG GAA GAT AAT GAT ATT CCA	1104
	Trp Leu Lys Cys Leu Glu His Leu Asn Met Glu Asp Asn Asp Ile Pro	
	335 340 345	
5	GGC ATA AAA AGC AAT ATG TTC ACA GGA TTG ATA AAC CTG AAA TAC TTA	1152
	Gly Ile Lys Ser Asn Met Phe Thr Gly Leu Ile Asn Leu Lys Tyr Leu	
	350 355 360	
10	AGT CTA TCC AAC TCC TTT ACA AGT TTG CGA ACT TTG ACA AAT GAA ACA	1200
	Ser Leu Ser Asn Ser Phe Thr Ser Leu Arg Thr Leu Thr Asn Glu Thr	
	365 370 375	
15	TTT GTA TCA CTT GCT CAT TCT CCC TTA CAC ATA CTC AAC CTA ACC AAG	1248
	Phe Val Ser Leu Ala His Ser Pro Leu His Ile Leu Asn Leu Thr Lys	
	380 385 390 395	
	AAT AAA ATC TCA AAA ATA GAG AGT GAT GCT TTC TCT TGG TTG GGC CAC	1296
	Asn Lys Ile Ser Lys Ile Glu Ser Asp Ala Phe Ser Trp Leu Gly His	
	400 405 410	
20	CTA GAA GTA CTT GAC CTG GGC CTT AAT GAA ATT GGG CAA GAA CTC ACA	1344
	Leu Glu Val Leu Asp Leu Gly Leu Asn Glu Ile Gly Gln Glu Leu Thr	
	415 420 425	
25	GGC CAG GAA TGG AGA GGT CTA GAA AAT ATT TTC GAA ATC TAT CTT TCC	1392
	Gly Gln Glu Trp Arg Gly Leu Glu Asn Ile Phe Glu Ile Tyr Leu Ser	
	430 435 440	
	TAC AAC AAG TAC CTG CAG CTG ACT AGG AAC TCC TTT GCC TTG GTC CCA	1440
	Tyr Asn Lys Tyr Leu Gln Leu Thr Arg Asn Ser Phe Ala Leu Val Pro	
	445 450 455	
30	AGC CTT CAA CGA CTG ATG CTC CGA AGG GTG GCC CTT AAA AAT GTG GAT	1488
	Ser Leu Gln Arg Leu Met Leu Arg Arg Val Ala Leu Lys Asn Val Asp	
	460 465 470 475	
35	AGC TCT CCT TCA CCA TTC CAG CCT CTT CGT AAC TTG ACC ATT CTG GAT	1536
	Ser Ser Pro Ser Pro Phe Gln Pro Leu Arg Asn Leu Thr Ile Leu Asp	
	480 485 490	
	CTA AGC AAC AAC AAC ATA GCC AAC ATA AAT GAT GAC ATG TTG GAG GGT	1584
	Leu Ser Asn Asn Asn Ile Ala Asn Ile Asn Asp Asp Met Leu Glu Gly	
	495 500 505	
40	CTT GAG AAA CTA GAA ATT CTC GAT TTG CAG CAT AAC AAC TTA GCA CGG	1632
	Leu Glu Lys Leu Glu Ile Leu Asp Leu Gln His Asn Asn Leu Ala Arg	
	510 515 520	
45	CTC TGG AAA CAC GCA AAC CCT GGT GGT CCC ATT TAT TTC CTA AAG GGT	1680
	Leu Trp Lys His Ala Asn Pro Gly Gly Pro Ile Tyr Phe Leu Lys Gly	
	525 530 535	
50	CTG TCT CAC CTC CAC ATC CTT AAC TTG GAG TCC AAC GGC TTT GAC GAG	1728
	Leu Ser His Leu His Ile Leu Asn Leu Glu Ser Asn Gly Phe Asp Glu	
	540 545 550 555	
55	ATC CCA GTT GAG GTC TTC AAG GAT TTA TTT GAA CTA AAG ATC ATC GAT	1776
	Ile Pro Val Glu Val Phe Lys Asp Leu Phe Glu Leu Lys Ile Ile Asp	
	560 565 570	

# ES 2 340 210 T3

5	TTA GGA TTG AAT AAT TTA AAC ACA CTT CCA GCA TCT GTC TTT AAT AAT Leu Gly Leu Asn Asn Leu Asn Thr Leu Pro Ala Ser Val Phe Asn Asn 575 580 585	1824
10	CAG GTG TCT CTA AAG TCA TTG AAC CTT CAG AAG AAT CTC ATA ACA TCC Gln Val Ser Leu Lys Ser Leu Asn Leu Gln Lys Asn Leu Ile Thr Ser 590 595 600	1872
15	GTT GAG AAG AAG GTT TTC GGG CCA GCT TTC AGG AAC CTG ACT GAG TTA Val Glu Lys Lys Val Phe Gly Pro Ala Phe Arg Asn Leu Thr Glu Leu 605 610 615	1920
20	GAT ATG CGC TTT AAT CCC TTT GAT TGC ACG TGT GAA AGT ATT GCC TGG Asp Met Arg Phe Asn Pro Phe Asp Cys Thr Cys Glu Ser Ile Ala Trp 620 625 630 635	1968
25	TTT GTT AAT TGG ATT AAC GAG ACC CAT ACC AAC ATC CCT GAG CTG TCA Phe Val Asn Trp Ile Asn Glu Thr His Thr Asn Ile Pro Glu Leu Ser 640 645 650	2016
30	AGC CAC TAC CTT TGC AAC ACT CCA CCT CAC TAT CAT GGG TTC CCA GTG Ser His Tyr Leu Cys Asn Thr Pro Pro His Tyr His Gly Phe Pro Val 655 660 665	2064
35	AGA CTT TTT GAT ACA TCA TCT TGC AAA GAC AGT GCC CCC TTT GAA CTC Arg Leu Phe Asp Thr Ser Ser Cys Lys Asp Ser Ala Pro Phe Glu Leu 670 675 680	2112
40	TTT TTC ATG ATC AAT ACC AGT ATC CTG TTG ATT TTT ATC TTT ATT GTA Phe Phe Met Ile Asn Thr Ser Ile Leu Leu Ile Phe Ile Phe Ile Val 685 690 695	2160
45	CTT CTC ATC CAC TTT GAG GGC TGG AGG ATA TCT TTT TAT TGG AAT GTT Leu Leu Ile His Phe Glu Gly Trp Arg Ile Ser Phe Tyr Trp Asn Val 700 705 710 715	2208
50	TCA GTA CAT CGA GTT CTT GGT TTC AAA GAA ATA GAC AGA CAG ACA GAA Ser Val His Arg Val Leu Gly Phe Lys Glu Ile Asp Arg Gln Thr Glu 720 725 730	2256
55	CAG TTT GAA TAT GCA GCA TAT ATA ATT CAT GCC TAT AAA GAT AAG GAT Gln Phe Glu Tyr Ala Ala Tyr Ile Ile His Ala Tyr Lys Asp Lys Asp 735 740 745	2304
60	TGG GTC TGG GAA CAT TTC TCT TCA ATG GAA AAG GAA GAC CAA TCT CTC Trp Val Trp Glu His Phe Ser Ser Met Glu Lys Glu Asp Gln Ser Leu 750 755 760	2352
65	AAA TTT TGT CTG GAA GAA AGG GAC TTT GAG GCG GGT GTT TTT GAA CTA Lys Phe Cys Leu Glu Glu Arg Asp Phe Glu Ala Gly Val Phe Glu Leu 765 770 775	2400
70	GAA GCA ATT GTT AAC AGC ATC AAA AGA AGC AGA AAA ATT ATT TTT GTT Glu Ala Ile Val Asn Ser Ile Lys Arg Ser Arg Lys Ile Ile Phe Val 780 785 790 795	2448
75	ATA ACA CAC CAT CTA TTA AAA GAC CCA TTA TGC AAA AGA TTC AAG GTA Ile Thr His His Leu Leu Lys Asp Pro Leu Cys Lys Arg Phe Lys Val 800 805 810	2496
80	CAT CAT GCA GTT CAA CAA GCT ATT GAA CAA AAT CTG GAT TCC ATT ATA	2544



# ES 2 340 210 T3

	His	His	Ala	Val	Gln	Gln	Ala	Ile	Glu	Gln	Asn	Leu	Asp	Ser	Ile	Ile	
				815					820					825			
5	TTG	GTT	TTC	CTT	GAG	GAG	ATT	CCA	GAT	TAT	AAA	CTG	AAC	CAT	GCA	CTC	2592
	Leu	Val	Phe	Leu	Glu	Glu	Ile	Pro	Asp	Tyr	Lys	Leu	Asn	His	Ala	Leu	
				830				835					840				
10	TGT	TTG	CGA	AGA	GGA	ATG	TTT	AAA	TCT	CAC	TGC	ATC	TTG	AAC	TGG	CCA	2640
	Cys	Leu	Arg	Arg	Gly	Met	Phe	Lys	Ser	His	Cys	Ile	Leu	Asn	Trp	Pro	
				845			850					855					
15	GTT	CAG	AAA	GAA	CGG	ATA	GGT	GCC	TTT	CGT	CAT	AAA	TTG	CAA	GTA	GCA	2688
	Val	Gln	Lys	Glu	Arg	Ile	Gly	Ala	Phe	Arg	His	Lys	Leu	Gln	Val	Ala	
				860			865				870					875	
	CTT	GGA	TCC	AAA	AAC	TCT	GTA	CAT	TAA								2715
	Leu	Gly	Ser	Lys	Asn	Ser	Val	His									
							880										

## (2) INFORMACIÓN PARA EL SEQ ID NO:6:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 904 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

### (ii) TIPO DE MOLÉCULA: proteína

### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:6:

35	Met	Arg	Gln	Thr	Leu	Pro	Cys	Ile	Tyr	Phe	Trp	Gly	Gly	Leu	Leu	Pro	
	-21	-20					-15					-10					
	Phe	Gly	Met	Leu	Cys	Ala	Ser	Ser	Thr	Thr	Lys	Cys	Thr	Val	Ser	His	
	-5					1				5					10		
40	Glu	Val	Ala	Asp	Cys	Ser	His	Leu	Lys	Leu	Thr	Gln	Val	Pro	Asp	Asp	
				15					20					25			
	Leu	Pro	Thr	Asn	Ile	Thr	Val	Leu	Asn	Leu	Thr	His	Asn	Gln	Leu	Arg	
				30				35					40				
45	Arg	Leu	Pro	Ala	Ala	Asn	Phe	Thr	Arg	Tyr	Ser	Gln	Leu	Thr	Ser	Leu	
		45				50						55					
	Asp	Val	Gly	Phe	Asn	Thr	Ile	Ser	Lys	Leu	Glu	Pro	Glu	Leu	Cys	Gln	
	60				65						70					75	
50	Lys	Leu	Pro	Met	Leu	Lys	Val	Leu	Asn	Leu	Gln	His	Asn	Glu	Leu	Ser	
					80					85					90		
	Gln	Leu	Ser	Asp	Lys	Thr	Phe	Ala	Phe	Cys	Thr	Asn	Leu	Thr	Glu	Leu	
				95					100					105			
55	His	Leu	Met	Ser	Asn	Ser	Ile	Gln	Lys	Ile	Lys	Asn	Asn	Pro	Phe	Val	
					110				115				120				
	Lys	Gln	Lys	Asn	Leu	Ile	Thr	Leu	Asp	Leu	Ser	His	Asn	Gly	Leu	Ser	
60				125				130					135				

# ES 2 340 210 T3

	Ser	Thr	Lys	Leu	Gly	Thr	Gln	Val	Gln	Leu	Glu	Asn	Leu	Gln	Glu	Leu	140	145	150	155
5	Leu	Leu	Ser	Asn	Asn	Lys	Ile	Gln	Ala	Leu	Lys	Ser	Glu	Glu	Leu	Asp	160	165	170	
	Ile	Phe	Ala	Asn	Ser	Ser	Leu	Lys	Lys	Leu	Glu	Leu	Ser	Ser	Asn	Gln	175	180	185	
10	Ile	Lys	Glu	Phe	Ser	Pro	Gly	Cys	Phe	His	Ala	Ile	Gly	Arg	Leu	Phe	190	195	200	
	Gly	Leu	Phe	Leu	Asn	Asn	Val	Gln	Leu	Gly	Pro	Ser	Leu	Thr	Glu	Lys	205	210	215	
15	Leu	Cys	Leu	Glu	Leu	Ala	Asn	Thr	Ser	Ile	Arg	Asn	Leu	Ser	Leu	Ser	220	225	230	235
	Asn	Ser	Gln	Leu	Ser	Thr	Thr	Ser	Asn	Thr	Thr	Phe	Leu	Gly	Leu	Lys	240	245	250	
20	Trp	Thr	Asn	Leu	Thr	Met	Leu	Asp	Leu	Ser	Tyr	Asn	Asn	Leu	Asn	Val	255	260	265	
	Val	Gly	Asn	Asp	Ser	Phe	Ala	Trp	Leu	Pro	Gln	Leu	Glu	Tyr	Phe	Phe	270	275	280	
25	Leu	Glu	Tyr	Asn	Asn	Ile	Gln	His	Leu	Phe	Ser	His	Ser	Leu	His	Gly	285	290	295	
	Leu	Phe	Asn	Val	Arg	Tyr	Leu	Asn	Leu	Lys	Arg	Ser	Phe	Thr	Lys	Gln	300	305	310	315
30	Ser	Ile	Ser	Leu	Ala	Ser	Leu	Pro	Lys	Ile	Asp	Asp	Phe	Ser	Phe	Gln	320	325	330	
	Trp	Leu	Lys	Cys	Leu	Glu	His	Leu	Asn	Met	Glu	Asp	Asn	Asp	Ile	Pro	335	340	345	
35	Gly	Ile	Lys	Ser	Asn	Met	Phe	Thr	Gly	Leu	Ile	Asn	Leu	Lys	Tyr	Leu	350	355	360	
	Ser	Leu	Ser	Asn	Ser	Phe	Thr	Ser	Leu	Arg	Thr	Leu	Thr	Asn	Glu	Thr	365	370	375	
40	Phe	Val	Ser	Leu	Ala	His	Ser	Pro	Leu	His	Ile	Leu	Asn	Leu	Thr	Lys	380	385	390	395
	Asn	Lys	Ile	Ser	Lys	Ile	Glu	Ser	Asp	Ala	Phe	Ser	Trp	Leu	Gly	His	400	405	410	
45	Leu	Glu	Val	Leu	Asp	Leu	Gly	Leu	Asn	Glu	Ile	Gly	Gln	Glu	Leu	Thr	415	420	425	
	Gly	Gln	Glu	Trp	Arg	Gly	Leu	Glu	Asn	Ile	Phe	Glu	Ile	Tyr	Leu	Ser	430	435	440	
50	Tyr	Asn	Lys	Tyr	Leu	Gln	Leu	Thr	Arg	Asn	Ser	Phe	Ala	Leu	Val	Pro	445	450	455	
55	Ser	Leu	Gln	Arg	Leu	Met	Leu	Arg	Arg	Val	Ala	Leu	Lys	Asn	Val	Asp				

# ES 2 340 210 T3

	460		465		470		475
5	Ser Ser Pro Ser	Pro Phe Gln Pro Leu	Arg Asn Leu Thr Ile Leu Asp				
		480	485			490	
	Leu Ser Asn Asn Asn Ile Ala Asn Ile	Asn Asp Asp Met Leu Glu Gly					
		495	500			505	
10	Leu Glu Lys Leu Glu Ile Leu Asp Leu Gln His Asn Asn Leu Ala Arg						
		510	515			520	
	Leu Trp Lys His Ala Asn Pro Gly Gly Pro Ile Tyr Phe Leu Lys Gly						
		525	530			535	
15	Leu Ser His Leu His Ile Leu Asn Leu Glu Ser Asn Gly Phe Asp Glu						
		540	545			550	555
	Ile Pro Val Glu Val Phe Lys Asp Leu Phe Glu Leu Lys Ile Ile Asp						
		560	565			570	
20	Leu Gly Leu Asn Asn Leu Asn Thr Leu Pro Ala Ser Val Phe Asn Asn						
		575	580			585	
	Gln Val Ser Leu Lys Ser Leu Asn Leu Gln Lys Asn Leu Ile Thr Ser						
		590	595			600	
25	Val Glu Lys Lys Val Phe Gly Pro Ala Phe Arg Asn Leu Thr Glu Leu						
		605	610			615	
	Asp Met Arg Phe Asn Pro Phe Asp Cys Thr Cys Glu Ser Ile Ala Trp						
		620	625			630	635
30	Phe Val Asn Trp Ile Asn Glu Thr His Thr Asn Ile Pro Glu Leu Ser						
		640	645			650	
	Ser His Tyr Leu Cys Asn Thr Pro Pro His Tyr His Gly Phe Pro Val						
		655	660			665	
35	Arg Leu Phe Asp Thr Ser Ser Cys Lys Asp Ser Ala Pro Phe Glu Leu						
		670	675			680	
	Phe Phe Met Ile Asn Thr Ser Ile Leu Leu Ile Phe Ile Phe Ile Val						
		685	690			695	
40	Leu Leu Ile His Phe Glu Gly Trp Arg Ile Ser Phe Tyr Trp Asn Val						
		700	705			710	715
	Ser Val His Arg Val Leu Gly Phe Lys Glu Ile Asp Arg Gln Thr Glu						
		720	725			730	
45	Gln Phe Glu Tyr Ala Ala Tyr Ile Ile His Ala Tyr Lys Asp Lys Asp						
		735	740			745	
	Trp Val Trp Glu His Phe Ser Ser Met Glu Lys Glu Asp Gln Ser Leu						
		750	755			760	
50	Lys Phe Cys Leu Glu Glu Arg Asp Phe Glu Ala Gly Val Phe Glu Leu						
		765	770			775	
55	Glu Ala Ile Val Asn Ser Ile Lys Arg Ser Arg Lys Ile Ile Phe Val						
		780	785			790	795

# ES 2 340 210 T3

Ile Thr His His Leu Leu Lys Asp Pro Leu Cys Lys Arg Phe Lys Val  
 800 805 810  
 5 His His Ala Val Gln Gln Ala Ile Glu Gln Asn Leu Asp Ser Ile Ile  
 815 820 825  
 Leu Val Phe Leu Glu Glu Ile Pro Asp Tyr Lys Leu Asn His Ala Leu  
 830 835 840  
 10 Cys Leu Arg Arg Gly Met Phe Lys Ser His Cys Ile Leu Asn Trp Pro  
 845 850 855  
 Val Gln Lys Glu Arg Ile Gly Ala Phe Arg His Lys Leu Gln Val Ala  
 860 865 870 875  
 15 Leu Gly Ser Lys Asn Ser Val His  
 880

## 20 (2) INFORMACIÓN PARA EL SEQ ID NO:7:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- 25 (A) LONGITUD: 2400 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

### 30 (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- 35 (B) LOCALIZACIÓN: 1..2397

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:7:

ATG GAG CTG AAT TTC TAC AAA ATC CCC GAC AAC CTC CCC TTC TCA ACC 48  
 Met Glu Leu Asn Phe Tyr Lys Ile Pro Asp Asn Leu Pro Phe Ser Thr  
 1 5 10 15  
 AAG AAC CTG GAC CTG AGC TTT AAT CCC CTG AGG CAT TTA GGC AGC TAT 96  
 Lys Asn Leu Asp Leu Ser Phe Asn Pro Leu Arg His Leu Gly Ser Tyr  
 20 25 30  
 AGC TTC TTC AGT TTC CCA GAA CTG CAG GTG CTG GAT TTA TCC AGG TGT 144  
 Ser Phe Phe Ser Phe Pro Glu Leu Gln Val Leu Asp Leu Ser Arg Cys  
 35 40 45  
 GAA ATC CAG ACA ATT GAA GAT GGG GCA TAT CAG AGC CTA AGC CAC CTC 192  
 Glu Ile Gln Thr Ile Glu Asp Gly Ala Tyr Gln Ser Leu Ser His Leu  
 50 55 60  
 TCT ACC TTA ATA TTG ACA GGA AAC CCC ATC CAG AGT TTA GCC CTG GGA 240  
 Ser Thr Leu Ile Leu Thr Gly Asn Pro Ile Gln Ser Leu Ala Leu Gly  
 65 70 75 80  
 GCC TTT TCT GGA CTA TCA AGT TTA CAG AAG CTG GTG GCT GTG GAG ACA 288  
 Ala Phe Ser Gly Leu Ser Ser Leu Gln Lys Leu Val Ala Val Glu Thr  
 85 90 95

# ES 2 340 210 T3

5	AAT CTA GCA TCT CTA GAG AAC TTC CCC ATT GGA CAT CTC AAA ACT TTG Asn Leu Ala Ser Leu Glu Asn Phe Pro Ile Gly His Leu Lys Thr Leu 100 105 110	336
10	AAA GAA CTT AAT GTG GCT CAC AAT CTT ATC CAA TCT TTC AAA TTA CCT Lys Glu Leu Asn Val Ala His Asn Leu Ile Gln Ser Phe Lys Leu Pro 115 120 125	384
15	GAG TAT TTT TCT AAT CTG ACC AAT CTA GAG CAC TTG GAC CTT TCC AGC Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu His Leu Asp Leu Ser Ser 130 135 140	432
20	AAC AAG ATT CAA AGT ATT TAT TGC ACA GAC TTG CGG GTT CTA CAT CAA Asn Lys Ile Gln Ser Ile Tyr Cys Thr Asp Leu Arg Val Leu His Gln 145 150 155 160	480
25	ATG CCC CTA CTC AAT CTC TCT TTA GAC CTG TCC CTG AAC CCT ATG AAC Met Pro Leu Leu Asn Leu Ser Leu Asp Leu Ser Leu Asn Pro Met Asn 165 170 175	528
30	TTT ATC CAA CCA GGT GCA TTT AAA GAA ATT AGG CTT CAT AAG CTG ACT Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile Arg Leu His Lys Leu Thr 180 185 190	576
35	TTA AGA AAT AAT TTT GAT AGT TTA AAT GTA ATG AAA ACT TGT ATT CAA Leu Arg Asn Asn Phe Asp Ser Leu Asn Val Met Lys Thr Cys Ile Gln 195 200 205	624
40	GGT CTG GCT GGT TTA GAA GTC CAT CGT TTG GTT CTG GGA GAA TTT AGA Gly Leu Ala Gly Leu Glu Val His Arg Leu Val Leu Gly Glu Phe Arg 210 215 220	672
45	AAT GAA GGA AAC TTG GAA AAG TTT GAC AAA TCT GCT CTA GAG GGC CTG Asn Glu Gly Asn Leu Glu Lys Phe Asp Lys Ser Ala Leu Glu Gly Leu 225 230 235 240	720
50	TGC AAT TTG ACC ATT GAA GAA TTC CGA TTA GCA TAC TTA GAC TAC TAC Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Ala Tyr Leu Asp Tyr Tyr 245 250 255	768
55	CTC GAT GAT ATT ATT GAC TTA TTT AAT TGT TTG ACA AAT GTT TCT TCA Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Thr Asn Val Ser Ser 260 265 270	816
60	TTT TCC CTG GTG AGT GTG ACT ATT GAA AGG GTA AAA GAC TTT TCT TAT Phe Ser Leu Val Ser Val Thr Ile Glu Arg Val Lys Asp Phe Ser Tyr 275 280 285	864
65	AAT TTC GGA TGG CAA CAT TTA GAA TTA GTT AAC TGT AAA TTT GGA CAG Asn Phe Gly Trp Gln His Leu Glu Leu Val Asn Cys Lys Phe Gly Gln 290 295 300	912
70	TTT CCC ACA TTG AAA CTC AAA TCT CTC AAA AGG CTT ACT TTC ACT TCC Phe Pro Thr Leu Lys Leu Lys Ser Leu Lys Arg Leu Thr Phe Thr Ser 305 310 315 320	960
75	AAC AAA GGT GGG AAT GCT TTT TCA GAA GTT GAT CTA CCA AGC CTT GAG Asn Lys Gly Gly Asn Ala Phe Ser Glu Val Asp Leu Pro Ser Leu Glu 325 330 335	1008
80	TTT CTA GAT CTC AGT AGA AAT GGC TTG AGT TTC AAA GGT TGC TGT TCT	1056

# ES 2 340 210 T3

	Phe	Leu	Asp	Leu	Ser	Arg	Asn	Gly	Leu	Ser	Phe	Lys	Gly	Cys	Cys	Ser	
				340					345					350			
5	CAA	AGT	GAT	TTT	GGG	ACA	ACC	AGC	CTA	AAG	TAT	TTA	GAT	CTG	AGC	TTC	1104
	Gln	Ser	Asp	Phe	Gly	Thr	Thr	Ser	Leu	Lys	Tyr	Leu	Asp	Leu	Ser	Phe	
			355					360					365				
10	AAT	GGT	GTT	ATT	ACC	ATG	AGT	TCA	AAC	TTC	TTG	GGC	TTA	GAA	CAA	CTA	1152
	Asn	Gly	Val	Ile	Thr	Met	Ser	Ser	Asn	Phe	Leu	Gly	Leu	Glu	Gln	Leu	
			370				375					380					
15	GAA	CAT	CTG	GAT	TTC	CAG	CAT	TCC	AAT	TTG	AAA	CAA	ATG	AGT	GAG	TTT	1200
	Glu	His	Leu	Asp	Phe	Gln	His	Ser	Asn	Leu	Lys	Gln	Met	Ser	Glu	Phe	
			385			390					395					400	
20	TCA	GTA	TTC	CTA	TCA	CTC	AGA	AAC	CTC	ATT	TAC	CTT	GAC	ATT	TCT	CAT	1248
	Ser	Val	Phe	Leu	Ser	Leu	Arg	Asn	Leu	Ile	Tyr	Leu	Asp	Ile	Ser	His	
					405					410					415		
25	ACT	CAC	ACC	AGA	GTT	GCT	TTC	AAT	GGC	ATC	TTC	AAT	GGC	TTG	TCC	AGT	1296
	Thr	His	Thr	Arg	Val	Ala	Phe	Asn	Gly	Ile	Phe	Asn	Gly	Leu	Ser	Ser	
				420					425					430			
30	CTC	GAA	GTC	TTG	AAA	ATG	GCT	GGC	AAT	TCT	TTC	CAG	GAA	AAC	TTC	CTT	1344
	Leu	Glu	Val	Leu	Lys	Met	Ala	Gly	Asn	Ser	Phe	Gln	Glu	Asn	Phe	Leu	
			435					440					445				
35	CCA	GAT	ATC	TTC	ACA	GAG	CTG	AGA	AAC	TTG	ACC	TTC	CTG	GAC	CTC	TCT	1392
	Pro	Asp	Ile	Phe	Thr	Glu	Leu	Arg	Asn	Leu	Thr	Phe	Leu	Asp	Leu	Ser	
			450				455						460				
40	CAG	TGT	CAA	CTG	GAG	CAG	TTG	TCT	CCA	ACA	GCA	TTT	AAC	TCA	CTC	TCC	1440
	Gln	Cys	Gln	Leu	Glu	Gln	Leu	Ser	Pro	Thr	Ala	Phe	Asn	Ser	Leu	Ser	
						470					475					480	
45	AGT	CTT	CAG	GTA	CTA	AAT	ATG	AGC	CAC	AAC	AAC	TTC	TTT	TCA	TTG	GAT	1488
	Ser	Leu	Gln	Val	Leu	Asn	Met	Ser	His	Asn	Asn	Phe	Phe	Ser	Leu	Asp	
					485					490					495		
50	ACG	TTT	CCT	TAT	AAG	TGT	CTG	AAC	TCC	CTC	CAG	GTT	CTT	GAT	TAC	AGT	1536
	Thr	Phe	Pro	Tyr	Lys	Cys	Leu	Asn	Ser	Leu	Gln	Val	Leu	Asp	Tyr	Ser	
				500					505					510			
55	CTC	AAT	CAC	ATA	ATG	ACT	TCC	AAA	AAA	CAG	GAA	CTA	CAG	CAT	TTT	CCA	1584
	Leu	Asn	His	Ile	Met	Thr	Ser	Lys	Lys	Gln	Glu	Leu	Gln	His	Phe	Pro	
			515					520					525				
60	AGT	AGT	CTA	GCT	TTC	TTA	AAT	CTT	ACT	CAG	AAT	GAC	TTT	GCT	TGT	ACT	1632
	Ser	Ser	Leu	Ala	Phe	Leu	Asn	Leu	Thr	Gln	Asn	Asp	Phe	Ala	Cys	Thr	
							535					540					
65	TGT	GAA	CAC	CAG	AGT	TTC	CTG	CAA	TGG	ATC	AAG	GAC	CAG	AGG	CAG	CTC	1680
	Cys	Glu	His	Gln	Ser	Phe	Leu	Gln	Trp	Ile	Lys	Asp	Gln	Arg	Gln	Leu	
						550					555					560	
70	TTG	GTG	GAA	GTT	GAA	CGA	ATG	GAA	TGT	GCA	ACA	CCT	TCA	GAT	AAG	CAG	1728
	Leu	Val	Glu	Val	Glu	Arg	Met	Glu	Cys	Ala	Thr	Pro	Ser	Asp	Lys	Gln	
					565					570					575		
75	GGC	ATG	CCT	GTG	CTG	AGT	TTG	AAT	ATC	ACC	TGT	CAG	ATG	AAT	AAG	ACC	1776
	Gly	Met	Pro	Val	Leu	Ser	Leu	Asn	Ile	Thr	Cys	Gln	Met	Asn	Lys	Thr	

# ES 2 340 210 T3

	580	585	590	
5	ATC ATT GGT GTG TCG GTC CTC AGT GTG CTT GTA GTA TCT GTT GTA GCA Ile Ile Gly Val Ser Val Leu Ser Val Leu Val Val Ser Val Val Ala 595 600 605			1824
10	GTT CTG GTC TAT AAG TTC TAT TTT CAC CTG ATG CTT CTT GCT GGC TGC Val Leu Val Tyr Lys Phe Tyr Phe His Leu Met Leu Leu Ala Gly Cys 610 615 620			1872
15	ATA AAG TAT GGT AGA GGT GAA AAC ATC TAT GAT GCC TTT GTT ATC TAC Ile Lys Tyr Gly Arg Phe Glu Asn Ile Tyr Asp Ala Phe Val Ile Tyr 625 630 635 640			1920
20	TCA AGC CAG GAT GAG GAC TGG GTA AGG AAT GAG CTA GTA AAG AAT TTA Ser Ser Gln Asp Glu Asp Trp Val Arg Asn Glu Leu Val Lys Asn Leu 645 650 655			1968
25	GAA GAA GGG GTG CCT CCA TTT CAG CTC TGC CTT CAC TAC AGA GAC TTT Glu Glu Gly Val Pro Pro Phe Gln Leu Cys Leu His Tyr Arg Asp Phe 660 665 670			2016
30	ATT CCC GGT GTG GCC ATT GCT GCC AAC ATC ATC CAT GAA GGT TTC CAT Ile Pro Gly Val Ala Ile Ala Ala Asn Ile Ile His Glu Gly Phe His 675 680 685			2064
35	AAA AGC CGA AAG GTG ATT GTT GTG GTG TCC CAG CAC TTC ATC CAG AGC Lys Ser Arg Lys Val Ile Val Val Val Ser Gln His Phe Ile Gln Ser 690 695 700			2112
40	CGC TGG TGT ATC TTT GAA TAT GAG ATT GCT CAG ACC TGG CAG TTT CTG Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala Gln Thr Trp Gln Phe Leu 705 710 715 720			2160
45	AGC AGT CGT GCT GGT ATC ATC TTC ATT GTC CTG CAG AAG GTG CAG AAG Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys Val Glu Lys 725 730 735			2208
50	ACC CTG CTC AGG CAG CAG GTG GAG CTG TAC CGC CTT CTC AGC AGG AAC Thr Leu Leu Arg Gln Gln Val Glu Leu Tyr Arg Leu Leu Ser Arg Asn 740 745 750			2256
55	ACT TAC CTG GAG TGG GAG GAC AGT GTC CTG GGG CGG CAC ATC TTC TGG Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu Gly Arg His Ile Phe Trp 755 760 765			2304
60	AGA CGA CTC AGA AAA GCC CTG CTG GAT GGT AAA TCA TGG AAT CCA GAA Arg Arg Leu Arg Lys Ala Leu Leu Asp Gly Lys Ser Trp Asn Pro Glu 770 775 780			2352
65	GGA ACA GTG GGT ACA GGA TGC AAT TGG CAG GAA GCA ACA TCT ATC Gly Thr Val Gly Thr Gly Cys Asn Trp Gln Glu Ala Thr Ser Ile 785 790 795			2397
70	TGA			2400

## (2) INFORMACIÓN PARA EL SEQ ID NO:8:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 799 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:8:

# ES 2 340 210 T3

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



# ES 2 340 210 T3

	Asn	Phe	Gly	Trp	Gln	His	Leu	Glu	Leu	Val	Asn	Cys	Lys	Phe	Gly	Gln	
	290						295					300					
5	Phe	Pro	Thr	Leu	Lys	Leu	Lys	Ser	Leu	Lys	Arg	Leu	Thr	Phe	Thr	Ser	
	305					310					315					320	
	Asn	Lys	Gly	Gly	Asn	Ala	Phe	Ser	Glu	Val	Asp	Leu	Pro	Ser	Leu	Glu	
					325					330					335		
10	Phe	Leu	Asp	Leu	Ser	Arg	Asn	Gly	Leu	Ser	Phe	Lys	Gly	Cys	Cys	Ser	
				340					345					350			
	Gln	Ser	Asp	Phe	Gly	Thr	Thr	Ser	Leu	Lys	Tyr	Leu	Asp	Leu	Ser	Phe	
			355					360					365				
15	Asn	Gly	Val	Ile	Thr	Met	Ser	Ser	Asn	Phe	Leu	Gly	Leu	Glu	Gln	Leu	
	370						375					380					
	Glu	His	Leu	Asp	Phe	Gln	His	Ser	Asn	Leu	Lys	Gln	Met	Ser	Glu	Phe	
	385					390					395					400	
20	Ser	Val	Phe	Leu	Ser	Leu	Arg	Asn	Leu	Ile	Tyr	Leu	Asp	Ile	Ser	His	
					405					410					415		
	Thr	His	Thr	Arg	Val	Ala	Phe	Asn	Gly	Ile	Phe	Asn	Gly	Leu	Ser	Ser	
				420					425					430			
25	Leu	Glu	Val	Leu	Lys	Met	Ala	Gly	Asn	Ser	Phe	Gln	Glu	Asn	Phe	Leu	
			435					440					445				
	Pro	Asp	Ile	Phe	Thr	Glu	Leu	Arg	Asn	Leu	Thr	Phe	Leu	Asp	Leu	Ser	
		450					455					460					
30	Gln	Cys	Gln	Leu	Glu	Gln	Leu	Ser	Pro	Thr	Ala	Phe	Asn	Ser	Leu	Ser	
	465					470					475					480	
	Ser	Leu	Gln	Val	Leu	Asn	Met	Ser	His	Asn	Asn	Phe	Phe	Ser	Leu	Asp	
					485					490					495		
35	Thr	Phe	Pro	Tyr	Lys	Cys	Leu	Asn	Ser	Leu	Gln	Val	Leu	Asp	Tyr	Ser	
				500					505					510			
	Leu	Asn	His	Ile	Met	Thr	Ser	Lys	Lys	Gln	Glu	Leu	Gln	His	Phe	Pro	
			515					520					525				
	Ser	Ser	Leu	Ala	Phe	Leu	Asn	Leu	Thr	Gln	Asn	Asp	Phe	Ala	Cys	Thr	
		530					535					540					
45	Cys	Glu	His	Gln	Ser	Phe	Leu	Gln	Trp	Ile	Lys	Asp	Gln	Arg	Gln	Leu	
	545					550					555					560	
	Leu	Val	Glu	Val	Glu	Arg	Met	Glu	Cys	Ala	Thr	Pro	Ser	Asp	Lys	Gln	
					565					570					575		
50	Gly	Met	Pro	Val	Leu	Ser	Leu	Asn	Ile	Thr	Cys	Gln	Met	Asn	Lys	Thr	
				580					585					590			
	Ile	Ile	Gly	Val	Ser	Val	Leu	Ser	Val	Leu	Val	Val	Ser	Val	Val	Ala	
			595				600						605				
55	Val	Leu	Val	Tyr	Lys	Phe	Tyr	Phe	His	Leu	Met	Leu	Leu	Ala	Gly	Cys	
60																	
65																	

# ES 2 340 210 T3

	610	615	620
5	Ile Lys Tyr Gly Arg Glu Asn Ile Tyr Asp Ala Phe Val Ile Tyr 625 630 635 640		
	Ser Ser Gln Asp Glu Asp Trp Val Arg Asn Glu Leu Val Lys Asn Leu 645 650 655		
10	Glu Glu Gly Val Pro Pro Phe Gln Leu Cys Leu His Tyr Arg Asp Phe 660 665 670		
	Ile Pro Gly Val Ala Ile Ala Ala Asn Ile Ile His Glu Gly Phe His 675 680 685		
15	Lys Ser Arg Lys Val Ile Val Val Val Ser Gln His Phe Ile Gln Ser 690 695 700		
	Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala Gln Thr Trp Gln Phe Leu 705 710 715 720		
20	Ser Ser Arg Ala Gly Ile Ile Phe Ile Val Leu Gln Lys Val Glu Lys 725 730 735		
	Thr Leu Leu Arg Gln Gln Val Glu Leu Tyr Arg Leu Leu Ser Arg Asn 740 745 750		
25	Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu Gly Arg His Ile Phe Trp 755 760 765		
	Arg Arg Leu Arg Lys Ala Leu Leu Asp Gly Lys Ser Trp Asn Pro Glu 770 775 780		
30	Gly Thr Val Gly Thr Gly Cys Asn Trp Gln Glu Ala Thr Ser Ile 785 790 795		

35

## (2) INFORMACIÓN PARA EL SEQ ID NO:9:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

40

- (A) LONGITUD: 1275 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

45

### (ii) TIPO DE MOLÉCULA: ADNc

### (ix) RASGO:

50

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..1095

### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:9:

55

TGT TGG GAT GTT TTT GAG GGA CTT TCT CAT CTT CAA GTT CTG TAT TTG	48
Cys Trp Asp Val Phe Glu Gly Leu Ser His Leu Gln Val Leu Tyr Leu	
1 5 10 15	
AAT CAT AAC TAT CTT AAT TCC CTT CCA CCA GGA GTA TTT AGC CAT CTG	96
Asn His Asn Tyr Leu Asn Ser Leu Pro Pro Gly Val Phe Ser His Leu	
20 25 30	

65

# ES 2 340 210 T3

	ACT GCA TTA AGG GGA CTA AGC CTC AAC TCC AAC AGG CTG ACA GTT CTT	144
	Thr Ala Leu Arg Gly Leu Ser Leu Asn Ser Asn Arg Leu Thr Val Leu	
	35 40 45	
5	TCT CAC AAT GAT TTA CCT GCT AAT TTA GAG ATC CTG GAC ATA TCC AGG	192
	Ser His Asn Asp Leu Pro Ala Asn Leu Glu Ile Leu Asp Ile Ser Arg	
	50 55 60	
10	AAC CAG CTC CTA GCT CCT AAT CCT GAT GTA TTT GTA TCA CTT AGT GTC	240
	Asn Gln Leu Leu Ala Pro Asn Pro Asp Val Phe Val Ser Leu Ser Val	
	65 70 75 80	
15	TTG GAT ATA ACT CAT AAC AAG TTC ATT TGT GAA TGT GAA CTT AGC ACT	288
	Leu Asp Ile Thr His Asn Lys Phe Ile Cys Glu Cys Glu Leu Ser Thr	
	85 90 95	
	TTT ATC AAT TGG CTT AAT CAC ACC AAT GTC ACT ATA GCT GGG CCT CCT	336
	Phe Ile Asn Trp Leu Asn His Thr Asn Val Thr Ile Ala Gly Pro Pro	
	100 105 110	
20	GCA GAC ATA TAT TGT GTG TAC CCT GAC TCG TTC TCT GGG GTT TCC CTC	384
	Ala Asp Ile Tyr Cys Val Tyr Pro Asp Ser Phe Ser Gly Val Ser Leu	
	115 120 125	
25	TTC TCT CTT TCC ACG GAA GGT TGT GAT GAA GAG GAA GTC TTA AAG TCC	432
	Phe Ser Leu Ser Thr Glu Gly Cys Asp Glu Glu Glu Val Leu Lys Ser	
	130 135 140	
	CTA AAG TTC TCC CTT TTC ATT GTA TGC ACT GTC ACT CTG ACT CTG TTC	480
	Leu Lys Phe Ser Leu Phe Ile Val Cys Thr Val Thr Leu Thr Leu Phe	
	145 150 155 160	
30	CTC ATG ACC ATC CTC ACA GTC ACA AAG TTC CGG GGC TTC TGT TTT ATC	528
	Leu Met Thr Ile Leu Thr Val Thr Lys Phe Arg Gly Phe Cys Phe Ile	
	165 170 175	
35	TGT TAT AAG ACA GCC CAG AGA CTG GTG TTC AAG GAC CAT CCC CAG GGC	576
	Cys Tyr Lys Thr Ala Gln Arg Leu Val Phe Lys Asp His Pro Gln Gly	
	180 185 190	
	ACA GAA CCT GAT ATG TAC AAA TAT GAT GCC TAT TTG TGC TTC AGC AGC	624
	Thr Glu Pro Asp Met Tyr Lys Tyr Asp Ala Tyr Leu Cys Phe Ser Ser	
	195 200 205	
40	AAA GAC TTC ACA TGG GTG CAG AAT GCT TTG CTC AAA CAC CTG GAC ACT	672
	Lys Asp Phe Thr Trp Val Gln Asn Ala Leu Leu Lys His Leu Asp Thr	
	210 215 220	
45	CAA TAC AGT GAC CAA AAC AGA TTC AAC CTG TGC TTT GAA GAA AGA GAC	720
	Gln Tyr Ser Asp Gln Asn Arg Phe Asn Leu Cys Phe Glu Glu Arg Asp	
	225 230 235 240	
50	TTT GTC CCA GGA GAA AAC CGC ATT GCC AAT ATC CAG GAT GCC ATC TGG	768
	Phe Val Pro Gly Glu Asn Arg Ile Ala Asn Ile Gln Asp Ala Ile Trp	
	245 250 255	
	AAC AGT AGA AAG ATC GTT TGT CTT GTG AGC AGA CAC TTC CTT AGA GAT	816
	Asn Ser Arg Lys Ile Val Cys Leu Val Ser Arg His Phe Leu Arg Asp	
	260 265 270	
55	GGC TGG TGC CTT GAA GCC TTC AGT TAT GCC CAG GGC AGG TGC TTA TCT	864

# ES 2 340 210 T3

	Gly	Trp	Cys	Leu	Glu	Ala	Phe	Ser	Tyr	Ala	Gln	Gly	Arg	Cys	Leu	Ser	
			275					280					285				
5	GAC	CTT	AAC	AGT	GCT	CTC	ATC	ATG	GTG	GTG	GTT	GGG	TCC	TTG	TCC	CAG	912
	Asp	Leu	Asn	Ser	Ala	Leu	Ile	Met	Val	Val	Val	Gly	Ser	Leu	Ser	Gln	
		290					295					300					
	TAC	CAG	TTG	ATG	AAA	CAT	CAA	TCC	ATC	AGA	GGC	TTT	GTA	CAG	AAA	CAG	960
10	Tyr	Gln	Leu	Met	Lys	His	Gln	Ser	Ile	Arg	Gly	Phe	Val	Gln	Lys	Gln	
	305					310					315					320	
	CAG	TAT	TTG	AGG	TGG	CCT	GAG	GAT	CTC	CAG	GAT	GTT	GGC	TGG	TTT	CTT	1008
	Gln	Tyr	Leu	Arg	Trp	Pro	Glu	Asp	Leu	Gln	Asp	Val	Gly	Trp	Phe	Leu	
				325						330					335		
15	CAT	AAA	CTC	TCT	CAA	CAG	ATA	CTA	AAG	AAA	GAA	AAG	GAA	AAG	AAG	AAA	1056
	His	Lys	Leu	Ser	Gln	Gln	Ile	Leu	Lys	Lys	Glu	Lys	Glu	Lys	Lys	Lys	
				340					345						350		
20	GAC	AAT	AAC	ATT	CCG	TTG	CAA	ACT	GTA	GCA	ACC	ATC	TCC	TAATCAAAGG			1105
	Asp	Asn	Asn	Ile	Pro	Leu	Gln	Thr	Val	Ala	Thr	Ile	Ser				
			355					360					365				
	AGCAATTTCC	AACTTATCTC	AAGCCACAAA	TAATCTTTCA	CTTTGTATTT	GCACCAAGTT											1165
25	ATCATTTTGG	GGTCCTCTCT	GGAGGTTTTT	TTTTTCTTTT	TGCTACTATG	AAAACAACAT											1225
	AAATCTCTCA	ATTTTCGTAT	CAAAAAAAAA	AAAAAAAAAA	TGGCGGCCGC												1275

30 (2) INFORMACIÓN PARA EL SEQ ID NO:10:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 365 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:10:

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

# ES 2 340 210 T3

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

## (2) INFORMACIÓN PARA EL SEQ ID NO:11:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 3138 pares de bases
  - (B) TIPO: ácido nucleico
  - (C) CUALIDAD DE LA HEBRA: sencilla
  - (D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: ADNc
- (ix) RASGO:
  - (A) NOMBRE/CLAVE: CDS
  - (B) LOCALIZACIÓN: 1..3135
- (ix) RASGO:
  - (A) NOMBRE/CLAVE: mat\_péptidos
  - (B) LOCALIZACIÓN: 67..3135

# ES 2 340 210 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:11:

5	ATG TGG ACA CTG AAG AGA CTA ATT CTT ATC CTT TTT AAC ATA ATC CTA Met Trp Thr Leu Lys Arg Leu Ile Leu Ile Leu Phe Asn Ile Ile Leu -22 -20 -15 -10	48
10	ATT TCC AAA CTC CTT GGG GCT AGA TGG TTT CCT AAA ACT CTG CCC TGT Ile Ser Lys Leu Leu Gly Ala Arg Trp Phe Pro Lys Thr Leu Pro Cys -5 1 5 10	96
15	GAT GTC ACT CTG GAT GTT CCA AAG AAC CAT GTG ATC GTG GAC TGC ACA Asp Val Thr Leu Asp Val Pro Lys Asn His Val Ile Val Asp Cys Thr 15 20 25	144
20	GAC AAG CAT TTG ACA GAA ATT CCT GGA GGT ATT CCC ACG AAC ACC ACG Asp Lys His Leu Thr Glu Ile Pro Gly Gly Ile Pro Thr Asn Thr Thr 30 35 40	192
25	AAC CTC ACC CTC ACC ATT AAC CAC ATA CCA GAC ATC TCC CCA GCG TCC Asn Leu Thr Leu Thr Ile Asn His Ile Pro Asp Ile Ser Pro Ala Ser 45 50 55	240
30	TTT CAC AGA CTG GAC CAT CTG GTA GAG ATC GAT TTC AGA TGC AAC TGT Phe His Arg Leu Asp His Leu Val Glu Ile Asp Phe Arg Cys Asn Cys 60 65 70	288
35	GTA CCT ATT CCA CTG GGG TCA AAA AAC AAC ATG TGC ATC AAG AGG CTG Val Pro Ile Pro Leu Ser Lys Asn Asn Met Cys Ile Lys Arg Leu 75 80 85 90	336
40	CAG ATT AAA CCC AGA AGC TTT AGT GGA CTC ACT TAT TTA AAA TCC CTT Gln Ile Lys Pro Arg Ser Phe Ser Gly Leu Thr Tyr Leu Lys Ser Leu 95 100 105	384
45	TAC CTG GAT GGA AAC CAG CTA CTA GAG ATA CCG CAG GGC CTC CCG CCT Tyr Leu Asp Gly Asn Gln Leu Leu Glu Ile Pro Gln Gly Leu Pro Pro 110 115 120	432
50	AGC TTA CAG CTT CTC AGC CTT GAG GCC AAC AAC ATC TTT TCC ATC AGA Ser Leu Gln Leu Leu Ser Leu Glu Ala Asn Asn Ile Phe Ser Ile Arg 125 130 135	480
55	AAA GAG AAT CTA ACA GAA CTG GCC AAC ATA GAA ATA CTC TAC CTG GGC Lys Glu Asn Leu Thr Glu Leu Ala Asn Ile Glu Ile Leu Tyr Leu Gly 140 145 150	528
60	CAA AAC TGT TAT TAT CGA AAT CCT TGT TAT GTT TCA TAT TCA ATA GAG Gln Asn Cys Tyr Tyr Arg Asn Pro Cys Tyr Val Ser Tyr Ser Ile Glu 155 160 165 170	576
65	AAA GAT GCC TTC CTA AAC TTG ACA AAG TTA AAA GTG CTC TCC CTG AAA Lys Asp Ala Phe Leu Asn Leu Thr Lys Leu Lys Val Leu Ser Leu Lys 175 180 185	624

# ES 2 340 210 T3

	GAT AAC AAT GTC ACA GCC GTC CCT ACT GTT TTG CCA TCT ACT TTA ACA	672
	Asp Asn Asn Val Thr Ala Val Pro Thr Val Leu Pro Ser Thr Leu Thr	
	190 195 200	
5	GAA CTA TAT CTC TAC AAC AAC ATG ATT GCA AAA ATC CAA GAA GAT GAT	720
	Glu Leu Tyr Leu Tyr Asn Asn Met Ile Ala Lys Ile Gln Glu Asp Asp	
	205 210 215	
10	TTT AAT AAC CTC AAC CAA TTA CAA ATT CTT GAC CTA AGT GGA AAT TGC	768
	Phe Asn Asn Leu Asn Gln Leu Gln Ile Leu Asp Leu Ser Gly Asn Cys	
	220 225 230	
15	CCT CGT TGT TAT AAT GCC CCA TTT CCT TGT GCG CCG TGT AAA AAT AAT	816
	Pro Arg Cys Tyr Asn Ala Pro Phe Pro Cys Ala Pro Cys Lys Asn Asn	
	235 240 245 250	
	TCT CCC CTA CAG ATC CCT GTA AAT GCT TTT GAT GCG CTG ACA GAA TTA	864
	Ser Pro Leu Gln Ile Pro Val Asn Ala Phe Asp Ala Leu Thr Glu Leu	
	255 260 265	
20	AAA GTT TTA CGT CTA CAC AGT AAC TCT CTT CAG CAT GTG CCC CCA AGA	912
	Lys Val Leu Arg Leu His Ser Asn Ser Leu Gln His Val Pro Pro Arg	
	270 275 280	
25	TGG TTT AAG AAC ATC AAC AAA CTC CAG GAA CTG GAT CTG TCC CAA AAC	960
	Trp Phe Lys Asn Ile Asn Lys Leu Gln Glu Leu Asp Leu Ser Gln Asn	
	285 290 295	
30	TTC TTG GCC AAA GAA ATT GGG GAT GCT AAA TTT CTG CAT TTT CTC CCC	1008
	Phe Leu Ala Lys Glu Ile Gly Asp Ala Lys Phe Leu His Phe Leu Pro	
	300 305 310	
35	AGC CTC ATC CAA TTG GAT CTG TCT TTC AAT TTT GAA CTT CAG GTC TAT	1056
	Ser Leu Ile Gln Leu Asp Leu Ser Phe Asn Phe Glu Leu Gln Val Tyr	
	315 320 325 330	
40	CGT GCA TCT ATG AAT CTA TCA CAA GCA TTT TCT TCA CTG AAA AGC CTG	1104
	Arg Ala Ser Met Asn Leu Ser Gln Ala Phe Ser Ser Leu Lys Ser Leu	
	335 340 345	
45	AAA ATT CTG CGG ATC AGA GGA TAT GTC TTT AAA GAG TTG AAA AGC TTT	1152
	Lys Ile Leu Arg Ile Arg Gly Tyr Val Phe Lys Glu Leu Lys Ser Phe	
	350 355 360	
50	AAC CTC TCG CCA TTA CAT AAT CTT CAA AAT CTT GAA GTT CTT GAT CTT	1200
	Asn Leu Ser Pro Leu His Asn Leu Gln Asn Leu Glu Val Leu Asp Leu	
	365 370 375	
55	GGC ACT AAC TTT ATA AAA ATT GCT AAC CTC AGC ATG TTT AAA CAA TTT	1248
	Gly Thr Asn Phe Ile Lys Ile Ala Asn Leu Ser Met Phe Lys Gln Phe	
	380 385 390	
60	AAA AGA CTG AAA GTC ATA GAT CTT TCA GTG AAT AAA ATA TCA CCT TCA	1296
	Lys Arg Leu Lys Val Ile Asp Leu Ser Val Asn Lys Ile Ser Pro Ser	
	395 400 405 410	
65	GGA GAT TCA AGT GAA GTT GGC TTC TGC TCA AAT GCC AGA ACT TCT GTA	1344
	Gly Asp Ser Ser Glu Val Gly Phe Cys Ser Asn Ala Arg Thr Ser Val	
	415 420 425	

# ES 2 340 210 T3

		GAA AGT TAT GAA CCC CAG GTC CTG GAA CAA TTA CAT TAT TTC AGA TAT	1392
		Glu Ser Tyr Glu Pro Gln Val Leu Glu Gln Leu His Tyr Phe Arg Tyr	
		430 435 440	
5		GAT AAG TAT GCA AGG AGT TGC AGA TTC AAA AAC AAA GAG GCT TCT TTC	1440
		Asp Lys Tyr Ala Arg Ser Cys Arg Phe Lys Asn Lys Glu Ala Ser Phe	
		445 450 455	
10		ATG TCT GTT AAT GAA AGC TGC TAC AAG TAT GGG CAG ACC TTG GAT CTA	1488
		Met Ser Val Asn Glu Ser Cys Tyr Lys Tyr Gly Gln Thr Leu Asp Leu	
		460 465 470	
15		AGT AAA AAT AGT ATA TTT TTT GTC AAG TCC TCT GAT TTT CAG CAT CTT	1536
		Ser Lys Asn Ser Ile Phe Phe Val Lys Ser Ser Asp Phe Gln His Leu	
		475 480 485 490	
		TCT TTC CTC AAA TGC CTG AAT CTG TCA GGA AAT CTC ATT AGC CAA ACT	1584
		Ser Phe Leu Lys Cys Leu Asn Leu Ser Gly Asn Leu Ile Ser Gln Thr	
		495 500 505	
20		CTT AAT GGC AGT GAA TTC CAA CCT TTA GCA GAG CTG AGA TAT TTG GAC	1632
		Leu Asn Gly Ser Glu Phe Gln Pro Leu Ala Glu Leu Arg Tyr Leu Asp	
		510 515 520	
25		TTC TCC AAC AAC CGG CTT GAT TTA CTC CAT TCA ACA GCA TTT GAA GAG	1680
		Phe Ser Asn Asn Arg Leu Asp Leu Leu His Ser Thr Ala Phe Glu Glu	
		525 530 535	
		CTT CAC AAA CTG GAA GTT CTG GAT ATA AGC AGT AAT AGC CAT TAT TTT	1728
		Leu His Lys Leu Glu Val Leu Asp Ile Ser Ser Asn Ser His Tyr Phe	
		540 545 550	
30		CAA TCA GAA GGA ATT ACT CAT ATG CTA AAC TTT ACC AAG AAC CTA AAG	1776
		Gln Ser Glu Gly Ile Thr His Met Leu Asn Phe Thr Lys Asn Leu Lys	
		555 560 565 570	
35		GTT CTG CAG AAA CTG ATG ATG AAC GAC AAT GAC ATC TCT TCC TCC ACC	1824
		Val Leu Gln Lys Leu Met Met Asn Asp Asn Asp Ile Ser Ser Ser Thr	
		575 580 585	
40		AGC AGG ACC ATG GAG AGT GAG TCT CTT AGA ACT CTG GAA TTC AGA GGA	1872
		Ser Arg Thr Met Glu Ser Glu Ser Leu Arg Thr Leu Glu Phe Arg Gly	
		590 595 600	
		AAT CAC TTA GAT GTT TTA TGG AGA GAA GGT GAT AAC AGA TAC TTA CAA	1920
		Asn His Leu Asp Val Leu Trp Arg Glu Gly Asp Asn Arg Tyr Leu Gln	
		605 610 615	
45		TTA TTC AAG AAT CTG CTA AAA TTA GAG GAA TTA GAC ATC TCT AAA AAT	1968
		Leu Phe Lys Asn Leu Leu Lys Leu Glu Glu Leu Asp Ile Ser Lys Asn	
		620 625 630	
50		TCC CTA AGT TTC TTG CCT TCT GGA GTT TTT GAT GGT ATG CCT CCA AAT	2016
		Ser Leu Ser Phe Leu Pro Ser Gly Val Phe Asp Gly Met Pro Pro Asn	
		635 640 645 650	
		CTA AAG AAT CTC TCT TTG GCC AAA AAT GGG CTC AAA TCT TTC AGT TGG	2064
		Leu Lys Asn Leu Ser Leu Ala Lys Asn Gly Leu Lys Ser Phe Ser Trp	
		655 660 665	
55		AAG AAA CTC CAG TGT CTA AAG AAC CTG GAA ACT TTG GAC CTC AGC CAC	2112



# ES 2 340 210 T3

	Lys	Lys	Leu	Gln	Cys	Leu	Lys	Asn	Leu	Glu	Thr	Leu	Asp	Leu	Ser	His	
				670					675					680			
5	AAC	CAA	CTG	ACC	ACT	GTC	CCT	GAG	AGA	TTA	TCC	AAC	TGT	TCC	AGA	AGC	2160
	Asn	Gln	Leu	Thr	Thr	Val	Pro	Glu	Arg	Leu	Ser	Asn	Cys	Ser	Arg	Ser	
			685					690					695				
10	CTC	AAG	AAT	CTG	ATT	CTT	AAG	AAT	AAT	CAA	ATC	AGG	AGT	CTG	ACG	AAG	2208
	Leu	Lys	Asn	Leu	Ile	Leu	Lys	Asn	Asn	Gln	Ile	Arg	Ser	Leu	Thr	Lys	
		700					705					710					
15	TAT	TTT	CTA	CAA	GAT	GCC	TTC	CAG	TTG	CGA	TAT	CTG	GAT	CTC	AGC	TCA	2256
	Tyr	Phe	Leu	Gln	Asp	Ala	Phe	Gln	Leu	Arg	Tyr	Leu	Asp	Leu	Ser	Ser	
	715					720					725					730	
20	AAT	AAA	ATC	CAG	ATG	ATC	CAA	AAG	ACC	AGC	TTC	CCA	GAA	AAT	GTC	CTC	2304
	Asn	Lys	Ile	Gln	Met	Ile	Gln	Lys	Thr	Ser	Phe	Pro	Glu	Asn	Val	Leu	
				735						740					745		
25	AAC	AAT	CTG	AAG	ATG	TTG	CTT	TTG	CAT	CAT	AAT	CGG	TTT	CTG	TGC	ACC	2352
	Asn	Asn	Leu	Lys	Met	Leu	Leu	Leu	His	His	Asn	Arg	Phe	Leu	Cys	Thr	
				750					755					760			
30	TGT	GAT	GCT	GTG	TGG	TTT	GTC	TGG	TGG	GTT	AAC	CAT	ACG	GAG	GTG	ACT	2400
	Cys	Asp	Ala	Val	Trp	Phe	Val	Trp	Trp	Val	Asn	His	Thr	Glu	Val	Thr	
			765				770						775				
35	ATT	CCT	TAC	CTG	GCC	ACA	GAT	GTG	ACT	TGT	GTG	GGG	CCA	GGA	GCA	CAC	2448
	Ile	Pro	Tyr	Leu	Ala	Thr	Asp	Val	Thr	Cys	Val	Gly	Pro	Gly	Ala	His	
		780					785					790					
40	AAG	GGC	CAA	AGT	GTG	ATC	TCC	CTG	GAT	CTG	TAC	ACC	TGT	GAG	TTA	GAT	2496
	Lys	Gly	Gln	Ser	Val	Ile	Ser	Leu	Asp	Leu	Tyr	Thr	Cys	Glu	Leu	Asp	
	795					800					805					810	
45	CTG	ACT	AAC	CTG	ATT	CTG	TTC	TCA	CTT	TCC	ATA	TCT	GTA	TCT	CTC	TTT	2544
	Leu	Thr	Asn	Leu	Ile	Leu	Phe	Ser	Leu	Ser	Ile	Ser	Val	Ser	Leu	Phe	
				815						820					825		
50	CTC	ATG	GTG	ATG	ATG	ACA	GCA	AGT	CAC	CTC	TAT	TTC	TGG	GAT	GTG	TGG	2592
	Leu	Met	Val	Met	Met	Thr	Ala	Ser	His	Leu	Tyr	Phe	Trp	Asp	Val	Trp	
				830					835					840			
55	TAT	ATT	TAC	CAT	TTC	TGT	AAG	GCC	AAG	ATA	AAG	GGG	TAT	CAG	CGT	CTA	2640
	Tyr	Ile	Tyr	His	Phe	Cys	Lys	Ala	Lys	Ile	Lys	Gly	Tyr	Gln	Arg	Leu	
			845					850					855				
60	ATA	TCA	CCA	GAC	TGT	TGC	TAT	GAT	GCT	TTT	ATT	GTG	TAT	GAC	ACT	AAA	2688
	Ile	Ser	Pro	Asp	Cys	Cys	Tyr	Asp	Ala	Phe	Ile	Val	Tyr	Asp	Thr	Lys	
		860					865					870					
65	GAC	CCA	GCT	GTG	ACC	GAG	TGG	GTT	TTG	GCT	GAG	CTG	GTG	GCC	AAA	CTG	2736
	Asp	Pro	Ala	Val	Thr	Glu	Trp	Val	Leu	Ala	Glu	Leu	Val	Ala	Lys	Leu	
						880					885					890	
70	GAA	GAC	CCA	AGA	GAG	AAA	CAT	TTT	AAT	TTA	TGT	CTC	GAG	GAA	AGG	GAC	2784
	Glu	Asp	Pro	Arg	Glu	Lys	His	Phe	Asn	Leu	Cys	Leu	Glu	Glu	Arg	Asp	
				895						900					905		
75	TGG	TTA	CCA	GGG	CAG	CCA	GTT	CTG	GAA	AAC	CTT	TCC	CAG	AGC	ATA	CAG	2832
	Trp	Leu	Pro	Gly	Gln	Pro	Val	Leu	Glu	Asn	Leu	Ser	Gln	Ser	Ile	Gln	

# ES 2 340 210 T3

	910	915	920	
5	CTT AGC AAA AAG ACA GTG TTT GTG ATG ACA GAC AAG TAT GCA AAG ACT Leu Ser Lys Lys Thr Val Phe Val Met Thr Asp Lys Tyr Ala Lys Thr 925 930 935	2880		
10	GAA AAT TTT AAG ATA GCA TTT TAC TTG TCC CAT CAG AGG CTC ATG GAT Glu Asn Phe Lys Ile Ala Phe Tyr Leu Ser His Gln Arg Leu Met Asp 940 945 950	2928		
15	GAA AAA GTT GAT GTG ATT ATC TTG ATA TTT CTT GAG AAG CCC TTT CAG Glu Lys Val Asp Val Ile Ile Leu Ile Phe Leu Glu Lys Pro Phe Gln 955 960 965 970	2976		
20	AAG TCC AAG TTC CTC CAG CTC CGG AAA AGG CTC TGT GGG AGT TCT GTC Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg Leu Cys Gly Ser Ser Val 975 980 985	3024		
25	CTT GAG TGG CCA ACA AAC CCG CAA GCT CAC CCA TAC TTC TGG CAG TGT Leu Glu Trp Pro Thr Asn Pro Gln Ala His Pro Tyr Phe Trp Gln Cys 990 995 1000	3072		
30	CTA AAG AAC GCC CTG GCC ACA GAC AAT CAT GTG GCC TAT AGT CAG GTG Leu Lys Asn Ala Leu Ala Thr Asp Asn His Val Ala Tyr Ser Gln Val 1005 1010 1015	3120		
35	TTC AAG GAA ACG GTC TAG Phe Lys Glu Thr Val 1020	3138		

30 (2) INFORMACIÓN PARA EL SEQ ID NO:12:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1045 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:12:

45	Met Trp Thr Leu Lys Arg Leu Ile Leu Ile Leu Phe Asn Ile Ile Leu -22 -20 -15 -10
50	Ile Ser Lys Leu Leu Gly Ala Arg Trp Phe Pro Lys Thr Leu Pro Cys -5 1 5 10
55	Asp Val Thr Leu Asp Val Pro Lys Asn His Val Ile Val Asp Cys Thr 15 20 25
60	Asp Lys His Leu Thr Glu Ile Pro Gly Gly Ile Pro Thr Asn Thr Thr 30 35 40
65	Asn Leu Thr Leu Thr Ile Asn His Ile Pro Asp Ile Ser Pro Ala Ser 45 50 55
	Phe His Arg Leu Asp His Leu Val Glu Ile Asp Phe Arg Cys Asn Cys 60 65 70
	Val Pro Ile Pro Leu Gly Ser Lys Asn Asn Met Cys Ile Lys Arg Leu

# ES 2 340 210 T3

	75					80						85				90
	Gln	Ile	Lys	Pro	Arg	Ser	Phe	Ser	Gly	Leu	Thr	Tyr	Leu	Lys	Ser	Leu
5					95					100					105	
	Tyr	Leu	Asp	Gly	Asn	Gln	Leu	Leu	Glu	Ile	Pro	Gln	Gly	Leu	Pro	Pro
				110					115					120		
10	Ser	Leu	Gln	Leu	Leu	Ser	Leu	Glu	Ala	Asn	Asn	Ile	Phe	Ser	Ile	Arg
			125					130					135			
	Lys	Glu	Asn	Leu	Thr	Glu	Leu	Ala	Asn	Ile	Glu	Ile	Leu	Tyr	Leu	Gly
		140					145					150				
15	Gln	Asn	Cys	Tyr	Tyr	Arg	Asn	Pro	Cys	Tyr	Val	Ser	Tyr	Ser	Ile	Glu
	155					160					165					170
	Lys	Asp	Ala	Phe	Leu	Asn	Leu	Thr	Lys	Leu	Lys	Val	Leu	Ser	Leu	Lys
					175					180					185	
20	Asp	Asn	Asn	Val	Thr	Ala	Val	Pro	Thr	Val	Leu	Pro	Ser	Thr	Leu	Thr
				190					195					200		
	Glu	Leu	Tyr	Leu	Tyr	Asn	Asn	Met	Ile	Ala	Lys	Ile	Gln	Glu	Asp	Asp
			205					210					215			
25	Phe	Asn	Asn	Leu	Asn	Gln	Leu	Gln	Ile	Leu	Asp	Leu	Ser	Gly	Asn	Cys
	220					225						230				
	Pro	Arg	Cys	Tyr	Asn	Ala	Pro	Phe	Pro	Cys	Ala	Pro	Cys	Lys	Asn	Asn
	235				240						245					250
30	Ser	Pro	Leu	Gln	Ile	Pro	Val	Asn	Ala	Phe	Asp	Ala	Leu	Thr	Glu	Leu
					255					260					265	
	Lys	Val	Leu	Arg	Leu	His	Ser	Asn	Ser	Leu	Gln	His	Val	Pro	Pro	Arg
				270					275					280		
35	Trp	Phe	Lys	Asn	Ile	Asn	Lys	Leu	Gln	Glu	Leu	Asp	Leu	Ser	Gln	Asn
			285					290					295			
	Phe	Leu	Ala	Lys	Glu	Ile	Gly	Asp	Ala	Lys	Phe	Leu	His	Phe	Leu	Pro
		300					305					310				
40	Ser	Leu	Ile	Gln	Leu	Asp	Leu	Ser	Phe	Asn	Phe	Glu	Leu	Gln	Val	Tyr
	315					320					325					330
	Arg	Ala	Ser	Met	Asn	Leu	Ser	Gln	Ala	Phe	Ser	Ser	Leu	Lys	Ser	Leu
					335					340					345	
45	Lys	Ile	Leu	Arg	Ile	Arg	Gly	Tyr	Val	Phe	Lys	Glu	Leu	Lys	Ser	Phe
				350					355					360		
	Asn	Leu	Ser	Pro	Leu	His	Asn	Leu	Gln	Asn	Leu	Glu	Val	Leu	Asp	Leu
50				365				370					375			
	Gly	Thr	Asn	Phe	Ile	Lys	Ile	Ala	Asn	Leu	Ser	Met	Phe	Lys	Gln	Phe
		380					385					390				
55	Lys	Arg	Leu	Lys	Val	Ile	Asp	Leu	Ser	Val	Asn	Lys	Ile	Ser	Pro	Ser
	395					400					405					410

# ES 2 340 210 T3

	Gly Asp Ser Ser Glu Val Gly Phe Cys Ser Asn Ala Arg Thr Ser Val	415	420	425
5	Glu Ser Tyr Glu Pro Gln Val Leu Glu Gln Leu His Tyr Phe Arg Tyr	430	435	440
	Asp Lys Tyr Ala Arg Ser Cys Arg Phe Lys Asn Lys Glu Ala Ser Phe	445	450	455
10	Met Ser Val Asn Glu Ser Cys Tyr Lys Tyr Gly Gln Thr Leu Asp Leu	460	465	470
	Ser Lys Asn Ser Ile Phe Phe Val Lys Ser Ser Asp Phe Gln His Leu	475	480	485
15	Ser Phe Leu Lys Cys Leu Asn Leu Ser Gly Asn Leu Ile Ser Gln Thr	495	500	505
	Leu Asn Gly Ser Glu Phe Gln Pro Leu Ala Glu Leu Arg Tyr Leu Asp	510	515	520
20	Phe Ser Asn Asn Arg Leu Asp Leu Leu His Ser Thr Ala Phe Glu Glu	525	530	535
	Leu His Lys Leu Glu Val Leu Asp Ile Ser Ser Asn Ser His Tyr Phe	540	545	550
25	Gln Ser Glu Gly Ile Thr His Met Leu Asn Phe Thr Lys Asn Leu Lys	555	560	565
	Val Leu Gln Lys Leu Met Met Asn Asp Asn Asp Ile Ser Ser Ser Thr	575	580	585
30	Ser Arg Thr Met Glu Ser Glu Ser Leu Arg Thr Leu Glu Phe Arg Gly	590	595	600
	Asn His Leu Asp Val Leu Trp Arg Glu Gly Asp Asn Arg Tyr Leu Gln	605	610	615
35	Leu Phe Lys Asn Leu Leu Lys Leu Glu Glu Leu Asp Ile Ser Lys Asn	620	625	630
	Ser Leu Ser Phe Leu Pro Ser Gly Val Phe Asp Gly Met Pro Pro Asn	635	640	645
40	Leu Lys Asn Leu Ser Leu Ala Lys Asn Gly Leu Lys Ser Phe Ser Trp	655	660	665
	Lys Lys Leu Gln Cys Leu Lys Asn Leu Glu Thr Leu Asp Leu Ser His	670	675	680
45	Asn Gln Leu Thr Thr Val Pro Glu Arg Leu Ser Asn Cys Ser Arg Ser	685	690	695
	Leu Lys Asn Leu Ile Leu Lys Asn Asn Gln Ile Arg Ser Leu Thr Lys	700	705	710
50	Tyr Phe Leu Gln Asp Ala Phe Gln Leu Arg Tyr Leu Asp Leu Ser Ser	715	720	725
				730

# ES 2 340 210 T3

Asn Lys Ile Gln Met Ile Gln Lys Thr Ser Phe Pro Glu Asn Val Leu  
 735 740 745  
 5 Asn Asn Leu Lys Met Leu Leu Leu His His Asn Arg Phe Leu Cys Thr  
 750 755 760  
 Cys Asp Ala Val Trp Phe Val Trp Trp Val Asn His Thr Glu Val Thr  
 765 770 775  
 10 Ile Pro Tyr Leu Ala Thr Asp Val Thr Cys Val Gly Pro Gly Ala His  
 780 785 790  
 Lys Gly Gln Ser Val Ile Ser Leu Asp Leu Tyr Thr Cys Glu Leu Asp  
 795 800 805 810  
 15 Leu Thr Asn Leu Ile Leu Phe Ser Leu Ser Ile Ser Val Ser Leu Phe  
 815 820 825  
 Leu Met Val Met Met Thr Ala Ser His Leu Tyr Phe Trp Asp Val Trp  
 830 835 840  
 20 Tyr Ile Tyr His Phe Cys Lys Ala Lys Ile Lys Gly Tyr Gln Arg Leu  
 845 850 855  
 Ile Ser Pro Asp Cys Cys Tyr Asp Ala Phe Ile Val Tyr Asp Thr Lys  
 860 865 870  
 Asp Pro Ala Val Thr Glu Trp Val Leu Ala Glu Leu Val Ala Lys Leu  
 875 880 885 890  
 30 Glu Asp Pro Arg Glu Lys His Phe Asn Leu Cys Leu Glu Glu Arg Asp  
 895 900 905  
 Trp Leu Pro Gly Gln Pro Val Leu Glu Asn Leu Ser Gln Ser Ile Gln  
 910 915 920  
 35 Leu Ser Lys Lys Thr Val Phe Val Met Thr Asp Lys Tyr Ala Lys Thr  
 925 930 935  
 Glu Asn Phe Lys Ile Ala Phe Tyr Leu Ser His Gln Arg Leu Met Asp  
 940 945 950  
 40 Glu Lys Val Asp Val Ile Ile Leu Ile Phe Leu Glu Lys Pro Phe Gln  
 955 960 965 970  
 Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg Leu Cys Gly Ser Ser Val  
 975 980 985  
 45 Leu Glu Trp Pro Thr Asn Pro Gln Ala His Pro Tyr Phe Trp Gln Cys  
 990 995 1000  
 Leu Lys Asn Ala Leu Ala Thr Asp Asn His Val Ala Tyr Ser Gln Val  
 1005 1010 1015  
 Phe Lys Glu Thr Val  
 1020

## (2) INFORMACIÓN PARA EL SEQ ID NO:13:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 180 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

### (ii) TIPO DE MOLÉCULA: ADNc

ES 2 340 210 T3

(ix) RASGO:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 1..177

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:13:

CTT GGA AAA CCT CTT CAG AAG TCT AAG TTT CTT CAG CTC AGG AAG AGA 48  
Leu Gly Lys Pro Leu Gln Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg  
1 5 10 15

CTC TGC AGG AGC TCT GTC CTT GAG TGG CCT GCA AAT CCA CAG GCT CAC 96  
Leu Cys Arg Ser Ser Val Leu Glu Trp Pro Ala Asn Pro Gln Ala His  
20 25 30

CCA TAC TTC TGG CAG TGC CTG AAA AAT GCC CTG ACC ACA GAC AAT CAT 144  
Pro Tyr Phe Trp Gln Cys Leu Lys Asn Ala Leu Thr Thr Asp Asn His  
35 40 45

GTG GCT TAT AGT CAA ATG TTC AAG GAA ACA GTC TAG 180  
Val Ala Tyr Ser Gln Met Phe Lys Glu Thr Val  
50 55

25 (2) INFORMACIÓN PARA EL SEQ ID NO:14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 59 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:14:

Leu Gly Lys Pro Leu Gln Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg  
1 5 10 15

Leu Cys Arg Ser Ser Val Leu Glu Trp Pro Ala Asn Pro Gln Ala His  
20 25 30

Pro Tyr Phe Trp Gln Cys Leu Lys Asn Ala Leu Thr Thr Asp Asn His  
35 40 45

Val Ala Tyr Ser Gln Met Phe Lys Glu Thr Val  
50 55

<sup>50</sup> (2) INFORMACIÓN PARA EL SEQ ID NO:15:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 990 pares de bases

(B) TIPO: ácido nucleico

(C) CUALIDAD DE LA HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(ix) RASGO:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 2..988

# ES 2 340 210 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:15:

5	G AAT TCC AGA CTT ATA AAC TTG AAA AAT CTC TAT TTG GCC TGG AAC	46
	Asn Ser Arg Leu Ile Asn Leu Lys Asn Leu Tyr Leu Ala Trp Asn	
	1 5 10 15	
10	TGC TAT TTT AAC AAA GTT TGC GAG AAA ACT AAC ATA GAA GAT GGA GTA	94
	Cys Tyr Phe Asn Lys Val Cys Glu Lys Thr Asn Ile Glu Asp Gly Val	
	20 25 30	
15	TTT GAA ACG CTG ACA AAT TTG GAG TTG CTA TCA CTA TCT TTC AAT TCT	142
	Phe Glu Thr Leu Thr Asn Leu Glu Leu Ser Leu Ser Phe Asn Ser	
	35 40 45	
20	CTT TCA CAT GTG CCA CCC AAA CTG CCA AGC TCC CTA CGC AAA CTT TTT	190
	Leu Ser His Val Pro Pro Lys Leu Pro Ser Ser Leu Arg Lys Leu Phe	
	50 55 60	
25	CTG AGC AAC ACC CAG ATC AAA TAC ATT AGT GAA GAA GAT TTC AAG GGA	238
	Leu Ser Asn Thr Gln Ile Lys Tyr Ile Ser Glu Glu Asp Phe Lys Gly	
	65 70 75	
30	TTG ATA AAT TTA ACA TTA CTA GAT TTA AGC GGG AAC TGT CCG AGG TGC	286
	Leu Ile Asn Leu Thr Leu Leu Asp Leu Ser Gly Asn Cys Pro Arg Cys	
	80 85 90 95	
35	TTC AAT GCC CCA TTT CCA TGC GTG CCT TGT GAT GGT GGT GCT TCA ATT	334
	Phe Asn Ala Pro Phe Pro Cys Val Pro Cys Asp Gly Gly Ala Ser Ile	
	100 105 110	
40	AAT ATA GAT CGT TTT GCT TTT CAA AAC TTG ACC CAA CTT CGA TAC CTA	382
	Asn Ile Asp Arg Phe Ala Phe Gln Asn Leu Thr Gln Leu Arg Tyr Leu	
	115 120 125	
45	AAC CTC TCT AGC ACT TCC CTC AGG AAG ATT AAT GCT GCC TGG TTT AAA	430
	Asn Leu Ser Ser Thr Ser Leu Arg Lys Ile Asn Ala Ala Trp Phe Lys	
	130 135 140	
50	AAT ATG CCT CAT CTG AAG GTG CTG GAT CTT GAA TTC AAC TAT TTA GTG	478
	Asn Met Pro His Leu Lys Val Leu Asp Leu Glu Phe Asn Tyr Leu Val	
	145 150 155	
55	GGA GAA ATA GCC TCT GGG GCA TTT TTA ACG ATG CTG CCC CGC TTA GAA	526
	Gly Glu Ile Ala Ser Gly Ala Phe Leu Thr Met Leu Pro Arg Leu Glu	
	160 165 170 175	
60	ATA CTT GAC TTG TCT TTT AAC TAT ATA AAG GGG AGT TAT CCA CAG CAT	574
	Ile Leu Asp Leu Ser Phe Asn Tyr Ile Lys Gly Ser Tyr Pro Gln His	
	180 185 190	
65	ATT AAT ATT TCC AGA AAC TTC TCT AAA CTT TTG TCT CTA CGG GCA TTG	622
	Ile Asn Ile Ser Arg Asn Phe Ser Lys Leu Leu Ser Leu Arg Ala Leu	

# ES 2 340 210 T3

	195	200	205	
5	CAT TTA AGA GGT TAT GTG TTC CAG GAA CTC AGA GAA GAT GAT TTC CAG His Leu Arg Gly Tyr Val Phe Gln Glu Leu Arg Glu Asp Asp Phe Gln 210 215 220	670		
10	CCC CTG ATG CAG CTT CCA AAC TTA TCG ACT ATC AAC TTG GGT ATT AAT Pro Leu Met Gln Leu Pro Asn Leu Ser Thr Ile Asn Leu Gly Ile Asn 225 230 235	718		
15	TTT ATT AAG CAA ATC GAT TTC AAA CTT TTC CAA AAT TTC TCC AAT CTG Phe Ile Lys Gln Ile Asp Phe Lys Leu Phe Gln Asn Phe Ser Asn Leu 240 245 250 255	766		
20	GAA ATT ATT TAC TTG TCA GAA AAC AGA ATA TCA CCG TTG GTA AAA GAT Glu Ile Ile Tyr Leu Ser Glu Asn Arg Ile Ser Pro Leu Val Lys Asp 260 265 270	814		
25	ACC CGG CAG AGT TAT GCA AAT AGT TCC TCT TTT CAA CGT CAT ATC CGG Thr Arg Gln Ser Tyr Ala Asn Ser Ser Ser Phe Gln Arg His Ile Arg 275 280 285	862		
30	AAA CGA CGC TCA ACA GAT TTT GAG TTT GAC CCA CAT TCG AAC TTT TAT Lys Arg Arg Ser Thr Asp Phe Glu Phe Asp Pro His Ser Asn Phe Tyr 290 295 300	910		
35	CAT TTC ACC CGT CCT TTA ATA AAG CCA CAA TGT GCT GCT TAT GGA AAA His Phe Thr Arg Pro Leu Ile Lys Pro Gln Cys Ala Ala Tyr Gly Lys 305 310 315	958		
40	GCC TTA GAT TTA AGC CTC AAC AGT ATT TTC TT Ala Leu Asp Leu Ser Leu Asn Ser Ile Phe 320 325	990		

## (2) INFORMACIÓN PARA EL SEQ ID NO:16:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 329 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

### (ii) TIPO DE MOLÉCULA: proteína

### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:16:

Asn Ser Arg Leu Ile Asn Leu Lys Asn Leu Tyr Leu Ala Trp Asn Cys  
1 5 10 15

Tyr Phe Asn Lys Val Cys Glu Lys Thr Asn Ile Glu Asp Gly Val Phe  
20 25 30

Glu Thr Leu Thr Asn Leu Glu Leu Leu Ser Leu Ser Phe Asn Ser Leu  
35 40 45

Ser His Val Pro Pro Lys Leu Pro Ser Ser Leu Arg Lys Leu Phe Leu  
50 55 60

Ser Asn Thr Gln Ile Lys Tyr Ile Ser Glu Glu Asp Phe Lys Gly Leu  
65 70 75 80



## ES 2 340 210 T3

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

### (2) INFORMACIÓN PARA EL SEQ ID NO:17:

- |    |  |
|----|--|
|    | (i) CARACTERÍSTICAS DE LA SECUENCIA:   |
| 50 | (A) LONGITUD: 1557 pares de bases  |
|    | (B) TIPO: ácido nucleico   |
|    | (C) CUALIDAD DE LA HEBRA: sencilla   |
| 55 | (D) TOPOLOGÍA: lineal  |
|    | (ii) TIPO DE MOLÉCULA: ADNc  |
|    | (ix) RASGO:  |
| 60 | (A) NOMBRE/CLAVE: CDS  |
|    | (B) LOCALIZACIÓN: 1..513   |
|    | (ix) RASGO:  |
| 65 | (A) NOMBRE/CLAVE: rasgo_misc   |
|    | (B) LOCALIZACIÓN: 27 8   |
|    | (D) OTRA INFORMACIÓN: /nota= "el nucleótido 278 denominado G, puede ser G o C" |

## ES 2 340 210 T3

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 445

(D) OTRA INFORMACIÓN: /nota= "el nucleótido 445 denominado A, puede ser A o T"

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 572

(D) OTRA INFORMACIÓN: /nota= "los nucleótidos 572, 593, 600, 607, 617, 622, 625, 631, 640, 646, 653, 719, 775, y 861 denominados C; pueden ser cada A, C, G, o T"

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:17:

	<b>CAG TCT CTT TCC ACA TCC CAA ACT TTC TAT GAT GCT TAC ATT TCT TAT</b>	<b>48</b>
	<b>Gln Ser Leu Ser Thr Ser Gln Thr Phe Tyr Asp Ala Tyr Ile Ser Tyr</b>	
	<div style="display: flex; justify-content: space-between; width: 80%;"> <span>1</span> <span>5</span> <span>10</span> <span>15</span> </div>	
20	<b>GAC ACC AAA GAT GCC TCT GTT ACT GAC TGG GTG ATA AAT GAG CTG CGC</b>	<b>96</b>
	<b>Asp Thr Lys Asp Ala Ser Val Thr Asp Trp Val Ile Asn Glu Leu Arg</b>	
	<div style="display: flex; justify-content: space-between; width: 80%;"> <span>20</span> <span>25</span> <span>30</span> </div>	
25	<b>TAC CAC CTT GAA GAG AGC CGA GAC AAA AAC GTT CTC CTT TGT CTA GAG</b>	<b>144</b>
	<b>Tyr His Leu Glu Glu Ser Arg Asp Lys Asn Val Leu Leu Cys Leu Glu</b>	
	<div style="display: flex; justify-content: space-between; width: 80%;"> <span>35</span> <span>40</span> <span>45</span> </div>	
	<b>GAG AGG GAT TGG GAC CCG GGA TTG GCC ATC ATC GAC AAC CTC ATG CAG</b>	<b>192</b>
	<b>Glu Arg Asp Trp Asp Pro Gly Leu Ala Ile Ile Asp Asn Leu Met Gln</b>	
30	<div style="display: flex; justify-content: space-between; width: 80%;"> <span>50</span> <span>55</span> <span>60</span> </div>	
	<b>AGC ATC AAC CAA AGC AAG AAA ACA GTA TTT GTT TTA ACC AAA AAA TAT</b>	<b>240</b>
	<b>Ser Ile Asn Gln Ser Lys Lys Thr Val Phe Val Leu Thr Lys Lys Tyr</b>	
	<div style="display: flex; justify-content: space-between; width: 80%;"> <span>65</span> <span>70</span> <span>75</span> <span>80</span> </div>	
35	<b>GCA AAA AGC TGG AAC TTT AAA ACA GCT TTT TAC TTG GGC TTG CAG AGG</b>	<b>288</b>
	<b>Ala Lys Ser Trp Asn Phe Lys Thr Ala Phe Tyr Leu Gly Leu Gln Arg</b>	
	<div style="display: flex; justify-content: space-between; width: 80%;"> <span>85</span> <span>90</span> <span>95</span> </div>	
	<b>CTA ATG GGT GAG AAC ATG GAT GTG ATT ATA TTT ATC CTG CTG GAG CCA</b>	<b>336</b>
40	<b>Leu Met Gly Glu Asn Met Asp Val Ile Ile Phe Ile Leu Leu Glu Pro</b>	
	<div style="display: flex; justify-content: space-between; width: 80%;"> <span>100</span> <span>105</span> <span>110</span> </div>	
	<b>GTG TTA CAG CAT TCT CCG TAT TTG AGG CTA CGG CAG CGG ATC TGT AAG</b>	<b>384</b>
	<b>Val Leu Gln His Ser Pro Tyr Leu Arg Leu Arg Gln Arg Ile Cys Lys</b>	
	<div style="display: flex; justify-content: space-between; width: 80%;"> <span>115</span> <span>120</span> <span>125</span> </div>	
45	<b>AGC TCC ATC CTC CAG TGG CCT GAC AAC CCG AAG GCA GAA AGG TTG TTT</b>	<b>432</b>
	<b>Ser Ser Ile Leu Gln Trp Pro Asp Asn Pro Lys Ala Glu Arg Leu Phe</b>	
	<div style="display: flex; justify-content: space-between; width: 80%;"> <span>130</span> <span>135</span> <span>140</span> </div>	
50	<b>TGG CAA ACT CTG AGA AAT GTG GTC TTG ACT GAA AAT GAT TCA CGG TAT</b>	<b>480</b>

# ES 2 340 210 T3

	Trp	Gln	Thr	Leu	Arg	Asn	Val	Val	Leu	Thr	Glu	Asn	Asp	Ser	Arg	Tyr		
	145					150					155					160		
5	AAC	AAT	ATG	TAT	GTC	GAT	TCC	ATT	AAG	CAA	TAC	TA	ACTG	ACGT	TA	AGTC	ATGA	533
	Asn	Asn	Met	Tyr	Val	Asp	Ser	Ile	Lys	Gln	Tyr							
					165					170								
	TTTC	GC	GCCA	TAATA	AAGAT	GCAA	AGGAAT	GACATT	TTCCG	TATTAG	TTAT	CTATT	GCTAC					593
10	GGTA	ACCAAA	TTACT	CCCAA	AAAC	CTTACG	TCGG	TTTCAA	AACA	ACCACA	TTCT	GTCT	GGC					653
	CCCAC	AGTTT	TTGAG	GGTCA	GGAG	TCCAGG	CCCAG	CATAA	CTGG	GCTTC	TGCTT	CAGGG						713
	TGTCT	CCAGA	GGCT	GCAATG	TAGGT	GTTCA	CCAG	AGACAT	AGGC	ATCACT	GGGG	TACAC						773
15	TCCAT	GTGGT	TGTTTT	CTGG	ATTCA	ATTCC	TCCT	GGGCTA	TTGG	CCAAAG	GCTATA	CTCA						833
	TGTA	AGCCAT	GCGAG	CCTAT	CCCACA	ACGG	CAGCT	TTGCTT	CATCAG	AGCT	AGCA	AAAAAAG						893
	AGAG	GTGCT	AGCA	AGATGA	AGTC	ACAATC	TTTT	TGTAATC	GAAT	CAAAAA	AGTG	ATATCT						953
20	CATCA	CTTTG	GCCAT	ATTCT	ATTTG	TTAGA	AGTA	AACCAC	AGGT	CCCACC	AGCT	CCATGG						1013
	GAGT	GACCAC	CTCAG	TCCAG	GGAAA	ACAGC	TGAAG	ACCAA	GATGG	TGAGC	TCTG	ATTGCT						1073
	TCAG	TTGGTC	ATCA	ACTATT	TTCC	CTTGAC	TGCT	GTCTG	GGAT	GGCCGG	CTAT	CTTGAT						1133
25	GGAT	AGATTG	TGAAT	ATCAG	GAGG	CCAGGG	ATCA	CTGTGG	ACCAT	CTTAG	CAGTT	GACCT						1193
	AACAC	ATCTT	CTTTT	CAATA	TCTA	AGAACT	TTTG	CCACTG	TGACT	AATGG	TCCT	AATATT						1253
30	AAGCT	GTGTG	TTATAT	TTTAT	CATAT	ATCTA	TGGC	TACATG	GTTAT	ATTAT	GCTG	TGGTTG						1313
	CGT	TCGGTTT	TATTT	ACAGT	TGCT	TTTACA	AATATT	TTGCT	GTAAC	ATTTG	ACTT	CTAAGG						1373
	TTTAG	ATGCC	ATTTA	AGAAC	TGAG	ATGGAT	AGCT	TTTAAA	GCAT	CTTTTA	CTTCT	TACCA						1433
35	TTTTT	TAAAA	GTATG	CAGCT	AAATT	CGAAG	CTTT	TGGTCT	ATATT	GTAA	TGCC	ATTGC						1493
	TGTAA	ATCTT	AAAAT	GAAATG	AATAA	AAATG	TTTC	ATTTTA	AAAAA	AAAAA	AAAAA	AAAAA						1553
	AAAA																	1557
40																		

## (2) INFORMACIÓN PARA EL SEQ ID NO:18:

### 45 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 171 aminoácidos

(B) TIPO: aminoácido

50 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:18:

55	Gln	Ser	Leu	Ser	Thr	Ser	Gln	Thr	Phe	Tyr	Asp	Ala	Tyr	Ile	Ser	Tyr	
	1				5					10				15			
60	Asp	Thr	Lys	Asp	Ala	Ser	Val	Thr	Asp	Trp	Val	Ile	Asn	Glu	Leu	Arg	
				20					25					30			

65

## ES 2 340 210 T3

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

### (2) INFORMACIÓN PARA EL SEQ ID NO:19:

- 30 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 629 pares de bases  
 (B) TIPO: ácido nucleico  
 35 (C) CUALIDAD DE LA HEBRA: sencilla  
 (D) TOPOLOGÍA: lineal  
 (ii) TIPO DE MOLÉCULA: ADNc  
 40 (ix) RASGO:  
 (A) NOMBRE/CLAVE: CDS  
 (B) LOCALIZACIÓN: 1..486  
 45 (ix) RASGO:  
 (A) NOMBRE/CLAVE: rasgo\_misc  
 (B) LOCALIZACIÓN: 144  
 50 (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 144 y 225 denominados C; pueden ser C o T"  
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 19:

55

60

65

# ES 2 340 210 T3

	AAT GAA TTG ATC CCC AAT CTA GAG AAG GAA GAT GGT TCT ATC TTG ATT	48
	Asn Glu Leu Ile Pro Asn Leu Glu Lys Glu Asp Gly Ser Ile Leu Ile	
	1 5 10 15	
5	TGC CTT TAT GAA AGC TAC TTT GAC CCT GGC AAA AGC ATT AGT GAA AAT	96
	Cys Leu Tyr Glu Ser Tyr Phe Asp Pro Gly Lys Ser Ile Ser Glu Asn	
	20 25 30	
10	ATT GTA AGC TTC ATT GAG AAA AGC TAT AAG TCC ATC TTT GTT TTG TCC	144
	Ile Val Ser Phe Ile Glu Lys Ser Tyr Lys Ser Ile Phe Val Leu Ser	
	35 40 45	
15	CCC AAC TTT GTC CAG AAT GAG TGG TGC CAT TAT GAA TTC TAC TTT GCC	192
	Pro Asn Phe Val Gln Asn Glu Trp Cys His Tyr Glu Phe Tyr Phe Ala	
	50 55 60	
20	CAC CAC AAT CTC TTC CAT GAA AAT TCT GAT CAC ATA ATT CTT ATC TTA	240
	His His Asn Leu Phe His Glu Asn Ser Asp His Ile Ile Leu Ile Leu	
	65 70 75 80	
25	CTG GAA CCC ATT CCA TTC TAT TGC ATT CCC ACC AGG TAT CAT AAA CTG	288
	Leu Glu Pro Ile Pro Phe Tyr Cys Ile Pro Thr Arg Tyr His Lys Leu	
	85 90 95	
	GAA GCT CTC CTG GAA AAA AAA GCA TAC TTG GAA TGG CCC AAG GAT AGG	336
	Glu Ala Leu Leu Glu Lys Lys Ala Tyr Leu Glu Trp Pro Lys Asp Arg	
	100 105 110	
30	CGT AAA TGT GGG CTT TTC TGG GCA AAC CTT CGA GCT GCT GTT AAT GTT	384
	Arg Lys Cys Gly Leu Phe Trp Ala Asn Leu Arg Ala Ala Val Asn Val	
	115 120 125	
35	AAT GTA TTA GCC ACC AGA GAA ATG TAT GAA CTG CAG ACA TTC ACA GAG	432
	Asn Val Leu Ala Thr Arg Glu Met Tyr Glu Leu Gln Thr Phe Thr Glu	
	130 135 140	
	TTA AAT GAA GAG TCT CGA GGT TCT ACA ATC TCT CTG ATG AGA ACA GAC	480
	Leu Asn Glu Glu Ser Arg Gly Ser Thr Ile Ser Leu Met Arg Thr Asp	
	145 150 155 160	
40	TGT CTA TAAAATCCCA CAGTCCTTGG GAAGTTGGGG ACCACATACA CTGTTGGGAT	536
	Cys Leu	
45	GTACATTGAT ACAACCTTTA TGATGGCAAT TTGACAATAT TTATTAAAAT AAAAAATGGT	596
	TATTCCTTC AAAAAAAAAA AAAAAAAAAA AAA	629

50 (2) INFORMACIÓN PARA EL SEQ ID NO:20:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

55 (A) LONGITUD: 162 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

60 (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:20:

65

## ES 2 340 210 T3

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

### (2) INFORMACIÓN PARA EL SEQ ID NO:21:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 427 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..426

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:21:

# ES 2 340 210 T3

	AAG AAC TCC AAA GAA AAC CTC CAG TTT CAT GCT TTT ATT TCA TAT AGT	48
	Lys Asn Ser Lys Glu Asn Leu Gln Phe His Ala Phe Ile Ser Tyr Ser	
	1 5 10 15	
5	GAA CAT GAT TCT GCC TGG GTG AAA AGT GAA TTG GTA CCT TAC CTA GAA	96
	Glu His Asp Ser Ala Trp Val Lys Ser Glu Leu Val Pro Tyr Leu Glu	
	20 25 30	
10	AAA GAA GAT ATA CAG ATT TGT CTT CAT GAG AGA AAC TTT GTC CCT GGC	144
	Lys Glu Asp Ile Gln Ile Cys Leu His Glu Arg Asn Phe Val Pro Gly	
	35 40 45	
15	AAG AGC ATT GTG GAA AAT ATC ATC AAC TGC ATT GAG AAG AGT TAC AAG	192
	Lys Ser Ile Val Glu Asn Ile Ile Asn Cys Ile Glu Lys Ser Tyr Lys	
	50 55 60	
20	TCC ATC TTT GTT TTG TCT CCC AAC TTT GTC CAG AGT GAG TGG TGC CAT	240
	Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Ser Glu Trp Cys His	
	65 70 75 80	
25	TAC GAA CTC TAT TTT GCC CAT CAC AAT CTC TTT CAT GAA GGA TCT AAT	288
	Tyr Glu Leu Tyr Phe Ala His His Asn Leu Phe His Glu Gly Ser Asn	
	85 90 95	
30	AAC TTA ATC CTC ATC TTA CTG GAA CCC ATT CCA CAG AAC AGC ATT CCC	336
	Asn Leu Ile Leu Ile Leu Leu Glu Pro Ile Pro Gln Asn Ser Ile Pro	
	100 105 110	
35	AAC AAG TAC CAC AAG CTG AAG GCT CTC ATG ACG CAG CGG ACT TAT TTG	384
	Asn Lys Tyr His Lys Leu Lys Ala Leu Met Thr Gln Arg Thr Tyr Leu	
	115 120 125	
40	CAG TGG CCC AAG GAG AAA AGC AAA CGT GGG CTC TTT TGG GCT	426
	Gln Trp Pro Lys Glu Lys Ser Lys Arg Gly Leu Phe Trp Ala	
	130 135 140	
45	A	427

40 (2) INFORMACIÓN PARA EL SEQ ID NO:22:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 142 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:22:

## ES 2 340 210 T3

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

### (2) INFORMACIÓN PARA EL SEQ ID NO:23:

#### 30 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 662 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- 35 (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### 40 (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..627

#### 45 (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 54
- (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 54, 103, y 345 son denominados A; cada uno puede ser A o G"

50

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 313
- (D) OTRA INFORMACIÓN: /nota= "el nucleótido 313 denominado G, puede ser G o T"

55

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 316
- (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 316, 380, 407, y 408 denominados C; pueden ser cada uno A, C, G, o T"

60

#### 65 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:23:



## ES 2 340 210 T3

	GCT TCC ACC TGT GCC TGG CCT GGC TTC CCT GGC GGG GGC GGC AAA GTG	48
	Ala Ser Thr Cys Ala Trp Pro Gly Phe Pro Gly Gly Gly Gly Lys Val	
	1 5 10 15	
5	GGC GAA ATG AGG ATG CCC TGC CCT ACG ATG CCT TCG TGG TCT TCG ACA	96
	Gly Glu Met Arg Met Pro Cys Pro Thr Met Pro Ser Trp Ser Ser Thr	
	20 25 30	
10	AAA CGC AGA GCG CAG TGG CAG ACT GGG TGT ACA ACG AGC TTC GGG GGC	144
	Lys Arg Arg Ala Gln Trp Gln Thr Gly Cys Thr Thr Ser Phe Gly Gly	
	35 40 45	
15	AGC TGG AGG AGT GCC GTG GGC GCT GGG CAC TCC GCC TGT GCC TGG AGG	192
	Ser Trp Arg Ser Ala Val Gly Ala Gly His Ser Ala Cys Ala Trp Arg	
	50 55 60	
20	AAC GCG ACT GGC TGC CTG GCA AAA CCC TCT TTG AGA ACC TGT GGG CCT	240
	Asn Ala Thr Gly Cys Leu Ala Lys Pro Ser Leu Arg Thr Cys Gly Pro	
	65 70 75 80	
25	CGG TCT ATG GCA GCC GCA AGA CGC TGT TTG TGC TGG CCC ACA CGG ACC	288
	Arg Ser Met Ala Ala Ala Arg Arg Cys Leu Cys Trp Pro Thr Arg Thr	
	85 90 95	
30	GGG TCA GTG GTC TCT TGC GCG CCA GTT CTC CTG CTG GCC CAG CAG CGC	336
	Gly Ser Val Val Ser Cys Ala Pro Val Leu Leu Leu Ala Gln Gln Arg	
	100 105 110	
35	CTG CTG GAA GAC CGC AAG GAC GTC GTG GTG CTG GTG ATC CTA ACG CCT	384
	Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu Val Ile Leu Thr Pro	
	115 120 125	
40	GAC GGC CAA GCC TCC CGA CTA CCC GAT GCG CTG ACC AGC GCC TCT GCC	432
	Asp Gly Gln Ala Ser Arg Leu Pro Asp Ala Leu Thr Ser Ala Ser Ala	
	130 135 140	
45	GCC AGA GTG TCC TCC TCT GGC CCC ACC AGC CCA GTG GTC GCG CAG CTT	480
	Ala Arg Val Ser Ser Ser Gly Pro Thr Ser Pro Val Val Ala Gln Leu	
	145 150 155 160	
50	CTG AGG CCA GCA TGC ATG GCC CTG ACC AGG GAC AAC CAC CAC TTC TAT	528
	Leu Arg Pro Ala Cys Met Ala Leu Thr Arg Asp Asn His His Phe Tyr	
	165 170 175	
55	AAC CGG AAC TTC TGC CAG GGA ACC CAC GGC CGA ATA GCC GTG AGC CGG	576
	Asn Arg Asn Phe Cys Gln Gly Thr His Gly Arg Ile Ala Val Ser Arg	
	180 185 190	
60	AAT CCT GCA CGG TGC CAC CTC CAC ACA CAC CTA ACA TAT GCC TGC CTG	624
	Asn Pro Ala Arg Cys His Leu His Thr His Leu Thr Tyr Ala Cys Leu	
	195 200 205	
65	ATC TGACCAACAC ATGCTCGCCA CCCTCACCAC ACACC	662
	Ile	

### (2) INFORMACIÓN PARA EL SEQ ID NO:24:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 209 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:24:

## ES 2 340 210 T3

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

### (2) INFORMACIÓN PARA EL SEQ ID NO:25:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 4865 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 107..2617

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: mat\_péptido
- (B) LOCALIZACIÓN: 173..2617

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 81

(D) OTRA INFORMACIÓN: /nota= "los nucleótidos 81, 3144, 3205, y 3563 denominados A, pueden ser cada uno A, C, G, o T"

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 84

## ES 2 340 210 T3

(D) OTRA INFORMACIÓN: /nota= “el nucleótido 84 denominado C, puede ser C o G”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 739

(D) OTRA INFORMACIÓN: /nota= “el nucleótido 739 denominado C, puede ser C o T”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 3132

(D) OTRA INFORMACIÓN: /nota= “los nucleótidos 3132, 3532, 3538, y 3553 denominados G, pueden ser cada uno G o T”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 3638

(D) OTRA INFORMACIÓN: /nota= “el nucleótido 3638 denominado A, puede ser A o T”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 3677

(D) OTRA INFORMACIÓN: /nota= “los nucleótidos 3677, 3685, y 3736 denominados C, pueden ser cada uno A o C”

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:25:

AAAATACTCC	CTTGCTCAA	AAACTGCTCG	GTCAAACGGT	GATAGCAAAC	CACGCATTCA	60
CAGGGCCACT	GCTGCTCACA	AAACCAGTGA	GGATGATGCC	AGGATG	ATG TCT GCC	115
					Met Ser Ala	
					-22 -20	
TCG CGC CTG GCT GGG ACT CTG ATC CCA GCC ATG GCC TTC CTC TCC TGC						163
Ser Arg Leu Ala Gly Thr Leu Ile Pro Ala Met Ala Phe Leu Ser Cys						
					-15 -10 -5	
GTG AGA CCA GAA AGC TGG GAG CCC TGC GTG GAG GTT CCT AAT ATT ACT						211
Val Arg Pro Glu Ser Trp Glu Pro Cys Val Glu Val Pro Asn Ile Thr						
					1 5 10	
TAT CAA TGC ATG GAG CTG AAT TTC TAC AAA ATC CCC GAC AAC CTC CCC						259
Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys Ile Pro Asp Asn Leu Pro						
					15 20 25	
TTC TCA ACC AAG AAC CTG GAC CTG AGC TTT AAT CCC CTG AGG CAT TTA						307
Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe Asn Pro Leu Arg His Leu						
					30 35 40 45	
GGC AGC TAT AGC TTC TTC AGT TTC CCA GAA CTG CAG GTG CTG GAT TTA						355
Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu Leu Gln Val Leu Asp Leu						
					50 55 60	
TCC AGG TGT GAA ATC CAG ACA ATT GAA GAT GGG GCA TAT CAG AGC CTA						403
Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp Gly Ala Tyr Gln Ser Leu						
					65 70 75	
AGC CAC CTC TCT ACC TTA ATA TTG ACA GGA AAC CCC ATC CAG AGT TTA						451
Ser His Leu Ser Thr Leu Ile Leu Thr Gly Asn Pro Ile Gln Ser Leu						
					80 85 90	
GCC CTG GGA GCC TTT TCT GGA CTA TCA AGT TTA CAG AAG CTG GTG GCT						499
Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser Leu Gln Lys Leu Val Ala						
					95 100 105	
GTG GAG ACA AAT CTA GCA TCT CTA GAG AAC TTC CCC ATT GGA CAT CTC						547
Val Glu Thr Asn Leu Ala Ser Leu Glu Asn Phe Pro Ile Gly His Leu						
					110 115 120 125	
AAA ACT TTG AAA GAA CTT AAT GTG GCT CAC AAT CTT ATC CAA TCT TTC						595
Lys Thr Leu Lys Glu Leu Asn Val Ala His Asn Leu Ile Gln Ser Phe						
					130 135 140	

# ES 2 340 210 T3

5	AAA TTA CCT GAG TAT TTT TCT AAT CTG ACC AAT CTA GAG CAC TTG GAC Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu His Leu Asp 145 150 155	643
10	CTT TCC AGC AAC AAG ATT CAA AGT ATT TAT TGC ACA GAC TTG CGG GTT Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr Cys Thr Asp Leu Arg Val 160 165 170	691
15	CTA CAT CAA ATG CCC CTA CTC AAT CTC TCT TTA GAC CTG TCC CTG AAC Leu His Gln Met Pro Leu Leu Asn Leu Ser Leu Asp Leu Ser Leu Asn 175 180 185	739
20	CCT ATG AAC TTT ATC CAA CCA GGT GCA TTT AAA GAA ATT AGG CTT CAT Pro Met Asn Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile Arg Leu His 190 195 200 205	787
25	AAG CTG ACT TTA AGA AAT AAT TTT GAT AGT TTA AAT GTA ATG AAA ACT Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser Leu Asn Val Met Lys Thr 210 215 220	835
30	TGT ATT CAA GGT CTG GCT GGT TTA GAA GTC CAT CGT TTG GTT CTG GGA Cys Ile Gln Gly Leu Ala Gly Leu Glu Val His Arg Leu Val Leu Gly 225 230 235	883
35	GAA TTT AGA AAT GAA GGA AAC TTG GAA AAG TTT GAC AAA TCT GCT CTA Glu Phe Arg Asn Glu Gly Asn Leu Glu Lys Phe Asp Lys Ser Ala Leu 240 245 250	931
40	GAG GGC CTG TGC AAT TTG ACC ATT GAA GAA TTC CGA TTA GCA TAC TTA Glu Gly Leu Cys Asn Leu Thr Ile Glu Glu Phe Arg Leu Ala Tyr Leu 255 260 265	979
45	GAC TAC TAC CTC GAT GAT ATT ATT GAC TTA TTT AAT TGT TTG ACA AAT Asp Tyr Tyr Leu Asp Asp Ile Ile Asp Leu Phe Asn Cys Leu Thr Asn 270 275 280 285	1027
50	GTT TCT TCA TTT TCC CTG GTG AGT GTG ACT ATT GAA AGG GTA AAA GAC Val Ser Ser Phe Ser Leu Val Ser Val Thr Ile Glu Arg Val Lys Asp 290 295 300	1075
55	TTT TCT TAT AAT TTC GGA TGG CAA CAT TTA GAA TTA GTT AAC TGT AAA Phe Ser Tyr Asn Phe Gly Trp Gln His Leu Glu Leu Val Asn Cys Lys 305 310 315	1123
60	TTT GGA CAG TTT CCC ACA TTG AAA CTC AAA TCT CTC AAA AGG CTT ACT Phe Gly Gln Phe Pro Thr Leu Lys Leu Lys Ser Leu Lys Arg Leu Thr 320 325 330	1171
65	TTC ACT TCC AAC AAA GGT GGG AAT GCT TTT TCA GAA GTT GAT CTA CCA Phe Thr Ser Asn Lys Gly Gly Asn Ala Phe Ser Glu Val Asp Leu Pro 335 340 345	1219
70	AGC CTT GAG TTT CTA GAT CTC AGT AGA AAT GGC TTG AGT TTC AAA GGT Ser Leu Glu Phe Leu Asp Leu Ser Arg Asn Gly Leu Ser Phe Lys Gly 350 355 360 365	1267
75	TGC TGT TCT CAA AGT GAT TTT GGG ACA ACC AGC CTA AAG TAT TTA GAT Cys Cys Ser Gln Ser Asp Phe Gly Thr Thr Ser Leu Lys Tyr Leu Asp 370 375 380	1315

# ES 2 340 210 T3

5	CTG AGC TTC AAT GGT GTT ATT ACC ATG AGT TCA AAC TTC TTG GGC TTA Leu Ser Phe Asn Gly Val Ile Thr Met Ser Ser Asn Phe Leu Gly Leu 385 390 395	1363
10	GAA CAA CTA GAA CAT CTG GAT TTC CAG CAT TCC AAT TTG AAA CAA ATG Glu Gln Leu Glu His Leu Asp Phe Gln His Ser Asn Leu Lys Gln Met 400 405 410	1411
15	AGT GAG TTT TCA GTA TTC CTA TCA CTC AGA AAC CTC ATT TAC CTT GAC Ser Glu Phe Ser Val Phe Leu Ser Leu Arg Asn Leu Ile Tyr Leu Asp 415 420 425	1459
20	ATT TCT CAT ACT CAC ACC AGA GTT GCT TTC AAT GGC ATC TTC AAT GGC Ile Ser His Thr His Thr Arg Val Ala Phe Asn Gly Ile Phe Asn Gly 430 435 440 445	1507
25	TTG TCC AGT CTC GAA GTC TTG AAA ATG GCT GGC AAT TCT TTC CAG GAA Leu Ser Ser Leu Glu Val Leu Lys Met Ala Gly Asn Ser Phe Gln Glu 450 455 460	1555
30	AAC TTC CTT CCA GAT ATC TTC ACA GAG CTG AGA AAC TTG ACC TTC CTG Asn Phe Leu Pro Asp Ile Phe Thr Glu Leu Arg Asn Leu Thr Phe Leu 465 470 475	1603
35	GAC CTC TCT CAG TGT CAA CTG GAG CAG TTG TCT CCA ACA GCA TTT AAC Asp Leu Ser Gln Cys Gln Leu Glu Gln Leu Ser Pro Thr Ala Phe Asn 480 485 490	1651
40	TCA CTC TCC AGT CTT CAG GTA CTA AAT ATG AGC CAC AAC AAC TTC TTT Ser Leu Ser Ser Leu Gln Val Leu Asn Met Ser His Asn Asn Phe Phe 495 500 505	1699
45	TCA TTG GAT ACG TTT CCT TAT AAG TGT CTG AAC TCC CTC CAG GTT CTT Ser Leu Asp Thr Phe Pro Tyr Lys Cys Leu Asn Ser Leu Gln Val Leu 510 515 520 525	1747
50	GAT TAC AGT CTC AAT CAC ATA ATG ACT TCC AAA AAA CAG GAA CTA CAG Asp Tyr Ser Leu Asn His Ile Met Thr Ser Lys Lys Gln Glu Leu Gln 530 535 540	1795
55	CAT TTT CCA AGT AGT CTA GCT TTC TTA AAT CTT ACT CAG AAT GAC TTT His Phe Pro Ser Ser Leu Ala Phe Leu Asn Leu Thr Gln Asn Asp Phe 545 550 555	1843
60	GCT TGT ACT TGT GAA CAC CAG AGT TTC CTG CAA TGG ATC AAG GAC CAG Ala Cys Thr Cys Glu His Gln Ser Phe Leu Gln Trp Ile Lys Asp Gln 560 565 570	1891
65	AGG CAG CTC TTG GTG GAA GTT GAA CGA ATG GAA TGT GCA ACA CCT TCA Arg Gln Leu Leu Val Glu Val Glu Arg Met Glu Cys Ala Thr Pro Ser 575 580 585	1939
70	GAT AAG CAG GGC ATG CCT GTG CTG AGT TTG AAT ATC ACC TGT CAG ATG Asp Lys Gln Gly Met Pro Val Leu Ser Leu Asn Ile Thr Cys Gln Met 590 595 600 605	1987
75	AAT AAG ACC ATC ATT GGT GTG TCG GTC CTC AGT GTG CTT GTA GTA TCT Asn Lys Thr Ile Ile Gly Val Ser Val Leu Ser Val Leu Val Val Ser 610 615 620	2035
80	GTT GTA GCA GTT CTG GTC TAT AAG TTC TAT TTT CAC CTG ATG CTT CTT	2083

# ES 2 340 210 T3

	Val	Val	Ala	Val	Leu	Val	Tyr	Lys	Phe	Tyr	Phe	His	Leu	Met	Leu	Leu	
				625					630					635			
5	GCT	GGC	TGC	ATA	AAG	TAT	GGT	AGA	GGT	GAA	AAC	ATC	TAT	GAT	GCC	TTT	2131
	Ala	Gly	Cys	Ile	Lys	Tyr	Gly	Arg	Gly	Glu	Asn	Ile	Tyr	Asp	Ala	Phe	
			640					645					650				
	GTT	ATC	TAC	TCA	AGC	CAG	GAT	GAG	GAC	TGG	GTA	AGG	AAT	GAG	CTA	GTA	2179
10	Val	Ile	Tyr	Ser	Ser	Gln	Asp	Glu	Asp	Trp	Val	Arg	Asn	Glu	Leu	Val	
		655				660						665					
	AAG	AAT	TTA	GAA	GAA	GGG	GTG	CCT	CCA	TTT	CAG	CTC	TGC	CTT	CAC	TAC	2227
	Lys	Asn	Leu	Glu	Glu	Gly	Val	Pro	Pro	Phe	Gln	Leu	Cys	Leu	His	Tyr	
	670					675					680					685	
15	AGA	GAC	TTT	ATT	CCC	GGT	GTG	GCC	ATT	GCT	GCC	AAC	ATC	ATC	CAT	GAA	2275
	Arg	Asp	Phe	Ile		Gly	Val	Ala	Ile	Ala	Ala	Asn	Ile	Ile	His	Glu	
					690					695					700		
	GGT	TTC	CAT	AAA	AGC	CGA	AAG	GTG	ATT	GTT	GTG	GTG	TCC	CAG	CAC	TTC	2323
20	Gly	Phe	His		Lys	Ser	Arg	Lys	Val	Ile	Val	Val	Val	Ser	Gln	His	
				705					710						715		
	ATC	CAG	AGC	CGC	TGG	TGT	ATC	TTT	GAA	TAT	GAG	ATT	GCT	CAG	ACC	TGG	2371
	Ile	Gln	Ser	Arg	Trp	Cys	Ile	Phe	Glu	Tyr	Glu	Ile	Ala	Gln	Thr	Trp	
25			720					725					730				
	CAG	TTT	CTG	AGC	AGT	CGT	GCT	GGT	ATC	ATC	TTC	ATT	GTC	CTG	CAG	AAG	2419
	Gln	Phe	Leu	Ser	Ser	Arg	Ala	Gly	Ile	Ile	Phe	Ile	Val	Leu	Gln	Lys	
		735				740						745					
30	GTG	GAG	AAG	ACC	CTG	CTC	AGG	CAG	CAG	GTG	GAG	CTG	TAC	CGC	CTT	CTC	2467
	Val	Glu	Lys	Thr	Leu	Arg	Gln	Gln	Val	Glu	Leu	Tyr	Arg	Leu	Leu		
	750					755				760					765		
	AGC	AGG	AAC	ACT	TAC	CTG	GAG	TGG	GAG	GAC	AGT	GTC	CTG	GGG	CGG	CAC	2515
35	Ser	Arg	Asn	Thr	Tyr	Leu	Glu	Trp	Glu	Asp	Ser	Val	Leu	Gly	Arg	His	
				770					775						780		
	ATC	TTC	TGG	AGA	CGA	CTC	AGA	AAA	GCC	CTG	CTG	GAT	GGT	AAA	TCA	TGG	2563
	Ile	Phe	Trp	Arg	Arg	Leu	Arg	Lys	Ala	Leu	Leu	Asp	Gly	Lys	Ser	Trp	
				785				790						795			
40	AAT	CCA	GAA	GGA	ACA	GTG	GGT	ACA	GGA	TGC	AAT	TGG	CAG	GAA	GCA	ACA	2611
	Asn	Pro	Glu	Gly	Thr	Val	Gly	Thr	Gly	Cys	Asn	Trp	Gln	Glu	Ala	Thr	
		800					805						810				
45	TCT	ATC	TGA	AAGAGGAA	AAATAAAAAC	CTCCTGAGGC	ATTCTTGCC	CAGCTGGGTC									2667
	Ser	Ile															
		815															
	CAACACTTGT	TCAGTTAATA	AGTATTAAAT	GCTGCCACAT	GTCAGGCCTT	ATGCTAAGGG											2727
	TCAGTAATTC	CATGGTGCAC	TAGATATGCA	GGGCTGCTAA	TCTCAAGGAG	CTTCCAGTGC											2787
50	AGAGGGAATA	AATGCTAGAC	TAAAATACAG	AGTCTTCCAG	GTGGGCATTT	CAACCAACTC											2847
	AGTCAAGGAA	CCCATGACAA	AGAAAGTCAT	TTCAACTCTT	ACCTCATCAA	GTTGAATAAA											2907
55	GACAGAGAAA	ACAGAAAGAG	ACATTGTTCT	TTTCTGAGT	CTTTTGAATG	GAAATTGTAT											2967

# ES 2 340 210 T3

	TATGTTATAG	CCATCATAAA	ACCATTTTGG	TAGTTTGTGAC	TGAACTGGGT	GTTCACTTTT	3027
5	TCCTTTTGA	TTGAATACAA	TTTAAATTCT	ACTTGATGAC	TGCAGTCGTC	AAGGGGCTCC	3087
	TGATGCAAGA	TGCCCCTTCC	ATTTTAAGTC	TGTCTCCTTA	CAGAGGTTAA	AGTCTAATGG	3147
	CTAATTCCTA	AGGAAACCTG	ATTAACACAT	GCTCACAACC	ATCCTGGTCA	TTCTCGAACA	3207
10	TGTTCTATTT	TTTAACTAAT	CACCCCTGAT	ATATTTTAT	TTTATATAT	CCAGTTTCA	3267
	TTTTTTTACG	TCTTGCCTAT	AAGCTAATAT	CATAAATAAG	GTTGTTTAAG	ACGTGCTTCA	3327
	AATATCCATA	TTAACCACTA	TTTTTCAAGG	AAGTATGGAA	AAGTACACTC	TGTCACCTTG	3387
15	TCACTCGATG	TCATTCCAAA	GTTATTGCCT	ACTAAGTAAT	GACTGTGATG	AAAGCAGCAT	3447
	TGAAATAATT	TGTTTAAAGG	GGGCACTCTT	TTAAACGGGA	AGAAAATTTT	CGCTTCCTGG	3507
	TCTTATCATG	GACAATTGCG	GCTAGAGGCA	GGAAGGAAGT	GGGATGACCT	CAGGAAGTCA	3567
20	CCTTTTCTTG	ATTCCAGAAA	CATATGGGCT	GATAAACCCG	GGGTGACCTC	ATGAAATGAG	3627
	TTGCAGCAGA	AGTTTATTTT	TTTCAGAACA	AGTGATGTTT	GATGGACCTC	TGAATCTCTT	3687
25	TAGGGAGACA	CAGATGGCTG	GGATCCCTCC	CCTGTACCCT	TCTCACTGCC	AGGAGAACTA	3747
	CGTGTGAAGG	TATTCAAGGC	AGGGAGTATA	CATTGCTGTT	TCCTGTTGGG	CAATGCTCCT	3807
	TGACCACATT	TTGGGAAGAG	TGGATGTTAT	CATTGAGAAA	ACAATGTGTC	TGGAATTAAT	3867
30	GGGGTTCTTA	TAAAGAAGGT	TCCCAGAAAA	GAATGTTTAT	TCCAGCTTCT	TCAGGAACA	3927
	GGAACATTCA	AGGAAAAGGA	CAATCAGGAT	GTCATCAGGG	AAATGAAAAT	AAAAACCACA	3987
	ATGAGATATC	ACCTTATACC	AGGTAGATGG	CTACTATAAA	AAAATGAAGT	GTCATCAAGG	4047
35	ATATAGAGAA	ATTGGAACCC	TTCTTCACTG	CTGGAGGGAA	TGGAAAATGG	TGTAGCCGTT	4107
	ATGAAAAACA	GTACGGAGGT	TTCTCAAAAA	TTAAAAATAG	AACTGCTATA	TGATCCAGCA	4167
	ATCTCACTTC	TGTATATATA	CCCAAAATAA	TTGAAATCAG	AATTTCAAGA	AAATATTTAC	4227
40	ACTCCCATGT	TCATTGTGGC	ACTCTTCACA	ATCACTGTTT	CCAAAGTTAT	GGAAACAACC	4287
	CAAATTTCCA	TTGGAAAATA	AATGGACAAA	GGAAATGTGC	ATATAACGTA	CAATGGGGAT	4347
	ATTATTCAGC	CTAAAAAAG	GGGGGATCCT	GTTATTTATG	ACAACATGAA	TAAACCCGGA	4407
45	GGCCATTATG	CTATGTAAAA	TGAGCAAGTA	ACAGAAAGAC	AAATACTGCC	TGATTTTCATT	4467
	TATATGAGGT	TCTAAAATAG	TCAAACTCAT	AGAAGCAGAG	AATAGAACAG	TGGTTCCTAG	4527
	GGAAAAGGAG	GAAGGGAGAA	ATGAGGAAAT	AGGGAGTTGT	CTAATTGGTA	TAAAATTATA	4587
50	GTATGCAAGA	TGAATTAGCT	CTAAAGATCA	GCTGTATAGC	AGAGTTCGTA	TAATGAACAA	4647
	TACTGTATTA	TGCACTTAAC	ATTTTGTAA	GAGGGTACCT	CTCATGTTAA	GTGTTCTTAC	4707
	CATATACATA	TACACAAGGA	AGCTTTTGGG	GGTGATGGAT	ATATTTATTA	CCTTGATGTG	4767
55	GGTGATGGTT	TGACAGGTAT	GTGACTATGT	CTAAACTCAT	CAAATGTGAT	ACATTAAATA	4827
	TATGCAGTTT	TATAATATCA	AAAAAAAAAA	AAAAAAAAAA			4865

## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:26.

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 837 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:26:

```

Met Ser Ala Ser Arg Leu Ala Gly Thr Leu Ile Pro Ala Met Ala Phe
-22      -20              -15              -10

Leu Ser Cys Val Arg Pro Glu Ser Trp Glu Pro Cys Val Glu Val Pro
-5              1              5              10

Asn Ile Thr Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys Ile Pro Asp
15              20              25

Asn Leu Pro Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe Asn Pro Leu
30              35              40

Arg His Leu Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu Leu Gln Val
45              50              55

Leu Asp Leu Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp Gly Ala Tyr
60              65              70

Gln Ser Leu Ser His Leu Ser Thr Leu Ile Leu Thr Gly Asn Pro Ile
75              80              85              90

Gln Ser Leu Ala Leu Gly Ala Phe Ser Gly Leu Ser Ser Leu Gln Lys
95              100              105

Leu Val Ala Val Glu Thr Asn Leu Ala Ser Leu Glu Asn Phe Pro Ile
110              115              120

Gly His Leu Lys Thr Leu Lys Glu Leu Asn Val Ala His Asn Leu Ile
125              130              135

Gln Ser Phe Lys Leu Pro Glu Tyr Phe Ser Asn Leu Thr Asn Leu Glu
140              145              150

His Leu Asp Leu Ser Ser Asn Lys Ile Gln Ser Ile Tyr Cys Thr Asp
155              160              165              170

Leu Arg Val Leu His Gln Met Pro Leu Leu Asn Leu Ser Leu Asp Leu
175              180              185

Ser Leu Asn Pro Met Asn Phe Ile Gln Pro Gly Ala Phe Lys Glu Ile
190              195              200

Arg Leu His Lys Leu Thr Leu Arg Asn Asn Phe Asp Ser Leu Asn Val
205              210              215

Met Lys Thr Cys Ile Gln Gly Leu Ala Gly Leu Glu Val His Arg Leu

```



# ES 2 340 210 T3

	220					225						230					
5	Val	Leu	Gly	Glu	Phe	Arg	Asn	Glu	Gly	Asn	Leu	Glu	Lys	Phe	Asp	Lys	
	235					240					245					250	
	Ser	Ala	Leu	Glu	Gly	Leu	Cys	Asn	Leu	Thr	Ile	Glu	Glu	Phe	Arg	Leu	
					255					260					265		
10	Ala	Tyr	Leu	Asp	Tyr	Tyr	Leu	Asp	Asp	Ile	Ile	Asp	Leu	Phe	Asn	Cys	
				270					275					280			
	Leu	Thr	Asn	Val	Ser	Ser	Phe	Ser	Leu	Val	Ser	Val	Thr	Ile	Glu	Arg	
			285					290					295				
15	Val	Lys	Asp	Phe	Ser	Tyr	Asn	Phe	Gly	Trp	Gln	His	Leu	Glu	Leu	Val	
	300						305					310					
	Asn	Cys	Lys	Phe	Gly	Gln	Phe	Pro	Thr	Leu	Lys	Leu	Lys	Ser	Leu	Lys	
	315					320					325					330	
20	Arg	Leu	Thr	Phe	Thr	Ser	Asn	Lys	Gly	Gly	Asn	Ala	Phe	Ser	Glu	Val	
					335					340					345		
	Asp	Leu	Pro	Ser	Leu	Glu	Phe	Leu	Asp	Leu	Ser	Arg	Asn	Gly	Leu	Ser	
				350					355					360			
25	Phe	Lys	Gly	Cys	Cys	Ser	Gln	Ser	Asp	Phe	Gly	Thr	Thr	Ser	Leu	Lys	
			365					370					375				
	Tyr	Leu	Asp	Leu	Ser	Phe	Asn	Gly	Val	Ile	Thr	Met	Ser	Ser	Asn	Phe	
	380						385					390					
30	Leu	Gly	Leu	Glu	Gln	Leu	Glu	His	Leu	Asp	Phe	Gln	His	Ser	Asn	Leu	
	395					400					405					410	
	Lys	Gln	Met	Ser	Glu	Phe	Ser	Val	Phe	Leu	Ser	Leu	Arg	Asn	Leu	Ile	
					415					420					425		
35	Tyr	Leu	Asp	Ile	Ser	His	Thr	His	Thr	Arg	Val	Ala	Phe	Asn	Gly	Ile	
			430						435					440			
	Phe	Asn	Gly	Leu	Ser	Ser	Leu	Glu	Val	Leu	Lys	Met	Ala	Gly	Asn	Ser	
			445					450					455				
40	Phe	Gln	Glu	Asn	Phe	Leu	Pro	Asp	Ile	Phe	Thr	Glu	Leu	Arg	Asn	Leu	
		460					465					470					
	Thr	Phe	Leu	Asp	Leu	Ser	Gln	Cys	Gln	Leu	Glu	Gln	Leu	Ser	Pro	Thr	
	475					480					485					490	
45	Ala	Phe	Asn	Ser	Leu	Ser	Ser	Leu	Gln	Val	Leu	Asn	Met	Ser	His	Asn	
					495					500					505		
	Asn	Phe	Phe	Ser	Leu	Asp	Thr	Phe	Pro	Tyr	Lys	Cys	Leu	Asn	Ser	Leu	
				510					515					520			
50	Gln	Val	Leu	Asp	Tyr	Ser	Leu	Asn	His	Ile	Met	Thr	Ser	Lys	Lys	Gln	
			525					530					535				
	Glu	Leu	Gln	His	Phe	Pro	Ser	Ser	Leu	Ala	Phe	Leu	Asn	Leu	Thr	Gln	
55		540					545					550					

## ES 2 340 210 T3

	Asn 555	Asp	Phe	Ala	Cys	Thr 560	Cys	Glu	His	Gln	Ser 565	Phe	Leu	Gln	Trp	Ile 570
5	Lys	Asp	Gln	Arg	Gln 575	Leu	Leu	Val	Glu	Val 580	Glu	Arg	Met	Glu	Cys 585	Ala
	Thr	Pro	Ser	Asp 590	Lys	Gln	Gly	Met	Pro 595	Val	Leu	Ser	Leu	Asn 600	Ile	Thr
10	Cys	Gln	Met	Asn 605	Lys	Thr	Ile	Ile 610	Gly	Val	Ser	Val	Leu	Ser 615	Val	Leu
	Val	Val	Ser	Val	Val	Ala	Val 620	Leu	Val	Tyr	Lys	Phe 630	Tyr	Phe	His	Leu
15	Met	Leu	Leu	Ala	Gly 640	Cys	Ile	Lys	Tyr	Gly	Arg 645	Gly	Glu	Asn	Ile	Tyr 650
	Asp	Ala	Phe	Val	Ile 655	Tyr	Ser	Ser	Gln	Asp 660	Glu	Asp	Trp	Val	Arg 665	Asn
20	Glu	Leu	Val	Lys 670	Asn	Leu	Glu	Glu	Gly 675	Val	Pro	Pro	Phe	Gln 680	Leu	Cys
	Leu	His	Tyr	Arg 685	Asp	Phe	Ile	Pro 690	Gly	Val	Ala	Ile	Ala 695	Ala	Asn	Ile
25	Ile	His	Glu	Gly	Phe	His	Lys 700	Ser	Arg	Lys 705	Val	Ile 710	Val	Val	Val	Ser
	Gln	His	Phe	Ile	Gln	Ser	Arg 720	Trp	Cys	Ile	Phe 725	Glu	Tyr	Glu	Ile	Ala 730
30	Gln	Thr	Trp	Gln	Phe 735	Leu	Ser	Ser	Arg	Ala 740	Gly	Ile	Ile	Phe	Ile 745	Val
	Leu	Gln	Lys	Val	Glu	Lys	Thr 750	Leu	Leu	Arg 755	Gln	Gln	Val	Glu	Leu	Tyr 760
35	Arg	Leu	Leu	Ser	Arg	Asn	Thr 765	Tyr 770	Leu	Glu	Trp	Glu	Asp 775	Ser	Val	Leu
	Gly	Arg	His	Ile	Phe	Trp	Arg 780	Arg	Leu	Arg	Lys 785	Ala 790	Leu	Leu	Asp	Gly
40	Lys	Ser	Trp	Asn	Pro	Glu	Gly 800	Thr	Val	Gly	Thr 805	Gly	Cys	Asn	Trp	Gln 810
45	Glu	Ala	Thr	Ser	Ile		815									

### (2) INFORMACIÓN PARA EL SEQ ID NO:27:

- |    |                                      |  |
|----|--------------------------------------|--|
| 50 | (i) CARACTERÍSTICAS DE LA SECUENCIA: |  |
|    | (A) LONGITUD: 300 pares de bases     |  |
|    | (B) TIPO: ácido nucleico             |  |
| 55 | (C) CUALIDAD DE LA HEBRA: sencilla   |  |
|    | (D) TOPOLOGÍA: lineal                |  |
|    | (ii) TIPO DE MOLÉCULA: ADNc          |  |
| 60 | (ix) RASGO:                          |  |
|    | (A) NOMBRE/CLAVE: CDS                |  |
|    | (B) LOCALIZACIÓN: 1..300             |  |
| 65 | (ix) RASGO:                          |  |
|    | (A) NOMBRE/CLAVE: rasgo_misc         |  |

## ES 2 340 210 T3

(B) LOCALIZACIÓN: 186

(D) OTRA INFORMACIÓN: /nota= "los nucleótidos 186, 196, 217, 276, y 300 denominados C, pueden ser cada uno A, C, G, o T"

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:27:

TCC TAT TCT ATG GAA AAA GAT GCT TTC CTA TTT ATG AGA AAT TTG AAG	48
Ser Tyr Ser Met Glu Lys Asp Ala Phe Leu Phe Met Arg Asn Leu Lys	
1 5 10 15	
GTT CTC TCA CTA AAA GAT AAC AAT GTC ACA GCT GTC CCC ACC ACT TTG	96
Val Leu Ser Leu Lys Asp Asn Asn Val Thr Ala Val Pro Thr Thr Leu	
20 25 30	
CCA CCT AAT TTA CTA GAG CTC TAT CTT TAT AAC AAT ATC ATT AAG AAA	144
Pro Pro Asn Leu Leu Glu Leu Tyr Leu Tyr Asn Asn Ile Ile Lys Lys	
35 40 45	
ATC CAA GAA AAT GAT TTC AAT AAC CTC AAT GAG TTG CAA GTC CTT GAC	192
Ile Gln Glu Asn Asp Phe Asn Asn Leu Asn Glu Leu Gln Val Leu Asp	
50 55 60	
CTA CGT GGA AAT TGC CCT CGA TGT CAT AAT GTC CCA TAT CCG TGT ACA	240
Leu Arg Gly Asn Cys Pro Arg Cys His Asn Val Pro Tyr Pro Cys Thr	
65 70 75 80	
CCG TGT GAA AAT AAT TCC CCC TTA CAG ATC CAT GAC AAT GCT TTC AAT	288
Pro Cys Glu Asn Asn Ser Pro Leu Gln Ile His Asp Asn Ala Phe Asn	
85 90 95	
TCA TCG ACA GAC	300
Ser Ser Thr Asp	
100	

(2) INFORMACIÓN PARA EL SEQ ID NO:28:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 100 aminoácidos

(B) TIPO: aminoácido (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:28:

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

## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:29:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1756 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..1182

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 1643
- (D) OTRA INFORMACIÓN: /nota= “el nucleótido 1643 denominado A, puede ser A o G”

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 1664
- (D) OTRA INFORMACIÓN: /nota= “el nucleótido 1664 denominado C, puede ser A, C, G, o T”

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 1680
- (D) OTRA INFORMACIÓN: /nota= “los nucleótidos 1680 y 1735 denominados G, pueden ser G o T”

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 1719
- (D) OTRA INFORMACIÓN: /nota= “el nucleótido 1719 denominado C, puede ser C o T”

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 1727
- (D) OTRA INFORMACIÓN: /nota= “el nucleótido 1727 denominado A, puede ser A, G, o T”

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:29:

# ES 2 340 210 T3

	TCT	CCA	GAA	ATT	CCC	TGG	AAT	TCC	TTG	CCT	CCT	GAG	GTT	TTT	GAG	GGT	48
	Ser	Pro	Glu	Ile	Pro	Trp	Asn	Ser	Leu	Pro	Pro	Glu	Val	Phe	Glu	Gly	
	1				5					10					15		
5	ATG	CCG	CCA	AAT	CTA	AAG	AAT	CTC	TCC	TTG	GCC	AAA	AAT	GGG	CTC	AAA	96
	Met	Pro	Pro	Asn	Leu	Lys	Asn	Leu	Ser	Leu	Ala	Lys	Asn	Gly	Leu	Lys	
				20					25					30			
10	TCT	TTC	TTT	TGG	GAC	AGA	CTC	CAG	TTA	CTG	AAG	CAT	TTG	GAA	ATT	TTG	144
	Ser	Phe	Phe	Trp	Asp	Arg	Leu	Gln	Leu	Leu	Lys	His	Leu	Glu	Ile	Leu	
			35				40						45				
15	GAC	CTC	AGC	CAT	AAC	CAG	CTG	ACA	AAA	GTA	CCT	GAG	AGA	TTG	GCC	AAC	192
	Asp	Leu	Ser	His	Asn	Gln	Thr	Lys	Val	Pro	Pro	Glu	Arg	Leu	Ala	Asn	
	50					55						60					
20	TGT	TCC	AAA	AGT	CTC	ACA	ACA	CTG	ATT	CTT	AAG	CAT	AAT	CAA	ATC	AGG	240
	Cys	Ser	Lys	Ser	Leu	Thr	Leu	Ile	Leu	Lys	His	Asn	Gln	Ile	Arg		
	65				70				75					80			
25	CAA	TTG	ACA	AAA	TAT	TTT	CTA	GAA	GAT	GCT	TTG	CAA	TTG	CGC	TAT	CTA	288
	Gln	Leu	Thr	Lys	Tyr	Phe	Leu	Glu	Asp	Leu	Gln	Leu	Arg	Tyr	Leu		
				85					90					95			
30	GAC	ATC	AGT	TCA	AAT	AAA	ATC	CAG	GTC	ATT	CAG	AAG	ACT	AGC	TTC	CCA	336
	Asp	Ile	Ser	Ser	Asn	Lys	Ile	Gln	Val	Ile	Gln	Lys	Thr	Ser	Phe	Pro	
				100				105						110			
35	GAA	AAT	GTC	CTC	AAC	AAT	CTG	GAG	ATG	TTG	GTT	TTA	CAT	CAC	AAT	CGC	384
	Glu	Asn	Val	Leu	Asn	Asn	Leu	Glu	Met	Leu	Val	Leu	His	His	Asn	Arg	
			115				120						125				
40	TTT	CTT	TGC	AAC	TGT	GAT	GCT	GTG	TGG	TTT	GTC	TGG	TGG	GTT	AAC	CAT	432
	Phe	Leu	Cys	Asn	Cys	Asp	Ala	Val	Trp	Phe	Val	Trp	Trp	Val	Asn	His	
		130				135						140					
45	ACA	GAT	GTT	ACT	ATT	CCA	TAC	CTG	GCC	ACT	GAT	GTG	ACT	TGT	GTA	GGT	480
	Thr	Asp	Val	Thr	Ile	Pro	Tyr	Leu	Ala	Thr	Asp	Val	Thr	Cys	Val	Gly	
	145				150						155				160		
50	CCA	GGA	GCA	CAC	AAA	GGT	CAA	AGT	GTC	ATA	TCC	CTT	GAT	CTG	TAT	ACG	528
	Pro	Gly	Ala	His	Lys	Gly	Gln	Ser	Val	Ile	Ser	Leu	Asp	Leu	Tyr	Thr	
				165					170						175		
55	TGT	GAG	TTA	GAT	CTC	ACA	AAC	CTG	ATT	CTG	TTC	TCA	GTT	TCC	ATA	TCA	576
	Cys	Glu	Leu	Asp	Leu	Thr	Asn	Leu	Ile	Leu	Phe	Ser	Val	Ser	Ile	Ser	
				180				185						190			
60	TCA	GTC	CTC	TTT	CTT	ATG	GTA	GTT	ATG	ACA	ACA	AGT	CAC	CTC	TTT	TTC	624
	Ser	Val	Leu	Phe	Leu	Met	Val	Val	Met	Thr	Thr	Ser	His	Leu	Phe	Phe	
			195				200						205				
65	TGG	GAT	ATG	TGG	TAC	ATT	TAT	TAT	TTT	TGG	AAA	GCA	AAG	ATA	AAG	GGG	672
	Trp	Asp	Met	Trp	Tyr	Ile	Tyr	Tyr	Phe	Trp	Lys	Ala	Lys	Ile	Lys	Gly	
		210				215						220					
70	TAT	CCA	GCA	TCT	GCA	ATC	CCA	TGG	AGT	CCT	TGT	TAT	GAT	GCT	TTT	ATT	720
	Tyr	Pro	Ala	Ser	Ala	Ile	Pro	Trp	Ser	Pro	Cys	Tyr	Asp	Ala	Phe	Ile	
	225				230						235					240	

## ES 2 340 210 T3

	GTG TAT GAC ACT AAA AAC TCA GCT GTG ACA GAA TGG GTT TTG CAG GAG	768
	Val Tyr Asp Thr Lys Asn Ser Ala Val Thr Glu Trp Val Leu Gln Glu	
	245 250 255	
5	CTG GTG GCA AAA TTG GAA GAT CCA AGA GAA AAA CAC TTC AAT TTG TGT	816
	Leu Val Ala Lys Leu Glu Asp Pro Arg Glu Lys His Phe Asn Leu Cys	
	260 265 270	
10	CTA GAA GAA AGA GAC TGG CTA CCA GGA CAG CCA GTT CTA GAA AAC CTT	864
	Leu Glu Glu Arg Asp Trp Leu Pro Gly Gln Pro Val Leu Glu Asn Leu	
	275 280 285	
15	TCC CAG AGC ATA CAG CTC AGC AAA AAG ACA GTG TTT GTG ATG ACA CAG	912
	Ser Gln Ser Ile Gln Leu Ser Lys Lys Thr Val Phe Val Met Thr Gln	
	290 295 300	
20	AAA TAT GCT AAG ACT GAG AGT TTT AAG ATG GCA TTT TAT TTG TCT CAT	960
	Lys Tyr Ala Lys Thr Glu Ser Phe Lys Met Ala Phe Tyr Leu Ser His	
	305 310 315 320	
25	CAG AGG CTC CTG GAT GAA AAA GTG GAT GTG ATT ATC TTG ATA TTC TTG	1008
	Gln Arg Leu Leu Asp Glu Lys Val Asp Val Ile Ile Leu Ile Phe Leu	
	325 330 335	
30	GAA AGA CCT CTT CAG AAG TCT AAG TTT CTT CAG CTC AGG AAG AGA CTC	1056
	Glu Arg Pro Leu Gln Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg Leu	
	340 345 350	
35	TGC AGG AGC TCT GTC CTT GAG TGG CCT GCA AAT CCA CAG GCT CAC CCA	1104
	Cys Arg Ser Ser Val Leu Glu Trp Pro Ala Asn Pro Gln Ala His Pro	
	355 360 365	
40	TAC TTC TGG CAG TGC CTG AAA AAT GCC CTG ACC ACA GAC AAT CAT GTG	1152
	Tyr Phe Trp Gln Cys Leu Lys Asn Ala Leu Thr Thr Asp Asn His Val	
	370 375 380	
45	GCT TAT AGT CAA ATG TTC AAG GAA ACA GTC TAGCTCTCTG AAGAATGTCA	1202
	Ala Tyr Ser Gln Met Phe Lys Glu Thr Val	
	385 390	
50	CCACCTAGGA CATGCCTTGG TACCTGAAGT TTTCATAAAG GTTCCATAA ATGAAGGTCT	1262
	GAATTTTTC TAACAGTTGT CATGGCTCAG ATTGGTGGGA AATCATCAAT ATATGGCTAA	1322
	GAAATTAAGA AGGGGAGACT GATAGAAGAT AATTTCTTTC TTCATGTGCC ATGCTCAGTT	1382
	AAATATTTCC CCTAGCTCAA ATCTGAAAAA CTGTGCCTAG GAGACAACAC AAGGCTTTGA	1442
	TTTATCTGCA TACAATTGAT AAGAGCCACA CATCTGCCCT GAAGAAGTAC TAGTAGTTT	1502
	AGTAGTAGGG TAAAAATTAC ACAAGCTTTC TCTCTCTCTG ATACTGAACT GTACCAGAGT	1562
	TCAATGAAAT AAAAGCCCAG AGAACTTCTC AGTAAATGGT TTCATTATCA TGTAGTATCC	1622
	ACCATGCAAT ATGCCACAAA ACCGCTACTG GTACAGGACA GCTGGTAGCT GCTTCAAGGC	1682
	CTCTTATCAT TTTCTTGGGG CCCATGGAGG GGTCTCTCTG GAAAAAGGGA AGGTTTTTTT	1742
	TGGCCATCCA TGAA	1756

### (2) INFORMACIÓN PARA EL SEQ ID NO:30:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 394 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

# ES 2 340 210 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:30:

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

## ES 2 340 210 T3

		Gln	Arg	Leu	Leu	Asp	Glu	Lys	Val	Asp	Val	Ile	Ile	Leu	Ile	Phe	Leu
						325					330					335	
5		Glu	Arg	Pro	Leu	Gln	Lys	Ser	Lys	Phe	Leu	Gln	Leu	Arg	Lys	Arg	Leu
					340					345					350		
		Cys	Arg	Ser	Ser	Val	Leu	Glu	Trp	Pro	Ala	Asn	Pro	Gln	Ala	His	Pro
				355					360					365			
10		Tyr	Phe	Trp	Gln	Cys	Leu	Lys	Asn	Ala	Leu	Thr	Thr	Asp	Asn	His	Val
		370						375					380				
		Ala	Tyr	Ser	Gln	Met	Phe	Lys	Glu	Thr	Val						
15		385					390										

### (2) INFORMACIÓN PARA EL SEQ ID NO:31:

- 20 (i) CARACTERÍSTICAS DE LA SECUENCIA:
  - (A) LONGITUD: 999 pares de bases
  - (B) TIPO: ácido nucleico
  - 25 (C) CUALIDAD DE LA HEBRA: sencilla
  - (D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: ADNc
- 30 (ix) RASGO:
  - (A) NOMBRE/CLAVE: CDS
  - (B) LOCALIZACIÓN: 2..847
- 35 (ix) RASGO:
  - (A) NOMBRE/CLAVE: rasgo\_misc
  - (B) LOCALIZACIÓN: 4
  - 40 (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 4 y 23 denominados C, pueden ser cada uno A, C, G, o T"
- (ix) RASGO:
  - (A) NOMBRE/CLAVE: rasgo\_misc
  - 45 (B) LOCALIZACIÓN: 650
  - (D) OTRA INFORMACIÓN: /nota= "el nucleótido 650 denominado G, puede ser A o G"
- (ix) RASGO:
  - (A) NOMBRE/CLAVE: rasgo\_misc
  - 50 (B) LOCALIZACIÓN: 715
  - (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 715, 825, y 845 denominados C, pueden ser cada uno C o T"
- 55 (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:31:



# ES 2 340 210 T3

	C TCC GAT GCC AAG ATT CGG CAC CAG GCA TAT TCA GAG GTC ATG ATG	46
	Ser Asp Ala Lys Ile Arg His Gln Ala Tyr Ser Glu Val Met Met	
	1 5 10 15	
5	GTT GGA TGG TCA GAT TCA TAC ACC TGT GAA TAC CCT TTA AAC CTA AGG	94
	Val Gly Trp Ser Asp Ser Tyr Thr Cys Glu Tyr Pro Leu Asn Leu Arg	
	20 25 30	
10	GGA ACT AGG TTA AAA GAC GTT CAT CTC CAC GAA TTA TCT TGC AAC ACA	142
	Gly Thr Arg Leu Lys Asp Val His Leu His Glu Leu Ser Cys Asn Thr	
	35 40 45	
15	GCT CTG TTG ATT GTC ACC ATT GTG GTT ATT ATG CTA GTT CTG GGG TTG	190
	Ala Leu Leu Ile Val Thr Ile Val Val Ile Met Leu Val Leu Gly Leu	
	50 55 60	
20	GCT GTG GCC TTC TGC TGT CTC CAC TTT GAT CTG CCC TGG TAT CTC AGG	238
	Ala Val Ala Phe Cys Cys Leu His Phe Asp Leu Pro Trp Tyr Leu Arg	
	65 70 75	
25	ATG CTA GGT CAA TGC ACA CAA ACA TGG CAC AGG GTT AGG AAA ACA ACC	286
	Met Leu Gly Gln Cys Thr Gln Thr Trp His Arg Val Arg Lys Thr Thr	
	80 85 90 95	
30	CAA GAA CAA CTC AAG AGA AAT GTC CGA TTC CAC GCA TTT ATT TCA TAC	334
	Gln Glu Gln Leu Lys Arg Asn Val Arg Phe His Ala Phe Ile Ser Tyr	
	100 105 110	
35	AGT GAA CAT GAT TCT CTG TGG GTG AAG AAT GAA TTG ATC CCC AAT CTA	382
	Ser Glu His Asp Ser Leu Trp Val Lys Asn Glu Leu Ile Pro Asn Leu	
	115 120 125	
40	GAG AAG GAA GAT GGT TCT ATC TTG ATT TGC CTT TAT GAA AGC TAC TTT	430
	Glu Lys Glu Asp Gly Ser Ile Leu Ile Cys Leu Tyr Glu Ser Tyr Phe	
	130 135 140	
45	GAC CCT GGC AAA AGC ATT AGT GAA AAT ATT GTA AGC TTC ATT GAG AAA	478
	Asp Pro Gly Lys Ser Ile Ser Glu Asn Ile Val Ser Phe Ile Glu Lys	
	145 150 155	
50	AGC TAT AAG TCC ATC TTT GTT TTG TCT CCC AAC TTT GTC CAG AAT GAG	526
	Ser Tyr Lys Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Asn Glu	
	160 165 170 175	
55	TGG TGC CAT TAT GAA TTC TAC TTT GCC CAC CAC AAT CTC TTC CAT GAA	574
	Trp Cys His Tyr Glu Phe Tyr Phe Ala His His Asn Leu Phe His Glu	
	180 185 190	
60	AAT TCT GAT CAC ATA ATT CTT ATC TTA CTG GAA CCC ATT CCA TTC TAT	622
	Asn Ser Asp His Ile Ile Leu Ile Leu Leu Glu Pro Ile Pro Phe Tyr	
	195 200 205	
65	TGC ATT CCC ACC AGG TAT CAT AAA CTG GAA GCT CTC CTG GAA AAA AAA	670
	Cys Ile Pro Thr Arg Tyr His Lys Leu Glu Ala Leu Glu Lys Lys	
	210 215 220	
70	GCA TAC TTG GAA TGG CCC AAG GAT AGG CGT AAA TGT GGG CTT TTC TGG	718
	Ala Tyr Leu Glu Trp Pro Lys Asp Arg Arg Lys Cys Gly Leu Phe Trp	
	225 230 235	

## ES 2 340 210 T3

	GCA AAC CTT CGA GCT GCT GTT AAT GTT AAT GTA TTA GCC ACC AGA GAA Ala Asn Leu Arg Ala Ala Val Asn Val Asn Val Leu Ala Thr Arg Glu 240 245 250 255	766
5	ATG TAT GAA CTG CAG ACA TTC ACA GAG TTA AAT GAA GAG TCT CGA GGT Met Tyr Glu Leu Gln Thr Phe Thr Glu Leu Asn Glu Glu Ser Arg Gly 260 265 270	814
10	TCT ACA ATC TCT CTG ATG AGA ACA GAC TGT CTA TAAATCCCA CAGTCCTGG Ser Thr Ile Ser Leu Met Arg Thr Asp Cys Leu 275 280	867
	GAAGTTGGGG ACCACATACA CTGTTGGGAT GTACATTGAT ACAACCTTTA TGATGGCAAT	927
15	TTGACAATAT TTATTAAAAT AAAAAATGGT TATTCCTTC AAAAAAAAAA AAAAAAAAAA	987
	AAAAAAAAAA AA	999

### (2) INFORMACIÓN PARA EL SEQ ID NO:32:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 282 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:32:

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

ES 2 340 210 T3

[illegible]

(2) INFORMACIÓN PARA EL SEQ ID NO:33:

- 25 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 1173 pares de bases
- (B) TIPO: ácido nucleico
- 30 (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: ADNc
- 35 (ix) RASGO:
- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..1008
- 40 (ix) RASGO:
- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 854
- (D) OTRA INFORMACIÓN: /nota= “el nucleótido 854 denominado A, puede ser A o T”
- 45 (ix) RASGO:
- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 1171
- 50 (D) OTRA INFORMACIÓN: /nota= “los nucleótidos 1171 y 1172 denominados C, pueden ser cada uno A, C, G, o T”
- (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:33:

55	CTG	CCT	GCT	GGC	ACC	CGG	CTC	CGG	AGG	CTG	GAT	GTC	AGC	TGC	AAC	AGC	48
	Leu	Pro	Ala	Gly	Thr	Arg	Leu	Arg	Arg	Leu	Asp	Val	Ser	Cys	Asn	Ser	
	1				5					10					15		
60	ATC	AGC	TTC	GTG	GCC	CCC	GGC	TTC	TTT	TCC	AAG	GCC	AAG	GAG	CTG	CGA	96
	Ile	Ser	Phe	Val	Ala	Pro	Gly	Phe	Phe	Ser	Lys	Ala	Lys	Glu	Leu	Arg	
				20					25					30			

# ES 2 340 210 T3

5	GAG CTC AAC CTT AGC GCC AAC GCC CTC AAG ACA GTG GAC CAC TCC TGG Glu Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp 35 40 45	144
10	TTT GGG CCC CTG GCG AGT GCC CTG CAA ATA CTA GAT GTA AGC GCC AAC Phe Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn 50 55 60	192
15	CCT CTG CAC TGC GCC TGT GGG GCG GCC TTT ATG GAC TTC CTG CTG GAG Pro Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Glu 65 70 75 80	240
20	GTG CAG GCT GCC GTG CCC GGT CTG CCC AGC CGG GTG AAG TGT GGC AGT Val Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser 85 90 95	288
25	CCG GGC CAG CTC CAG GGC CTC AGC ATC TTT GCA CAG GAC CTG CGC CTC Pro Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala Gln Asp Leu Arg Leu 100 105 110	336
30	TGC CTG GAT GAG GCC CTC TCC TGG GAC TGT TTC GCC CTC TCG CTG CTG Cys Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu 115 120 125	384
35	GCT GTG GCT CTG GGC CTG GGT GTG CCC ATG CTG CAT CAC CTC TGT GGC Ala Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly 130 135 140	432
40	TGG GAC CTC TGG TAC TGC TTC CAC CTG TGC CTG GCC TGG CTT CCC TGG Trp Asp Leu Trp Tyr Cys Phe His Leu Cys Leu Ala Trp Leu Pro Trp 145 150 155 160	480
45	CGG GGG CGG CAA AGT GGG CGA GAT GAG GAT GCC CTG CCC TAC GAT GCC Arg Gly Arg Gln Ser Gly Arg Asp Glu Asp Ala Leu Pro Tyr Asp Ala 165 170 175	528
50	TTC GTG GTC TTC GAC AAA ACG CAG AGC GCA GTG GCA GAC TGG GTG TAC Phe Val Val Phe Asp Lys Thr Gln Ser Ala Val Ala Asp Trp Val Tyr 180 185 190	576
55	AAC GAG CTT CGG GGG CAG CTG GAG GAG TGC CGT GGG CGC TGG GCA CTC Asn Glu Leu Arg Gly Gln Leu Glu Glu Cys Arg Gly Arg Trp Ala Leu 195 200 205	624
60	CGC CTG TGC CTG GAG GAA CGC GAC TGG CTG CCT GGC AAA ACC CTC TTT Arg Leu Cys Leu Glu Glu Arg Asp Trp Leu Pro Gly Lys Thr Leu Phe 210 215 220	672
65	GAG AAC CTG TGG GCC TCG GTC TAT GGC AGC CGC AAG ACG CTG TTT GTG Glu Asn Leu Trp Ala Ser Val Tyr Gly Ser Arg Lys Thr Leu Phe Val 225 230 235 240	720
70	CTG GCC CAC ACG GAC CGG GTC AGT GGT CTC TTG CGC GCC AGC TTC CTG Leu Ala His Thr Asp Arg Val Ser Gly Leu Leu Arg Ala Ser Phe Leu 245 250 255	768
75	CTG GCC CAG CAG CGC CTG CTG GAG GAC CGC AAG GAC GTC GTG GTG CTG Leu Ala Gln Gln Arg Leu Leu Glu Asp Arg Lys Asp Val Val Val Leu 260 265 270	816
80	GTG ATC CTG AGC CCT GAC GGC CGC CGC TCC CGC TAC GAG CGG CTG CGC	864

# ES 2 340 210 T3

	Val	Ile	Leu	Ser	Pro	Asp	Gly	Arg	Arg	Ser	Arg	Tyr	Glu	Arg	Leu	Arg	
			275					280					285				
5	CAG	CGC	CTC	TGC	CGC	CAG	AGT	GTC	CTC	CTC	TGG	CCC	CAC	CAG	CCC	AGT	912
	Gln	Arg	Leu	Cys	Arg	Gln	Ser	Val	Leu	Leu	Trp	Pro	His	Gln	Pro	Ser	
			290				295					300					
	GGT	CAG	CGC	AGC	TTC	TGG	GCC	CAG	CTG	GGC	ATG	GCC	CTG	ACC	AGG	GAC	960
10	Gly	Gln	Arg	Ser	Phe	Trp	Ala	Gln	Leu	Gly	Met	Ala	Leu	Thr	Arg	Asp	
			305			310				315						320	
	AAC	CAC	CAC	TTC	TAT	AAC	CGG	AAC	TTC	TGC	CAG	GGA	CCC	ACG	GCC	GAA	1008
	Asn	His	His	Phe	Tyr	Asn	Arg	Asn	Phe	Cys	Gln	Gly	Pro	Thr	Ala	Glu	
				325					330						335		
15	TAGCCGTGAG CCGGAATCCT GCACGGTGCC ACCTCCACAC TCACCTCACC TCTGCCTGCC																1068
	TGGTCTGACC CTCCCCTGCT CGCCTCCCTC ACCCCACACC TGACACAGAG CAGGCACTCA																1128
20	ATAAATGCTA CCGAAGGCTA AAAAAAAAAA AAAAAAAAAA AACCA																1173

## (2) INFORMACIÓN PARA EL SEQ ID NO:34:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 336 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: proteína
- (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:34:

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

# ES 2 340 210 T3

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

35

## (2) INFORMACIÓN PARA EL SEQ ID NO:35:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

40

- (A) LONGITUD: 4 97 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

45

### (ii) TIPO DE MOLÉCULA: ADNc

### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO: 35:

50	TGGCCACAC GGACCGGTC AGTGGCCTCC TGCACACCAG CTTCTGCTG GCTCAGCAGC	60
	GCCTGTGGA AGACCGCAAG GACGTGGTGG TGTGGTGAT CCTGCGTCCG GATGCCCCAC	120
	CGTCCCGCTA TGTGCGACTG CGCCAGCGTC TCTGCCGCCA GAGTGTGCTC TTCTGGCCCC	180
55	AGCGACCCAA CGGGCAGGGG GGCTTCTGGG CCCAGCTGAG TACAGCCCTG ACTAGGGACA	240
	ACCGCCACTT CTATAACCAG AACTTCTGCC GGGGACCTAC AGCAGAATAG CTCAGAGCAA	300
60	CAGCTGGAAG CAGCTGCATC TTCATGTCTG GTTCCCGAGT TGCTCTGCCT GCCTTGCTCT	360
	GTCTTACTAC ACCGCTATTT GGCAAGTGCG CAATATATGC TACCAAGCCA CCAGGCCAC	420
	GGAGCAAAGG TTGGCTGTAA AGGGTAGTTT TCTTCCCATG CATCTTTCAG GAGAGTGAAG	480
65	ATAGACACCA AACCAC	497

## ES 2 340 210 T3

La invención también proporciona las siguientes realizaciones definidas en las cláusulas de más abajo:

- 5 C1. Una proteína o péptido de DTLR2 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 4.
- 10 C2. Una proteína o péptido de DTLR3 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 6.
- 15 C3. Una proteína o péptido de DTLR4 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 26.
- 20 C4. Una proteína o péptido de DTLR5 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 10.
- 25 C5. Una proteína o péptido de DTLR6 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 12.
- 30 C6. Una proteína o péptido de DTLR7 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 16 o 18.
- 35 C7. Una proteína o péptido de DTLR8 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 32.
- 40 C8. Una proteína o péptido de DTLR9 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 22.
- 45 C9. Una proteína o péptido de DTLR10 esencialmente pura o recombinante que muestra una identidad de secuencia de al menos aproximadamente 85% a lo largo de una longitud de al menos aproximadamente 12 aminoácidos con el SEQ ID NO: 34.
- 50 C10. Una proteína de fusión que comprende la proteína o péptido de cualquiera de C1 a 9.
- C11. Un compuesto de unión que se une específicamente a la proteína o péptido de cualquiera de C1 a 9.
- C12. El compuesto de unión de C11 que es un anticuerpo o un fragmento de anticuerpo.
- 55 C13. Un ácido nucleico que codifica la proteína o péptido de cualquiera de C1 a 9.
- C14. Un vector de expresión que comprende el ácido nucleico de C13.
- 60 C15. Una célula anfitriona que comprende el vector de C14.
- 65 C16. Un procedimiento para producir recombinantemente un polipéptido que comprende cultivar la célula anfitriona de C15 en condiciones en las que se expresa el polipéptido.

## ES 2 340 210 T3

### Propuesta de secuencias

- SEQ ID NO: 1 proporciona la secuencia de nucleótidos de DTLR1 de primate.
- 5 SEQ ID NO: 2 proporciona la secuencia de polipéptidos de DTLR1 primate.
- SEQ ID NO: 3 proporciona la secuencia de nucleótidos de DTLR2 de primate.
- SEQ ID NO: 4 proporciona la secuencia de polipéptidos de DTLR2 de primate.
- 10 SEQ ID NO: 5 proporciona la secuencia de nucleótidos de DTLR3 de primate.
- SEQ ID NO: 6 proporciona la secuencia de polipéptidos de DTLR3 de primate.
- 15 SEQ ID NO: 7 proporciona la secuencia de nucleótidos de DTLR4 de primate.
- SEQ ID NO: 8 proporciona la secuencia de polipéptidos de DTLR4 de primate.
- SEQ ID NO: 9 proporciona la secuencia de nucleótidos de DTLR5 de primate.
- 20 SEQ ID NO: 10 proporciona la secuencia de polipéptidos de DTLR5 de primate.
- SEQ ID NO: 11 proporciona la secuencia de nucleótidos de DTLR6 de primate.
- 25 SEQ ID NO: 12 proporciona la secuencia de polipéptidos de DTLR6 de primate.
- SEQ ID NO: 13 proporciona la secuencia de nucleótidos de DTLR6 de roedor.
- SEQ ID NO: 14 proporciona la secuencia de polipéptidos de DTLR6 de roedor.
- 30 SEQ ID NO: 15 proporciona la secuencia de nucleótidos de DTLR7 de primate.
- SEQ ID NO: 16 proporciona la secuencia de polipéptidos de DTLR7 de primate.
- 35 SEQ ID NO: 17 proporciona la secuencia de nucleótidos de DTLR7 de primate.
- SEQ ID NO: 18 proporciona la secuencia de polipéptidos de DTLR7 de primate.
- SEQ ID NO: 19 proporciona la secuencia de nucleótidos de DTLR8 de primate.
- 40 SEQ ID NO: 20 proporciona la secuencia de polipéptidos de DTLR8 de primate.
- SEQ ID NO: 21 proporciona la secuencia de nucleótidos de DTLR9 de primate.
- 45 SEQ ID NO: 22 proporciona la secuencia de polipéptidos de DTLR9 de primate.
- SEQ ID NO: 23 proporciona la secuencia de nucleótidos de DTLR10 de primate.
- SEQ ID NO: 24 proporciona la secuencia de polipéptidos de DTLR10 de primate.
- 50 SEQ ID NO: 25 proporciona la secuencia de nucleótidos de DTLR4 de primate.
- SEQ ID NO: 26 proporciona la secuencia de polipéptidos de DTLR4 de primate.
- 55 SEQ ID NO: 27 proporciona la secuencia de nucleótidos de DTLR6 de roedor.
- SEQ ID NO: 28 proporciona la secuencia de polipéptidos de DTLR6 de roedor.
- SEQ ID NO: 29 proporciona la secuencia de nucleótidos de DTLR6 de roedor.
- 60 SEQ ID NO: 30 proporciona la secuencia de polipéptidos de DTLR6 de roedor.
- SEQ ID NO: 31 proporciona la secuencia de nucleótidos de DTLR8 de primate.
- 65 SEQ ID NO: 32 proporciona la secuencia de polipéptidos de DTLR8 de primate.
- SEQ ID NO: 33 proporciona la secuencia de nucleótidos de DTLR10 de primate.



## ES 2 340 210 T3

SEQ ID NO: 34 proporciona la secuencia de polipéptidos de DTLR10 de primate.

SEQ ID NO: 35 proporciona la secuencia de nucleótidos de DTLR10 de roedor.

5

### (1) INFORMACIÓN GENERAL:

- 10 (i) SOLICITANTE: Hardiman, Gerard T.  
Rock, Fernando L.  
Bazan, J. Fernando  
Kastelein, Robert A.
- 15 (ii) TÍTULO DE LA INVENCIÓN: PROTEÍNAS RECEPTORAS HUMANAS; REACTIVOS Y MÉTODOS RELACIONADOS
- (ii) NÚMERO DE SECUENCIAS: 35
- (iv) DESTINATARIO DE CORRESPONDENCIA:
- 20 (A) DESTINATARIO: DNAX Research Institute  
(B) CALLE: 901 California Avenue  
(C) CIUDAD: Palo Alto  
(D) ESTADO: California  
25 (E) PAÍS: USA  
(F) ZIP: 94304-1104
- (v) FORMA LEGIBLE CON ORDENADOR:
- 30 (A) TIPO MEDIO: Disco flexible  
(B) ORDENADOR: PC compatible con IBM  
(C) SISTEMA OPERATIVO: PC-DOS/MS-DOS  
35 (D) SOPORTE LÓGICO: PatentIn Release núm. 1.0, Versión núm. 1.30
- (vi) DATOS DE LA SOLICITUD ACTUAL:
- 40 (A) NÚMERO DE SOLICITUD:  
(B) FECHA DE PRESENTACIÓN:  
(C) CLASIFICACIÓN:
- (viii) INFORMACIÓN DEL ABOGADO/AGENTE:
- 45 (A) NOMBRE: Ching, Edwin P.  
(B) NÚMERO DE REGISTRO: 34,090  
(C) NÚMERO DE REFERENCIA/ABOGADO: DX0724X
- (ix) INFORMACIÓN DE TELECOMUNICACIONES:
- 50 (A) TELÉFONO: 650-852-9196  
(B) TELEFAX: 650-496-1200
- (2) INFORMACIÓN PARA EL SEQ ID NO:1:
- 55 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 2367 pares de bases  
(B) TIPO: ácido nucleico  
60 (C) CUALIDAD DE LA HEBRA: sencilla  
(D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: ADNc
- 65 (ix) RASGO:
- (A) NOMBRE/CLAVE: CDS

# ES 2 340 210 T3

(B) LOCALIZACIÓN: 1..2358

(ix) RASGO:

(A) NOMBRE/CLAVE: mat\_péptido

(B) LOCALIZACIÓN: 67..2358

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:1:

10	ATG ACT AGC ATC TTC CAT TTT GCC ATT ATC TTC ATG TTA ATA CTT CAG Met Thr Ser Ile Phe His Phe Ala Ile Ile Phe Met Leu Ile Leu Gln -22 -20 -15 -10	48
15	ATC AGA ATA CAA TTA TCT GAA GAA AGT GAA TTT TTA GTT GAT AGG TCA Ile Arg Ile Gln Leu Ser Glu Glu Ser Glu Phe Leu Val Asp Arg Ser -5 1 5 10	96
20	AAA AAC GGT CTC ATC CAC GTT CCT AAA GAC CTA TCC CAG AAA ACA ACA Lys Asn Gly Leu Ile His Val Pro Lys Asp Leu Ser Gln Lys Thr Thr 15 20 25	144
25	ATC TTA AAT ATA TCG CAA AAT TAT ATA TCT GAG CTT TGG ACT TCT GAC Ile Leu Asn Ile Ser Gln Asn Tyr Ile Ser Glu Leu Trp Thr Ser Asp 30 35 40	192
30	ATC TTA TCA CTG TCA AAA CTG AGG ATT TTG ATA ATT TCT CAT AAT AGA Ile Leu Ser Leu Ser Lys Leu Arg Ile Leu Ile Ile Ser His Asn Arg 45 50 55	240
35	ATC CAG TAT CTT GAT ATC AGT GTT TTC AAA TTC AAC CAG GAA TTG GAA Ile Gln Tyr Leu Asp Ile Ser Val Phe Lys Phe Asn Gln Glu Leu Glu 60 65 70	288
40	TAC TTG GAT TTG TCC CAC AAC AAG TTG GTG AAG ATT TCT TGC CAC CCT Tyr Leu Asp Leu Ser His Asn Lys Leu Val Lys Ile Ser Cys His Pro 75 80 85 90	336
45	ACT GTG AAC CTC AAG CAC TTG GAC CTG TCA TTT AAT GCA TTT GAT GCC Thr Val Asn Leu Lys His Leu Asp Leu Ser Phe Asn Ala Phe Asp Ala 95 100 105	384
50	CTG CCT ATA TGC AAA GAG TTT GGC AAT ATG TCT CAA CTA AAA TTT CTG	432

# ES 2 340 210 T3

	Leu	Pro	Ile	Cys 110	Lys	Glu	Phe	Gly	Asn 115	Met	Ser	Gln	Leu	Lys 120	Phe	Leu	
5	GGG Gly	TTG Leu	AGC Ser 125	ACC Thr	ACA Thr	CAC His	TTA Leu	GAA Glu 130	AAA Lys	TCT Ser	AGT Ser	GTG Val	CTG Leu 135	CCA Pro	ATT Ile	GCT Ala	480
	CAT His	TTG Leu 140	AAT Asn	ATC Ile	AGC Ser	AAG Lys	GTC Val 145	TTG Leu	CTG Leu	GTC Val	TTA Leu	GGA Gly 150	GAG Glu	ACT Thr	TAT Tyr	GGG Gly	528
10	GAA Glu 155	AAA Lys	GAA Glu	GAC Asp	CCT Pro	GAG Glu 160	GGC Gly	CTT Leu	CAA Gln	GAC Asp	TTT Phe 165	AAC Asn	ACT Thr	GAG Glu	AGT Ser	CTG Leu 170	576
15	CAC His	ATT Ile	GTG Val	TTC Phe 175	CCC Pro	ACA Thr	AAC Asn	AAA Lys	GAA Glu	TTT Phe 180	CAT His	TTT Phe	ATT Ile	TTG Leu	GAT Asp 185	GTG Val	624
	TCA Ser	GTC Val	AAG Lys	ACT Thr 190	GTA Val	GCA Ala	AAT Asn	CTG Leu	GAA Glu 195	CTA Leu	TCT Ser	AAT Asn	ATC Ile 200	AAA Lys	TGT Cys	GTG Val	672
20	CTA Leu	GAA Glu	GAT Asp 205	AAC Asn	AAA Lys	TGT Cys	TCT Ser	TAC Tyr 210	TTC Phe	CTA Leu	AGT Ser	ATT Ile	CTG Leu 215	GCG Ala	AAA Lys	CTT Leu	720
25	CAA Gln 220	ACA Asn	AAT Pro	CCA Lys	AAG Lys	TTA Leu	TCA Asn 225	AGT Ser	CTT Leu	ACC Thr	TTA Leu	AAC Asn 230	AAC Asn	ATT Ile	GAA Glu	ACA Thr	768
	ACT Thr 235	TGG Trp	AAT Asn	TCT Ser	TTC Phe	ATT Ile 240	AGG Arg	ATC Ile	CTC Leu	CAA Gln	CTA Leu 245	GTT Val	TGG Trp	CAT His	ACA Thr	ACT Thr 250	816
30	GTA Val	TGG Trp	TAT Tyr	TTC Phe	TCA Ser 255	ATT Ile	TCA Ser	AAC Asn	GTG Val 260	AAG Lys	CTA Leu	CAG Gln	GGT Gly	CAG Gln	CTG Leu 265	GAC Asp	864
	TTC Phe	AGA Arg	GAT Asp	TTT Phe 270	GAT Asp	TAT Tyr	TCT Ser	GGC Gly	ACT Thr 275	TCC Ser	TTG Leu	AAG Lys	GCC Ala 280	TTG Leu	TCT Ser	ATA Ile	912
35	CAC His	CAA Gln 285	GTT Val	GTC Val	AGC Ser	GAT Asp	GTG Val	TTC Phe 290	GGT Gly	TTT Phe	CCG Pro	CAA Gln	AGT Ser 295	TAT Tyr	ATC Ile	TAT Tyr	960
40	GAA Glu 300	ATC Ile	TTT Phe	TCG Ser	AAT Asn	ATG Met	AAC Asn 305	ATC Ile	AAA Lys	AAT Asn	TTT Phe	ACA Thr 310	GTG Val	TCT Ser	GGT Gly	ACA Thr	1008
45	CGC Arg 315	ATG Met	GTC Val	CAC His	ATG Met	CTT Leu 320	TGC Cys	CCA Pro	TCC Ser	AAA Lys	ATT Ile 325	AGC Ser	CCG Pro	TTC Phe	CTG Leu 330	CAT His 330	1056
	TTG Leu	GAT Asp	TTT Phe	TCC Ser 335	AAT Asn	AAT Asn	CTC Leu	TTA Leu	ACA Thr 340	GAC Asp	ACG Thr	GTT Val	TTT Phe	GAA Glu 345	AAT Asn 345	TGT Cys	1104
50	GGG Gly	CAC His	CTT Leu	ACT Thr 350	GAG Glu	TTG Leu	GAG Glu	ACA Thr 355	CTT Leu	ATT Ile	TTA Leu	CAA Gln	ATG Met	AAT Asn 360	CAA Gln	TTA Leu	1152
	AAA Lys	GAA Glu	CTT Leu 365	TCA Ser	AAA Lys	ATA Ile	GCT Ala	GAA Glu 370	ATG Met	ACT Thr	ACA Thr	CAG Gln	ATG Met 375	AAG Lys	TCT Ser	CTG Leu	1200
55	CAA Leu	CAA Leu	TTG Leu	GAT Leu	ATT Leu	AGC Leu	CAG Leu	AAT Leu	TCT Leu	GTA Leu	AGC Leu	TAT Leu	GAT Leu	GAA Leu	AAG Leu	AAA Leu	1248
60																	
65																	

# ES 2 340 210 T3

	Gln	Gln	Leu	Asp	Ile	Ser	Gln	Asn	Ser	Val	Ser	Tyr	Asp	Glu	Lys	Lys	
	380						385					390					
5	GGA Gly 395	GAC Asp	TGT Cys	TCT Ser	TGG Trp	ACT Thr 400	AAA Lys	AGT Ser	TTA Leu	TTA Leu	AGT Ser 405	TTA Leu	AAT Asn	ATG Met	TCT Ser	TCA Ser 410	1296
	AAT Asn	ATA Ile	CTT Leu	ACT Thr	GAC His 415	ACT Thr	ATT Ile	TTC Phe	AGA Arg	TGT Cys 420	TTA Leu	CCT Pro	CCC Pro	AGG Arg	ATC Ile 425	AAG Lys	1344
10	GTA Val	CTT Leu	GAT Asp	CTT Leu 430	CAC His	AGC Ser	AAT Asn	AAA Lys	ATA Ile 435	AAG Lys	AGC Ser	ATT Ile	CCT Pro	AAA Lys 440	CAA Gln	GTC Val	1392
15	GTA Val	AAA Lys	CTG Leu 445	GAA Glu	GCT Ala	TTG Leu	CAA Gln	GAA Glu 450	CTC Leu	AAT Asn	GTT Val	GCT Ala 455	TTC Phe	AAT Asn	TCT Ser	TTA Leu	1440
	ACT Thr	GAC Asp 460	CTT Leu	CCT Pro	GGA Gly	TGT Cys	GGC Gly 465	AGC Ser	TTT Phe	AGC Ser	AGC Ser	CTT Leu 470	TCT Ser	GTA Val	TTG Leu	ATC Ile	1488
20	ATT Ile 475	GAT Asp	CAC His	AAT Asn	TCA Ser	GTT Val 480	TCC Ser	CAC His	CCA Pro	TCA Ser	GCT Ala 485	GAT Asp	TTC Phe	TTC Phe	CAG Gln	AGC Ser 490	1536
25	TGC Cys	CAG Gln	AAG Lys	ATG Met	AGG Arg 495	TCA Ser	ATA Ile	AAA Lys	GCA Ala	GGG Gly 500	GAC Asp	AAT Asn	CCA Pro	TTC Phe	CAA Gln 505	TGT Cys	1584
	ACC Thr	TGT Cys	GAG Glu	CTC Leu 510	GGA Gly	GAA Glu	TTT Phe	GTC Val	AAA Lys 515	AAT Asn	ATA Ile	GAC Asp	CAA Gln	GTA Val 520	TCA Ser	AGT Ser	1632
30	GAA Glu	GTG Val	TTA Leu 525	GAG Glu	GGC Gly	TGG Trp	CCT Pro	GAT Asp 530	TCT Ser	TAT Tyr	AAG Lys	TGT Cys	GAC Asp 535	TAC Tyr	CCG Pro	GAA Glu	1680
35	AGT Ser	TAT Tyr 540	AGA Arg	GGA Gly	ACC Thr	CTA Leu	CTA Leu 545	AAG Lys	GAC Asp	TTT Phe	CAC His	ATG Met 550	TCT Ser	GAA Glu	TTA Leu	TCC Ser	1728
	TGC Cys 555	AAC Asn	ATA Ile	ACT Thr	CTG Leu	CTG Leu 560	ATC Ile	GTC Val	ACC Thr	ATC Ile	GTT Val 565	GCC Ala	ACC Thr	ATG Met	CTG Leu	GTG Val 570	1776
40	TTG Leu	GCT Ala	GTG Val	ACT Thr	GTG Val 575	ACC Thr	TCC Ser	CTC Leu	TGC Cys	ATC Ile 580	TAC Tyr	TTG Leu	GAT Asp	CTG Leu	CCC Pro	TGG Trp	1824
45	TAT Tyr	CTC Leu	AGG Arg	ATG Met 590	GTG Val	TGC Cys	CAG Gln	TGG Trp	ACC Thr 595	CAG Gln	ACC Thr	CGG Arg	CGC Arg	AGG Arg 600	GCC Ala	AGG Arg	1872
	AAC Asn	ATA Ile	CCC Pro 605	TTA Leu	GAA Glu	GAA Glu	CTC Leu	CAA Gln 610	AGA Arg	AAT Asn	CTC Leu	CAG Gln	TTT Phe 615	CAT His	GCA Ala	TTT Phe	1920
50	ATT Ile	TCA Ser 620	TAT Tyr	AGT Ser	GGG Gly	CAC His	GAT Asp 625	TCT Ser	TTC Phe	TGG Trp	GTG Val	AAG Lys 630	AAT Asn	GAA Glu	TTA Leu	TTG Leu	1968
55	CCA Pro 635	AAC Asn	CTA Leu	GAG Glu	AAA Lys	GAA Glu 640	GGT Gly	ATG Met	CAG Gln	ATT Ile	TGC Cys 645	CTT Leu	CAT His	GAG Glu	AGA Arg	AAC Asn 650	2016
	TTT GTT	CCT GTC	GGC AAG	AGC ATT	GTG GAA	AAT ATC	ATC ATC	ACC TGC	ATT GAG								2064

## ES 2 340 210 T3

	Phe	Val	Pro	Gly	Lys	Ser	Ile	Val	Glu	Asn	Ile	Ile	Thr	Cys	Ile	Glu	
					655					660					665		
5	AAG	AGT	TAC	AAG	TCC	ATC	TTT	GTT	TTG	TCT	CCC	AAC	TTT	GTC	CAG	AGT	2112
	Lys	Ser	Tyr	Lys	Ser	Ile	Phe	Val	Leu	Ser	Pro	Asn	Phe	Val	Gln	Ser	
				670					675					680			
	GAA	TGG	TGC	CAT	TAT	GAA	CTC	TAC	TTT	GCC	CAT	CAC	AAT	CTC	TTT	CAT	2160
10	Glu	Trp	Cys	His	Tyr	Glu	Leu	Tyr	Phe	Ala	His	His	Asn	Leu	Phe	His	
			685					690					695				
	GAA	GGA	TCT	AAT	AGC	TTA	ATC	CTG	ATC	TTG	CTG	GAA	CCC	ATT	CCG	CAG	2208
	Glu	Gly	Ser	Asn	Ser	Leu	Ile	Leu	Ile	Leu	Leu	Glu	Pro	Ile	Pro	Gln	
			700				705					710					
15	TAC	TCC	ATT	CCT	AGC	AGT	TAT	CAC	AAG	CTC	AAA	AGT	CTC	ATG	GCC	AGG	2256
	Tyr	Ser	Ile	Pro	Ser	Ser	Tyr	His	Lys	Leu	Lys	Ser	Leu	Met	Ala	Arg	
						720				725						730	
	AGG	ACT	TAT	TTG	GAA	TGG	CCC	AAG	GAA	AAG	AGC	AAA	CGT	GGC	CTT	TTT	2304
20	Arg	Thr	Tyr	Leu	Glu	Trp	Pro	Lys	Glu	Lys	Ser	Lys	Arg	Gly	Leu	Phe	
					735					740					745		
	TGG	GCT	AAC	TTA	AGG	GCA	GCC	ATT	AAT	ATT	AAG	CTG	ACA	GAG	CAA	GCA	2352
	Trp	Ala	Asn	Leu	Arg	Ala	Ala	Ile	Asn	Ile	Lys	Leu	Thr	Glu	Gln	Ala	
					750				755					760			
25	AAG	AAA	TAGTCTAGA														2367
	Lys	Lys															

### (2) INFORMACIÓN PARA EL SEQ ID NO:2:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 786 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:2:

	Met	Thr	Ser	Ile	Phe	His	Phe	Ala	Ile	Ile	Phe	Met	Leu	Ile	Leu	Gln	
	-22		-20					-15					-10				
45	Ile	Arg	Ile	Gln	Leu	Ser	Glu	Glu	Ser	Glu	Phe	Leu	Val	Asp	Arg	Ser	
		-5					1				5					10	
	Lys	Asn	Gly	Leu	Ile	His	Val	Pro	Lys	Asp	Leu	Ser	Gln	Lys	Thr	Thr	
				15						20					25		
50	Ile	Leu	Asn	Ile	Ser	Gln	Asn	Tyr	Ile	Ser	Glu	Leu	Trp	Thr	Ser	Asp	
			30					35						40			
	Ile	Leu	Ser	Leu	Ser	Lys	Leu	Arg	Ile	Leu	Ile	Ile	Ser	His	Asn	Arg	
			45					50					55				
55	Ile	Gln	Tyr	Leu	Asp	Ile	Ser	Val	Phe	Lys	Phe	Asn	Gln	Glu	Leu	Glu	
		60				65						70					
	Tyr	Leu	Asp	Leu	Ser	His	Asn	Lys	Leu	Val	Lys	Ile	Ser	Cys	His	Pro	
		75				80					85				90		
60	Thr	Val	Asn	Leu	Lys	His	Leu	Asp	Leu	Ser	Phe	Asn	Ala	Phe	Asp	Ala	
				95						100					105		
	Leu	Pro	Ile	Cys	Lys	Glu	Phe	Gly	Asn	Met	Ser	Gln	Leu	Lys	Phe	Leu	

# ES 2 340 210 T3

	110					115					120					
5	Gly	Leu	Ser	Thr	Thr	His	Leu	Glu	Lys	Ser	Ser	Val	Leu	Pro	Ile	Ala
			125					130					135			
	His	Leu	Asn	Ile	Ser	Lys	Val	Leu	Leu	Val	Leu	Gly	Glu	Thr	Tyr	Gly
		140					145					150				
10	Glu	Lys	Glu	Asp	Pro	Glu	Gly	Leu	Gln	Asp	Phe	Asn	Thr	Glu	Ser	Leu
	155					160					165					170
	His	Ile	Val	Phe	Pro	Thr	Asn	Lys	Glu	Phe	His	Phe	Ile	Leu	Asp	Val
					175					180					185	
15	Ser	Val	Lys	Thr	Val	Ala	Asn	Leu	Glu	Leu	Ser	Asn	Ile	Lys	Cys	Val
				190					195					200		
	Leu	Glu	Asp	Asn	Lys	Cys	Ser	Tyr	Phe	Leu	Ser	Ile	Leu	Ala	Lys	Leu
			205					210					215			
20	Gln	Thr	Asn	Pro	Lys	Leu	Ser	Ser	Leu	Thr	Leu	Asn	Asn	Ile	Glu	Thr
		220					225						230			
	Thr	Trp	Asn	Ser	Phe	Ile	Arg	Ile	Leu	Gln	Leu	Val	Trp	His	Thr	Thr
	235					240					245					250
25	Val	Trp	Tyr	Phe	Ser	Ile	Ser	Asn	Val	Lys	Leu	Gln	Gly	Gln	Leu	Asp
				255						260					265	
	Phe	Arg	Asp	Phe	Asp	Tyr	Ser	Gly	Thr	Ser	Leu	Lys	Ala	Leu	Ser	Ile
				270					275						280	
30	His	Gln	Val	Val	Ser	Asp	Val	Phe	Gly	Phe	Pro	Gln	Ser	Tyr	Ile	Tyr
			285					290					295			
	Glu	Ile	Phe	Ser	Asn	Met	Asn	Ile	Lys	Asn	Phe	Thr	Val	Ser	Gly	Thr
		300					305					310				
35	Arg	Met	Val	His	Met	Leu	Cys	Pro	Ser	Lys	Ile	Ser	Pro	Phe	Leu	His
	315					320					325					330
	Leu	Asp	Phe	Ser	Asn	Leu	Leu	Thr	Asp	Thr	Val	Phe	Glu	Asn	Cys	
				335					340					345		
40	Gly	His	Leu	Thr	Glu	Leu	Glu	Thr	Leu	Ile	Leu	Gln	Met	Asn	Gln	Leu
				350					355					360		
	Lys	Glu	Leu	Ser	Lys	Ile	Ala	Glu	Met	Thr	Thr	Gln	Met	Lys	Ser	Leu
			365					370					375			
45	Gln	Gln	Leu	Asp	Ile	Ser	Gln	Asn	Ser	Val	Ser	Tyr	Asp	Glu	Lys	Lys
		380					385					390				
	Gly	Asp	Cys	Ser	Trp	Thr	Lys	Ser	Leu	Leu	Ser	Leu	Asn	Met	Ser	Ser
	395					400					405					410
50	Asn	Ile	Leu	Thr	Asp	Thr	Ile	Phe	Arg	Cys	Leu	Pro	Pro	Arg	Ile	Lys
				415						420					425	
	Val	Leu	Asp	Leu	His	Ser	Asn	Lys	Ile	Lys	Ser	Ile	Pro	Lys	Gln	Val
				430					435					440		
55	Val	Lys	Leu	Glu	Ala	Leu	Gln	Glu	Leu	Asn	Val	Ala	Phe	Asn	Ser	Leu
			445					450					455			
	Thr	Asp	Leu	Pro	Gly	Cys	Gly	Ser	Phe	Ser	Ser	Leu	Ser	Val	Leu	Ile
		460					465					470				

## ES 2 340 210 T3

5           Ile Asp His Asn Ser Val Ser His Pro Ser Ala Asp Phe Phe Gln Ser  
           475                           480                           485                           490  
           Cys Gln Lys Met Arg Ser Ile Lys Ala Gly Asp Asn Pro Phe Gln Cys  
                                   495                           500                           505  
           Thr Cys Glu Leu Gly Glu Phe Val Lys Asn Ile Asp Gln Val Ser Ser  
                                   510                           515                           520  
 10       Glu Val Leu Glu Gly Trp Pro Asp Ser Tyr Lys Cys Asp Tyr Pro Glu  
                                   525                           530                           535  
           Ser Tyr Arg Gly Thr Leu Leu Lys Asp Phe His Met Ser Glu Leu Ser  
                                   540                           545                           550  
 15       Cys Asn Ile Thr Leu Leu Ile Val Thr Ile Val Ala Thr Met Leu Val  
           555                           560                           565  
           Leu Ala Val Thr Val Thr Ser Leu Cys Ile Tyr Leu Asp Leu Pro Trp  
                                   575                           580                           585  
 20       Tyr Leu Arg Met Val Cys Gln Trp Thr Gln Thr Arg Arg Arg Ala Arg  
                                   590                           595                           600  
           Asn Ile Pro Leu Glu Glu Leu Gln Arg Asn Leu Gln Phe His Ala Phe  
                                   605                           610                           615  
 25       Ile Ser Tyr Ser Gly His Asp Ser Phe Trp Val Lys Asn Glu Leu Leu  
           620                           625                           630  
           Pro Asn Leu Glu Lys Glu Gly Met Gln Ile Cys Leu His Glu Arg Asn  
           635                           640                           645                           650  
 30       Phe Val Pro Gly Lys Ser Ile Val Glu Asn Ile Ile Thr Cys Ile Glu  
                                   655                           660                           665  
           Lys Ser Tyr Lys Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Ser  
                                   670                           675                           680  
 35       Glu Trp Cys His Tyr Glu Leu Tyr Phe Ala His His Asn Leu Phe His  
           685                           690                           695  
           Glu Gly Ser Asn Ser Leu Ile Leu Ile Leu Leu Glu Pro Ile Pro Gln  
           700                           705                           710  
 40       Tyr Ser Ile Pro Ser Ser Tyr His Lys Leu Lys Ser Leu Met Ala Arg  
           715                           720                           725                           730  
           Arg Thr Tyr Leu Glu Trp Pro Lys Glu Lys Ser Lys Arg Gly Leu Phe  
                                   735                           740                           745  
 45       Trp Ala Asn Leu Arg Ala Ala Ile Asn Ile Lys Leu Thr Glu Gln Ala  
                                   750                           755                           760  
           Lys Lys

### (2) INFORMACIÓN PARA EL SEQ ID NO:3:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- 55           (A) LONGITUD: 2355 pares de bases  
           (B) TIPO: ácido nucleico  
           (C) CUALIDAD DE LA HEBRA: sencilla  
 60           (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNC

#### (ix) RASGO:

- 65           (A) NOMBRE/CLAVE: CDS  
           (B) LOCALIZACIÓN: 1..2352

# ES 2 340 210 T3

(ix) RASGO:

(A) NOMBRE/CLAVE: mat\_péptido

(B) LOCALIZACIÓN: 67..2352

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:3:

5	ATG CCA CAT ACT TTG TGG ATG GTG TGG GTC TTG GGG GTC ATC ATC AGC	48
10	Met Pro His Thr Leu Trp Met Val Trp Val Leu Gly Val Ile Ile Ser	
	-22 -20 -15 -10	
15	CTC TCC AAG GAA GAA TCC TCC AAT CAG GCT TCT CTG TCT TGT GAC CGC	96
	Leu Ser Lys Glu Glu Ser Ser Asn Gln Ala Ser Leu Ser Cys Asp Arg	
	-5 1 5 10	
20	AAT GGT ATC TGC AAG GGC AGC TCA GGA TCT TTA AAC TCC ATT CCC TCA	144
	Asn Gly Ile Cys Lys Gly Ser Ser Gly Ser Leu Asn Ser Ile Pro Ser	
	15 20 25	
25	GGG CTC ACA GAA GCT GTA AAA AGC CTT GAC CTG TCC AAC AAC AGG ATC	192
	Gly Leu Thr Glu Ala Val Lys Ser Leu Asp Leu Ser Asn Asn Arg Ile	
	30 35 40	
30	ACC TAC ATT AGC AAC AGT GAC CTA CAG AGG TGT GTG AAC CTC CAG GCT	240
	Thr Tyr Ile Ser Asn Ser Asp Leu Gln Arg Cys Val Asn Leu Gln Ala	
	45 50 55	
35	CTG GTG CTG ACA TCC AAT GGA ATT AAC ACA ATA GAG GAA GAT TCT TTT	288
	Leu Val Leu Thr Ser Asn Gly Ile Asn Thr Ile Glu Glu Asp Ser Phe	
	60 65 70	
40	TCT TCC CTG GGC AGT CTT GAA CAT TTA GAC TTA TCC TAT AAT TAC TTA	336
	Ser Ser Leu Gly Ser Leu Glu His Leu Asp Leu Ser Tyr Asn Tyr Leu	
	75 80 85 90	
45	TCT AAT TTA TCG TCT TCC TGG TTC AAG CCC CTT TCT TCT TTA ACA TTC	384
	Ser Asn Leu Ser Ser Trp Phe Lys Pro Leu Ser Ser Ser Leu Thr Phe	
	95 100 105	
50	TTA AAC TTA CTG GGA AAT CCT TAC AAA ACC CTA GGG GAA ACA TCT CTT	432
	Leu Asn Leu Leu Gly Asn Pro Tyr Lys Thr Leu Gly Glu Thr Ser Leu	
	110 115 120	
55	TTT TCT CAT CTC ACA AAA TTG CAA ATC CTG AGA GTG GGA AAT ATG GAC	480
	Phe Ser His Leu Thr Lys Leu Gln Ile Leu Arg Val Gly Asn Met Asp	
	125 130 135	
60	ACC TTC ACT AAG ATT CAA AGA AAA GAT TTT GCT GGA CTT ACC TTC CTT	528
	Thr Phe Thr Lys Ile Gln Arg Lys Asp Phe Ala Gly Leu Thr Phe Leu	
	140 145 150	
65	GAG GAA CTT GAG ATT GAT GCT TCA GAT CTA CAG AGC TAT GAG CCA AAA	576
	Glu Glu Leu Glu Ile Asp Ala Ser Asp Leu Gln Ser Tyr Glu Pro Lys	
	155 160 165 170	
70	AGT TTG AAG TCA ATT CAG AAC GTA AGT CAT CTG ATC CTT CAT ATG AAG	624
	Ser Leu Lys Ser Ile Gln Asn Val Ser His Leu Ile Leu His Met Lys	
	175 180 185	
75	CAG CAT ATT TTA CTG CTG GAG ATT TTT GTA GAT GTT ACA AGT TCC GTG	672
	Gln His Ile Leu Leu Glu Ile Phe Val Asp Val Thr Ser Ser Val	
	190 195 200	
80	GAA TGT TTG GAA CTG CGA GAT ACT GAT TTG GAC ACT TTC CAT TTT TCA	720



# ES 2 340 210 T3

	Glu	Cys	Leu	Glu	Leu	Arg	Asp	Thr	Asp	Leu	Asp	Thr	Phe	His	Phe	Ser	
			205					210					215				
5	GAA	CTA	TCC	ACT	GGT	GAA	ACA	AAT	TCA	TTG	ATT	AAA	AAG	TTT	ACA	TTT	768
	Glu	Leu	Ser	Thr	Gly	Glu	Thr	Asn	Ser	Leu	Ile	Lys	Lys	Phe	Thr	Phe	
		220					225					230					
	AGA	AAT	GTG	AAA	ATC	ACC	GAT	GAA	AGT	TTG	TTT	CAG	GTT	ATG	AAA	CTT	816
	Arg	Asn	Val	Lys	Ile	Thr	Asp	Glu	Ser	Leu	Phe	Gln	Val	Met	Lys	Leu	
	235					240					245					250	
10	TTG	AAT	CAG	ATT	TCT	GGA	TTG	TTA	GAA	TTA	GAG	TTT	GAT	GAC	TGT	ACC	864
	Leu	Asn	Gln	Ile	Ser	Gly	Leu	Leu	Glu	Leu	Glu	Phe	Asp	Asp	Cys	Thr	
					255					260					265		
15	CTT	AAT	GGA	GTT	GGT	AAT	TTT	AGA	GCA	TCT	GAT	AAT	GAC	AGA	GTT	ATA	912
	Leu	Asn	Gly	Val	Gly	Asn	Phe	Arg	Ala	Ser	Asp	Asn	Asp	Arg	Val	Ile	
				270					275					280			
	GAT	CCA	GGT	AAA	GTG	GAA	ACG	TTA	ACA	ATC	CGG	AGG	CTG	CAT	ATT	CCA	960
	Asp	Pro	Gly	Lys	Val	Glu	Thr	Leu	Thr	Ile	Arg	Arg	Leu	His	Ile	Pro	
			285					290					295				
20	AGG	TTT	TAC	TTA	TTT	TAT	GAT	CTG	AGC	ACT	TTA	TAT	TCA	CTT	ACA	GAA	1008
	Arg	Phe	Tyr	Leu	Phe	Tyr	Asp	Leu	Ser	Thr	Leu	Tyr	Ser	Leu	Thr	Glu	
		300					305					310					
25	AGA	GTT	AAA	AGA	ATC	ACA	GTA	GAA	AAC	AGT	AAA	GTT	TTT	CTG	GTT	CCT	1056
	Arg	Val	Lys	Arg	Ile	Thr	Val	Glu	Asn	Ser	Lys	Val	Phe	Leu	Val	Pro	
	315					320					325					330	
	TGT	TTA	CTT	TCA	CAA	CAT	TTA	AAA	TCA	TTA	GAA	TAC	TTG	GAT	CTC	AGT	1104
	Cys	Leu	Leu	Ser	Gln	His	Leu	Lys	Ser	Leu	Glu	Tyr	Leu	Asp	Leu	Ser	
					335					340					345		
30	GAA	AAT	TTG	ATG	GTT	GAA	GAA	TAC	TTG	AAA	AAT	TCA	GCC	TGT	GAG	GAT	1152
	Glu	Asn	Leu	Met	Val	Glu	Glu	Tyr	Leu	Lys	Asn	Ser	Ala	Cys	Glu	Asp	
				350					355					360			
	GCC	TGG	CCC	TCT	CTA	CAA	ACT	TTA	ATT	TTA	AGG	CAA	AAT	CAT	TTG	GCA	1200
35	Ala	Trp	Pro	Ser	Leu	Gln	Thr	Leu	Ile	Leu	Arg	Gln	Asn	His	Leu	Ala	
			365					370					375				
	TCA	TTG	GAA	AAA	ACC	GGA	GAG	ACT	TTG	CTC	ACT	CTG	AAA	AAC	TTG	ACT	1248
	Ser	Leu	Glu	Lys	Thr	Gly	Glu	Thr	Leu	Leu	Thr	Leu	Lys	Asn	Leu	Thr	
		380					385					390					
40	AAC	ATT	GAT	ATC	AGT	AAG	AAT	AGT	TTT	CAT	TCT	ATG	CCT	GAA	ACT	TGT	1296
	Asn	Ile	Asp	Ile	Ser	Lys	Asn	Ser	Phe	His	Ser	Met	Pro	Glu	Thr	Cys	
						400					405					410	
	CAG	TGG	CCA	GAA	AAG	ATG	AAA	TAT	TTG	AAC	TTA	TCC	AGC	ACA	CGA	ATA	1344
45	Gln	Trp	Pro	Glu	Lys	Met	Lys	Tyr	Leu	Asn	Leu	Ser	Ser	Thr	Arg	Ile	
					415					420					425		
	CAC	AGT	GTA	ACA	GGC	TGC	ATT	CCC	AAG	ACA	CTG	GAA	ATT	TTA	GAT	GTT	1392
	His	Ser	Val	Thr	Gly	Cys	Ile	Pro	Lys	Thr	Leu	Glu	Ile	Leu	Asp	Val	
				430					435					440			
50	AGC	AAC	AAC	AAT	CTC	AAT	TTA	TTT	TCT	TTG	AAT	TTG	CCG	CAA	CTC	AAA	1440
	Ser	Asn	Asn	Asn	Leu	Asn	Leu	Phe	Ser	Leu	Asn	Leu	Pro	Gln	Leu	Lys	
			445					450					455				
	GAA	CTT	TAT	ATT	TCC	AGA	AAT	AAG	TTG	ATG	ACT	CTA	CCA	GAT	GCC	TCC	1488
55	Glu	Leu	Tyr	Ile	Ser	Arg	Asn	Lys	Leu	Met	Thr	Leu	Pro	Asp	Ala	Ser	
		460					465					470					
	CTC	TTA	CCC	ATG	TTA	CTA	GTA	TTG	AAA	ATC	AGT	AGG	AAT	GCA	ATA	ACT	1536

# ES 2 340 210 T3

	Leu 475	Leu	Pro	Met	Leu 480	Leu	Val	Leu	Lys	Ile	Ser 485	Arg	Asn	Ala	Ile	Thr 490	
5	ACG Thr	TTT Phe	TCT Ser	AAG Lys	GAG Glu 495	CAA Gln	CTT Leu	GAC Asp	TCA Ser	TTT Phe 500	CAC His	ACA Thr	CTG Leu	AAG Lys	ACT Thr 505	TTG Leu	1584
	GAA Glu	GCT Ala	GGT Gly	GGC Gly 510	AAT Asn	AAC Asn	TTC Phe	ATT Ile	TGC Cys 515	TCC Ser	TGT Cys	GAA Glu	TTC Phe	CTC Leu 520	TCC Ser	TTC Phe	1632
10	ACT Thr	CAG Gln	GAG Gln 525	CAG Gln	CAA Gln	GCA Ala	CTG Leu	GCC Ala 530	AAA Lys	GTC Val	TTG Leu	ATT Ile	GAT Asp 535	TGG Trp	CCA Pro	GCA Ala	1680
15	AAT Asn	TAC Tyr 540	CTG Leu	TGT Cys	GAC Asp	TCT Ser	CCA Pro 545	TCC Ser	CAT His	GTG Val	CGT Arg	GGC Gly 550	CAG Gln	CAG Gln	GTT Val	CAG Gln	1728
	GAT Asp 555	GTC Val	CGC Arg	CTC Leu	TCG Ser	GTG Val 560	TCG Ser	GAA Glu	TGT Cys	CAC His	AGG Arg 565	ACA Thr	GCA Ala	CTG Leu	GTG Val	TCT Ser 570	1776
20	GGC Gly	ATG Met	TGC Cys	TGT Cys	GCT Ala 575	CTG Leu	TTC Phe	CTG Leu	CTG Leu	ATC Ile 580	CTG Leu	CTC Leu	ACG Thr	GGG Gly	GTG Val 585	CTG Leu	1824
25	TGC Cys	CAC His	CGT Arg	TTC Phe 590	CAT His	GGC Gly	CTG Leu	TGG Trp	TAT Tyr 595	ATG Met	AAA Lys	ATG Met	ATG Met	TGG Trp 600	GCC Ala	TGG Trp	1872
	CTC Leu	CAG Gln	GCC Ala 605	AAA Lys	AGG Arg	AAG Lys	CCC Pro	AGG Arg 610	AAA Lys	GCT Ala	CCC Pro	AGC Ser	AGG Arg 615	AAC Asn	ATC Ile	TGC Cys	1920
30	TAT Tyr	GAT Asp 620	GCA Ala	TTT Phe	GTT Val	TCT Ser	TAC Tyr 625	AGT Ser	GAG Glu	CGG Arg	GAT Asp	GCC Ala 630	TAC Tyr	TGG Trp	GTG Val	GAG Glu	1968
35	AAC Asn 635	CTT Leu	ATG Met	GTC Val	CAG Gln	GAG Glu 640	CTG Leu	GAG Glu	AAC Asn	TTC Phe	AAT Asn 645	CCC Pro	CCC Pro	TTC Phe	AAG Lys	TTG Leu 650	2016
	TGT Cys	CTT Leu	CAT His	AAG Lys	CGG Arg 655	GAC Asp	TTC Phe	ATT Ile	CCT Pro	GGC Gly 660	AAG Lys	TGG Trp	ATC Ile	ATT Ile 665	GAC Asp 665	AAT Asn	2064
40	ATC Ile	ATT Ile	GAC Asp	TCC Ser 670	ATT Ile	GAA Glu	AAG Lys	AGC Ser	CAC His 675	AAA Lys	ACT Thr	GTC Val	TTT Phe	GTG Val 680	CTT Leu	TCT Ser	2112
45	GAA Glu	AAC Asn	TTT Phe 685	GTG Val	AAG Lys	AGT Ser	GAG Glu	TGG Trp 690	TGC Cys	AAG Lys	TAT Tyr	GAA Glu	CTG Leu 695	GAC Asp	TTC Phe	TCC Ser	2160
	CAT His	TTC Phe 700	CGT Arg	CTT Leu	TTT Phe	GAA Glu 705	GAG Glu 705	AAC Asn	AAT Asn	GAT Asp	GCT Ala	GCC Ala 710	ATT Ile	CTC Leu	ATT Ile	CTT Leu	2208
50	CTG Leu 715	GAG Glu	CCC Pro	ATT Ile	GAG Glu	AAA Lys 720	AAA Lys	GCC Ala	ATT Ile	CCC Pro	CAG Gln 725	CGC Arg	TTC Phe	TGC Cys	AAG Lys	CTG Leu 730	2256
55	CGG Arg	AAG Lys	ATA Ile	ATG Met	AAC Asn 735	ACC Thr	AAG Lys	ACC Thr	TAC Tyr	CTG Leu 740	GAG Glu	TGG Trp	CCC Pro	ATG Met	GAC Asp 745	GAG Glu	2304
	GCT	CAG	CGG	GAA	GGA	TTT	TGG	GTA	AAT	CTG	AGA	GCT	GCG	ATA	AAG	TCC	2352
60	Ala	Gln	Arg	Glu 750	Gly	Phe	Trp	Val	Asn 755	Leu	Arg	Ala	Ala	Ile 760	Lys	Ser	
	TAG																2355

## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:4:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 784 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:4:

```

Met Pro His Thr Leu Trp Met Val Trp Val Leu Gly Val Ile Ile Ser
-22   -20           -15           -10
Leu Ser Lys Glu Glu Ser Ser Asn Gln Ala Ser Leu Ser Cys Asp Arg
-5           1           5           10
Asn Gly Ile Cys Lys Gly Ser Ser Gly Ser Leu Asn Ser Ile Pro Ser
15           20           25
Gly Leu Thr Glu Ala Val Lys Ser Leu Asp Leu Ser Asn Asn Arg Ile
30           35           40
Thr Tyr Ile Ser Asn Ser Asp Leu Gln Arg Cys Val Asn Leu Gln Ala
45           50           55
Leu Val Leu Thr Ser Asn Gly Ile Asn Thr Ile Glu Glu Asp Ser Phe
60           65           70
Ser Ser Leu Gly Ser Leu Glu His Leu Asp Leu Ser Tyr Asn Tyr Leu
75           80           85           90
Ser Asn Leu Ser Ser Ser Trp Phe Lys Pro Leu Ser Ser Leu Thr Phe
95           100          105
Leu Asn Leu Leu Gly Asn Pro Tyr Lys Thr Leu Gly Glu Thr Ser Leu
110          115          120
Phe Ser His Leu Thr Lys Leu Gln Ile Leu Arg Val Gly Asn Met Asp
125          130          135
Thr Phe Thr Lys Ile Gln Arg Lys Asp Phe Ala Gly Leu Thr Phe Leu
140          145          150
Glu Glu Leu Glu Ile Asp Ala Ser Asp Leu Gln Ser Tyr Glu Pro Lys
155          160          165          170
Ser Leu Lys Ser Ile Gln Asn Val Ser His Leu Ile Leu His Met Lys
175          180          185
Gln His Ile Leu Leu Leu Glu Ile Phe Val Asp Val Thr Ser Ser Val
190          195          200
Glu Cys Leu Glu Leu Arg Asp Thr Asp Leu Asp Thr Phe His Phe Ser
205          210          215
Glu Leu Ser Thr Gly Glu Thr Asn Ser Leu Ile Lys Lys Phe Thr Phe
220          225          230
Arg Asn Val Lys Ile Thr Asp Glu Ser Leu Phe Gln Val Met Lys Leu
235          240          245          250

```

# ES 2 340 210 T3

Leu Asn Gln Ile Ser Gly Leu Leu Glu Leu Glu Phe Asp Asp Cys Thr  
 255 260 265  
 5 Leu Asn Gly Val Gly Asn Phe Arg Ala Ser Asp Asn Asp Arg Val Ile  
 270 275 280  
 Asp Pro Gly Lys Val Glu Thr Leu Thr Ile Arg Arg Leu His Ile Pro  
 285 290 295  
 10 Arg Phe Tyr Leu Phe Tyr Asp Leu Ser Thr Leu Tyr Ser Leu Thr Glu  
 300 305 310  
 Arg Val Lys Arg Ile Thr Val Glu Asn Ser Lys Val Phe Leu Val Pro  
 315 320 325 330  
 15 Cys Leu Leu Ser Gln His Leu Lys Ser Leu Glu Tyr Leu Asp Leu Ser  
 335 340 345  
 Glu Asn Leu Met Val Glu Glu Tyr Leu Lys Asn Ser Ala Cys Glu Asp  
 350 355 360  
 20 Ala Trp Pro Ser Leu Gln Thr Leu Ile Leu Arg Gln Asn His Leu Ala  
 365 370 375  
 Ser Leu Glu Lys Thr Gly Glu Thr Leu Leu Thr Leu Lys Asn Leu Thr  
 380 385 390  
 25 Asn Ile Asp Ile Ser Lys Asn Ser Phe His Ser Met Pro Glu Thr Cys  
 395 400 405 410  
 Gln Trp Pro Glu Lys Met Lys Tyr Leu Asn Leu Ser Ser Thr Arg Ile  
 415 420 425  
 30 His Ser Val Thr Gly Cys Ile Pro Lys Thr Leu Glu Ile Leu Asp Val  
 430 435 440  
 Ser Asn Asn Asn Leu Asn Leu Phe Ser Leu Asn Leu Pro Gln Leu Lys  
 445 450 455  
 35 Glu Leu Tyr Ile Ser Arg Asn Lys Leu Met Thr Leu Pro Asp Ala Ser  
 460 465 470  
 Leu Leu Pro Met Leu Leu Val Leu Lys Ile Ser Arg Asn Ala Ile Thr  
 475 480 485 490  
 40 Thr Phe Ser Lys Glu Gln Leu Asp Ser Phe His Thr Leu Lys Thr Leu  
 495 500 505  
 Glu Ala Gly Gly Asn Asn Phe Ile Cys Ser Cys Glu Phe Leu Ser Phe  
 510 515 520  
 45 Thr Gln Glu Gln Gln Ala Leu Ala Lys Val Leu Ile Asp Trp Pro Ala  
 525 530 535  
 Asn Tyr Leu Cys Asp Ser Pro Ser His Val Arg Gly Gln Gln Val Gln  
 540 545 550  
 50 Asp Val Arg Leu Ser Val Ser Glu Cys His Arg Thr Ala Leu Val Ser  
 555 560 565 570  
 Gly Met Cys Cys Ala Leu Phe Leu Leu Ile Leu Leu Thr Gly Val Leu  
 575 580 585  
 55 Cys His Arg Phe His Gly Leu Trp Tyr Met Lys Met Met Trp Ala Trp  
 590 595 600  
 Leu Gln Ala Lys Arg Lys Pro Arg Lys Ala Pro Ser Arg Asn Ile Cys  
 605 610 615

## ES 2 340 210 T3

		Tyr	Asp	Ala	Phe	val	Ser	Tyr	Ser	Glu	Arg	Asp	Ala	Tyr	Trp	val	Glu
		620						625					630				
5		Asn	Leu	Met	val	Gln	Glu	Leu	Glu	Asn	Phe	Asn	Pro	Pro	Phe	Lys	Leu
		635					640					645					650
		Cys	Leu	His	Lys	Arg	Asp	Phe	Ile	Pro	Gly	Lys	Trp	Ile	Ile	Asp	Asn
						655					660					665	
10		Ile	Ile	Asp	Ser	Ile	Glu	Lys	Ser	His	Lys	Thr	val	Phe	val	Leu	Ser
					670					675					680		
		Glu	Asn	Phe	val	Lys	Ser	Glu	Trp	Cys	Lys	Tyr	Glu	Leu	Asp	Phe	Ser
				685					690					695			
15		His	Phe	Arg	Leu	Phe	Glu	Glu	Asn	Asn	Asp	Ala	Ala	Ile	Leu	Ile	Leu
		700					705						710				
		Leu	Glu	Pro	Ile	Glu	Lys	Lys	Ala	Ile	Pro	Gln	Arg	Phe	Cys	Lys	Leu
		715				720						725					730
20		Arg	Lys	Ile	Met	Asn	Thr	Lys	Thr	Tyr	Leu	Glu	Trp	Pro	Met	Asp	Glu
					735						740					745	
		Ala	Gln	Arg	Glu	Gly	Phe	Trp	Val	Asn	Leu	Arg	Ala	Ala	Ile	Lys	Ser
					750					755					760		

### (2) INFORMACIÓN PARA EL SEQ ID NO:5:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 2715 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..2712

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: mat\_péptido
- (B) LOCALIZACIÓN: 64..2712

#### (xi) DESCRIPCIÓN DE LA SECUENCIA: SEQ ID NO:5:

		ATG	AGA	CAG	ACT	TTG	CCT	TGT	ATC	TAC	TTT	TGG	GGG	GGC	CTT	TTG	CCC	
		Met	Arg	Gln	Thr	Leu	Pro	Cys	Ile	Tyr	Phe	Trp	Gly	Gly	Leu	Leu	Pro	
		-21	-20					-15					-10					
50		TTT	GGG	ATG	CTG	TGT	GCA	TCC	TCC	ACC	ACC	AAG	TGC	ACT	GTT	AGC	CAT	48
		Phe	Gly	Met	Leu	Cys	Ala	Ser	Ser	Thr	Thr	Lys	Cys	Thr	Val	Ser	His	
		-5					1				5					10		
55		GAA	GTT	GCT	GAC	TGC	AGC	CAC	CTG	AAG	TTG	ACT	CAG	GTA	CCC	GAT	GAT	144
		Glu	Val	Ala	Asp	Cys	Ser	His	Leu	Lys	Leu	Thr	Gln	Val	Pro	Asp	Asp	
				15					20					25				
60		CTA	CCC	ACA	AAC	ATA	ACA	GTG	TTG	AAC	CTT	ACC	CAT	AAT	CAA	CTC	AGA	192
		Leu	Pro	Thr	Asn	Ile	Thr	Val	Leu	Asn	Leu	Thr	His	Asn	Gln	Leu	Arg	
				30					35					40				
65		AGA	TTA	CCA	GCC	GCC	AAC	TTC	ACA	AGG	TAT	AGC	CAG	CTA	ACT	AGC	TTG	240

# ES 2 340 210 T3

	Arg	Leu	Pro	Ala	Ala	Asn	Phe	Thr	Arg	Tyr	Ser	Gln	Leu	Thr	Ser	Leu	
	45						50					55					
5	GAT	GTA	GGA	TTT	AAC	ACC	ATC	TCA	AAA	CTG	GAG	CCA	GAA	TTG	TGC	CAG	288
	Asp	Val	Gly	Phe	Asn	Thr	Ile	Ser	Lys	Leu	Glu	Pro	Glu	Leu	Cys	Gln	
	60					65				70						75	
	AAA	CTT	CCC	ATG	TTA	AAA	GTT	TTG	AAC	CTC	CAG	CAC	AAT	GAG	CTA	TCT	336
	Lys	Leu	Pro	Met	Leu	Lys	Val	Leu	Asn	Leu	Gln	His	Asn	Glu	Leu	Ser	
10					80					85					90		
	CAA	CTT	TCT	GAT	AAA	ACC	TTT	GCC	TTC	TGC	ACG	AAT	TTG	ACT	GAA	CTC	384
	Gln	Leu	Ser	Asp	Lys	Thr	Phe	Ala	Phe	Cys	Thr	Asn	Leu	Thr	Glu	Leu	
				95				100						105			
15	CAT	CTC	ATG	TCC	AAC	TCA	ATC	CAG	AAA	ATT	AAA	AAT	AAT	CCC	TTT	GTC	432
	His	Leu	Met	Ser	Asn	Ser	Ile	Gln	Lys	Ile	Lys	Asn	Asn	Pro	Phe	Val	
			110					115					120				
	AAG	CAG	AAG	AAT	TTA	ATC	ACA	TTA	GAT	CTG	TCT	CAT	AAT	GGC	TTG	TCA	480
	Lys	Gln	Lys	Asn	Leu	Ile	Thr	Leu	Asp	Leu	Ser	His	Asn	Gly	Leu	Ser	
20		125					130					135					
	TCT	ACA	AAA	TTA	GGA	ACT	CAG	GTT	CAG	CTG	GAA	AAT	CTC	CAA	GAG	CTT	528
	Ser	Thr	Lys	Leu	Gly	Thr	Gln	Val	Gln	Leu	Glu	Asn	Leu	Gln	Glu	Leu	
	140				145					150						155	
25	CTA	TTA	TCA	AAC	AAT	AAA	ATT	CAA	GCG	CTA	AAA	AGT	GAA	GAA	CTG	GAT	576
	Leu	Leu	Ser	Asn	Asn	Lys	Ile	Gln	Ala	Leu	Lys	Ser	Glu	Glu	Leu	Asp	
				160					165						170		
	ATC	TTT	GCC	AAT	TCA	TCT	TTA	AAA	AAA	TTA	GAG	TTG	TCA	TCG	AAT	CAA	624
	Ile	Phe	Ala	Asn	Ser	Ser	Leu	Lys	Lys	Leu	Glu	Leu	Ser	Ser	Asn	Gln	
				175					180					185			
30	ATT	AAA	GAG	TTT	TCT	CCA	GGG	TGT	TTT	CAC	GCA	ATT	GGA	AGA	TTA	TTT	672
	Ile	Lys	Glu	Phe	Ser	Pro	Gly	Cys	Phe	His	Ala	Ile	Gly	Arg	Leu	Phe	
			190					195					200				
	GGC	CTC	TTT	CTG	AAC	AAT	GTC	CAG	CTG	GGT	CCC	AGC	CTT	ACA	GAG	AAG	720
35	Gly	Leu	Phe	Leu	Asn	Asn	Val	Gln	Leu	Gly	Pro	Ser	Leu	Thr	Glu	Lys	
		205					210					215					
	CTA	TGT	TTG	GAA	TTA	GCA	AAC	ACA	AGC	ATT	CGG	AAT	CTG	TCT	CTG	AGT	768
	Leu	Cys	Leu	Glu	Leu	Ala	Asn	Thr	Ser	Ile	Arg	Asn	Leu	Ser	Leu	Ser	
		220				225					230					235	
40	AAC	AGC	CAG	CTG	TCC	ACC	ACC	AGC	AAT	ACA	ACT	TTC	TTG	GGA	CTA	AAG	816
	Asn	Ser	Gln	Leu	Ser	Thr	Thr	Ser	Asn	Thr	Thr	Phe	Leu	Gly	Leu	Lys	
					240					245					250		
45	TGG	ACA	AAT	CTC	ACT	ATG	CTC	GAT	CTT	TCC	TAC	AAC	AAC	TTA	AAT	GTG	864
	Trp	Thr	Asn	Leu	Thr	Met	Leu	Asp	Leu	Ser	Tyr	Asn	Asn	Leu	Asn	Val	
			255					260						265			
	GTT	GGT	AAC	GAT	TCC	TTT	GCT	TGG	CTT	CCA	CAA	CTA	GAA	TAT	TTC	TTC	912
	Val	Gly	Asn	Asp	Ser	Phe	Ala	Trp	Leu	Pro	Gln	Leu	Glu	Tyr	Phe	Phe	
			270					275					280				
50	CTA	GAG	TAT	AAT	AAT	ATA	CAG	CAT	TTG	TTT	TCT	CAC	TCT	TTG	CAC	GGG	960
	Leu	Glu	Tyr	Asn	Asn	Ile	Gln	His	Leu	Phe	Ser	His	Ser	Leu	His	Gly	
		285					290					295					
	CTT	TTC	AAT	GTG	AGG	TAC	CTG	AAT	TTG	AAA	CGG	TCT	TTT	ACT	AAA	CAA	1008
55	Leu	Phe	Asn	Val	Arg	Tyr	Leu	Asn	Leu	Lys	Arg	Ser	Phe	Thr	Lys	Gln	
		300				305					310					315	
	AGT	ATT	TCC	CTT	GCC	TCA	CTC	CCC	AAG	ATT	GAT	GAT	TTT	TCT	TTT	CAG	1056

# ES 2 340 210 T3

	Ser	Ile	Ser	Leu	Ala 320	Ser	Leu	Pro	Lys	Ile 325	Asp	Asp	Phe	Ser	Phe 330	Gln	
5	TGG Trp	CTA Leu	AAA Lys	TGT Cys 335	TTG Leu	GAG Glu	CAC His	CTT Leu	AAC Asn 340	ATG Met	GAA Glu	GAT Asp	AAT Asn	GAT Asp 345	ATT Ile	CCA Pro	1104
	GGC Gly	ATA Ile	AAA Lys 350	AGC Ser	AAT Asn	ATG Met	TTC Phe	ACA Thr 355	GGA Gly	TTG Leu	ATA Ile	AAC Asn	CTG Leu 360	AAA Lys	TAC Tyr	TTA Leu	1152
10	AGT Ser	CTA Leu 365	TCC Ser	AAC Asn	TCC Ser	TTT Phe	ACA Thr 370	AGT Ser	TTG Leu	CGA Arg	ACT Thr	TTG Leu 375	ACA Thr	AAT Asn	GAA Glu	ACA Thr	1200
15	TTT Phe 380	GTA Val	TCA Ser	CTT Leu	GCT Ala	CAT His 385	TCT Ser	CCC Pro	TTA Leu	CAC His	ATA Ile 390	CTC Leu	AAC Asn	CTA Leu	ACC Thr	AAG Lys 395	1248
	AAT Asn	AAA Lys	ATC Ile	TCA Ser	AAA Lys 400	ATA Ile	GAG Glu	AGT Ser	GAT Asp	GCT Ala 405	TTC Phe	TCT Ser	TGG Trp	TTG Leu	GGC Gly 410	CAC His	1296
20	CTA Leu	GAA Glu	GTA Val	CTT Leu 415	GAC Asp	CTG Leu	GGC Gly	CTT Leu	AAT Asn 420	GAA Glu	ATT Ile	GGG Gly	CAA Gln	GAA Glu 425	CTC Leu	ACA Thr	1344
25	GGC Gly	CAG Gln	GAA Glu 430	TGG Trp	AGA Arg	GGT Gly	CTA Leu	GAA Glu 435	AAT Asn	ATT Ile	TTC Phe	GAA Glu	ATC Ile 440	TAT Tyr	CTT Leu	TCC Ser	1392
	TAC Tyr	AAC Asn 445	AAG Lys	TAC Tyr	CTG Leu	CAG Gln	CTG Leu 450	ACT Thr	AGG Arg	AAC Asn	TCC Ser	TTT Phe 455	GCC Ala	TTG Leu	GTC Val	CCA Pro	1440
30	AGC Ser 460	CTT Leu	CAA Gln	CGA Arg	CTG Leu	ATG Met 465	CTC Leu	CGA Arg	AGG Arg	GTG Val	GCC Ala 470	CTT Leu	AAA Lys	AAT Asn	GTG Val	GAT Asp 475	1488
35	AGC Ser	TCT Ser	CCT Pro	TCA Ser	CCA Pro 480	TTC Phe	CAG Gln	CCT Pro	CTT Leu	CGT Arg 485	AAC Asn	TTG Leu	ACC Thr	ATT Ile	CTG Leu 490	GAT Asp	1536
	CTA Leu	AGC Ser	AAC Asn	AAC Asn 495	ATA Ile	GCC Ala	AAC Asn	ATA Ile 500	AAT Asn	GAT Asp	GAC Asp	ATG Met	TTG Leu 505	GAG Glu	GGT Gly		1584
40	CTT Leu	GAG Glu	AAA Lys 510	CTA Leu	GAA Glu	ATT Ile	CTC Leu	GAT Asp 515	TTG Leu	CAG Gln	CAT His	AAC Asn	AAC Asn 520	TTA Leu	GCA Ala	CGG Arg	1632
45	CTC Leu 525	TGG Trp	AAA Lys	CAC His	GCA Ala	AAC Asn	CCT Pro 530	GGT Gly	GGT Gly	CCC Pro	ATT Ile	TAT Tyr 535	TTC Phe	CTA Leu	AAG Lys	GGT Gly	1680
	CTG Leu 540	TCT Ser	CAC His	CTC Leu	CAC His	ATC Ile 545	CTT Leu	AAC Asn	TTG Leu	GAG Glu	TCC Ser 550	AAC Asn	GGC Gly	TTT Phe	GAC Asp	GAG Glu 555	1728
50	ATC Ile	CCA Pro	GTT Val	GAG Glu 560	GTC Val	TTC Phe	AAG Lys	GAT Asp	TTA Leu	TTT Phe 565	GAA Glu	CTA Leu	AAG Lys	ATC Ile 570	ATC Ile	GAT Asp	1776
55	TTA Leu	GGA Gly	TTG Leu	AAT Asn 575	AAT Asn	TTA Leu	AAC Asn	ACA Thr	CTT Leu 580	CCA Pro	GCA Ala	TCT Ser	GTC Val	TTT Phe 585	AAT Asn	AAT Asn	1824
	CAG Glu	GTG Val	TCT Ser	CTA Leu	AAG Lys	TCA Ser	TTG Leu	AAC Asn	CTT Leu	CAG Glu	AAG Lys	AAT Asn	CTC Leu	ATA Ile	ACA Thr	TCC Ser	1872

# ES 2 340 210 T3

	Gln	Val	Ser	Leu	Lys	Ser	Leu	Asn	Leu	Gln	Lys	Asn	Leu	Ile	Thr	Ser	
			590					595					600				
5	GTT	GAG	AAG	AAG	GTT	TTC	GGG	CCA	GCT	TTC	AGG	AAC	CTG	ACT	GAG	TTA	1920
	Val	Glu	Lys	Lys	Val	Phe	Gly	Pro	Ala	Phe	Arg	Asn	Leu	Thr	Glu	Leu	
		605					610					615					
	GAT	ATG	CGC	TTT	AAT	CCC	TTT	GAT	TGC	ACG	TGT	GAA	AGT	ATT	GCC	TGG	1968
	Asp	Met	Arg	Phe	Asn	Pro	Phe	Asp	Cys	Thr	Cys	Glu	Ser	Ile	Ala	Trp	
		620				625					630					635	
10	TTT	GTT	AAT	TGG	ATT	AAC	GAG	ACC	CAT	ACC	AAC	ATC	CCT	GAG	CTG	TCA	2016
	Phe	Val	Asn	Trp	Ile	Asn	Glu	Thr	His	Thr	Asn	Ile	Pro	Glu	Leu	Ser	
					640					645					650		
15	AGC	CAC	TAC	CTT	TGC	AAC	ACT	CCA	CCT	CAC	TAT	CAT	GGG	TTT	CCA	GTG	2064
	Ser	His	Tyr	Leu	Cys	Asn	Thr	Pro	Pro	His	Tyr	His	Gly	Phe	Pro	Val	
				655					660					665			
	AGA	CTT	TTT	GAT	ACA	TCA	TCT	TGC	AAA	GAC	AGT	GCC	CCC	TTT	GAA	CTC	2112
	Arg	Leu	Phe	Asp	Thr	Ser	Ser	Cys	Lys	Asp	Ser	Ala	Pro	Phe	Glu	Leu	
			670					675					680				
20	TTT	TTC	ATG	ATC	AAT	ACC	AGT	ATC	CTG	TTG	ATT	TTT	ATC	TTT	ATT	GTA	2160
	Phe	Phe	Met	Ile	Asn	Thr	Ser	Ile	Leu	Leu	Ile	Phe	Ile	Phe	Ile	Val	
		685					690					695					
25	CTT	CTC	ATC	CAC	TTT	GAG	GGC	TGG	AGG	ATA	TCT	TTT	TAT	TGG	AAT	GTT	2208
	Leu	Leu	Ile	His	Phe	Glu	Gly	Trp	Arg	Ile	Ser	Phe	Tyr	Trp	Asn	Val	
					705						710					715	
	TCA	GTA	CAT	CGA	GTT	CTT	GGT	TTC	AAA	GAA	ATA	GAC	AGA	CAG	ACA	GAA	2256
	Ser	Val	His	Arg	Val	Leu	Gly	Phe	Lys	Glu	Ile	Asp	Arg	Gln	Thr	Glu	
					720					725				730			
30	CAG	TTT	GAA	TAT	GCA	GCA	TAT	ATA	ATT	CAT	GCC	TAT	AAA	GAT	AAG	GAT	2304
	Gln	Phe	Glu	Tyr	Ala	Ala	Tyr	Ile	Ile	His	Ala	Tyr	Lys	Asp	Lys	Asp	
				735				740						745			
	TGG	GTC	TGG	GAA	CAT	TTC	TCT	TCA	ATG	GAA	AAG	GAA	GAC	CAA	TCT	CTC	2352
	Trp	Val	Trp	Glu	His	Phe	Ser	Ser	Met	Glu	Lys	Glu	Asp	Gln	Ser	Leu	
			750					755					760				
35	AAA	TTT	TGT	CTG	GAA	GAA	AGG	GAC	TTT	GAG	GCG	GGT	GTT	TTT	GAA	CTA	2400
	Lys	Phe	Cys	Leu	Glu	Glu	Arg	Asp	Phe	Glu	Ala	Gly	Val	Phe	Glu	Leu	
		765					770					775					
40	GAA	GCA	ATT	GTT	AAC	AGC	ATC	AAA	AGA	AGC	AGA	AAA	ATT	ATT	TTT	GTT	2448
	Glu	Ala	Ile	Val	Asn	Ser	Ile	Lys	Arg	Ser	Arg	Lys	Ile	Ile	Phe	Val	
		780				785					790					795	
	ATA	ACA	CAC	CAT	CTA	TTA	AAA	GAC	CCA	TTA	TGC	AAA	AGA	TTC	AAG	GTA	2496
	Ile	Thr	His	His	Leu	Leu	Lys	Asp	Pro	Leu	Cys	Lys	Arg	Phe	Lys	Val	
					800					805					810		
	CAT	CAT	GCA	GTT	CAA	CAA	GCT	ATT	GAA	CAA	AAT	CTG	GAT	TCC	ATT	ATA	2544
	His	His	Ala	Val	Gln	Gln	Ala	Ile	Glu	Gln	Asn	Leu	Asp	Ser	Ile	Ile	
				815					820					825			
50	TTG	GTT	TTC	CTT	GAG	GAG	ATT	CCA	GAT	TAT	AAA	CTG	AAC	CAT	GCA	CTC	2592
	Leu	Val	Phe	Leu	Glu	Glu	Ile	Pro	Asp	Tyr	Lys	Leu	Asn	His	Ala	Leu	
			830					835					840				
	TGT	TTG	CGA	AGA	GGA	ATG	TTT	AAA	TCT	CAC	TGC	ATC	TTG	AAC	TGG	CCA	2640
	Cys	Leu	Arg	Arg	Gly	Met	Phe	Lys	Ser	His	Cys	Ile	Leu	Asn	Trp	Pro	
		845					850					855					
55	GTT	CAG	AAA	GAA	CGG	ATA	GGT	GCC	TTT	CGT	CAT	AAA	TTG	CAA	GTA	GCA	2688
60	Val	Gln	Lys	Glu	Arg	Ile	Gly	Ala	Phe	Arg	His	Lys	Leu	Gln	Val	Ala	
						865					870					875	
	CTT	GGA	TCC	AAA	AAC	TCT	GTA	CAT	TAA								2715
	Leu	Gly	Ser	Lys	Asn	Ser	Val	His									
					880												
65																	



## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:6:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 904 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:6:

```

Met Arg Gln Thr Leu Pro Cys Ile Tyr Phe Trp Gly Gly Leu Leu Pro
-21 -20 -15 -10
Phe Gly Met Leu Cys Ala Ser Ser Thr Thr Lys Cys Thr Val Ser His
-5 1 5 10
Glu Val Ala Asp Cys Ser His Leu Lys Leu Thr Gln Val Pro Asp Asp
15 20 25
Leu Pro Thr Asn Ile Thr Val Leu Asn Leu Thr His Asn Gln Leu Arg
30 35 40
Arg Leu Pro Ala Ala Asn Phe Thr Arg Tyr Ser Gln Leu Thr Ser Leu
45 50 55
Asp Val Gly Phe Asn Thr Ile Ser Lys Leu Glu Pro Glu Leu Cys Gln
60 65 70 75
Lys Leu Pro Met Leu Lys Val Leu Asn Leu Gln His Asn Glu Leu Ser
80 85 90
Gln Leu Ser Asp Lys Thr Phe Ala Phe Cys Thr Asn Leu Thr Glu Leu
95 100 105
His Leu Met Ser Asn Ser Ile Gln Lys Ile Lys Asn Asn Pro Phe Val
110 115 120
Lys Gln Lys Asn Leu Ile Thr Leu Asp Leu Ser His Asn Gly Leu Ser
125 130 135
Ser Thr Lys Leu Gly Thr Gln Val Gln Leu Glu Asn Leu Gln Glu Leu
140 145 150 155
Leu Leu Ser Asn Asn Lys Ile Gln Ala Leu Lys Ser Glu Glu Leu Asp
160 165 170
Ile Phe Ala Asn Ser Ser Leu Lys Lys Leu Glu Leu Ser Ser Asn Gln
175 180 185
Ile Lys Glu Phe Ser Pro Gly Cys Phe His Ala Ile Gly Arg Leu Phe
190 195 200
Gly Leu Phe Leu Asn Asn Val Gln Leu Gly Pro Ser Leu Thr Glu Lys
205 210 215
Leu Cys Leu Glu Leu Ala Asn Thr Ser Ile Arg Asn Leu Ser Leu Ser
220 225 230 235
Asn Ser Gln Leu Ser Thr Thr Ser Asn Thr Thr Phe Leu Gly Leu Lys

```

# ES 2 340 210 T3

	240										245										250											
5	Trp	Thr	Asn	Leu	Thr	Met	Leu	Asp	Leu	Ser	Tyr	Asn	Asn	Leu	Asn	Val																
				255					260																							
	Val	Gly	Asn	Asp	Ser	Phe	Ala	Trp	Leu	Pro	Gln	Leu	Glu	Tyr	Phe	Phe																
			270					275					280																			
	Leu	Glu	Tyr	Asn	Asn	Ile	Gln	His	Leu	Phe	Ser	His	Ser	Leu	His	Gly																
		285					290					295																				
10	Leu	Phe	Asn	Val	Arg	Tyr	Leu	Asn	Leu	Lys	Arg	Ser	Phe	Thr	Lys	Gln																
	300					305					310					315																
	Ser	Ile	Ser	Leu	Ala	Ser	Leu	Pro	Lys	Ile	Asp	Asp	Phe	Ser	Phe	Gln																
					320					325						330																
15	Trp	Leu	Lys	Cys	Leu	Glu	His	Leu	Asn	Met	Glu	Asp	Asn	Asp	Ile	Pro																
				335					340																							
	Gly	Ile	Lys	Ser	Asn	Met	Phe	Thr	Gly	Leu	Ile	Asn	Leu	Lys	Tyr	Leu																
			350					355					360																			
20	Ser	Leu	Ser	Asn	Ser	Phe	Thr	Ser	Leu	Arg	Thr	Leu	Thr	Asn	Glu	Thr																
		365					370					375																				
	Phe	Val	Ser	Leu	Ala	His	Ser	Pro	Leu	His	Ile	Leu	Asn	Leu	Thr	Lys																
	380					385					390					395																
25	Asn	Lys	Ile	Ser	Lys	Ile	Glu	Ser	Asp	Ala	Phe	Ser	Trp	Leu	Gly	His																
					400					405					410																	
	Leu	Glu	Val	Leu	Asp	Leu	Gly	Leu	Asn	Glu	Ile	Gly	Gln	Glu	Leu	Thr																
				415					420					425																		
30	Gly	Gln	Glu	Trp	Arg	Gly	Leu	Glu	Asn	Ile	Phe	Glu	Ile	Tyr	Leu	Ser																
			430					435					440																			
	Tyr	Asn	Lys	Tyr	Leu	Gln	Leu	Thr	Arg	Asn	Ser	Phe	Ala	Leu	Val	Pro																
		445					450					455																				
35	Ser	Leu	Gln	Arg	Leu	Met	Leu	Arg	Arg	Val	Ala	Leu	Lys	Asn	Val	Asp																
	460					465					470					475																
	Ser	Ser	Pro	Ser	Pro	Phe	Gln	Pro	Leu	Arg	Asn	Leu	Thr	Ile	Leu	Asp																
					480					485					490																	
40	Leu	Ser	Asn	Asn	Asn	Ile	Ala	Asn	Ile	Asn	Asp	Asp	Met	Leu	Glu	Gly																
				495					500					505																		
	Leu	Glu	Lys	Leu	Glu	Ile	Leu	Asp	Leu	Gln	His	Asn	Asn	Leu	Ala	Arg																
			510					515					520																			
45	Leu	Trp	Lys	His	Ala	Asn	Pro	Gly	Gly	Pro	Ile	Tyr	Phe	Leu	Lys	Gly																
		525					530					535																				
	Leu	Ser	His	Leu	His	Ile	Leu	Asn	Leu	Glu	Ser	Asn	Gly	Phe	Asp	Glu																
		540				545					550					555																
50	Ile	Pro	Val	Glu	Val	Phe	Lys	Asp	Leu	Phe	Glu	Leu	Lys	Ile	Ile	Asp																
					560					565					570																	
	Leu	Gly	Leu	Asn	Asn	Leu	Asn	Thr	Leu	Pro	Ala	Ser	Val	Phe	Asn	Asn																
				575					580					585																		
55	Gln	Val	Ser	Leu	Lys	Ser	Leu	Asn	Leu	Gln	Lys	Asn	Leu	Ile	Thr	Ser																
			590					595					600																			

## ES 2 340 210 T3

	Val	Glu	Lys	Lys	Val	Phe	Gly	Pro	Ala	Phe	Arg	Asn	Leu	Thr	Glu	Leu
		605					610					615				
5	Asp	Met	Arg	Phe	Asn	Pro	Phe	Asp	Cys	Thr	Cys	Glu	Ser	Ile	Ala	Trp
	620					625					630					635
	Phe	Val	Asn	Trp	Ile	Asn	Glu	Thr	His	Thr	Asn	Ile	Pro	Glu	Leu	Ser
					640					645					650	
10	Ser	His	Tyr	Leu	Cys	Asn	Thr	Pro	Pro	His	Tyr	His	Gly	Phe	Pro	Val
				655					660					665		
	Arg	Leu	Phe	Asp	Thr	Ser	Ser	Cys	Lys	Asp	Ser	Ala	Pro	Phe	Glu	Leu
			670					675					680			
15	Phe	Phe	Met	Ile	Asn	Thr	Ser	Ile	Leu	Leu	Ile	Phe	Ile	Phe	Ile	Val
		685					690					695				
	Leu	Leu	Ile	His	Phe	Glu	Gly	Trp	Arg	Ile	Ser	Phe	Tyr	Trp	Asn	Val
	700					705					710					715
20	Ser	Val	His	Arg	Val	Leu	Gly	Phe	Lys	Glu	Ile	Asp	Arg	Gln	Thr	Glu
					720					725					730	
	Gln	Phe	Glu	Tyr	Ala	Ala	Tyr	Ile	Ile	His	Ala	Tyr	Lys	Asp	Lys	Asp
				735					740					745		
25	Trp	Val	Trp	Glu	His	Phe	Ser	Ser	Met	Glu	Lys	Glu	Asp	Gln	Ser	Leu
			750					755					760			
	Lys	Phe	Cys	Leu	Glu	Glu	Arg	Asp	Phe	Glu	Ala	Gly	Val	Phe	Glu	Leu
		765					770					775				
30	Glu	Ala	Ile	Val	Asn	Ser	Ile	Lys	Arg	Ser	Arg	Lys	Ile	Ile	Phe	Val
	780					785					790					795
	Ile	Thr	His	His	Leu	Leu	Lys	Asp	Pro	Leu	Cys	Lys	Arg	Phe	Lys	Val
					800					805					810	
35	His	His	Ala	Val	Gln	Gln	Ala	Ile	Glu	Gln	Asn	Leu	Asp	Ser	Ile	Ile
				815					820					825		
	Leu	Val	Phe	Leu	Glu	Glu	Ile	Pro	Asp	Tyr	Lys	Leu	Asn	His	Ala	Leu
			830					835					840			
40	Cys	Leu	Arg	Arg	Gly	Met	Phe	Lys	Ser	His	Cys	Ile	Leu	Asn	Trp	Pro
		845					850					855				
	Val	Gln	Lys	Glu	Arg	Ile	Gly	Ala	Phe	Arg	His	Lys	Leu	Gln	Val	Ala
	860					865				870						875
45	Leu	Gly	Ser	Lys	Asn	Ser	Val	His								
					880											

50 (2) INFORMACIÓN PARA EL SEQ ID NO:7:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 55 (A) LONGITUD: 2400 pares de bases  
 (B) TIPO: ácido nucleico  
 (C) CUALIDAD DE LA HEBRA: sencilla  
 (D) TOPOLOGÍA: lineal

60 (ii) TIPO DE MOLÉCULA: ADNc

(ix) RASGO:

- (A) NOMBRE/CLAVE: CDS  
 65 (B) LOCALIZACIÓN: 1..2397

# ES 2 340 210 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:7:

5	ATG Met 1	GAG Glu	CTG Leu	AAT Asn	TTC Phe 5	TAC Tyr	AAA Lys	ATC Ile	CCC Pro	GAC Asp 10	AAC Asn	CTC Leu	CCC Pro	TTC Phe 15	TCA Ser	ACC Thr	48
	AAG Lys	AAC Asn	CTG Leu	GAC Asp 20	CTG Leu	AGC Ser	TTT Phe	AAT Asn	CCC Pro 25	CTG Leu	AGG Arg	CAT His	TTA Leu	GGC Gly 30	AGC Ser	TAT Tyr	96
10	AGC Ser	TTC Phe	TTC Phe 35	AGT Ser	TTC Phe	CCA Pro	GAA Glu	CTG Leu 40	CAG Gln	GTG Val	CTG Leu	GAT Asp	TTA Leu 45	TCC Ser	AGG Arg	TGT Cys	144
15	GAA Glu 50	ATC Ile	CAG Gln	ACA Thr	ATT Ile	GAA Glu	GAT Asp 55	GGG Gly	GCA Ala	TAT Tyr	CAG Gln	AGC Ser 60	CTA Leu	AGC Ser	CAC His	CTC Leu	192
	TCT Ser 65	ACC Thr	TTA Leu	ATA Ile	TTG Leu	ACA Thr 70	GGA Gly	AAC Asn	CCC Pro	ATC Ile	CAG Gln 75	AGT Ser	TTA Leu	GCC Ala	CTG Leu	GGA Gly 80	240
20	GCC Ala	TTT Phe	TCT Ser	GGA Gly	CTA Leu 85	TCA Ser	AGT Ser	TTA Leu	CAG Gln 90	AAG Lys	CTG Leu	GTG Val	GCT Ala	GTG Val	GAG Glu 95	ACA Thr	288
25	AAT Asn	CTA Leu	GCA Ala	TCT Ser 100	CTA Leu	GAG Glu	AAC Asn	TTT Phe	CCC Pro 105	ATT Ile	GGA Gly	CAT His	CTC Leu	AAA Lys 110	ACT Thr	TTG Leu	336
	AAA Lys	GAA Glu	CTT Leu 115	AAT Asn	GTG Val	GCT Ala	CAC His	AAT Asn 120	CTT Leu	ATC Ile	CAA Gln	TCT Ser	TTT Phe 125	AAA Lys	TTA Leu	CCT Pro	384
30	GAG Glu 130	TAT Tyr	TTT Phe	TCT Ser	AAT Asn	CTG Leu	ACC Thr 135	AAT Asn	CTA Leu	GAG Glu	CAC His	TTG Leu 140	GAC Asp	CTT Leu	TCC Ser	AGC Ser	432
35	AAC Asn 145	AAG Lys	ATT Ile	CAA Gln	AGT Ser	ATT Ile 150	TAT Tyr	TGC Cys	ACA Thr	GAC Asp 155	TTG Leu	CGG Arg	GTT Val	CTA Leu	CAT His	CAA Gln 160	480
	ATG Met	CCC Pro	CTA Leu	CTC Leu	AAT Asn 165	CTC Leu	TCT Ser	TTA Leu	GAC Asp 170	CTG Leu	TCC Ser	CTG Leu	AAC Asn	CCT Pro	ATG Met 175	AAC Asn	528
40	TTT Phe	ATC Ile	CAA Gln	CCA Pro 180	GGT Gly	GCA Ala	TTT Phe	AAA Lys	GAA Glu 185	ATT Ile	AGG Arg	CTT Leu	CAT His	AAG Lys 190	CTG Leu	ACT Thr	576
	TTA Leu	AGA Arg	AAT Asn 195	AAT Asn	TTT Phe	GAT Asp	AGT Ser	TTA Leu 200	AAT Asn	GTA Val	ATG Met	AAA Lys	ACT Thr 205	TGT Cys	ATT Ile	CAA Gln	624
45	GGT Gly 210	CTG Leu	GCT Ala	GGT Gly	TTA Leu	GAA Glu	GTC Val 215	CAT His	CGT Arg	TTG Leu	GTT Val	CTG Leu 220	GGA Gly	GAA Glu	TTT Phe	AGA Arg	672
50	AAT Asn 225	GAA Glu	GGA Gly	AAC Asn	TTG Leu	GAA Glu 230	AAG Lys	TTT Phe	GAC Asp	AAA Lys 235	TCT Ser	GCT Ala	CTA Leu	GAG Glu	GGC Gly	CTG Leu 240	720
	TGC Cys	AAT Asn	TTG Leu	ACC Thr	ATT Ile 245	GAA Glu	GAA Glu	TTT Phe	CGA Arg	TTA Leu 250	GCA Ala	TAC Tyr	TTA Leu	GAC Asp	TAC Tyr 255	TAC Tyr	768

# ES 2 340 210 T3

	CTC	GAT	GAT	ATT	ATT	GAC	TTA	TTT	AAT	TGT	TTG	ACA	AAT	GTT	TCT	TCA	816
	Leu	Asp	Asp	Ile	Ile	Asp	Leu	Phe	Asn	Cys	Leu	Thr	Asn	Val	Ser	Ser	
				260					265					270			
5	TTT	TCC	CTG	GTG	AGT	GTG	ACT	ATT	GAA	AGG	GTA	AAA	GAC	TTT	TCT	TAT	864
	Phe	Ser	Leu	Val	Ser	Val	Thr	Ile	Glu	Arg	Val	Lys	Asp	Phe	Ser	Tyr	
			275					280					285				
10	AAT	TTC	GGA	TGG	CAA	CAT	TTA	GAA	TTA	GTT	AAC	TGT	AAA	TTT	GGA	CAG	912
	Asn	Phe	Gly	Trp	Gln	His	Leu	Glu	Leu	Val	Asn	Cys	Lys	Phe	Gly	Gln	
		290					295					300					
	TTT	CCC	ACA	TTG	AAA	CTC	AAA	TCT	CTC	AAA	AGG	CTT	ACT	TTC	ACT	TCC	960
	Phe	Pro	Thr	Ile	Lys	Leu	Lys	Ser	Leu	Lys	Arg	Leu	Thr	Phe	Thr	Ser	
	305					310					315					320	
15	AAC	AAA	GGT	GGG	AAT	GCT	TTT	TCA	GAA	GTT	GAT	CTA	CCA	AGC	CTT	GAG	1008
	Asn	Lys	Gly	Gly	Asn	Ala	Phe	Ser	Glu	Val	Asp	Leu	Pro	Ser	Leu	Glu	
					325					330					335		
20	TTT	CTA	GAT	CTC	AGT	AGA	AAT	GGC	TTG	AGT	TTC	AAA	GGT	TGC	TGT	TCT	1056
	Phe	Leu	Asp	Leu	Ser	Arg	Asn	Gly	Leu	Ser	Phe	Lys	Gly	Cys	Cys	Ser	
				340				345					350				
	CAA	AGT	GAT	TTT	GGG	ACA	ACC	AGC	CTA	AAG	TAT	TTA	GAT	CTG	AGC	TTC	1104
	Gln	Ser	Asp	Phe	Gly	Thr	Thr	Ser	Leu	Lys	Tyr	Leu	Asp	Leu	Ser	Phe	
			355					360					365				
25	AAT	GGT	GTT	ATT	ACC	ATG	AGT	TCA	AAC	TTC	TTG	GGC	TTA	GAA	CAA	CTA	1152
	Asn	Gly	Val	Ile	Thr	Met	Ser	Ser	Asn	Phe	Leu	Gly	Leu	Glu	Gln	Leu	
		370					375					380					
30	GAA	CAT	CTG	GAT	TTC	CAG	CAT	TCC	AAT	TTG	AAA	CAA	ATG	AGT	GAG	TTT	1200
	Glu	His	Leu	Asp	Phe	Gln	His	Ser	Asn	Leu	Lys	Gln	Met	Ser	Glu	Phe	
	385					390					395					400	
	TCA	GTA	TTC	CTA	TCA	CTC	AGA	AAC	CTC	ATT	TAC	CTT	GAC	ATT	TCT	CAT	1248
	Ser	Val	Phe	Leu	Ser	Leu	Arg	Asn	Leu	Ile	Tyr	Leu	Asp	Ile	Ser	His	
					405					410					415		
35	ACT	CAC	ACC	AGA	GTT	GCT	TTC	AAT	GGC	ATC	TTC	AAT	GGC	TTG	TCC	AGT	1296
	Thr	His	Thr	Arg	Val	Ala	Phe	Asn	Gly	Ile	Phe	Asn	Gly	Leu	Ser	Ser	
				420					425					430			
40	CTC	GAA	GTC	TTG	AAA	ATG	GCT	GGC	AAT	TCT	TTC	CAG	GAA	AAC	TTC	CTT	1344
	Leu	Glu	Val	Leu	Lys	Met	Ala	Gly	Asn	Ser	Phe	Gln	Glu	Asn	Phe	Leu	
			435					440					445				
	CCA	GAT	ATC	TTC	ACA	GAG	CTG	AGA	AAC	TTG	ACC	TTC	CTG	GAC	CTC	TCT	1392
	Pro	Asp	Ile	Phe	Thr	Glu	Leu	Arg	Asn	Leu	Thr	Phe	Leu	Asp	Leu	Ser	
		450					455					460					
45	CAG	TGT	CAA	CTG	GAG	CAG	TTG	TCT	CCA	ACA	GCA	TTT	AAC	TCA	CTC	TCC	1440
	Gln	Cys	Gln	Leu	Glu	Gln	Leu	Ser	Pro	Thr	Ala	Phe	Asn	Ser	Leu	Ser	
	465					470					475					480	
50	AGT	CTT	CAG	GTA	CTA	AAT	ATG	AGC	CAC	AAC	AAC	TTC	TTT	TCA	TTG	GAT	1488
	Ser	Leu	Gln	Val	Leu	Asn	Met	Ser	His	Asn	Asn	Phe	Phe	Ser	Leu	Asp	
					485					490					495		
	ACG	TTT	CCT	TAT	AAG	TGT	CTG	AAC	TCC	CTC	CAG	GTT	CTT	GAT	TAC	AGT	1536
	Thr	Phe	Pro	Tyr	Lys	Cys	Leu	Asn	Ser	Leu	Gln	Val	Leu	Asp	Tyr	Ser	
				500					505					510			
55	CTC	AAT	CAC	ATA	ATG	ACT	TCC	AAA	AAA	CAG	GAA	CTA	CAG	CAT	TTT	CCA	1584
	Leu	Asn	His	Ile	Met	Thr	Ser	Lys	Lys	Gln	Glu	Leu	Gln	His	Phe	Pro	
			515					520					525				
60																	
65																	

# ES 2 340 210 T3

	AGT Ser	AGT Ser 530	CTA Leu	GCT Ala	TTC Phe	TTA Leu	AAT Asn 535	CTT Leu	ACT Thr	CAG Gln	AAT Asn 540	GAC Asp 540	TTT Phe	GCT Ala	TGT Cys	ACT Thr	1632
5	TGT Cys 545	GAA Glu	CAC His	CAG Gln	AGT Ser	TTC Phe 550	CTG Leu	CAA Gln	TGG Trp	ATC Ile	AAG Lys 555	GAC Asp	CAG Gln	AGG Arg	CAG Gln	CTC Leu 560	1680
10	TTG Leu	GTG Val	GAA Glu	GTT Val	GAA Glu 565	CGA Arg	ATG Met	GAA Glu	TGT Cys	GCA Ala 570	ACA Thr	CCT Pro	TCA Ser	GAT Asp	AAG Lys 575	CAG Gln	1728
	GGC Gly	ATG Met	CCT Pro	GTG Val 580	CTG Leu	AGT Ser	TTG Leu	AAT Asn	ATC Ile 585	ACC Thr	TGT Cys	CAG Gln	ATG Met	AAT Asn 590	AAG Lys	ACC Thr	1776
15	ATC Ile 595	ATT Ile	GGT Gly	GTG Val	TGC Ser	GTC Val	CTC Leu	AGT Ser 600	GTG Val	CTT Leu	GTA Val	GTA Val	TCT Ser 605	GTT Val	GTA Val	GCA Ala	1824
20	GTT Val 610	CTG Leu	GTC Val	TAT Tyr	AAG Lys	TTC Phe	TAT Tyr 615	TTT Phe	CAC His	CTG Leu	ATG Met	CTT Leu 620	CTT Leu	GCT Ala	GGC Gly	TGC Cys	1872
	ATA Ile 625	AAG Lys	TAT Tyr	GGT Gly	AGA Arg	GGT Gly 630	GAA Glu	AAC Asn	ATC Ile	TAT Tyr	GAT Asp 635	GCC Ala	TTT Phe	GTT Val	ATC Ile	TAC Tyr 640	1920
25	TCA Ser	AGC Ser	CAG Gln	GAT Asp	GAG Glu 645	GAC Asp	TGG Trp	GTA Val	AGG Arg	AAT Asn 650	GAG Glu	CTA Leu	GTA Val	AAG Lys	AAT Asn 655	TTA Leu	1968
30	GAA Glu 660	GAA Glu	GGG Gly	GTG Val	CCT Pro	CCA Pro	TTT Phe	CAG Gln	CTC Leu 665	TGC Cys	CTT Leu	CAC His	TAC Tyr	AGA Arg 670	GAC Asp	TTT Phe	2016
	ATT Ile 675	CCC Pro	GGT Gly	GTG Val	GCC Ala	ATT Ile	GCT Ala	GCC Ala 680	AAC Asn	ATC Ile	ATC Ile	CAT His	GAA Glu 685	GGT Gly	TTC Phe	CAT His	2064
35	AAA Lys 690	AGC Ser	CGA Arg	AAG Lys	GTG Val	ATT Ile 695	GTT Val	GTG Val	TCC Ser	CAG Gln	CAC His 700	TTT Phe	ATC Ile	CAG Gln	AGC Ser		2112
40	CGC Arg 705	TGG Trp	TGT Cys	ATC Ile	TTT Phe	GAA Glu 710	TAT Tyr	GAG Glu	ATT Ile	GCT Ala 715	CAG Gln	ACC Thr	TGG Trp	CAG Gln	TTT Phe	CTG Leu 720	2160
	AGC Ser	AGT Ser	CGT Arg	GCT Ala	GGT Gly 725	ATC Ile	ATC Ile	TTC Phe	ATT Ile	GTC Val 730	CTG Leu	CAG Gln	AAG Lys	GTG Val	GAG Glu 735	AAG Lys	2208
45	ACC Thr	CTG Leu	CTC Leu	AGG Arg 740	CAG Gln	CAG Gln	GTG Val	GAG Glu	CTG Leu 745	TAC Tyr	CGC Arg	CTT Leu	CTC Leu	AGC Ser 750	AGG Arg	AAC Asn	2256
50	ACT Thr	TAC Tyr	CTG Leu 755	GAG Glu	TGG Trp	GAG Glu	GAC Asp	AGT Ser 760	GTC Val	CTG Leu	GGG Gly	CGG Arg	CAC His 765	ATC Ile	TTC Phe	TGG Trp	2304
	AGA Arg 770	CGA Arg	CTC Leu	AGA Arg	AAA Lys	GCC Ala	CTG Leu 775	CTG Leu	GAT Asp	GGT Gly	AAA Lys	TCA Ser 780	TGG Trp	AAT Asn	CCA Pro	GAA Glu	2352
55	GGA Gly 785	ACA Thr	GTG Val	GGT Gly	ACA Thr	GGA Gly 790	TGC Cys	AAT Asn	TGG Trp	CAG Gln	GAA Glu 795	GCA Ala	ACA Thr	TCT Ser	ATC Ile		2397
60	TGA																2400

## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:8:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 799 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:8:

```

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

```

# ES 2 340 210 T3

	Asn	Phe	Gly	Trp	Gln	His	Leu	Glu	Leu	Val	Asn	Cys	Lys	Phe	Gly	Gln
		290					295					300				
5	Phe	Pro	Thr	Leu	Lys	Leu	Lys	Ser	Leu	Lys	Arg	Leu	Thr	Phe	Thr	Ser
	305					310					315					320
	Asn	Lys	Gly	Gly	Asn	Ala	Phe	Ser	Glu	Val	Asp	Leu	Pro	Ser	Leu	Glu
					325					330					335	
10	Phe	Leu	Asp	Leu	Ser	Arg	Asn	Gly	Leu	Ser	Phe	Lys	Gly	Cys	Cys	Ser
				340					345					350		
	Gln	Ser	Asp	Phe	Gly	Thr	Thr	Ser	Leu	Lys	Tyr	Leu	Asp	Leu	Ser	Phe
			355					360					365			
15	Asn	Gly	Val	Ile	Thr	Met	Ser	Ser	Asn	Phe	Leu	Gly	Leu	Glu	Gln	Leu
	370						375					380				
	Glu	His	Leu	Asp	Phe	Gln	His	Ser	Asn	Leu	Lys	Gln	Met	Ser	Glu	Phe
	385					390					395					400
20	Ser	Val	Phe	Leu	Ser	Leu	Arg	Asn	Leu	Ile	Tyr	Leu	Asp	Ile	Ser	His
					405					410					415	
	Thr	His	Thr	Arg	Val	Ala	Phe	Asn	Gly	Ile	Phe	Asn	Gly	Leu	Ser	Ser
				420					425					430		
25	Leu	Glu	Val	Leu	Lys	Met	Ala	Gly	Asn	Ser	Phe	Gln	Glu	Asn	Phe	Leu
			435					440					445			
	Pro	Asp	Ile	Phe	Thr	Glu	Leu	Arg	Asn	Leu	Thr	Phe	Leu	Asp	Leu	Ser
		450					455					460				
30	Gln	Cys	Gln	Leu	Glu	Gln	Leu	Ser	Pro	Thr	Ala	Phe	Asn	Ser	Leu	Ser
	465					470					475					480
	Ser	Leu	Gln	Val	Leu	Asn	Met	Ser	His	Asn	Asn	Phe	Phe	Ser	Leu	Asp
					485					490					495	
35	Thr	Phe	Pro	Tyr	Lys	Cys	Leu	Asn	Ser	Leu	Gln	Val	Leu	Asp	Tyr	Ser
				500					505					510		
	Leu	Asn	His	Ile	Met	Thr	Ser	Lys	Lys	Gln	Glu	Leu	Gln	His	Phe	Pro
			515					520					525			
40	Ser	Ser	Leu	Ala	Phe	Leu	Asn	Leu	Thr	Gln	Asn	Asp	Phe	Ala	Cys	Thr
		530					535					540				
	Cys	Glu	His	Gln	Ser	Phe	Leu	Gln	Trp	Ile	Lys	Asp	Gln	Arg	Gln	Leu
	545					550					555					560
45	Leu	Val	Glu	Val	Glu	Arg	Met	Glu	Cys	Ala	Thr	Pro	Ser	Asp	Lys	Gln
					565					570					575	
	Gly	Met	Pro	Val	Leu	Ser	Leu	Asn	Ile	Thr	Cys	Gln	Met	Asn	Lys	Thr
				580					585					590		
50	Ile	Ile	Gly	Val	Ser	Val	Leu	Ser	Val	Leu	Val	Val	Ser	Val	Val	Ala
			595					600					605			
	Val	Leu	Val	Tyr	Lys	Phe	Tyr	Phe	His	Leu	Met	Leu	Leu	Ala	Gly	Cys
		610					615					620				
55	Ile	Lys	Tyr	Gly	Arg	Gly	Glu	Asn	Ile	Tyr	Asp	Ala	Phe	Val	Ile	Tyr
	625					630					635					640
	Ser	Ser	Gln	Asp	Glu	Asp	Trp	Val	Arg	Asn	Glu	Leu	Val	Lys	Asn	Leu
					645					650					655	



## ES 2 340 210 T3

		Glu	Glu	Gly	Val	Pro	Pro	Phe	Gln	Leu	Cys	Leu	His	Tyr	Arg	Asp	Phe
					660					665					670		
5		Ile	Pro	Gly	Val	Ala	Ile	Ala	Ala	Asn	Ile	Ile	His	Glu	Gly	Phe	His
				675					680					685			
		Lys	Ser	Arg	Lys	Val	Ile	Val	Val	Val	Ser	Gln	His	Phe	Ile	Gln	Ser
		690						695					700				
10		Arg	Trp	Cys	Ile	Phe	Glu	Tyr	Glu	Ile	Ala	Gln	Thr	Trp	Gln	Phe	Leu
		705					710					715					720
		Ser	Ser	Arg	Ala	Gly	Ile	Ile	Phe	Ile	Val	Leu	Gln	Lys	Val	Glu	Lys
						725					730					735	
15		Thr	Leu	Leu	Arg	Gln	Gln	Val	Glu	Leu	Tyr	Arg	Leu	Leu	Ser	Arg	Asn
					740					745					750		
		Thr	Tyr	Leu	Glu	Trp	Glu	Asp	Ser	Val	Leu	Gly	Arg	His	Ile	Phe	Trp
				755					760					765			
20		Arg	Arg	Leu	Arg	Lys	Ala	Leu	Leu	Asp	Gly	Lys	Ser	Trp	Asn	Pro	Glu
		770						775					780				
		Gly	Thr	Val	Gly	Thr	Gly	Cys	Asn	Trp	Gln	Glu	Ala	Thr	Ser	Ile	
		785					790					795					

### (2) INFORMACIÓN PARA EL SEQ ID NO:9:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1275 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..1095

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:9:

		TGT	TGG	GAT	GTT	TTT	GAG	GGA	CTT	TCT	CAT	CTT	CAA	GTT	CTG	TAT	TTG	
		Cys	Trp	Asp	Val	Phe	Glu	Gly	Leu	Ser	His	Leu	Gln	Val	Leu	Tyr	Leu	48
		1				5					10					15		
50		AAT	CAT	AAC	TAT	CTT	AAT	TCC	CTT	CCA	CCA	GGA	GTA	TTT	AGC	CAT	CTG	96
		Asn	His	Asn	Tyr	Leu	Asn	Ser	Leu	Pro	Pro	Gly	Val	Phe	Ser	His	Leu	
					20					25					30			
55		ACT	GCA	TTA	AGG	GGA	CTA	AGC	CTC	AAC	TCC	AAC	AGG	CTG	ACA	GTT	CTT	144
		Thr	Ala	Leu	Arg	Gly	Leu	Ser	Leu	Asn	Ser	Asn	Arg	Leu	Thr	Val	Leu	
				35					40					45				
		TCT	CAC	AAT	GAT	TTA	CCT	GCT	AAT	TTA	GAG	ATC	CTG	GAC	ATA	TCC	AGG	192
		Ser	His	Asn	Asp	Leu	Pro	Ala	Asn	Leu	Glu	Ile	Leu	Asp	Ile	Ser	Arg	
				50				55					60					
60		AAC	CAG	CTC	CTA	GCT	CCT	AAT	CCT	GAT	GTA	TTT	GTA	TCA	CTT	AGT	GTC	240
		Asn	Gln	Leu	Leu	Ala	Pro	Asn	Pro	Asp	Val	Phe	Val	Ser	Leu	Ser	Val	
		65					70					75					80	
65		TTG	GAT	ATA	ACT	CAT	AAC	AAG	TTC	ATT	TGT	GAA	TGT	GAA	CTT	AGC	ACT	288
		Leu	Asp	Ile	Thr	His	Asn	Lys	Phe	Ile	Cys	Glu	Cys	Glu	Leu	Ser	Thr	

# ES 2 340 210 T3

	85								90								95															
	TTT	ATC	AAT	TGG	CTT	AAT	CAC	ACC	AAT	GTC	ACT	ATA	GCT	GGG	CCT	CCT																336
	Phe	Ile	Asn	Trp	Leu	Asn	His	Thr	Asn	Val	Thr	Ile	Ala	Gly	Pro	Pro																
				100					105					110																		
5	GCA	GAC	ATA	TAT	TGT	GTG	TAC	CCT	GAC	TCG	TTC	TCT	GGG	GTT	TCC	CTC																384
	Ala	Asp	Ile	Tyr	Cys	Val	Tyr	Pro	Asp	Ser	Phe	Ser	Gly	Val	Ser	Leu																
			115					120					125																			
10	TTC	TCT	CTT	TCC	ACG	GAA	GGT	TGT	GAT	GAA	GAG	GAA	GTC	TTA	AAG	TCC																432
	Phe	Ser	Leu	Ser	Thr	Glu	Gly	Cys	Asp	Glu	Glu	Glu	Val	Leu	Lys	Ser																
		130					135					140																				
	CTA	AAG	TTC	TCC	CTT	TTC	ATT	GTA	TGC	ACT	GTC	ACT	CTG	ACT	CTG	TTC																480
	Leu	Lys	Phe	Ser	Leu	Phe	Ile	Val	Cys	Thr	Val	Thr	Leu	Thr	Leu	Phe																
		145				150					155					160																
15	CTC	ATG	ACC	ATC	CTC	ACA	GTC	ACA	AAG	TTC	CGG	GGC	TTC	TGT	TTT	ATC																528
	Leu	Met	Thr	Ile	Leu	Thr	Val	Thr	Lys	Phe	Arg	Gly	Phe	Cys	Phe	Ile																
				165					170					175																		
20	TGT	TAT	AAG	ACA	GCC	CAG	AGA	CTG	GTG	TTC	AAG	GAC	CAT	CCC	CAG	GGC																576
	Cys	Tyr	Lys	Thr	Ala	Gln	Arg	Leu	Val	Phe	Lys	Asp	His	Pro	Gln	Gly																
				180				185						190																		
	ACA	GAA	CCT	GAT	ATG	TAC	AAA	TAT	GAT	GCC	TAT	TTG	TGC	TTC	AGC	AGC																624
	Thr	Glu	Pro	Asp	Met	Tyr	Lys	Tyr	Asp	Ala	Tyr	Leu	Cys	Phe	Ser	Ser																
			195					200					205																			
25	AAA	GAC	TTC	ACA	TGG	GTG	CAG	AAT	GCT	TTG	CTC	AAA	CAC	CTG	GAC	ACT																672
	Lys	Asp	Phe	Thr	Trp	Val	Gln	Asn	Ala	Leu	Leu	Lys	His	Leu	Asp	Thr																
		210				215						220																				
30	CAA	TAC	AGT	GAC	CAA	AAC	AGA	TTC	AAC	CTG	TGC	TTT	GAA	GAA	AGA	GAC																720
	Gln	Tyr	Ser	Asp	Gln	Asn	Arg	Phe	Asn	Leu	Cys	Phe	Glu	Glu	Arg	Asp																
		225			230					235					240																	
	TTT	GTC	CCA	GGA	GAA	AAC	CGC	ATT	GCC	AAT	ATC	CAG	GAT	GCC	ATC	TGG																768
	Phe	Val	Pro	Gly	Glu	Asn	Arg	Ile	Ala	Asn	Ile	Gln	Asp	Ala	Ile	Trp																
				245					250					255																		
35	AAC	AGT	AGA	AAG	ATC	GTT	TGT	CTT	GTG	AGC	AGA	CAC	TTC	CTT	AGA	GAT																816
	Asn	Ser	Arg	Lys	Ile	Val	Cys	Leu	Val	Ser	Arg	His	Phe	Leu	Arg	Asp																
				260				265						270																		
40	GGC	TGG	TGC	CTT	GAA	GCC	TTC	AGT	TAT	GCC	CAG	GGC	AGG	TGC	TTA	TCT																864
	Gly	Trp	Cys	Leu	Glu	Ala	Phe	Ser	Tyr	Ala	Gln	Gly	Arg	Cys	Leu	Ser																
			275					280					285																			
	GAC	CTT	AAC	AGT	GCT	CTC	ATC	ATG	GTG	GTG	GTT	GGG	TCC	TTG	TCC	CAG																912
	Asp	Leu	Asn	Ser	Ala	Leu	Ile	Met	Val	Val	Val	Gly	Ser	Leu	Ser	Gln																
			290				295					300																				
45	TAC	CAG	TTG	ATG	AAA	CAT	CAA	TCC	ATC	AGA	GGC	TTT	GTA	CAG	AAA	CAG																960
	Tyr	Gln	Leu	Met	Lys	His	Gln	Ser	Ile	Arg	Gly	Phe	Val	Gln	Lys	Gln																
		305				310				315					320																	
	CAG	TAT	TTG	AGG	TGG	CCT	GAG	GAT	CTC	CAG	GAT	GTT	GGC	TGG	TTT	CTT																1008
	Gln	Tyr	Leu	Arg	Trp	Pro	Glu	Asp	Leu	Gln	Asp	Val	Gly	Trp	Phe	Leu																
				325				330						335																		
50	CAT	AAA	CTC	TCT	CAA	CAG	ATA	CTA	AAG	AAA	GAA	AAG	GAA	AAG	AAG	AAA																1056
	His	Lys	Leu	Ser	Gln	Gln	Ile	Leu	Lys	Lys	Glu	Lys	Glu	Lys	Lys	Lys																
			340					345					350																			
55	GAC	AAT	AAC	ATT	CCG	TTG	CAA	ACT	GTA	GCA	ACC	ATC	TCC	TAATCAAAGG																		1105
	Asp	Asn	Asn	Ile	Pro	Leu	Gln	Thr	Val	Ala	Thr	Ile	Ser																			
			355					360					365																			
60	AGCAATTTCC	AACTTATCTC AAGCCACAAA TAACTCTTCA CTTTGTATTT GCACCAAGTT															1165															
	ATCATTTTGG	GGTCCTCTCT GGAGGTTTTT TTTTCTTTT TGCTACTATG AAAACAACAT															1225															
	AAATCTCTCA	ATTTTCGTAT CAAAAAAAAA AAAAAAAAAA TGGCGGCCGC															1275															

## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:10:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 365 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:10:

```

Cys Trp Asp Val Phe Glu Gly Leu Ser His Leu Gln Val Leu Tyr Leu
 1      5      10
Asn His Asn Tyr Leu Asn Ser Leu Pro Pro Gly Val Phe Ser His Leu
 20      25
Thr Ala Leu Arg Gly Leu Ser Leu Asn Ser Asn Arg Leu Thr Val Leu
 35      40      45
Ser His Asn Asp Leu Pro Ala Asn Leu Glu Ile Leu Asp Ile Ser Arg
 50      55      60
Asn Gln Leu Leu Ala Pro Asn Pro Asp Val Phe Val Ser Leu Ser Val
 65      70      75      80
Leu Asp Ile Thr His Asn Lys Phe Ile Cys Glu Cys Glu Leu Ser Thr
 85      90      95
Phe Ile Asn Trp Leu Asn His Thr Asn Val Thr Ile Ala Gly Pro Pro
100      105      110
Ala Asp Ile Tyr Cys Val Tyr Pro Asp Ser Phe Ser Gly Val Ser Leu
115      120      125
Phe Ser Leu Ser Thr Glu Gly Cys Asp Glu Glu Glu Val Leu Lys Ser
130      135      140
Leu Lys Phe Ser Leu Phe Ile Val Cys Thr Val Thr Leu Thr Leu Phe
145      150      155      160
Leu Met Thr Ile Leu Thr Val Thr Lys Phe Arg Gly Phe Cys Phe Ile
165      170      175
Cys Tyr Lys Thr Ala Gln Arg Leu Val Phe Lys Asp His Pro Gln Gly
180      185      190
Thr Glu Pro Asp Met Tyr Lys Tyr Asp Ala Tyr Leu Cys Phe Ser Ser
195      200      205
Lys Asp Phe Thr Trp Val Gln Asn Ala Leu Leu Lys His Leu Asp Thr
210      215      220
Gln Tyr Ser Asp Gln Asn Arg Phe Asn Leu Cys Phe Glu Glu Arg Asp
225      230      235      240
Phe Val Pro Gly Glu Asn Arg Ile Ala Asn Ile Gln Asp Ala Ile Trp
245      250      255

```

## ES 2 340 210 T3

Asn Ser Arg Lys Ile Val Cys Leu Val Ser Arg His Phe Leu Arg Asp  
 260 265 270  
 Gly Trp Cys Leu Glu Ala Phe Ser Tyr Ala Gln Gly Arg Cys Leu Ser  
 275 280 285  
 Asp Leu Asn Ser Ala Leu Ile Met Val Val Val Gly Ser Leu Ser Gln  
 290 295 300  
 Tyr Gln Leu Met Lys His Gln Ser Ile Arg Gly Phe Val Gln Lys Gln  
 305 310 315 320  
 Gln Tyr Leu Arg Trp Pro Glu Asp Leu Gln Asp Val Gly Trp Phe Leu  
 325 330 335  
 His Lys Leu Ser Gln Gln Ile Leu Lys Lys Glu Lys Glu Lys Lys Lys  
 340 345 350  
 Asp Asn Asn Ile Pro Leu Gln Thr Val Ala Thr Ile Ser  
 355 360 365

20

### (2) INFORMACIÓN PARA EL SEQ ID NO:11:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 3138 pares de bases
  - (B) TIPO: ácido nucleico
  - (C) CUALIDAD DE LA HEBRA: sencilla
  - (D) TOPOLOGÍA: lineal

### (ii) TIPO DE MOLÉCULA: ADNc

- (ix) RASGO:
- (A) NOMBRE/CLAVE: CDS
  - (B) LOCALIZACIÓN: 1..3135

- (ix) RASGO:
- (A) NOMBRE/CLAVE: mat\_péptido
  - (B) LOCALIZACIÓN: 67..3135

### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:11:

ATG TGG ACA CTG AAG AGA CTA ATT CTT ATC CTT TTT AAC ATA ATC CTA 48  
 Met Trp Thr Leu Lys Arg Leu Ile Leu Ile Leu Phe Asn Ile Ile Leu  
 -22 -20 -15 -10  
 ATT TCC AAA CTC CTT GGG GCT AGA TGG TTT CCT AAA ACT CTG CCC TGT 96  
 Ile Ser Lys Leu Leu Gly Ala Arg Trp Phe Pro Lys Thr Leu Pro Cys  
 -5 1 5 10  
 GAT GTC ACT CTG GAT GTT CCA AAG AAC CAT GTG ATC GTG GAC TGC ACA 144  
 Asp Val Thr Leu Asp Val Pro Lys Asn His Val Ile Val Asp Cys Thr  
 15 20 25  
 GAC AAG CAT TTG ACA GAA ATT CCT GGA GGT ATT CCC ACG AAC ACC ACG 192  
 Asp Lys His Leu Thr Glu Ile Pro Gly Gly Ile Pro Thr Asn Thr Thr  
 30 35 40  
 AAC CTC ACC CTC ACC ATT AAC CAC ATA CCA GAC ATC TCC CCA GCG TCC 240  
 Asn Leu Thr Leu Thr Ile Asn His Ile Pro Asp Ile Ser Pro Ala Ser  
 45 50 55  
 TTT CAC AGA CTG GAC CAT CTG GTA GAG ATC GAT TTC AGA TGC AAC TGT 288  
 Phe His Arg Leu Asp His Leu Val Glu Ile Asp Phe Arg Cys Asn Cys  
 60 65 70  
 GTA CCT ATT CCA CTG GGG TCA AAA AAC AAC ATG TGC ATC AAG AGG CTG 336

65

# ES 2 340 210 T3

	Val 75	Pro	Ile	Pro	Leu	Gly 80	Ser	Lys	Asn	Asn	Met 85	Cys	Ile	Lys	Arg	Leu 90	
5	CAG Gln	ATT Ile	AAA Lys	CCC Pro	AGA Arg 95	AGC Ser	TTT Phe	AGT Ser	GGA Gly	CTC Leu 100	ACT Thr	TAT Tyr	TTA Leu	AAA Lys	TCC Ser 105	CTT Leu	384
	TAC Tyr	CTG Leu	GAT Asp	GGA Gly 110	AAC Asn	CAG Gln	CTA Leu	CTA Leu	GAG Glu 115	ATA Ile	CCG Pro	CAG Gln	GGC Gly	CTC Leu 120	CCG Pro	CCT Pro	432
10	AGC Ser	TTA Leu	CAG Gln 125	CTT Leu	CTC Leu	AGC Ser	CTT Leu	GAG Glu 130	GCC Ala	AAC Asn	AAC Asn	ATC Ile	TTT Phe 135	TCC Ser	ATC Ile	AGA Arg	480
15	AAA Lys	GAG Glu 140	AAT Asn	CTA Leu	ACA Thr	GAA Glu	CTG Leu 145	GCC Ala	AAC Asn	ATA Ile	GAA Glu 150	ATA Ile	CTC Leu	TAC Tyr	CTG Leu	GGC Gly	528
	CAA Gln 155	AAC Asn	TGT Cys	TAT Tyr	TAT Tyr	CGA Arg 160	AAT Asn	CCT Pro	TGT Cys	TAT Tyr	GTT Val 165	TCA Ser	TAT Tyr	TCA Ser	ATA Ile	GAG Glu 170	576
20	AAA Lys	GAT Asp	GCC Ala	TTC Phe	CTA Leu 175	AAC Asn	TTG Leu	ACA Thr	AAG Lys	TTA Leu 180	AAA Lys	GTG Val	CTC Leu	TCC Ser	CTG Leu 185	AAA Lys	624
25	GAT Asp	AAC Asn	AAT Asn	GTC Val 190	ACA Thr	GCC Ala	GTC Val	CCT Pro	ACT Thr 195	GTT Val	TTG Leu	CCA Pro	TCT Ser	ACT Thr 200	TTA Leu	ACA Thr	672
	GAA Glu	CTA Leu	TAT Tyr 205	CTC Leu	TAC Tyr	AAC Asn	AAC Asn	ATG Met 210	ATT Ile	GCA Ala	AAA Lys	ATC Ile	CAA Gln 215	GAA Glu	GAT Asp	GAT Asp	720
30	TTT Phe 220	AAT Asn	AAC Asn	CTC Leu	AAC Asn	CAA Gln	TTA Leu 225	CAA Gln	ATT Ile	CTT Leu	GAC Asp	CTA Leu 230	AGT Ser	GGA Gly	AAT Asn	TGC Cys	768
35	CCT Pro 235	CGT Arg	TGT Cys	TAT Tyr	AAT Asn	GCC Ala 240	CCA Pro	TTT Phe	CCT Pro	TGT Cys	GCG Ala 245	CCG Pro	TGT Cys	AAA Lys	AAT Asn	AAT Asn 250	816
	TCT Ser	CCC Pro	CTA Leu	CAG Gln	ATC Ile 255	CCT Pro	GTA Val	AAT Asn	GCT Ala	TTT Phe 260	GAT Asp	GCG Ala	CTG Leu	ACA Thr	GAA Glu 265	TTA Leu	864
40	AAA Lys	GTT Val	TTA Leu	CGT Arg 270	CTA Leu	CAC His	AGT Ser	AAC Asn	TCT Ser 275	CTT Leu	CAG Gln	CAT His	GTG Val	CCC Pro 280	CCA Pro	AGA Arg	912
45	TGG Trp	TTT Phe	AAG Lys 285	AAC Asn	ATC Ile	AAC Asn	AAA Lys	CTC Leu 290	CAG Gln	GAA Glu	CTG Leu	GAT Asp	CTG Leu 295	TCC Ser	CAA Gln	AAC Asn	960
	TTC Phe 300	TTG Leu	GCC Ala	AAA Lys	GAA Glu	ATT Ile	GGG Gly 305	GAT Asp	GCT Ala	AAA Lys	TTT Phe 310	CTG Leu	CAT His	TTT Phe	CTC Leu	CCC Pro	1008
50	AGC Ser 315	CTC Leu	ATC Ile	CAA Gln	TTG Leu	GAT Asp 320	CTG Leu	TCT Ser	TTC Phe	AAT Asn	TTT Phe 325	GAA Glu	CTT Leu	CAG Gln	GTC Val	TAT Tyr 330	1056
	CGT Arg	GCA Ala	TCT Ser	ATG Met	AAT Asn 335	CTA Leu	TCA Ser	CAA Gln	GCA Ala	TTT Phe 340	TCT Ser	TCA Ser	CTG Leu	AAA Lys	AGC Ser 345	CTG Leu	1104
55	AAA	ATT	CTG	CGG	ATC	AGA	GGA	TAT	GTC	TTT	AAA	GAG	TTG	AAA	AGC	TTT	1152

# ES 2 340 210 T3

	Lys	Ile	Leu	Arg 350	Ile	Arg	Gly	Tyr	Val 355	Phe	Lys	Glu	Leu	Lys 360	Ser	Phe	
5	AAC Asn	CTC Leu	TCG Ser 365	CCA Pro	TTA Leu	CAT His	AAT Asn	CTT Leu 370	CAA Gln	AAT Asn	CTT Leu	GAA Glu	GTT Val 375	CTT Leu	GAT Asp	CTT Leu	1200
	GGC Gly	ACT Thr 380	AAC Asn	TTT Phe	ATA Ile	AAA Lys	ATT Ile 385	GCT Ala	AAC Asn	CTC Leu	AGC Ser	ATG Met 390	TTT Phe	AAA Lys	CAA Gln	TTT Phe	1248
10	AAA Lys 395	AGA Arg	CTG Leu	AAA Lys	GTC Val	ATA Ile 400	GAT Asp	CTT Leu	TCA Ser	GTG Val	AAT Asn 405	AAA Lys	ATA Ile	TCA Ser	CCT Pro	TCA Ser 410	1296
15	GGA Gly	GAT Asp	TCA Ser	AGT Ser	GAA Glu 415	GTT Val	GGC Gly	TTC Phe	TGC Cys	TCA Ser 420	AAT Asn	GCC Ala	AGA Arg	ACT Thr	TCT Ser 425	GTA Val	1344
	GAA Glu	AGT Ser	TAT Tyr	GAA Pro 430	CCC Pro	CAG Gln	GTC Val	CTG Leu	GAA Glu 435	CAA Gln	TTA Leu	CAT His	TAT Tyr	TTT Phe 440	AGA Arg	TAT Tyr	1392
20	GAT Asp	AAG Lys	TAT Tyr 445	GCA Ala	AGG Arg	AGT Ser	TGC Cys	AGA Arg 450	TTC Phe	AAA Lys	AAC Asn	AAA Lys	GAG Glu 455	GCT Ala	TCT Ser	TTC Phe	1440
25	ATG Met 460	TCT Ser	GTT Val	AAT Asn	GAA Glu	AGC Ser	TGC Cys 465	TAC Tyr	AAG Lys	TAT Tyr	GGG Gly	CAG Gln 470	ACC Thr	TTG Leu	GAT Asp	CTA Leu	1488
	AGT Ser 475	AAA Lys	AAT Asn	AGT Ser	ATA Ile	TTT Phe 480	TTT Phe	GTC Val	AAG Lys	TCC Ser	TCT Ser 485	GAT Asp	TTT Phe	CAG Gln	CAT His	CTT Leu 490	1536
30	TCT Ser	TTC Phe	CTC Leu	AAA Lys	TGC Leu 495	CTG Leu	AAT Asn	CTG Leu	TCA Ser	GGA Gly 500	AAT Asn	CTC Leu	ATT Ile	AGC Ser	CAA Gln 505	ACT Thr	1584
35	CTT Leu	AAT Asn	GGC Gly	AGT Ser 510	GAA Glu	TTC Phe	CAA Gln	CCT Pro	TTA Leu 515	GCA Ala	GAG Glu	CTG Leu	AGA Arg	TAT Tyr 520	TTG Leu	GAC Asp	1632
	TTC Phe	TCC Ser	AAC Asn 525	AAC Asn	CGG Arg	CTT Leu	GAT Asp	TTA Leu 530	CTC Leu	CAT His	TCA Ser	ACA Thr	GCA Ala 535	TTT Phe	GAA Glu	GAG Glu	1680
40	CTT Leu	CAC His 540	AAA Lys	CTG Leu	GAA Glu	GTT Val	CTG Leu 545	GAT Asp	ATA Ile	AGC Ser	AGT Ser	AAT Asn 550	AGC Ser	CAT His	TAT Tyr	TTT Phe	1728
45	CAA Gln 555	TCA Ser	GAA Glu	GGA Gly	ATT Ile 560	ACT Thr	CAT His	ATG Met	CTA Leu	AAC Asn	TTT Phe 565	ACC Thr	AAG Lys	AAC Asn	CTA Leu	AAG Lys 570	1776
	GTT Val	CTG Leu	CAG Gln	AAA Lys	CTG Leu 575	ATG Met	ATG Met	AAC Asn	GAC Asp	AAT Asn 580	GAC Asp	ATC Ile	TCT Ser	TCC Ser	TCC Ser 585	ACC Thr	1824
50	AGC Ser	AGG Arg	ACC Thr	ATG Met 590	GAG Glu	AGT Ser	GAG Glu	TCT Ser	CTT Leu 595	AGA Arg	ACT Thr	CTG Leu	GAA Glu	TTC Phe 600	AGA Arg	GGA Gly	1872
	AAT Asn	CAC His	TTA Leu 605	GAT Asp	GTT Val	TTA Leu	TGG Trp 610	AGA Arg	GAA Glu	GGT Gly	GAT Asp	AAC Asn	AGA Arg 615	TAC Tyr	TTA Leu	CAA Gln	1920
55	TTA	TTC	AAG	AAT	CTG	CTA	AAA	TTA	GAG	GAA	TTA	GAC	ATC	TCT	AAA	AAT	1968

# ES 2 340 210 T3

	Leu	Phe	Lys	Asn	Leu	Leu	Lys	Leu	Glu	Glu	Leu	Asp	Ile	Ser	Lys	Asn	
	620						625					630					
5	TCC	CTA	AGT	TTC	TTG	CCT	TCT	GGA	GTT	TTT	GAT	GGT	ATG	CCT	CCA	AAT	2016
	Ser	Leu	Ser	Phe	Leu	Pro	Ser	Gly	Val	Phe	Asp	Gly	Met	Pro	Pro	Asn	
	635					640					645					650	
	CTA	AAG	AAT	CTC	TCT	TTG	GCC	AAA	AAT	GGG	CTC	AAA	TCT	TTC	AGT	TGG	2064
	Leu	Lys	Asn	Leu	Ser	Leu	Ala	Lys	Asn	Gly	Leu	Lys	Ser	Phe	Ser	Trp	
					655					660					665		
10	AAG	AAA	CTC	CAG	TGT	CTA	AAG	AAC	CTG	GAA	ACT	TTG	GAC	CTC	AGC	CAC	2112
	Lys	Lys	Leu	Gln	Cys	Leu	Lys	Asn	Leu	Glu	Thr	Leu	Asp	Leu	Ser	His	
				670					675					680			
15	AAC	CAA	CTG	ACC	ACT	GTC	CCT	GAG	AGA	TTA	TCC	AAC	TGT	TCC	AGA	AGC	2160
	Asn	Gln	Leu	Thr	Thr	Val	Pro	Glu	Arg	Leu	Ser	Asn	Cys	Ser	Arg	Ser	
			685					690					695				
	CTC	AAG	AAT	CTG	ATT	CTT	AAG	AAT	AAT	CAA	ATC	AGG	AGT	CTG	ACG	AAG	2208
	Leu	Lys	Asn	Leu	Ile	Leu	Lys	Asn	Asn	Gln	Ile	Arg	Ser	Leu	Thr	Lys	
		700					705					710					
20	TAT	TTT	CTA	CAA	GAT	GCC	TTC	CAG	TTG	CGA	TAT	CTG	GAT	CTC	AGC	TCA	2256
	Tyr	Phe	Leu	Gln	Asp	Ala	Phe	Gln	Leu	Arg	Tyr	Leu	Asp	Leu	Ser	Ser	
					720						725					730	
	AAT	AAA	ATC	CAG	ATG	ATC	CAA	AAG	ACC	AGC	TTC	CCA	GAA	AAT	GTC	CTC	2304
	Asn	Lys	Ile	Gln	Met	Ile	Gln	Lys	Thr	Ser	Phe	Pro	Glu	Asn	Val	Leu	
					735					740					745		
	AAC	AAT	CTG	AAG	ATG	TTG	CTT	TTG	CAT	CAT	AAT	CGG	TTT	CTG	TGC	ACC	2352
	Asn	Asn	Leu	Lys	Met	Leu	Leu	Leu	His	His	Asn	Arg	Phe	Leu	Cys	Thr	
				750					755					760			
30	TGT	GAT	GCT	GTG	TGG	TTT	GTC	TGG	TGG	GTT	AAC	CAT	ACG	GAG	GTG	ACT	2400
	Cys	Asp	Ala	Val	Trp	Phe	Val	Trp	Trp	Val	Asn	His	Thr	Glu	Val	Thr	
			765				770						775				
	ATT	CCT	TAC	CTG	GCC	ACA	GAT	GTG	ACT	TGT	GTG	GGG	CCA	GGA	GCA	CAC	2448
	Ile	Pro	Tyr	Leu	Ala	Thr	Asp	Val	Thr	Cys	Val	Gly	Pro	Gly	Ala	His	
		780					785					790					
	AAG	GGC	CAA	AGT	GTG	ATC	TCC	CTG	GAT	CTG	TAC	ACC	TGT	GAG	TTA	GAT	2496
	Lys	Gly	Gln	Ser	Val	Ile	Ser	Leu	Asp	Leu	Tyr	Thr	Cys	Glu	Leu	Asp	
		795				800					805					810	
40	CTG	ACT	AAC	CTG	ATT	CTG	TTC	TCA	CTT	TCC	ATA	TCT	GTA	TCT	CTC	TTT	2544
	Leu	Thr	Asn	Leu	Ile	Leu	Phe	Ser	Leu	Ser	Ile	Ser	Val	Ser	Leu	Phe	
				815						820					825		
	CTC	ATG	GTG	ATG	ATG	ACA	GCA	AGT	CAC	CTC	TAT	TTC	TGG	GAT	GTG	TGG	2592
	Leu	Met	Val	Met	Met	Thr	Ala	Ser	His	Leu	Tyr	Phe	Trp	Asp	Val	Trp	
				830					835					840			
	TAT	ATT	TAC	CAT	TTC	TGT	AAG	GCC	AAG	ATA	AAG	GGG	TAT	CAG	CGT	CTA	2640
	Tyr	Ile	Tyr	His	Phe	Cys	Lys	Ala	Lys	Ile	Lys	Gly	Tyr	Gln	Arg	Leu	
			845					850					855				
50	ATA	TCA	CCA	GAC	TGT	TGC	TAT	GAT	GCT	TTT	ATT	GTG	TAT	GAC	ACT	AAA	2688
	Ile	Ser	Pro	Asp	Cys	Cys	Tyr	Asp	Ala	Phe	Ile	Val	Tyr	Asp	Thr	Lys	
			860				865					870					
	GAC	CCA	GCT	GTG	ACC	GAG	TGG	GTT	TTG	GCT	GAG	CTG	GTG	GCC	AAA	CTG	2736
	Asp	Pro	Ala	Val	Thr	Glu	Trp	Val	Leu	Ala	Glu	Leu	Val	Ala	Lys	Leu	
						880					885					890	
55	GAA	GAC	CCA	AGA	GAG	AAA	CAT	TTT	AAT	TTA	TGT	CTC	GAG	GAA	AGG	GAC	2784

# ES 2 340 210 T3

	Glu	Asp	Pro	Arg	Glu	Lys	His	Phe	Asn	Leu	Cys	Leu	Glu	Glu	Arg	Asp	
					895					900					905		
5	TGG	TTA	CCA	GGG	CAG	CCA	GTT	CTG	GAA	AAC	CTT	TCC	CAG	AGC	ATA	CAG	2832
	Trp	Leu	Pro	Gly	Gln	Pro	Val	Leu	Glu	Asn	Leu	Ser	Gln	Ser	Ile	Gln	
				910					915					920			
	CTT	AGC	AAA	AAG	ACA	GTG	TTT	GTG	ATG	ACA	GAC	AAG	TAT	GCA	AAG	ACT	2880
10	Leu	Ser	Lys	Lys	Thr	Val	Phe	Val	Met	Thr	Asp	Lys	Tyr	Ala	Lys	Thr	
			925					930					935				
	GAA	AAT	TTT	AAG	ATA	GCA	TTT	TAC	TTG	TCC	CAT	CAG	AGG	CTC	ATG	GAT	2928
	Glu	Asn	Phe	Lys	Ile	Ala	Phe	Tyr	Leu	Ser	His	Gln	Arg	Leu	Met	Asp	
		940					945					950					
15	GAA	AAA	GTT	GAT	GTG	ATT	ATC	TTG	ATA	TTT	CTT	GAG	AAG	CCC	TTT	CAG	2976
	Glu	Lys	Val	Asp	Val	Ile	Ile	Leu	Ile	Phe	Leu	Glu	Lys	Pro	Phe	Gln	
	955					960					965					970	
	AAG	TCC	AAG	TTC	CTC	CAG	CTC	CGG	AAA	AGG	CTC	TGT	GGG	AGT	TCT	GTC	3024
20	Lys	Ser	Lys	Phe	Leu	Gln	Leu	Arg	Lys	Arg	Leu	Cys	Gly	Ser	Ser	Val	
					975					980					985		
	CTT	GAG	TGG	CCA	ACA	AAC	CCG	CAA	GCT	CAC	CCA	TAC	TTC	TGG	CAG	TGT	3072
	Leu	Glu	Trp	Pro	Thr	Asn	Pro	Gln	Ala	His	Pro	Tyr	Phe	Trp	Gln	Cys	
				990					995					1000			
25	CTA	AAG	AAC	GCC	CTG	GCC	ACA	GAC	AAT	CAT	GTG	GCC	TAT	AGT	CAG	GTG	3120
	Leu	Lys	Asn	Ala	Leu	Ala	Thr	Asp	Asn	His	Val	Ala	Tyr	Ser	Gln	Val	
			1005					1010					1015				
	TTC	AAG	GAA	ACG	GTC	TAG											3138
30	Phe	Lys	Glu	Thr	Val												
			1020														

## (2) INFORMACIÓN PARA EL SEQ ID NO:12:

### 35 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1045 aminoácidos

(B) TIPO: aminoácido

40 (D) TOPOLOGÍA: lineal

### (ii) TIPO DE MOLÉCULA: proteína

### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:12:

45	Met	Trp	Thr	Leu	Lys	Arg	Leu	Ile	Leu	Ile	Leu	Phe	Asn	Ile	Ile	Leu	
	-22		-20					-15					-10				
	Ile	Ser	Lys	Leu	Leu	Gly	Ala	Arg	Trp	Phe	Pro	Lys	Thr	Leu	Pro	Cys	
		-5					1				5					10	
50	Asp	Val	Thr	Leu	Asp	Val	Pro	Lys	Asn	His	Val	Ile	Val	Asp	Cys	Thr	
					15					20					25		
	Asp	Lys	His	Leu	Thr	Glu	Ile	Pro	Gly	Gly	Ile	Pro	Thr	Asn	Thr	Thr	
				30					35					40			
55	Asn	Leu	Thr	Leu	Thr	Ile	Asn	His	Ile	Pro	Asp	Ile	Ser	Pro	Ala	Ser	
			45					50					55				
	Phe	His	Arg	Leu	Asp	His	Leu	Val	Glu	Ile	Asp	Phe	Arg	Cys	Asn	Cys	
		60					65					70					
60	Val	Pro	Ile	Pro	Leu	Gly	Ser	Lys	Asn	Asn	Met	Cys	Ile	Lys	Arg	Leu	
		75				80					85					90	

65



# ES 2 340 210 T3

	Gln	Ile	Lys	Pro	Arg <sub>95</sub>	Ser	Phe	Ser	Gly	Leu <sub>100</sub>	Thr	Tyr	Leu	Lys	Ser <sub>105</sub>	Leu
5	Tyr	Leu	Asp	Gly <sub>110</sub>	Asn	Gln	Leu	Leu	Glu <sub>115</sub>	Ile	Pro	Gln	Gly	Leu <sub>120</sub>	Pro	Pro
	Ser	Leu	Gln <sub>125</sub>	Leu	Leu	Ser	Leu	Glu <sub>130</sub>	Ala	Asn	Asn	Ile	Phe <sub>135</sub>	Ser	Ile	Arg
10	Lys	Glu <sub>140</sub>	Asn	Leu	Thr	Glu	Leu <sub>145</sub>	Ala	Asn	Ile	Glu	Ile <sub>150</sub>	Leu	Tyr	Leu	Gly
	Gln <sub>155</sub>	Asn	Cys	Tyr	Tyr	Arg <sub>160</sub>	Asn	Pro	Cys	Tyr	Val <sub>165</sub>	Ser	Tyr	Ser	Ile	Glu <sub>170</sub>
15	Lys	Asp	Ala	Phe	Leu <sub>175</sub>	Asn	Leu	Thr	Lys	Leu <sub>180</sub>	Lys	Val	Leu	Ser	Leu <sub>185</sub>	Lys
	Asp	Asn	Asn	Val <sub>190</sub>	Thr	Ala	Val	Pro	Thr <sub>195</sub>	Val	Leu	Pro	Ser	Thr <sub>200</sub>	Leu	Thr
20	Glu	Leu	Tyr <sub>205</sub>	Leu	Tyr	Asn	Asn	Met <sub>210</sub>	Ile	Ala	Lys	Ile	Gln <sub>215</sub>	Glu	Asp	Asp
	Phe	Asn <sub>220</sub>	Asn	Leu	Asn	Gln <sub>225</sub>	Leu	Gln	Ile	Leu	Asp	Leu <sub>230</sub>	Ser	Gly	Asn	Cys
25	Pro	Arg <sub>235</sub>	Cys	Tyr	Asn	Ala <sub>240</sub>	Pro	Phe	Pro	Cys	Ala <sub>245</sub>	Pro	Cys	Lys	Asn	Asn <sub>250</sub>
	Ser	Pro	Leu	Gln	Ile <sub>255</sub>	Pro	Val	Asn	Ala	Phe <sub>260</sub>	Asp	Ala	Leu	Thr	Glu <sub>265</sub>	Leu
30	Lys	Val	Leu	Arg <sub>270</sub>	Leu	His	Ser	Asn	Ser <sub>275</sub>	Leu	Gln	His	Val	Pro <sub>280</sub>	Pro	Arg
	Trp	Phe	Lys <sub>285</sub>	Asn	Ile	Asn	Lys	Leu <sub>290</sub>	Gln	Glu	Leu	Asp	Leu <sub>295</sub>	Ser	Gln	Asn
35	Phe	Leu	Ala	Lys	Glu	Ile	Gly <sub>305</sub>	Asp	Ala	Lys	Phe	Leu <sub>310</sub>	His	Phe	Leu	Pro
	Ser	Leu	Ile	Gln	Leu	Asp <sub>320</sub>	Leu	Ser	Phe	Asn	Phe <sub>325</sub>	Glu	Leu	Gln	Val	Tyr <sub>330</sub>
40	Arg	Ala	Ser	Met	Asn <sub>335</sub>	Leu	Ser	Gln	Ala	Phe <sub>340</sub>	Ser	Ser	Leu	Lys	Ser <sub>345</sub>	Leu
	Lys	Ile	Leu	Arg <sub>350</sub>	Ile	Arg	Gly	Tyr	Val <sub>355</sub>	Phe	Lys	Glu	Leu	Lys <sub>360</sub>	Ser	Phe
45	Asn	Leu	Ser <sub>365</sub>	Pro	Leu	His	Asn	Leu <sub>370</sub>	Gln	Asn	Leu	Glu	Val <sub>375</sub>	Leu	Asp	Leu
	Gly	Thr	Asn	Phe	Ile	Lys	Ile <sub>385</sub>	Ala	Asn	Leu	Ser	Met <sub>390</sub>	Phe	Lys	Gln	Phe
50	Lys	Arg	Leu	Lys	Val	Ile <sub>400</sub>	Asp	Leu	Ser	Val	Asn <sub>405</sub>	Lys	Ile	Ser	Pro	Ser <sub>410</sub>
	Gly	Asp	Ser	Ser	Glu <sub>415</sub>	Val	Gly	Phe	Cys	Ser <sub>420</sub>	Asn	Ala	Arg	Thr	Ser <sub>425</sub>	Val
55	Glu	Ser	Tyr	Glu <sub>430</sub>	Pro	Gln	Val	Leu	Glu <sub>435</sub>	Gln	Leu	His	Tyr	Phe <sub>440</sub>	Arg	Tyr
	Asp	Lys	Tyr <sub>445</sub>	Ala	Arg	Ser	Cys	Arg <sub>450</sub>	Phe	Lys	Asn	Lys	Glu <sub>455</sub>	Ala	Ser	Phe
60																
65																

# ES 2 340 210 T3

	Met	Ser	Val	Asn	Glu	Ser	Cys	Tyr	Lys	Tyr	Gly	Gln	Thr	Leu	Asp	Leu
	460						465					470				
5	Ser	Lys	Asn	Ser	Ile	Phe	Phe	Val	Lys	Ser	Ser	Asp	Phe	Gln	His	Leu
	475					480					485					490
	Ser	Phe	Leu	Lys	Cys	Leu	Asn	Leu	Ser	Gly	Asn	Leu	Ile	Ser	Gln	Thr
					495					500					505	
10	Leu	Asn	Gly	Ser	Glu	Phe	Gln	Pro	Leu	Ala	Glu	Leu	Arg	Tyr	Leu	Asp
				510					515					520		
	Phe	Ser	Asn	Asn	Arg	Leu	Asp	Leu	Leu	His	Ser	Thr	Ala	Phe	Glu	Glu
			525					530					535			
15	Leu	His	Lys	Leu	Glu	Val	Leu	Asp	Ile	Ser	Ser	Asn	Ser	His	Tyr	Phe
	540						545					550				
	Gln	Ser	Glu	Gly	Ile	Thr	His	Met	Leu	Asn	Phe	Thr	Lys	Asn	Leu	Lys
	555					560					565					570
20	Val	Leu	Gln	Lys	Leu	Met	Met	Asn	Asp	Asn	Asp	Ile	Ser	Ser	Ser	Thr
					575					580					585	
	Ser	Arg	Thr	Met	Glu	Ser	Glu	Ser	Leu	Arg	Thr	Leu	Glu	Phe	Arg	Gly
				590					595					600		
25	Asn	His	Leu	Asp	Val	Leu	Trp	Arg	Glu	Gly	Asp	Asn	Arg	Tyr	Leu	Gln
			605					610					615			
	Leu	Phe	Lys	Asn	Leu	Leu	Lys	Leu	Glu	Glu	Leu	Asp	Ile	Ser	Lys	Asn
	620						625					630				
30	Ser	Leu	Ser	Phe	Leu	Pro	Ser	Gly	Val	Phe	Asp	Gly	Met	Pro	Pro	Asn
	635					640					645					650
	Leu	Lys	Asn	Leu	Ser	Leu	Ala	Lys	Asn	Gly	Leu	Lys	Ser	Phe	Ser	Trp
					655					660					665	
	Lys	Lys	Leu	Gln	Cys	Leu	Lys	Asn	Leu	Glu	Thr	Leu	Asp	Leu	Ser	His
				670					675					680		
40	Asn	Gln	Leu	Thr	Thr	Val	Pro	Glu	Arg	Leu	Ser	Asn	Cys	Ser	Arg	Ser
			685					690					695			
	Leu	Lys	Asn	Leu	Ile	Leu	Lys	Asn	Asn	Gln	Ile	Arg	Ser	Leu	Thr	Lys
	700						705					710				
45	Tyr	Phe	Leu	Gln	Asp	Ala	Phe	Gln	Leu	Arg	Tyr	Leu	Asp	Leu	Ser	Ser
	715					720					725					730
	Asn	Lys	Ile	Gln	Met	Ile	Gln	Lys	Thr	Ser	Phe	Pro	Glu	Asn	Val	Leu
					735					740					745	
50	Asn	Asn	Leu	Lys	Met	Leu	Leu	Leu	His	His	Asn	Arg	Phe	Leu	Cys	Thr
				750					755					760		
	Cys	Asp	Ala	Val	Trp	Phe	Val	Trp	Trp	Val	Asn	His	Thr	Glu	Val	Thr
			765					770					775			
55	Ile	Pro	Tyr	Leu	Ala	Thr	Asp	Val	Thr	Cys	Val	Gly	Pro	Gly	Ala	His
	780						785					790				
	Lys	Gly	Gln	Ser	Val	Ile	Ser	Leu	Asp	Leu	Tyr	Thr	Cys	Glu	Leu	Asp
	795					800					805					810
60	Leu	Thr	Asn	Leu	Ile	Leu	Phe	Ser	Leu	Ser	Ile	Ser	Val	Ser	Leu	Phe
65																

## ES 2 340 210 T3

	815	820	825
	Leu Met Val Met Met Thr Ala Ser His	Leu Tyr Phe Trp Asp Val Trp	
	830	835	840
5	Tyr Ile Tyr His Phe Cys Lys Ala Lys Ile Lys Gly Tyr Gln Arg Leu		
	845	850	855
	Ile Ser Pro Asp Cys Cys Tyr Asp Ala Phe Ile Val Tyr Asp Thr Lys		
	860	865	870
10	Asp Pro Ala Val Thr Glu Trp Val Leu Ala Glu Leu Val Ala Lys Leu		
	875	880	885
	Glu Asp Pro Arg Glu Lys His Phe Asn Leu Cys Leu Glu Glu Arg Asp		
	895	900	905
15	Trp Leu Pro Gly Gln Pro Val Leu Glu Asn Leu Ser Gln Ser Ile Gln		
	910	915	920
	Leu Ser Lys Lys Thr Val Phe Val Met Thr Asp Lys Tyr Ala Lys Thr		
	925	930	935
20	Glu Asn Phe Lys Ile Ala Phe Tyr Leu Ser His Gln Arg Leu Met Asp		
	940	945	950
	Glu Lys Val Asp Val Ile Ile Leu Ile Phe Leu Glu Lys Pro Phe Gln		
	955	960	965
25	Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg Leu Cys Gly Ser Ser Val		
	975	980	985
	Leu Glu Trp Pro Thr Asn Pro Gln Ala His Pro Tyr Phe Trp Gln Cys		
	990	995	1000
30	Leu Lys Asn Ala Leu Ala Thr Asp Asn His Val Ala Tyr Ser Gln Val		
	1005	1010	1015
	Phe Lys Glu Thr Val		
	1020		
35			

### (2) INFORMACIÓN PARA EL SEQ ID NO:13:

- 40 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 180 pares de bases
  - (B) TIPO: ácido nucleico
  - 45 (C) CUALIDAD DE LA HEBRA: sencilla
  - (D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: ADNC
- 50 (ix) RASGO:
- (A) NOMBRE/CLAVE: CDS
  - (B) LOCALIZACIÓN: 1..177
- 55 (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:13:

60

65

## ES 2 340 210 T3

	CTT GGA AAA CCT CTT CAG AAG TCT AAG TTT CTT CAG CTC AGG AAG AGA	48
	Leu Gly Lys Pro Leu Gln Lys Ser Lys Phe Leu Gln Leu Arg Lys Arg	
	1 5 10 15	
5	CTC TGC AGG AGC TCT GTC CTT GAG TGG CCT GCA AAT CCA CAG GCT CAC	96
	Leu Cys Arg Ser Ser Val Leu Glu Trp Pro Ala Asn Pro Gln Ala His	
	20 25 30	
10	CCA TAC TTC TGG CAG TGC CTG AAA AAT GCC CTG ACC ACA GAC AAT CAT	144
	Pro Tyr Phe Trp Gln Cys Leu Lys Asn Ala Leu Thr Thr Asp Asn His	
	35 40 45	
15	GTG GCT TAT AGT CAA ATG TTC AAG GAA ACA GTC TAG	180
	Val Ala Tyr Ser Gln Met Phe Lys Glu Thr Val	
	50 55	

20 (2) INFORMACIÓN PARA EL SEQ ID NO:14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 59 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

30 (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:14:

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

45 (2) INFORMACIÓN PARA EL SEQ ID NO:15:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 990 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

55 (ii) TIPO DE MOLÉCULA: ADNc

(ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 2..988

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:15:

65

# ES 2 340 210 T3

	G AAT TCC AGA CTT ATA AAC TTG AAA AAT CTC TAT TTG GCC TGG AAC	46
	Asn Ser Arg Leu Ile Asn Leu Lys Asn Leu Tyr Leu Ala Trp Asn	
	1 5 10 15	
5	TGC TAT TTT AAC AAA GTT TGC GAG AAA ACT AAC ATA GAA GAT GGA GTA	94
	Cys Tyr Phe Asn Lys Val Cys Glu Lys Thr Asn Ile Glu Asp Gly Val	
	20 25 30	
10	TTT GAA ACG CTG ACA AAT TTG GAG TTG CTA TCA CTA TCT TTC AAT TCT	142
	Phe Glu Thr Leu Thr Asn Leu Glu Leu Leu Ser Leu Ser Phe Asn Ser	
	35 40 45	
15	CTT TCA CAT GTG CCA CCC AAA CTG CCA AGC TCC CTA CGC AAA CTT TTT	190
	Leu Ser His Val Pro Pro Lys Leu Pro Ser Ser Leu Arg Lys Leu Phe	
	50 55 60	
20	CTG AGC AAC ACC CAG ATC AAA TAC ATT AGT GAA GAA GAT TTC AAG GGA	238
	Leu Ser Asn Thr Gln Ile Lys Tyr Ile Ser Glu Glu Asp Phe Lys Gly	
	65 70 75	
25	TTG ATA AAT TTA ACA TTA CTA GAT TTA AGC GGG AAC TGT CCG AGG TGC	286
	Leu Ile Asn Leu Thr Leu Leu Asp Leu Ser Gly Asn Cys Pro Arg Cys	
	80 85 90 95	
30	TTC AAT GCC CCA TTT CCA TGC GTG CCT TGT GAT GGT GGT GCT TCA ATT	334
	Phe Asn Ala Pro Phe Pro Cys Val Pro Cys Asp Gly Gly Ala Ser Ile	
	100 105 110	
35	AAT ATA GAT CGT TTT GCT TTT CAA AAC TTG ACC CAA CTT CGA TAC CTA	382
	Asn Ile Asp Arg Phe Ala Phe Gln Asn Leu Thr Gln Leu Arg Tyr Leu	
	115 120 125	
40	AAC CTC TCT AGC ACT TCC CTC AGG AAG ATT AAT GCT GCC TGG TTT AAA	430
	Asn Leu Ser Ser Thr Ser Leu Arg Lys Ile Asn Ala Ala Trp Phe Lys	
	130 135 140	
45	AAT ATG CCT CAT CTG AAG GTG CTG GAT CTT GAA TTC AAC TAT TTA GTG	478
	Asn Met Pro His Leu Lys Val Leu Asp Leu Glu Phe Asn Tyr Leu Val	
	145 150 155	
50	GGA GAA ATA GCC TCT GGG GCA TTT TTA ACG ATG CTG CCC CGC TTA GAA	526
	Gly Glu Ile Ala Ser Gly Ala Phe Leu Thr Met Leu Pro Arg Leu Glu	
	160 165 170 175	
55	ATA CTT GAC TTG TCT TTT AAC TAT ATA AAG GGG AGT TAT CCA CAG CAT	574
	Ile Leu Asp Leu Ser Phe Asn Tyr Ile Lys Gly Ser Tyr Pro Gln His	
	180 185 190	
60	ATT AAT ATT TCC AGA AAC TTC TCT AAA CTT TTG TCT CTA CGG GCA TTG	622
	Ile Asn Ile Ser Arg Asn Phe Ser Lys Leu Leu Ser Leu Arg Ala Leu	
	195 200 205	
65	CAT TTA AGA GGT TAT GTG TTC CAG GAA CTC AGA GAA GAT GAT TTC CAG	670
	His Leu Arg Gly Tyr Val Phe Gln Glu Leu Arg Glu Asp Asp Phe Gln	
	210 215 220	
70	CCC CTG ATG CAG CTT CCA AAC TTA TCG ACT ATC AAC TTG GGT ATT AAT	718
	Pro Leu Met Gln Leu Pro Asn Leu Ser Thr Ile Asn Leu Gly Ile Asn	
	225 230 235	

## ES 2 340 210 T3

	TTT ATT AAG CAA ATC GAT TTC AAA CTT TTC CAA AAT TTC TCC AAT CTG	766
	Phe Ile Lys Gln Ile Asp Phe Lys Leu Phe Gln Asn Phe Ser Asn Leu	
	240 245 250 255	
5	GAA ATT ATT TAC TTG TCA GAA AAC AGA ATA TCA CCG TTG GTA AAA GAT	814
	Glu Ile Ile Tyr Leu Ser Glu Asn Arg Ile Ser Pro Leu Val Lys Asp	
	260 265 270	
10	ACC CGG CAG AGT TAT GCA AAT AGT TCC TCT TTT CAA CGT CAT ATC CGG	862
	Thr Arg Gln Ser Tyr Ala Asn Ser Ser Phe Gln Arg His Ile Arg	
	275 280 285	
15	AAA CGA CGC TCA ACA GAT TTT GAG TTT GAC CCA CAT TCG AAC TTT TAT	910
	Lys Arg Arg Ser Thr Asp Phe Glu Phe Asp Pro His Ser Asn Phe Tyr	
	290 295 300	
20	CAT TTC ACC CGT CCT TTA ATA AAG CCA CAA TGT GCT GCT TAT GGA AAA	958
	His Phe Thr Arg Pro Leu Ile Lys Pro Gln Cys Ala Ala Tyr Gly Lys	
	305 310 315	
25	GCC TTA GAT TTA AGC CTC AAC AGT ATT TTC TT	990
	Ala Leu Asp Leu Ser Leu Asn Ser Ile Phe	
	320 325	

### (2) INFORMACIÓN PARA EL SEQ ID NO:16:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 329 aminoácidos
  - (B) TIPO: aminoácido
  - (D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: proteína
- (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:16:

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

ES 2 340 210 T3

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

(2) INFORMACIÓN PARA EL SEQ ID NO:17:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1557 pares de bases  
(B) TIPO: ácido nucleico  
(C) CUALIDAD DE LA HEBRA: sencilla  
(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADNC

(ix) RASGO:

(A) NOMBRE/CLAVE: CDS  
(B) LOCALIZACIÓN: 1..513

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc  
(B) LOCALIZACIÓN: 278  
(D) OTRA INFORMACIÓN: /nota= “el nucleótido 278 denominado G, puede ser G o C”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc  
(B) LOCALIZACIÓN: 445  
(D) OTRA INFORMACIÓN: /nota= “el nucleótido 445 denominado A, puede ser A o T”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc  
(B) LOCALIZACIÓN: 572

(D) OTRA INFORMACIÓN: /nota= “los nucleótidos 572, 593, 600, 607, 617, 622, 625, 631, 640, 646, 653, 719, 775, y 861 denominados C; pueden ser A, C, G, o T”

# ES 2 340 210 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:17:

5	CAG TCT CTT TCC ACA TCC CAA ACT TTC TAT GAT GCT TAC ATT TCT TAT Gln Ser Leu Ser Thr Ser Gln Thr Phe Tyr Asp Ala Tyr Ile Ser Tyr 1 5 10 15	48
	GAC ACC AAA GAT GCC TCT GTT ACT GAC TGG GTG ATA AAT GAG CTG CGC Asp Thr Lys Asp Ala Ser Val Thr Asp Trp Val Ile Asn Glu Leu Arg 20 25 30	96
10	TAC CAC CTT GAA GAG AGC CGA GAC AAA AAC GTT CTC CTT TGT CTA GAG Tyr His Leu Glu Glu Ser Arg Asp Lys Asn Val Leu Leu Cys Leu Glu 35 40 45	144
15	GAG AGG GAT TGG GAC CCG GGA TTG GCC ATC ATC GAC AAC CTC ATG CAG Glu Arg Asp Trp Asp Pro Gly Leu Ala Ile Ile Asp Asn Leu Met Gln 50 55 60	192
	AGC ATC AAC CAA AGC AAG AAA ACA GTA TTT GTT TTA ACC AAA AAA TAT Ser Ile Asn Gln Ser Lys Lys Thr Val Phe Val Leu Thr Lys Lys Tyr 65 70 75 80	240
20	GCA AAA AGC TGG AAC TTT AAA ACA GCT TTT TAC TTG GGC TTG CAG AGG Ala Lys Ser Trp Asn Phe Lys Thr Ala Phe Tyr Leu Gly Leu Gln Arg 85 90 95	288
25	CTA ATG GGT GAG AAC ATG GAT GTG ATT ATA TTT ATC CTG CTG GAG CCA Leu Met Gly Glu Asn Met Asp Val Ile Phe Ile Leu Leu Glu Pro 100 105 110	336
	GTG TTA CAG CAT TCT CCG TAT TTG AGG CTA CGG CAG CGG ATC TGT AAG Val Leu Gln His Ser Pro Tyr Leu Arg Leu Arg Gln Arg Ile Cys Lys 115 120 125	384
30	AGC TCC ATC CTC CAG TGG CCT GAC AAC CCG AAG GCA GAA AGG TTG TTT Ser Ser Ile Leu Gln Trp Pro Asp Asn Pro Lys Ala Glu Arg Leu Phe 130 135 140	432
	TGG CAA ACT CTG AGA AAT GTG GTC TTG ACT GAA AAT GAT TCA CGG TAT Trp Gln Thr Leu Arg Asn Val Val Leu Thr Glu Asn Asp Ser Arg Tyr 145 150 155 160	480
35	AAC AAT ATG TAT GTC GAT TCC ATT AAG CAA TAC TAACTGACGT TAAGTCATGA Asn Asn Met Tyr Val Asp Ser Ile Lys Gln Tyr 165 170	533
40	TTTCGCGCCA TAATAAAGAT GCAAAGGAAT GACATTCCG TATTAGTTAT CTATTGCTAC GGTAACCAAA TTAATCCCAA AAACCTTACG TCGGTTTCAA AACAACCACA TTCTGCTGGC CCCACAGTTT TTGAGGGTCA GGAGTCCAGG CCCAGCATAA CTGGGTCTTC TGCTTCAGGG TGCTCTCCAGA GGCTGCAATG TAGGTGTTCA CCAGAGACAT AGGCATCACT GGGGTACAC TCCATGTGGT TGTTTTCTGG ATTCAATTCC TCCTGGGCTA TTGGCCAAAG GCTATACTCA TGTAAGCCAT GCGAGCCTAT CCCACAACGG CAGCTTGCTT CATCAGAGCT AGCAAAAAAG AGAGGTTGCT AGCAAGATGA AGTCACAATC TTTTGTAAATC GAATCAAAAA AGTGATATCT CATCACTTTG GCCATATTCT ATTTGTTAGA AGTAAACCAC AGGTCCCACC AGCTCCATGG GAGTGACCAC CTCAGTCCAG GGAACACAGC TGAAGACCAA GATGGTGAGC TCTGATTGCT TCAGTTGGTC ATCAACTATT TTCCTTGAC TGCTGTCCTG GGATGGCCGG CTATCTTGAT GGATAGATTG TGAATATCAG GAGGCCAGGG ATCACTGTGG ACCATCTTAG CAGTTGACCT AACACATCTT CTTTTCAATA TCTAAGAACT TTTGCCACTG TGAATAATGG TCCTAATATT AAGCTGTTGT TTATATTTAT CATATATCTA TGGCTACATG GTTATATTAT GCTGTGGTTG CGTTCGGTTT TATTTACAGT TGCTTTTACA AATATTGCT GTAACATTTG ACTTCTAAGG TTTAGATGCC ATTTAAGAAC TGAGATGGAT AGCTTTTAAA GCATCTTTTA CTTCTTACCA TTTTTTAAAA GTATGCAGCT AAATTCGAAG CTTTGGTCT ATATTGTAA TTGCCATTGC TGTAATCTT AAAATGAATG AATAAAAATG TTTCATTTTA AAAAAAAAAA AAAAAAAAAA AAAA	593 653 713 773 833 893 953 1013 1073 1133 1193 1253 1313 1373 1433 1493 1553 1557



## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:18:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 171 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:18:

```

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

```

### (2) INFORMACIÓN PARA EL SEQ ID NO:19:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 629 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNC

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..486

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 144
- (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 144 y 225 denominado C; puede ser C o T"

## ES 2 340 210 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:19:

5	AAT Asn 1	GAA Glu	TTG Leu	ATC Ile	CCC Pro 5	AAT Asn	CTA Leu	GAG Glu	AAG Lys	GAA Glu 10	GAT Asp	GGT Gly	TCT Ser	ATC Ile	TTG Leu 15	ATT Ile	48	
	TGC Cys	CTT Leu	TAT Tyr	GAA Glu 20	AGC Ser	TAC Tyr	TTT Phe	GAC Asp	CCT Pro 25	GGC Gly	AAA Lys	AGC Ser	ATT Ile	AGT Ser 30	GAA Glu	AAT Asn	96	
10	ATT Ile	GTA Val	AGC Ser 35	TTC Phe	ATT Ile	GAG Glu	AAA Lys	AGC Ser 40	TAT Tyr	AAG Lys	TCC Ser	ATC Ile	TTT Phe 45	GTT Val	TTG Leu	TCC Ser	144	
15	CCC Pro	AAC Asn	TTT Phe	GTC Val	CAG Gln	AAT Asn	GAG Glu	TGG Trp	TGC Cys	CAT His	TAT Tyr	GAA Glu	TTC Phe	TAC Tyr	TTT Phe	GCC Ala	192	
	50			55			60											
20	CAC His 65	CAC His	AAT Asn	CTC Leu	TTC Phe	CAT His 70	GAA Glu	AAT Asn	TCT Ser	GAT Asp	CAC His 75	ATA Ile	ATT Ile	CTT Leu	ATC Ile 80	TTA Leu	240	
	CTG Leu	GAA Glu	CCC Pro	ATT Ile	CCA Pro 85	TTC Phe	TAT Tyr	TGC Cys	ATT Ile	CCC Pro 90	ACC Thr	AGG Arg	TAT Tyr	CAT His	AAA Lys 95	CTG Leu	288	
25	GAA Glu	GCT Ala	CTC Leu	CTG Leu 100	GAA Glu	AAA Lys	AAA Lys	GCA Ala 105	TAC Tyr 105	TTG Leu	GAA Glu	TGG Trp	CCC Pro	AAG Lys 110	GAT Asp	AGG Arg	336	
30	CGT Arg	AAA Lys	TGT Cys 115	GGG Gly	CTT Leu	TTC Phe	TGG Trp	GCA Ala 120	AAC Asn	CTT Leu	CGA Arg	GCT Ala	GCT Ala 125	GTT Val	AAT Asn	GTT Val	384	
	AAT Asn 130	GTA Val	TTA Leu	GCC Ala	ACC Thr	AGA Arg	GAA Glu 135	ATG Met	TAT Tyr	GAA Glu	CTG Leu	CAG Gln 140	ACA Thr	TTC Phe	ACA Thr	GAG Glu	432	
35	TTA Leu 145	AAT Asn	GAA Glu	GAG Glu	TCT Ser	CGA Arg 150	GGT Gly	TCT Ser	ACA Thr	ATC Ile	TCT Ser 155	CTG Leu	ATG Met	AGA Arg	ACA Thr	GAC Asp 160	480	
40	TGT Cys	CTA Leu	TAAAATCCCA CAGTCCTTGG GAAGTTGGGG ACCACATACA CTGTTGGGAT															536
	GTACATTGAT ACAACCTTTA TGATGGCAAT TTGACAATAT TTATTAAAAT AAAAAATGGT																	596
45	TATTCCTTC AAAAAAAAAA AAAAAAAAAA AAA																	629

(2) INFORMACIÓN PARA EL SEQ ID NO:20:

50 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 162 aminoácidos

(B) TIPO: aminoácido

55 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:20:

60

65

## ES 2 340 210 T3

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

### (2) INFORMACIÓN PARA EL SEQ ID NO:21:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 427 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..426

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:21:

## ES 2 340 210 T3

	AAG Lys 1	AAC Asn	TCC Ser	AAA Lys	GAA Glu 5	AAC Asn	CTC Leu	CAG Gln	TTT Phe	CAT His 10	GCT Ala	TTT Phe	ATT Ile	TCA Ser	TAT Tyr 15	AGT Ser	48
5	GAA Glu	CAT His	GAT Asp	TCT Ser 20	GCC Ala	TGG Trp	GTG Val	AAA Lys	AGT Ser 25	GAA Glu	TTG Leu	GTA Val	CCT Pro	TAC Tyr 30	CTA Leu	GAA Glu	96
	AAA Lys	GAA Glu	GAT Asp 35	ATA Ile	CAG Gln	ATT Ile	TGT Cys	CTT Leu 40	CAT His	GAG Glu	AGA Arg	AAC Asn	TTT Phe 45	GTC Val	CCT Pro	GGC Gly	144
10	AAG Lys	AGC Ser 50	ATT Ile	GTG Val	GAA Glu	AAT Asn	ATC Ile 55	ATC Ile	AAC Asn	TGC Cys	ATT Ile	GAG Glu 60	AAG Lys	AGT Ser	TAC Tyr	AAG Lys	192
15	TCC Ser 65	ATC Ile	TTT Phe	GTT Val	TTG Leu	TCT Ser 70	CCC Pro	AAC Asn	TTT Phe	GTC Val	CAG Gln 75	AGT Ser	GAG Glu	TGG Trp	TGC Cys	CAT His 80	240
	TAC Tyr	GAA Glu	CTC Leu	TAT Tyr	TTT Phe 85	GCC Ala	CAT His	CAC His	AAT Asn	CTC Leu 90	TTT Phe	CAT His	GAA Glu	GGA Gly	TCT Ser 95	AAT Asn	288
20	AAC Asn	TTA Leu	ATC Ile	CTC Leu 100	ATC Ile	TTA Leu	CTG Leu	GAA Glu	CCC Pro 105	ATT Ile	CCA Pro	CAG Gln	AAC Asn	AGC Ser 110	ATT Ile	CCC Pro	336
	AAC Asn	AAG Lys	TAC Tyr 115	CAC His	AAG Lys	CTG Leu	AAG Lys	GCT Ala 120	CTC Leu	ATG Met	ACG Thr	CAG Gln	CGG Arg 125	ACT Thr	TAT Tyr	TTG Leu	384
25	CAG Gln 130	TGG Trp	CCC Pro	AAG Lys	GAG Glu	AAA Lys	AGC Ser 135	AAA Lys	CGT Arg	GGG Gly	CTC Leu	TTT Phe 140	TGG Trp	GCT Ala			426
30	A																427

### (2) INFORMACIÓN PARA EL SEQ ID NO:22:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 142 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:22:

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

## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:23:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 662 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNC

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..627

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 54
- (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 54, 103, y 345 se denominan A; pueden ser cada uno A o G"

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 313
- (D) OTRA INFORMACIÓN: /nota= "el nucleótido 313 denominado G, puede ser G o T"

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 316
- (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 316, 380, 407, y 408 denominados C; pueden ser cada uno A, C, G, o T"

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:23:

## ES 2 340 210 T3

	GCT	TCC	ACC	TGT	GCC	TGG	CCT	GGC	TTC	CCT	GGC	GGG	GGC	GGC	AAA	GTG	48
	Ala	Ser	Thr	Cys	Ala	Trp	Pro	Gly	Phe	Pro	Gly	Gly	Gly	Gly	Lys	Val	
	1				5					10					15		
5	GGC	GAA	ATG	AGG	ATG	CCC	TGC	CCT	ACG	ATG	CCT	TCG	TGG	TCT	TCG	ACA	96
	Gly	Glu	Met	Arg	Met	Pro	Cys	Pro	Thr	Met	Pro	Ser	Trp	Ser	Ser	Thr	
				20					25					30			
10	AAA	CGC	AGA	GCG	CAG	TGG	CAG	ACT	GGG	TGT	ACA	ACG	AGC	TTC	GGG	GGC	144
	Lys	Arg	Arg	Ala	Gln	Trp	Gln	Thr	Gly	Cys	Thr	Thr	Ser	Phe	Gly	Gly	
			35					40					45				
	AGC	TGG	AGG	AGT	GCC	GTG	GGC	GCT	GGG	CAC	TCC	GCC	TGT	GCC	TGG	AGG	192
	Ser	Trp	Arg	Ser	Ala	Val	Gly	Ala	Gly	His	Ser	Ala	Cys	Ala	Trp	Arg	
		50				55						60					
15	AAC	GCG	ACT	GGC	TGC	CTG	GCA	AAA	CCC	TCT	TTG	AGA	ACC	TGT	GGG	CCT	240
	Asn	Ala	Thr	Gly	Cys	Leu	Ala	Lys	Pro	Ser	Leu	Arg	Thr	Cys	Gly	Pro	
	65					70					75					80	
20	CGG	TCT	ATG	GCA	GCC	GCA	AGA	CGC	TGT	TTG	TGC	TGG	CCC	ACA	CGG	ACC	288
	Arg	Ser	Met	Ala	Ala	Ala	Arg	Arg	Cys	Leu	Cys	Trp	Pro	Thr	Arg	Thr	
				85						90					95		
	GGG	TCA	GTG	GTC	TCT	TGC	GCG	CCA	GTT	CTC	CTG	CTG	GCC	CAG	CAG	CGC	336
	Gly	Ser	Val	Val	Ser	Cys	Ala	Pro	Val	Leu	Leu	Leu	Ala	Gln	Gln	Arg	
				100				105						110			
25	CTG	CTG	GAA	GAC	CGC	AAG	GAC	GTC	GTG	GTG	CTG	GTG	ATC	CTA	ACG	CCT	384
	Leu	Leu	Glu	Asp	Arg	Lys	Asp	Val	Val	Val	Leu	Val	Ile	Leu	Thr	Pro	
			115					120					125				
30	GAC	GGC	CAA	GCC	TCC	CGA	CTA	CCC	GAT	GCG	CTG	ACC	AGC	GCC	TCT	GCC	432
	Asp	Gly	Gln	Ala	Ser	Arg	Leu	Pro	Asp	Ala	Leu	Thr	Ser	Ala	Ser	Ala	
		130					135					140					
	GCC	AGA	GTG	TCC	TCC	TCT	GGC	CCC	ACC	AGC	CCA	GTG	GTC	GCG	CAG	CTT	480
	Ala	Arg	Val	Ser	Ser	Ser	Gly	Pro	Thr	Ser	Pro	Val	Val	Ala	Gln	Leu	
		145				150					155					160	
35	CTG	AGG	CCA	GCA	TGC	ATG	GCC	CTG	ACC	AGG	GAC	AAC	CAC	CAC	TTC	TAT	528
	Leu	Arg	Pro	Ala	Cys	Met	Ala	Leu	Thr	Arg	Asp	Asn	His	His	Phe	Tyr	
				165						170					175		
40	AAC	CGG	AAC	TTC	TGC	CAG	GGA	ACC	CAC	GGC	CGA	ATA	GCC	GTG	AGC	CGG	576
	Asn	Arg	Asn	Phe	Cys	Gln	Gly	Thr	His	Gly	Arg	Ile	Ala	Val	Ser	Arg	
				180				185						190			
	AAT	CCT	GCA	CGG	TGC	CAC	CTC	CAC	ACA	CAC	CTA	ACA	TAT	GCC	TGC	CTG	624
	Asn	Pro	Ala	Arg	Cys	His	Leu	His	Thr	His	Leu	Thr	Tyr	Ala	Cys	Leu	
			195				200						205				
45	ATC	TGACCAACAC	ATGCTCGCCA	CCCTCACCAC	ACACC												662
	Ile																

### (2) INFORMACIÓN PARA EL SEQ ID NO:24:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 209 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:24:

## ES 2 340 210 T3

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

### (2) INFORMACIÓN PARA EL SEQ ID NO:25:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 4865 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNC

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 107..2617

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: mat\_péptido
- (B) LOCALIZACIÓN: 173..2617

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 81

(D) OTRA INFORMACIÓN: /nota= "los nucleótidos 81, 3144, 3205, y 3563 denominados A, pueden ser cada uno A, C, G, o T"

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc

## ES 2 340 210 T3

(B) LOCALIZACIÓN: 84

(D) OTRA INFORMACIÓN: /nota= “el nucleótido 84 denominado C, puede ser C o G”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 739

(D) OTRA INFORMACIÓN: /nota= “el nucleótido 739 denominado C, puede ser C o T”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 3132

(D) OTRA INFORMACIÓN: /nota= “los nucleótidos 3132, 3532, 3538, y 3553 denominados G, pueden ser cada uno G o T”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 3638

(D) OTRA INFORMACIÓN: /nota= “el nucleótido 3638 denominado A, puede ser A o T”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 3677

(D) OTRA INFORMACIÓN: /nota= “los nucleótidos 3677, 3685, y 3736 denominados C, pueden ser cada uno A o C”

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:25:

35	AAAATACTCC CTTGCCTCAA AAAGTCTCG GTCAAACGGT GATAGCAAAC CACGCATTCA	60
	CAGGGCCACT GCTGCTCACA AAACCAGTGA GGATGATGCC AGGATG ATG TCT GCC	115
	Met Ser Ala -22 -20	
40	TCG CGC CTG GCT GGG ACT CTG ATC CCA GCC ATG GCC TTC CTC TCC TGC	163
	Ser Arg Leu Ala Gly Thr Leu Ile Pro Ala Met Ala Phe Leu Ser Cys	
	-15 -10 -5	
	GTG AGA CCA GAA AGC TGG GAG CCC TGC GTG GAG GTT CCT AAT ATT ACT	211
	Val Arg Pro Glu Ser Trp Glu Pro Cys Val Glu Val Pro Asn Ile Thr	
	1 5 10	
45	TAT CAA TGC ATG GAG CTG AAT TTC TAC AAA ATC CCC GAC AAC CTC CCC	259
	Tyr Gln Cys Met Glu Leu Asn Phe Tyr Lys Ile Pro Asp Asn Leu Pro	
	15 20 25	
50	TTC TCA ACC AAG AAC CTG GAC CTG AGC TTT AAT CCC CTG AGG CAT TTA	307
	Phe Ser Thr Lys Asn Leu Asp Leu Ser Phe Asn Pro Leu Arg His Leu	
	30 35 40 45	
	GGC AGC TAT AGC TTC TTC AGT TTC CCA GAA CTG CAG GTG CTG GAT TTA	355
	Gly Ser Tyr Ser Phe Phe Ser Phe Pro Glu Leu Gln Val Leu Asp Leu	
	50 55 60	
55	TCC AGG TGT GAA ATC CAG ACA ATT GAA GAT GGG GCA TAT CAG AGC CTA	403
	Ser Arg Cys Glu Ile Gln Thr Ile Glu Asp Gly Ala Tyr Gln Ser Leu	
	65 70 75	



# ES 2 340 210 T3

5	AGC Ser	CAC His	CTC Leu 80	TCT Ser	ACC Thr	TTA Leu	ATA Ile	TTG Leu 85	ACA Thr	GGA Gly	AAC Asn	CCC Pro	ATC Ile 90	CAG Gln	AGT Ser	TTA Leu	451
	GCC Ala 95	CTG Leu	GGA Gly	GCC Ala	TTT Phe	TCT Ser	GGA Gly 100	CTA Leu	TCA Ser	AGT Ser	TTA Leu	CAG Gln 105	AAG Lys	CTG Leu	GTG Val	GCT Ala	499
10	GTG Val 110	GAG Glu	ACA Thr	AAT Asn	CTA Leu	GCA Ala 115	TCT Ser	CTA Leu	GAG Glu	AAC Asn	TTC Phe 120	CCC Pro	ATT Ile	GGA Gly	CAT His	CTC Leu 125	547
	AAA Lys	ACT Thr	TTG Leu	AAA Lys	GAA Glu 130	CTT Leu	AAT Asn	GTG Val	GCT Ala	CAC His 135	AAT Asn	CTT Leu	ATC Ile	CAA Gln	TCT Ser 140	TTC Phe	595
15	AAA Lys	TTA Leu	CCT Pro	GAG Glu 145	TAT Tyr	TTT Phe	TCT Ser	AAT Asn	CTG Leu 150	ACC Thr	AAT Asn	CTA Leu	GAG Glu 155	CAC His	TTG Leu	GAC Asp	643
20	CTT Leu	TCC Ser	AGC Ser 160	AAC Asn	AAG Lys	ATT Ile	CAA Gln	AGT Ser 165	ATT Ile	TAT Tyr	TGC Cys	ACA Thr	GAC Asp 170	TTG Leu	CGG Arg	GTT Val	691
	CTA Leu	CAT His 175	CAA Gln	ATG Met	CCC Pro	CTA Leu	CTC Leu 180	AAT Asn	CTC Leu	TCT Ser	TTA Leu	GAC Asp 185	CTG Leu	TCC Ser	CTG Leu	AAC Asn	739
25	CCT Pro 190	ATG Met	AAC Asn	TTT Phe	ATC Ile	CAA Gln 195	CCA Pro	GGT Gly	GCA Ala	TTT Phe	AAA Lys 200	GAA Glu	ATT Ile	AGG Arg	CTT Leu	CAT His 205	787
30	AAG Lys	CTG Leu	ACT Thr	TTA Leu	AGA Arg 210	AAT Asn	AAT Asn	TTT Phe	GAT Asp	AGT Ser 215	TTA Leu	AAT Asn	GTA Val	ATG Met	AAA Lys 220	ACT Thr	835
	TGT Cys	ATT Ile	CAA Gln 225	GGT Gly	CTG Leu	GCT Ala	GGT Gly	TTA Leu	GAA Glu 230	GTC Val	CAT His	CGT Arg	TTG Leu	GTT Val 235	CTG Leu	GGA Gly	883
35	GAA Glu	TTT Phe	AGA Arg 240	AAT Asn	GAA Glu	GGA Gly	AAC Asn	TTG Leu 245	GAA Glu	AAG Lys	TTT Phe	GAC Asp 250	AAA Lys	TCT Ser	GCT Ala	CTA Leu	931
40	GAG Glu 255	GGC Gly	CTG Leu	TGC Cys	AAT Asn	TTG Leu	ACC Thr 260	ATT Ile	GAA Glu	GAA Glu	TTC Phe	CGA Arg 265	TTA Leu	GCA Ala	TAC Tyr	TTA Leu	979
	GAC Asp 270	TAC Tyr	TAC Tyr	CTC Leu	GAT Asp	GAT Asp 275	ATT Ile	ATT Ile	GAC Asp	TTA Leu	TTT Phe 280	AAT Asn	TGT Cys	TTG Leu	ACA Thr	AAT Asn 285	1027
45	GTT Val	TCT Ser	TCA Ser	TTT Phe 290	TCC Ser	CTG Leu	GTG Val	AGT Ser	GTG Val	ACT Thr 295	ATT Ile	GAA Glu	AGG Arg	GTA Val	AAA Lys 300	GAC Asp	1075
50	TTT Phe	TCT Ser	TAT Tyr	AAT Asn 305	TTC Phe	GGA Gly	TGG Trp	CAA Gln 310	CAT His	TTA Leu	GAA Glu	TTA Leu	GTT Val 315	AAC Asn	TGT Cys	AAA Lys	1123
	TTT Phe	GGA Gly	CAG Gln 320	TTT Phe	CCC Pro	ACA Thr	TTG Leu	AAA Lys 325	CTC Leu	AAA Lys	TCT Ser	CTC Leu	AAA Lys 330	AGG Arg	CTT Leu	ACT Thr	1171
55	TTC Phe 335	ACT Thr	TCC Ser	AAC Asn	AAA Lys	GGT Gly	GGG Gly 340	AAT Asn	GCT Ala	TTT Phe	TCA Ser	GAA Glu 345	GTT Val	GAT Asp	CTA Leu	CCA Pro	1219

# ES 2 340 210 T3

	AGC Ser 350	CTT Leu	GAG Glu	TTT Phe	CTA Leu	GAT Asp 355	CTC Leu	AGT Ser	AGA Arg	AAT Asn	GGC Gly 360	TTG Leu	AGT Ser	TTC Phe	AAA Lys	GGT Gly 365	1267
5	TGC Cys	TGT Cys	TCT Ser	CAA Gln	AGT Ser 370	GAT Asp	TTT Phe	GGG Gly	ACA Thr	ACC Thr 375	AGC Ser	CTA Leu	AAG Lys	TAT Tyr	TTA Leu 380	GAT Asp	1315
10	CTG Leu	AGC Ser	TTC Phe	AAT Asn 385	GGT Gly	GTT Val	ATT Ile	ACC Met 390	ATG Met	AGT Ser	TCA Ser	AAC Asn	TTC Phe	TTG Leu 395	GGC Gly	TTA Leu	1363
	GAA Glu	CAA Gln	CTA Leu 400	GAA Glu	CAT His	CTG Leu	GAT Asp	TTC Phe 405	CAG Gln	CAT His	TCC Ser	AAT Asn	TTG Leu 410	AAA Lys	CAA Gln	ATG Met	1411
15	AGT Ser	GAG Glu 415	TTT Phe	TCA Ser	GTA Val	TTC Phe	CTA Leu 420	TCA Ser	CTC Leu	AGA Arg	AAC Asn	CTC Leu 425	ATT Ile	TAC Tyr	CTT Leu	GAC Asp	1459
20	ATT Ile 430	TCT Ser	CAT His	ACT Thr	CAC His	ACC Thr 435	AGA Arg	GTT Val	GCT Ala	TTC Phe	AAT Asn 440	GGC Gly	ATC Ile	TTC Phe	AAT Asn	GGC Gly 445	1507
	TTG Leu	TCC Ser	AGT Ser	CTC Leu	GAA Glu 450	GTC Val	TTG Leu	AAA Lys	ATG Met	GCT Ala 455	GGC Gly	AAT Asn	TCT Ser	TTC Phe	CAG Gln 460	GAA Glu	1555
25	AAC Asn	TTC Phe	CTT Leu	CCA Pro 465	GAT Asp	ATC Ile	TTC Phe	ACA Thr	GAG Glu 470	CTG Leu	AGA Arg	AAC Asn	TTG Leu	ACC Thr 475	TTC Phe	CTG Leu	1603
30	GAC Asp	CTC Leu	TCT Ser 480	CAG Gln	TGT Cys	CAA Gln	CTG Leu	GAG Glu 485	CAG Gln	TTG Leu	TCT Ser	CCA Pro	ACA Thr 490	GCA Ala	TTT Phe	AAC Asn	1651
	TCA Ser	CTC Leu 495	TCC Ser	AGT Ser	CTT Leu	CAG Gln	GTA Val 500	CTA Leu	AAT Asn	ATG Met	AGC Ser	CAC His 505	AAC Asn	AAC Asn	TTC Phe	TTT Phe	1699
35	TCA Ser 510	TTG Leu	GAT Asp	ACG Thr	TTT Phe	CCT Pro 515	TAT Tyr	AAG Lys	TGT Cys	CTG Leu	AAC Asn 520	TCC Ser	CTC Leu	CAG Gln	GTT Val 525	CTT Leu	1747
40	GAT Asp	TAC Tyr	AGT Ser	CTC Leu	AAT Asn 530	CAC His	ATA Ile	ATG Met	ACT Thr	TCC Ser 535	AAA Lys	AAA Lys	CAG Gln	GAA Glu	CTA Leu 540	CAG Gln	1795
	CAT His	TTT Phe	CCA Pro	AGT Ser 545	AGT Ser	CTA Leu	GCT Ala	TTC Phe	TTA Leu 550	AAT Asn	CTT Leu	ACT Thr	CAG Gln	AAT Asn 555	GAC Asp	TTT Phe	1843
45	GCT Ala	TGT Cys	ACT Thr 560	TGT Cys	GAA Glu	CAC His	CAG Gln	AGT Ser 565	TTC Phe	CTG Leu	CAA Gln	TGG Trp	ATC Ile 570	AAG Lys	GAC Asp	CAG Gln	1891
50	AGG Arg	CAG Gln 575	CTC Leu	TTG Leu	GTG Val	GAA Glu	GTT Val 580	GAA Glu	CGA Arg	ATG Met	GAA Glu 585	TGT Cys	GCA Ala	ACA Thr	CCT Pro	TCA Ser	1939
	GAT Asp 590	AAG Lys	CAG Gln	GGC Gly	ATG Met	CCT Pro 595	GTG Val	CTG Leu	AGT Ser	TTG Leu	AAT Asn 600	ATC Ile	ACC Thr	TGT Cys	CAG Gln	ATG Met 605	1987
55	AAT Asn	AAG Lys	ACC Thr	ATC Ile	ATT Ile 610	GGT Gly	GTG Val	TCG Ser	GTC Val	CTC Leu 615	AGT Ser	GTG Val	CTT Leu	GTA Val	GTA Val 620	TCT Ser	2035

# ES 2 340 210 T3

	GTT Val	GTA Val	GCA Ala	GTT Val 625	CTG Leu	GTC Val	TAT Tyr	AAG Lys	TTC Phe 630	TAT Tyr	TTT Phe	CAC His	CTG Leu	ATG Met 635	CTT Leu	CTT Leu	2083	
5	GCT Ala	GGC Gly	TGC Cys 640	ATA Ile	AAG Lys	TAT Tyr	GGT Gly	AGA Arg 645	GGT Gly	GAA Glu	AAC Asn	ATC Ile	TAT Tyr 650	GAT Asp	GCC Ala	TTT Phe	2131	
	GTT Val	ATC Ile 655	TAC Tyr	TCA Ser	AGC Ser	CAG Gln	GAT Asp 660	GAG Glu	GAC Asp	TGG Trp	GTA Val	AGG Arg 665	AAT Asn	GAG Glu	CTA Leu	GTA Val	2179	
10	AAG Lys 670	AAT Asn	TTA Leu	GAA Glu	GAA Glu	GGG Gly 675	GTG Val	CCT Pro	CCA Pro	TTT Phe	CAG Gln 680	CTC Leu	TGC Cys	CTT Leu	CAC His	TAC Tyr 685	2227	
15	AGA Arg	GAC Asp	TTT Phe	ATT Ile	CCC Pro 690	GGT Gly	GTG Val	GCC Ala	ATT Ile	GCT Ala 695	GCC Ala	AAC Asn	ATC Ile	ATC Ile	CAT His 700	GAA Glu	2275	
	GGT Gly	TTC Phe	CAT His	AAA Lys 705	AGC Ser	CGA Arg	AAG Lys	GTG Val	ATT Ile 710	GTT Val	GTG Val	GTG Val	TCC Ser	CAG Gln 715	CAC His	TTC Phe	2323	
20	ATC Ile	CAG Gln	AGC Ser 720	CGC Arg	TGG Trp	TGT Cys	ATC Ile	TTT Phe 725	GAA Glu	TAT Tyr	GAG Glu	ATT Ile	GCT Ala 730	CAG Gln	ACC Thr	TGG Trp	2371	
25	CAG Gln	TTT Phe 735	CTG Leu	AGC Ser	AGT Ser	CGT Arg	GCT Ala 740	GGT Gly	ATC Ile	ATC Ile	TTC Phe	ATT Ile 745	GTC Val	CTG Leu	CAG Gln	AAG Lys	2419	
	GTG Val 750	GAG Glu	AAG Lys	ACC Thr	CTG Leu	CTC Leu 755	AGG Arg	CAG Gln	CAG Gln	GTG Val	GAG Glu 760	CTG Leu	TAC Tyr	CGC Arg	CTT Leu	CTC Leu 765	2467	
30	AGC Ser	AGG Arg	AAC Asn	ACT Thr	TAC Tyr 770	CTG Leu	GAG Glu	TGG Trp	GAG Glu	GAC Asp 775	AGT Ser	GTC Val	CTG Leu	GGG Gly	CGG Arg 780	CAC His	2515	
35	ATC Ile	TTC Phe	TGG Trp	AGA Arg 785	CGA Arg	CTC Leu	AGA Arg	AAA Lys	GCC Ala 790	CTG Leu	CTG Leu	GAT Asp	GGT Gly	AAA Lys 795	TCA Ser	TGG Trp	2563	
	AAT Asn	CCA Pro	GAA Glu 800	GGA Gly	ACA Thr	GTG Val	GGT Gly	ACA Thr 805	GGA Gly	TGC Cys	AAT Asn	TGG Trp	CAG Gln 810	GAA Glu	GCA Ala	ACA Thr	2611	
40	TCT Ser	ATC Ile 815	TGAAGAGGAA AAATAAAAAC CTCCTGAGGC ATTTCTTGCC CAGCTGGGTC														2667	
45	CAACACTTGT TCAGTTAATA AGTATTAAAT GCTGCCACAT GTCAGGCCTT ATGCTAAGGG																	2727
	TGAGTAATTC CATGGTGCAC TAGATATGCA GGGCTGCTAA TCTCAAGGAG CTTCCAGTGC																	2787
	AGAGGGAATA AATGCTAGAC TAAATACAG AGTCTTCCAG GTGGGCATTT CAACCAACTC																	2847
50	AGTCAAGGAA CCCATGACAA AGAAAGTCAT TTCAACTCTT ACCTCATCAA GTTGAATAAA																	2907
	GACAGAGAAA ACAGAAAGAG ACATTGTTCT TTTCTGAGT CTTTGAATG GAAATTGTAT																	2967
	TATGTTATAG CCATCATAAA ACCATTTTGG TAGTTTGGAC TGAAGTGGGT GTTCACTTTT																	3027
	TCCTTTTGA TTGAATACAA TTAAATTCT ACTTGATGAC TGCAGTCGTC AAGGGGCTCC																	3087
55	TGATGCAAGA TGCCCTTCC ATTTTAAAGTC TGTCTCCTTA CAGAGGTTAA AGTCTAATGG																	3147

## ES 2 340 210 T3

	CTAATTCCTA AGGAAACCTG ATTAACACAT GCTCACAACC ATCCTGGTCA TTCTCGAACA	3207
	TGTTCTATTT TTTAACTAAT CACCCCTGAT ATATTTTAT TTTTATATAT CCAGTTTTCA	3267
5	TTTTTTTACG TCTTGCCTAT AAGCTAATAT CATAAATAAG GTTGTTTAAG ACGTGCTTCA	3327
	AATATCCATA TTAACCACTA TTTTTC AAGG AAGTATGGAA AAGTACACTC TGCTACTTTG	3387
	TCACTCGATG TCATTCCAAA GTTATTGCCT ACTAAGTAAT GACTGTCATG AAAGCAGCAT	3447
10	TGAAATAATT TGTTTAAAGG GGGCACTCTT TTAAACGGGA AGAAAATTTT CGCTTCCTGG	3507
	TCTTATCATG GACAATTTGG GCTAGAGGCA GGAAGGAAGT GGGATGACCT CAGGAAGTCA	3567
	CCTTTTCTTG ATTCCAGAAA CATATGGGCT GATAAACCCG GGGTGACCTC ATGAAATGAG	3627
15	TTGCAGCAGA AGTTTATTTT TTTCAGAACAA AGTGATGTTT GATGGACCTC TGAATCTCTT	3687
	TAGGGAGACA CAGATGGCTG GGATCCCTCC CCTGTACCCT TCTCACTGCC AGGAGAACTA	3747
	CGTGTGAAGG TATTCAAGGC AGGGAGTATA CATTGCTGTT TCCTGTTGGG CAATGCTCCT	3807
20	TGACCACATT TTGGGAAGAG TGGATGTTAT CATTGAGAAA ACAATGTGTC TGGGAATTAAT	3867
	GGGGTTCTTA TAAAGAAGGT TCCCAGAAAA GAATGTTTCT TCCAGCTTCT TCAGGAAACA	3927
	GGAACATTCA AGGAAAAGGA CAATCAGGAT GTCATCAGGG AAATGAAAAT AAAAACCACA	3987
25	ATGAGATATC ACCTTATACC AGGTAGATGG CTACTATAAA AAAATGAAGT GTCATCAAGG	4047
	ATATAGAGAA ATTGGAACCC TTCTTCACTG CTGGAGGGAA TGGAAAATGG TGTAGCCGTT	4107
	ATGAAAAACA GTACGGAGGT TTCTCAAAAA TTAAAAATAG AACTGCTATA TGATCCAGCA	4167
30	ATCTCACTTC TGTATATATA CCCAAAAATA TTGAAATCAG AATTTC AAGA AAATATTTAC	4227
	ACTCCCATGT TCATTGTGGC ACTCTTCACA ATCACTGTTT CCAAAGTTAT GGAAACAACC	4287
	CAAATTTCCA TTGGAAAATA AATGGACAAA GGAATGTGTC ATATAACGTA CAATGGGGAT	4347
35	ATTATTCAGC CTAAAAAAG GGGGGATCCT GTTATTTATG ACAACATGAA TAAACCCGGA	4407
	GGCCATTATG CTATGTAAAA TGAGCAAGTA ACAGAAAGAC AAATACTGCC TGATTTTATT	4467
	TATATGAGGT TCTAAAATAG TCAAACCTCAT AGAAGCAGAG AATAGAACAG TGGTTCCTAG	4527
40	GGAAAAGGAG GAAGGGAGAA ATGAGGAAAT AGGGAGTTGT CTAATTGGTA TAAAATTATA	4587
	GTATGCAAGA TGAATTAGCT CTAAAGATCA GCTGTATAGC AGAGTTCGTA TAATGAACAA	4647
	TACTGTATTA TGCACTTAAC A'TTTGTAA GAGGGTACCT CTCATGTAA GTGTTCTTAC	4707
45	CATATACATA TACACAAGGA AGCTTTTGGG GGTGATGGAT ATATTTATTA CCTTGATTGT	4767
	GGTGATGGTT TGACAGGTAT GTGACTATGT CTAAACTCAT CAAATTGTAT ACATTAAATA	4827
50	TATGCAGTTT TATAATATCA AAAAAAAAAA AAAAAAAAAA	4865

### (2) INFORMACIÓN PARA EL SEQ ID NO:26:

- 55 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 837 aminoácidos
- (B) TIPO: aminoácido
- (D) TOPOLOGÍA: lineal
- 60 (ii) TIPO DE MOLÉCULA: proteína
- (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:26:

65

# ES 2 340 210 T3

	Met	Ser	Ala	Ser	Arg	Leu	Ala	Gly	Thr	Leu	Ile	Pro	Ala	Met	Ala	Phe
	-22		-20					-15					-10			
5	Leu	Ser	Cys	Val	Arg	Pro	Glu	Ser	Trp	Glu	Pro	Cys	Val	Glu	Val	Pro
		-5					1				5					10
	Asn	Ile	Thr	Tyr	Gln	Cys	Met	Glu	Leu	Asn	Phe	Tyr	Lys	Ile	Pro	Asp
					15					20					25	
10	Asn	Leu	Pro	Phe	Ser	Thr	Lys	Asn	Leu	Asp	Leu	Ser	Phe	Asn	Pro	Leu
				30					35					40		
	Arg	His	Leu	Gly	Ser	Tyr	Ser	Phe	Phe	Ser	Phe	Pro	Glu	Leu	Gln	Val
			45					50					55			
15	Leu	Asp	Leu	Ser	Arg	Cys	Glu	Ile	Gln	Thr	Ile	Glu	Asp	Gly	Ala	Tyr
		60					65					70				
	Gln	Ser	Leu	Ser	His	Leu	Ser	Thr	Leu	Ile	Leu	Thr	Gly	Asn	Pro	Ile
		75				80					85					90
20	Gln	Ser	Leu	Ala	Leu	Gly	Ala	Phe	Ser	Gly	Leu	Ser	Ser	Leu	Gln	Lys
					95					100					105	
	Leu	Val	Ala	Val	Glu	Thr	Asn	Leu	Ala	Ser	Leu	Glu	Asn	Phe	Pro	Ile
				110					115					120		
25	Gly	His	Leu	Lys	Thr	Leu	Lys	Glu	Leu	Asn	Val	Ala	His	Asn	Leu	Ile
			125					130					135			
	Gln	Ser	Phe	Lys	Leu	Pro	Glu	Tyr	Phe	Ser	Asn	Leu	Thr	Asn	Leu	Glu
		140					145					150				
30	His	Leu	Asp	Leu	Ser	Ser	Asn	Lys	Ile	Gln	Ser	Ile	Tyr	Cys	Thr	Asp
	155					160					165					170
	Leu	Arg	Val	Leu	His	Gln	Met	Pro	Leu	Leu	Asn	Leu	Ser	Leu	Asp	Leu
					175					180					185	
35	Ser	Leu	Asn	Pro	Met	Asn	Phe	Ile	Gln	Pro	Gly	Ala	Phe	Lys	Glu	Ile
				190					195					200		
	Arg	Leu	His	Lys	Leu	Thr	Leu	Arg	Asn	Asn	Phe	Asp	Ser	Leu	Asn	Val
			205					210					215			
40	Met	Lys	Thr	Cys	Ile	Gln	Gly	Leu	Ala	Gly	Leu	Glu	Val	His	Arg	Leu
		220					225					230				
	Val	Leu	Gly	Glu	Phe	Arg	Asn	Glu	Gly	Asn	Leu	Glu	Lys	Phe	Asp	Lys
		235				240					245					250
45	Ser	Ala	Leu	Glu	Gly	Leu	Cys	Asn	Leu	Thr	Ile	Glu	Glu	Phe	Arg	Leu
					255					260					265	
	Ala	Tyr	Leu	Asp	Tyr	Tyr	Leu	Asp	Asp	Ile	Ile	Asp	Leu	Phe	Asn	Cys
				270					275					280		
50	Leu	Thr	Asn	Val	Ser	Ser	Phe	Ser	Leu	Val	Ser	Val	Thr	Ile	Glu	Arg
				285				290					295			
	Val	Lys	Asp	Phe	Ser	Tyr	Asn	Phe	Gly	Trp	Gln	His	Leu	Glu	Leu	Val
		300					305					310				
55	Asn	Cys	Lys	Phe	Gly	Gln	Phe	Pro	Thr	Leu	Lys	Leu	Lys	Ser	Leu	Lys
	315					320					325					330

# ES 2 340 210 T3

	Arg	Leu	Thr	Phe	Thr 335	Ser	Asn	Lys	Gly	Gly 340	Asn	Ala	Phe	Ser	Glu 345	Val
5	Asp	Leu	Pro	Ser 350	Leu	Glu	Phe	Leu	Asp 355	Leu	Ser	Arg	Asn	Gly 360	Leu	Ser
	Phe	Lys	Gly 365	Cys	Cys	Ser	Gln	Ser 370	Asp	Phe	Gly	Thr	Thr 375	Ser	Leu	Lys
10	Tyr	Leu 380	Asp	Leu	Ser	Phe	Asn 385	Gly	Val	Ile	Thr	Met 390	Ser	Ser	Asn	Phe
	Leu	Gly 395	Leu	Glu	Gln	Leu	Glu	His	Leu	Asp	Phe 405	Gln	His	Ser	Asn	Leu 410
15	Lys	Gln	Met	Ser	Glu 415	Phe	Ser	Val	Phe	Leu 420	Ser	Leu	Arg	Asn	Leu 425	Ile
	Tyr	Leu	Asp	Ile 430	Ser	His	Thr	His	Thr 435	Arg	Val	Ala	Phe	Asn 440	Gly	Ile
20	Phe	Asn	Gly 445	Leu	Ser	Ser	Leu	Glu	Val 450	Leu	Lys	Met	Ala 455	Gly	Asn	Ser
	Phe	Gln	Glu	Asn	Phe	Leu	Pro 465	Asp	Ile	Phe	Thr	Glu 470	Leu	Arg	Asn	Leu
25	Thr	Phe	Leu	Asp	Leu	Ser 480	Gln	Cys	Gln	Leu	Glu 485	Gln	Leu	Ser	Pro	Thr 490
	Ala	Phe	Asn	Ser	Leu 495	Ser	Ser	Leu	Gln	Val 500	Leu	Asn	Met	Ser	His 505	Asn
30	Asn	Phe	Phe	Ser 510	Leu	Asp	Thr	Phe	Pro 515	Tyr	Lys	Cys	Leu	Asn 520	Ser	Leu
	Gln	Val	Leu 525	Asp	Tyr	Ser	Leu	Asn 530	His	Ile	Met	Thr	Ser 535	Lys	Lys	Gln
35	Glu	Leu 540	Gln	His	Phe	Pro	Ser 545	Ser	Leu	Ala	Phe	Leu 550	Asn	Leu	Thr	Gln
	Asn	Asp	Phe	Ala	Cys	Thr 560	Cys	Glu	His	Gln	Ser 565	Phe	Leu	Gln	Trp	Ile 570
40	Lys	Asp	Gln	Arg	Gln 575	Leu	Leu	Val	Glu	Val 580	Glu	Arg	Met	Glu	Cys 585	Ala
	Thr	Pro	Ser	Asp 590	Lys	Gln	Gly	Met	Pro 595	Val	Leu	Ser	Leu	Asn 600	Ile	Thr
45	Cys	Gln	Met 605	Asn	Lys	Thr	Ile	Ile 610	Gly	Val	Ser	Val	Leu 615	Ser	Val	Leu
	Val	Val	Ser	Val	Val	Ala	Val 625	Leu	Val	Tyr	Lys	Phe	Tyr	Phe	His	Leu
50	Met	Leu	Leu	Ala	Gly	Cys 640	Ile	Lys	Tyr	Gly	Arg 645	Gly	Glu	Asn	Ile	Tyr 650
	Asp	Ala	Phe	Val	Ile 655	Tyr	Ser	Ser	Gln	Asp 660	Glu	Asp	Trp	Val	Arg 665	Asn
55	Glu	Leu	Val	Lys 670	Asn	Leu	Glu	Glu	Gly 675	Val	Pro	Pro	Phe	Gln 680	Leu	Cys
	Leu	His	Tyr 685	Arg	Asp	Phe	Ile	Pro 690	Gly	Val	Ala	Ile	Ala 695	Ala	Asn	Ile

## ES 2 340 210 T3

5           Ile His Glu Gly Phe His Lys Ser Arg Lys Val Ile Val Val Val Ser  
               700                               705                               710  
 10       Gln His Phe Ile Gln Ser Arg Trp Cys Ile Phe Glu Tyr Glu Ile Ala  
               715                               720                               725                               730  
       Gln Thr Trp Gln Phe Leu Ser Ser Arg Ala Gly Ile Ile Phe Ile Val  
                               735                               740                               745  
 15       Leu Gln Lys Val Glu Lys Thr Leu Leu Arg Gln Gln Val Glu Leu Tyr  
                               750                               755                               760  
       Arg Leu Leu Ser Arg Asn Thr Tyr Leu Glu Trp Glu Asp Ser Val Leu  
                               765                               770                               775  
 20       Gly Arg His Ile Phe Trp Arg Arg Leu Arg Lys Ala Leu Leu Asp Gly  
               780                               785                               790  
       Lys Ser Trp Asn Pro Glu Gly Thr Val Gly Thr Gly Cys Asn Trp Gln  
               795                               800                               805                               810  
       Glu Ala Thr Ser Ile  
                               815

### (2) INFORMACIÓN PARA EL SEQ ID NO:27:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 300 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..300

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 186
- (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 186, 196, 217, 276, y 300 denominados C, pueden ser cada uno A, C, G, o T"

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:27:

## ES 2 340 210 T3

	TCC	TAT	TCT	ATG	GAA	AAA	GAT	GCT	TTC	CTA	TTT	ATG	AGA	AAT	TTG	AAG	48
	Ser	Tyr	Ser	Met	Glu	Lys	Asp	Ala	Phe	Leu	Phe	Met	Arg	Asn	Leu	Lys	
	1				5				10					15			
5	GTT	CTC	TCA	CTA	AAA	GAT	AAC	AAT	GTC	ACA	GCT	GTC	CCC	ACC	ACT	TTG	96
	Val	Leu	Ser	Leu	Lys	Asp	Asn	Asn	Val	Thr	Ala	Val	Pro	Thr	Thr	Leu	
				20					25				30				
	CCA	CCT	AAT	TTA	CTA	GAG	CTC	TAT	CTT	TAT	AAC	AAT	ATC	ATT	AAG	AAA	144
	Pro	Pro	Asn	Leu	Leu	Glu	Leu	Tyr	Leu	Tyr	Asn	Asn	Ile	Ile	Lys	Lys	
			35					40					45				
10	ATC	CAA	GAA	AAT	GAT	TTC	AAT	AAC	CTC	AAT	GAG	TTG	CAA	GTC	CTT	GAC	192
	Ile	Gln	Glu	Asn	Asp	Phe	Asn	Asn	Leu	Asn	Glu	Leu	Gln	Val	Leu	Asp	
		50					55					60					
	CTA	CGT	GGA	AAT	TGC	CCT	CGA	TGT	CAT	AAT	GTC	CCA	TAT	CCG	TGT	ACA	240
15	Leu	Arg	Gly	Asn	Cys	Pro	Arg	Cys	His	Asn	Val	Pro	Tyr	Pro	Cys	Thr	
	65					70					75					80	
	CCG	TGT	GAA	AAT	AAT	TCC	CCC	TTA	CAG	ATC	CAT	GAC	AAT	GCT	TTC	AAT	288
	Pro	Cys	Glu	Asn	Asn	Ser	Pro	Leu	Gln	Ile	His	Asp	Asn	Ala	Phe	Asn	
					85					90					95		
20	TCA	TCG	ACA	GAC													300
	Ser	Ser	Thr	Asp													
				100													

### (2) INFORMACIÓN PARA EL SEQ ID NO:28:

25

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 100 aminoácidos

(B) TIPO: aminoácido

30

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:28:

35

Ser Tyr Ser Met Glu Lys Asp Ala Phe Leu Phe Met Arg Asn Leu Lys  
1 5 10 15

Val Leu Ser Leu Lys Asp Asn Asn Val Thr Ala Val Pro Thr Thr Leu  
20 25 30

40

Pro Pro Asn Leu Leu Glu Leu Tyr Leu Tyr Asn Asn Ile Ile Lys Lys  
35 40 45

Ile Gln Glu Asn Asp Phe Asn Asn Leu Asn Glu Leu Gln Val Leu Asp  
50 55 60

45

Leu Arg Gly Asn Cys Pro Arg Cys His Asn Val Pro Tyr Pro Cys Thr  
65 70 75 80

Pro Cys Glu Asn Asn Ser Pro Leu Gln Ile His Asp Asn Ala Phe Asn  
85 90 95

50

Ser Ser Thr Asp  
100

### (2) INFORMACIÓN PARA EL SEQ ID NO:29:

55	(i)	CARACTERÍSTICAS DE LA SECUENCIA:
	(A)	LONGITUD: 1756 pares de bases
	(B)	TIPO: ácido nucleico
60	(C)	CUALIDAD DE LA HEBRA: sencilla
	(D)	TOPOLOGÍA: lineal
	(ii)	TIPO DE MOLÉCULA: ADNc
	(ix)	RASGO:
65	(A)	NOMBRE/CLAVE: CDS
	(B)	LOCALIZACIÓN: 1..1182



## ES 2 340 210 T3

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 1643

(D) OTRA INFORMACIÓN: /nota= "el nucleótido 1643 denominado A, puede ser A o G"

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 1664

(D) OTRA INFORMACIÓN: /nota= "el nucleótido 1664 denominado C, puede ser A, C, G, o T"

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 1680

(D) OTRA INFORMACIÓN: /nota= "los nucleótidos 1680 y 1735 denominado G, puede ser G o T"

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 1719

(D) OTRA INFORMACIÓN: /nota= "el nucleótido 1719 denominado C, puede ser C o T"

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 1727

(D) OTRA INFORMACIÓN: /nota= "el nucleótido 1727 denominado A, puede ser A, G, o T"

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:29:

TCT CCA GAA ATT CCC TGG AAT TCC TTG CCT CCT GAG GTT TTT GAG GGT	48
Ser Pro Glu Ile Pro Trp Asn Ser Leu Pro Pro Glu Val Phe Glu Gly	
1 5 10 15	
ATG CCG CCA AAT CTA AAG AAT CTC TCC TTG GCC AAA AAT GGG CTC AAA	96
Met Pro Pro Asn Leu Lys Asn Leu Ser Leu Ala Lys Asn Gly Leu Lys	
20 25 30	
TCT TTC TTT TGG GAC AGA CTC CAG TTA CTG AAG CAT TTG GAA ATT TTG	144
Ser Phe Phe Trp Asp Arg Leu Gln Leu Leu Lys His Leu Glu Ile Leu	
35 40 45	
GAC CTC AGC CAT AAC CAG CTG ACA AAA GTA CCT GAG AGA TTG GCC AAC	192
Asp Leu Ser His Asn Gln Thr Lys Val Pro Glu Arg Leu Ala Asn	
50 55 60	
TGT TCC AAA AGT CTC ACA ACA CTG ATT CTT AAG CAT AAT CAA ATC AGG	240
Cys Ser Lys Ser Leu Thr Thr Leu Ile Leu Lys His Asn Gln Ile Arg	
65 70 75 80	
CAA TTG ACA AAA TAT TTT CTA GAA GAT GCT TTG CAA TTG CGC TAT CTA	288
Gln Leu Thr Lys Tyr Phe Leu Glu Asp Ala Leu Gln Leu Arg Tyr Leu	
85 90 95	
GAC ATC AGT TCA AAT AAA ATC CAG GTC ATT CAG AAG ACT AGC TTC CCA	336
Asp Ile Ser Ser Asn Lys Ile Gln Val Ile Gln Lys Thr Ser Phe Pro	
100 105 110	
GAA AAT GTC CTC AAC AAT CTG GAG ATG TTG GTT TTA CAT CAC AAT CGC	384
Glu Asn Val Leu Asn Asn Leu Glu Met Leu Val Leu His His Asn Arg	
115 120 125	
TTT CTT TGC AAC TGT GAT GCT GTG TGG TTT GTC TGG TGG GTT AAC CAT	432
Phe Leu Cys Asn Cys Asp Ala Val Trp Phe Val Trp Trp Val Asn His	
130 135 140	
ACA GAT GTT ACT ATT CCA TAC CTG GCC ACT GAT GTG ACT TGT GTA GGT	480
Thr Asp Val Thr Ile Pro Tyr Leu Ala Thr Asp Val Thr Cys Val Gly	
145 150 155 160	
CCA GGA GCA CAC AAA GGT CAA AGT GTC ATA TCC CTT GAT CTG TAT ACG	528
Pro Gly Ala His Lys Gly Gln Ser Val Ile Ser Leu Asp Leu Tyr Thr	
165 170 175	
TGT GAG TTA GAT CTC ACA AAC CTG ATT CTG TTC TCA GTT TCC ATA TCA	576
Cys Glu Leu Asp Leu Thr Asn Leu Ile Leu Phe Ser Val Ser Ile Ser	

# ES 2 340 210 T3

	180	185	190	
5	TCA GTC CTC TTT CTT ATG GTA GTT ATG ACA ACA AGT CAC CTC TTT TTC Ser Val Leu 195 Phe Leu Met Val Val Met Thr Thr Ser His 205 Leu Phe Phe			624
	TGG GAT ATG TGG TAC ATT TAT TAT TTT TGG AAA GCA AAG ATA AAG GGG Trp Asp Met 210 Trp Tyr Ile Tyr Tyr Phe Trp Lys Ala Lys Ile Lys Gly			672
10	TAT CCA GCA TCT GCA ATC CCA TGG AGT CCT TGT TAT GAT GCT TTT ATT Tyr Pro Ala Ser Ala Ile Pro Trp Ser Pro Cys 235 Tyr Asp Ala Phe Ile 240			720
15	GTG TAT GAC ACT AAA AAC TCA GCT GTG ACA GAA TGG GTT TTG CAG GAG Val Tyr Asp Thr Lys 245 Asn Ser Ala Val Thr Glu Trp Val Leu Gln Glu 255			768
	CTG GTG GCA AAA TTG GAA GAT CCA AGA GAA AAA CAC TTC AAT TTG TGT Leu Val Ala 260 Leu Glu Asp Pro Arg 265 Glu Lys His Phe Asn 270 Leu Cys			816
20	CTA GAA GAA AGA GAC TGG CTA CCA GGA CAG CCA GTT CTA GAA AAC CTT Leu Glu 275 Arg Asp Trp Leu Pro 280 Gly Gln Pro Val 285 Leu Glu Asn Leu			864
25	TCC CAG AGC ATA CAG CTC AGC AAA AAG ACA GTG TTT GTG ATG ACA CAG Ser Gln Ser 290 Ile Gln Leu Ser Lys Lys Thr Val Phe Val Met Thr Gln 300			912
	AAA TAT GCT AAG ACT GAG AGT TTT AAG ATG GCA TTT TAT TTG TCT CAT Lys Tyr Ala 305 Lys Thr Glu Ser Phe Lys Met Ala Phe Tyr Leu Ser His 320			960
30	CAG AGG CTC CTG GAT GAA AAA GTG GAT GTG ATT ATC TTG ATA TTC TTG Gln Arg Leu 325 Asp Glu Lys Val Asp Val Ile Ile Leu Ile Phe Leu 335			1008
35	GAA AGA CCT CTT CAG AAG TCT AAG TTT CTT CAG CTC AGG AAG AGA CTC Glu Arg Pro 340 Gln Lys Ser Lys Phe 345 Leu Gln Leu Arg Lys Arg Leu 350			1056
	TGC AGG AGC TCT GTC CTT GAG TGG CCT GCA AAT CCA CAG GCT CAC CCA Cys Arg Ser 355 Ser Val Leu Glu Trp Pro Ala Asn Pro Gln Ala His Pro 365			1104
40	TAC TTC TGG CAG TGC CTG AAA AAT GCC CTG ACC ACA GAC AAT CAT GTG Tyr Phe Trp Gln Cys Leu Lys Asn Ala Leu Thr Thr Asp Asn His Val 370 375 380			1152
45	GCT TAT AGT CAA ATG TTC AAG GAA ACA GTC TAGCTCTCTG AAGAATGTCA Ala Tyr Ser Gln Met Phe Lys Glu Thr Val 385 390			1202
	CCACCTAGGA CATGCCTTGG TACCTGAAGT TTTCATAAAG GTTTCATAA ATGAAGGTCT			1262
	GAATTTTTC TAACAGTTGT CATGGCTCAG ATTGGTGGGA AATCATCAAT ATATGGCTAA			1322
50	GAAATTAAGA AGGGGAGACT GATAGAAGAT AATTCTTTTC TTCATGTGCC ATGCTCAGTT			1382
	AAATATTTTC CTTAGCTCAA ATCTGAAAAA CTGTGCCTAG GAGACAACAC AAGGCTTTGA			1442
	TTTATCTGCA TACAATTGAT AAGAGCCACA CATCTGCCCT GAAGAAGTAC TAGTAGTTTT			1502
55	AGTAGTAGGG TAAAAATTAC ACAAGCTTTC TCTCTCTCTG ATACTGAACT GTACCAGAGT			1562
	TCAATGAAAT AAAAGCCAG AGAACTTCTC AGTAAATGGT TTCATTATCA TGTAGTATCC			1622
	ACCATGCAAT ATGCCACAAA ACCGCTACTG GTACAGGACA GCTGGTAGCT GCTTCAAGGC			1682
60	CTCTTATCAT TTTCTTGGGG CCCATGGAGG GGTCTCTCTG GAAAAAGGGA AGGTTTTTTT			1742
	TGGCCATCCA TGAA			1756

## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:30:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 394 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:30:

```

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

```

ES 2 340 210 T3

		Leu	Glu	Glu 275	Arg	Asp	Trp	Leu	Pro 280	Gly	Gln	Pro	Val	Leu 285	Glu	Asn	Leu
5		Ser	Gln 290	Ser	Ile	Gln	Leu	Ser 295	Lys	Lys	Thr	Val	Phe 300	Val	Met	Thr	Gln
		Lys 305	Tyr	Ala	Lys	Thr	Glu 310	Ser	Phe	Lys	Met	Ala 315	Phe	Tyr	Leu	Ser	His 320
10		Gln	Arg	Leu	Leu	Asp 325	Glu	Lys	Val	Asp	Val 330	Ile	Ile	Leu	Ile	Phe 335	Leu
		Glu	Arg	Pro	Leu 340	Gln	Lys	Ser	Lys	Phe 345	Leu	Gln	Leu	Arg	Lys 350	Arg	Leu
15		Cys	Arg	Ser 355	Ser	Val	Leu	Glu	Trp 360	Pro	Ala	Asn	Pro	Gln 365	Ala	His	Pro
		Tyr	Phe 370	Trp	Gln	Cys	Leu	Lys 375	Asn	Ala	Leu	Thr	Thr 380	Asp	Asn	His	Val
20		Ala 385	Tyr	Ser	Gln	Met	Phe 390	Lys	Glu	Thr	Val						

25 (2) INFORMACIÓN PARA EL SEQ ID NO:31:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 999 pares de bases

(B) TIPO: ácido nucleico

(C) CUALIDAD DE LA HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADNc

(ix) RASGO:

(A) NOMBRE/CLAVE: CDS

(B) LOCALIZACIÓN: 2..847

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 4

(D) OTRA INFORMACIÓN: /nota= “los nucleótidos 4 y 23 denominados C, pueden ser cada uno A, C, G, o T”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 650

(D) OTRA INFORMACIÓN: /nota= “el nucleótido 650 denominado G, puede ser A o G”

(ix) RASGO:

(A) NOMBRE/CLAVE: rasgo\_misc

(B) LOCALIZACIÓN: 715

(D) OTRA INFORMACIÓN: /nota= “los nucleótidos 715, 825, y 845 denominados C, pueden ser cada uno C o T”

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:31:

# ES 2 340 210 T3

	C TCC GAT GCC AAG ATT CGG CAC CAG GCA TAT TCA GAG GTC ATG ATG	46
	Ser Asp Ala Lys Ile Arg His Gln Ala Tyr Ser Glu Val Met Met	
	1 5 10 15	
5	GTT GGA TGG TCA GAT TCA TAC ACC TGT GAA TAC CCT TTA AAC CTA AGG	94
	Val Gly Trp Ser Asp Ser Tyr Thr Cys Glu Tyr Pro Leu Asn Leu Arg	
	20 25 30	
10	GGA ACT AGG TTA AAA GAC GTT CAT CTC CAC GAA TTA TCT TGC AAC ACA	142
	Gly Thr Arg Leu Lys Asp Val His Leu His Glu Leu Ser Cys Asn Thr	
	35 40 45	
15	GCT CTG TTG ATT GTC ACC ATT GTG GTT ATT ATG CTA GTT CTG GGG TTG	190
	Ala Leu Leu Ile Val Thr Ile Val Val Ile Met Leu Val Leu Gly Leu	
	50 55 60	
20	GCT GTG GCC TTC TGC TGT CTC CAC TTT GAT CTG CCC TGG TAT CTC AGG	238
	Ala Val Ala Phe Cys Cys Leu His Phe Asp Leu Pro Trp Tyr Leu Arg	
	65 70 75	
25	ATG CTA GGT CAA TGC ACA CAA ACA TGG CAC AGG GTT AGG AAA ACA ACC	286
	Met Leu Gly Gln Cys Thr Gln Thr Trp His Arg Val Arg Lys Thr Thr	
	80 85 90 95	
30	CAA GAA CAA CTC AAG AGA AAT GTC CGA TTC CAC GCA TTT ATT TCA TAC	334
	Gln Glu Gln Leu Lys Arg Asn Val Arg Phe His Ala Phe Ile Ser Tyr	
	100 105 110	
35	AGT GAA CAT GAT TCT CTG TGG GTG AAG AAT GAA TTG ATC CCC AAT CTA	382
	Ser Glu His Asp Ser Leu Trp Val Lys Asn Glu Leu Ile Pro Asn Leu	
	115 120 125	
40	GAG AAG GAA GAT GGT TCT ATC TTG ATT TGC CTT TAT GAA AGC TAC TTT	430
	Glu Lys Glu Asp Gly Ser Ile Leu Ile Cys Leu Tyr Glu Ser Tyr Phe	
	130 135 140	
45	GAC CCT GGC AAA AGC ATT AGT GAA AAT ATT GTA AGC TTC ATT GAG AAA	478
	Asp Pro Gly Lys Ser Ile Ser Glu Asn Ile Val Ser Phe Ile Glu Lys	
	145 150 155	
50	AGC TAT AAG TCC ATC TTT GTT TTG TCT CCC AAC TTT GTC CAG AAT GAG	526
	Ser Tyr Lys Ser Ile Phe Val Leu Ser Pro Asn Phe Val Gln Asn Glu	
	160 165 170 175	
55	TGG TGC CAT TAT GAA TTC TAC TTT GCC CAC CAC AAT CTC TTC CAT GAA	574
	Trp Cys His Tyr Glu Phe Tyr Phe Ala His His Asn Leu Phe His Glu	
	180 185 190	
60	AAT TCT GAT CAC ATA ATT CTT ATC TTA CTG GAA CCC ATT CCA TTC TAT	622
	Asn Ser Asp His Ile Ile Leu Ile Leu Leu Glu Pro Ile Pro Phe Tyr	
	195 200 205	
65	TGC ATT CCC ACC AGG TAT CAT AAA CTG GAA GCT CTC CTG GAA AAA AAA	670
	Cys Ile Pro Thr Arg Tyr His Lys Leu Glu Ala Leu Leu Glu Lys Lys	
	210 215 220	
70	GCA TAC TTG GAA TGG CCC AAG GAT AGG CGT AAA TGT GGG CTT TTC TGG	718
	Ala Tyr Leu Glu Trp Pro Lys Asp Arg Arg Lys Cys Gly Leu Phe Trp	
	225 230 235	
75	GCA AAC CTT CGA GCT GCT GTT AAT GTT AAT GTA TTA GCC ACC AGA GAA	766
	Ala Asn Leu Arg Ala Ala Val Asn Val Asn Val Leu Ala Thr Arg Glu	
	240 245 250 255	
80	ATG TAT GAA CTG CAG ACA TTC ACA GAG TTA AAT GAA GAG TCT CGA GGT	814
	Met Tyr Glu Leu Gln Thr Phe Thr Glu Leu Asn Glu Glu Ser Arg Gly	
	260 265 270	
85	TCT ACA ATC TCT CTG ATG AGA ACA GAC TGT CTA TAAATCCCA CAGTCCTTGG	867
	Ser Thr Ile Ser Leu Met Arg Thr Asp Cys Leu	
	275 280	
90	GAAGTTGGGG ACCACATACA CTGTTGGGAT GTACATTGAT ACAACCTTTA TGATGGCAAT	927
	TTGACAATAT TTATTAAAAT AAAAAATGGT TATTCCTTC AAAAAAAAAA AAAAAAAAAA	987
	AAAAAAAAAA AA	999

## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:32:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 282 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:32:

```

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

```

## ES 2 340 210 T3

### (2) INFORMACIÓN PARA EL SEQ ID NO:33:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1173 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNc

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: CDS
- (B) LOCALIZACIÓN: 1..1008

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 854
- (D) OTRA INFORMACIÓN: /nota= "el nucleótido 854 denominado A, puede ser A o T"

#### (ix) RASGO:

- (A) NOMBRE/CLAVE: rasgo\_misc
- (B) LOCALIZACIÓN: 1171
- (D) OTRA INFORMACIÓN: /nota= "los nucleótidos 1171 y 1172 denominados C, pueden ser cada uno A, C, G, o T"

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:33:

35	CTG CCT GCT GGC ACC CGG CTC CGG AGG CTG GAT GTC AGC TGC AAC AGC Leu Pro Ala Gly Thr Arg Leu Arg Arg Leu Asp Val Ser Cys Asn Ser <div style="display: flex; justify-content: space-between; width: 80%; margin: 0 auto;"> <span>1</span> <span>5</span> <span>10</span> <span>15</span> </div>	48
	ATC AGC TTC GTG GCC CCC GGC TTC TTT TCC AAG GCC AAG GAG CTG CGA Ile Ser Phe Val Ala Pro Gly Phe Phe Ser Lys Ala Lys Glu Leu Arg <div style="display: flex; justify-content: space-between; width: 80%; margin: 0 auto;"> <span>20</span> <span>25</span> <span>30</span> </div>	96
40	GAG CTC AAC CTT AGC GCC AAC GCC CTC AAG ACA GTG GAC CAC TCC TGG Glu Leu Asn Leu Ser Ala Asn Ala Leu Lys Thr Val Asp His Ser Trp <div style="display: flex; justify-content: space-between; width: 80%; margin: 0 auto;"> <span>35</span> <span>40</span> <span>45</span> </div>	144
45	TTT GGG CCC CTG GCG AGT GCC CTG CAA ATA CTA GAT GTA AGC GCC AAC Phe Gly Pro Leu Ala Ser Ala Leu Gln Ile Leu Asp Val Ser Ala Asn <div style="display: flex; justify-content: space-between; width: 80%; margin: 0 auto;"> <span>50</span> <span>55</span> <span>60</span> </div>	192
	CCT CTG CAC TGC GCC TGT GGG GCG GCC TTT ATG GAC TTC CTG CTG GAG Pro Leu His Cys Ala Cys Gly Ala Ala Phe Met Asp Phe Leu Leu Glu <div style="display: flex; justify-content: space-between; width: 80%; margin: 0 auto;"> <span>65</span> <span>70</span> <span>75</span> <span>80</span> </div>	240
50	GTG CAG GCT GCC GTG CCC GGT CTG CCC AGC CGG GTG AAG TGT GGC AGT Val Gln Ala Ala Val Pro Gly Leu Pro Ser Arg Val Lys Cys Gly Ser <div style="display: flex; justify-content: space-between; width: 80%; margin: 0 auto;"> <span>85</span> <span>90</span> <span>95</span> </div>	288
55	CCG GGC CAG CTC CAG GGC CTC AGC ATC TTT GCA CAG GAC CTG CGC CTC Pro Gly Gln Leu Gln Gly Leu Ser Ile Phe Ala Gln Asp Leu Arg Leu <div style="display: flex; justify-content: space-between; width: 80%; margin: 0 auto;"> <span>100</span> <span>105</span> <span>110</span> </div>	336
	TGC CTG GAT GAG GCC CTC TCC TGG GAC TGT TTC GCC CTC TCG CTG CTG Cys Leu Asp Glu Ala Leu Ser Trp Asp Cys Phe Ala Leu Ser Leu Leu <div style="display: flex; justify-content: space-between; width: 80%; margin: 0 auto;"> <span>115</span> <span>120</span> <span>125</span> </div>	384
60	GCT GTG GCT CTG GGC CTG GGT GTG CCC ATG CTG CAT CAC CTC TGT GGC Ala Val Ala Leu Gly Leu Gly Val Pro Met Leu His His Leu Cys Gly <div style="display: flex; justify-content: space-between; width: 80%; margin: 0 auto;"> <span>130</span> <span>135</span> <span>140</span> </div>	432
65	TGG GAC CTC TGG TAC TGC TTC CAC CTG TGC CTG GCC TGG CTT CCC TGG 	480

## ES 2 340 210 T3

	Trp 145	Asp	Leu	Trp	Tyr	Cys 150	Phe	His	Leu	Cys	Leu 155	Ala	Trp <sup>+</sup>	Leu	Pro	Trp 160	
5	CGG Arg	GGG Gly	CGG Arg	CAA Gln	AGT Ser 165	GGG Gly	CGA Arg	GAT Asp	GAG Glu	GAT Asp 170	GCC Ala	CTG Leu	CCC Pro	TAC Tyr	GAT Asp 175	GCC Ala	528
10	TTC Phe	GTG Val	GTC Val	TTC Phe 180	GAC Asp	AAA Lys	ACG Thr	CAG Gln	AGC Ser 185	GCA Ala	GTG Val	GCA Ala	GAC Asp	TGG Trp 190	GTG Val	TAC Tyr	576
15	AAC Asn	GAG Glu	CTT Leu 195	CGG Arg	GGG Gly	CAG Gln	CTG Leu	GAG Glu 200	GAG Glu	TGC Cys	CGT Arg	GGG Gly 205	CGC Arg	TGG Trp	GCA Ala	CTC Leu	624
20	CGC Arg	CTG Leu 210	TGC Cys	CTG Leu	GAG Glu	GAA Glu 215	CGC Arg	GAC Asp	TGG Trp	CTG Leu	CCT Pro	GGC Gly 220	AAA Lys	ACC Thr	CTC Leu	TTT Phe	672
25	GAG Glu 225	AAC Asn	CTG Leu	TGG Trp	GCC Ala 230	TCG Ser 235	GTC Val	TAT Tyr	GGC Gly	AGC Ser	CGC Arg 235	AAG Lys	ACG Thr	CTG Leu	TTT Phe 240	GTG Val	720
30	CTG Leu	GCC Ala	CAC His	ACG Thr	GAC Asp 245	CGG Arg	GTC Val	AGT Ser	GGT Gly	CTC Leu 250	TTG Leu	CGC Arg	GCC Ala	AGC Ser	TTC Phe 255	CTG Leu	768
35	CTG Leu	GCC Ala	CAG Gln	CAG Gln 260	CGC Arg	CTG Leu	CTG Leu	GAG Glu	GAC Asp 265	CGC Arg	AAG Lys	GAC Asp	GTC Val 270	GTG Val 270	GTG Val	CTG Leu	816
40	GTG Val	ATC Ile	CTG Leu 275	AGC Ser	CCT Pro	GAC Asp	GGC Gly	CGC Arg 280	CGC Arg	TCC Ser	CGC Arg	TAC Tyr	GAG Glu 285	CGG Arg	CTG Leu	CGC Arg	864
45	CAG Gln 290	CGC Arg	CTC Leu	TGC Cys	CGC Arg	CAG Gln 295	AGT Ser	GTC Val	CTC Leu	CTC Leu	TGG Trp	CCC Pro 300	CAC His	CAG Gln	CCC Pro	AGT Ser	912
50	GGT Gly 305	CAG Gln	CGC Arg	AGC Ser	TTC Phe	TGG Trp 310	GCC Ala	CAG Gln	CTG Leu	GGC Gly	ATG Met 315	GCC Ala	CTG Leu	ACC Thr	AGG Arg	GAC Asp 320	960
55	AAC Asn	CAC His	CAC His	TTC Phe	TAT Tyr 325	AAC Asn	CGG Arg	AAC Asn	TTC Phe	TGC Cys 330	CAG Gln	GGA Gly	CCC Pro	ACG Thr	GCC Ala 335	GAA Glu	1008
60	TAGCCGTGAG CCGGAATCCT GCACGGTGCC ACCTCCACAC TCACCTCACC TCTGCCTGCC																1068
65	TGGTCTGACC CTCCCCTGCT CGCCTCCCTC ACCCCACACC TGACACAGAG CAGGCACTCA																1128
70	ATAAATGCTA CCGAAGGCTA AAAAAAAAAA AAAAAAAAAA AACCA																1173

### (2) INFORMACIÓN PARA EL SEQ ID NO:34:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 336 aminoácidos

(B) TIPO: aminoácido

(D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: proteína

#### (xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:34:



## ES 2 340 210 T3

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

### (2) INFORMACIÓN PARA EL SEQ ID NO:35:

#### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 4 97 pares de bases
- (B) TIPO: ácido nucleico
- (C) CUALIDAD DE LA HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

#### (ii) TIPO DE MOLÉCULA: ADNC

## ES 2 340 210 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA:SEQ ID NO:35:

	TGGCCACAC	GGACCGCGTC	AGTGGCCTCC	TGCGCACCAG	CTTCCTGCTG	GCTCAGCAGC	60
5	GCCTGTTGGA	AGACCGCAAG	GACGTGGTGG	TGTTGGTGAT	CCTGCGTCCG	GATGCCCCAC	120
	CGTCCCGCTA	TGTGCGACTG	CGCCAGCGTC	TCTGCCGCCA	GAGTGTGCTC	TTCTGGCCCC	180
	AGCGACCCAA	CGGGCAGGGG	GGCTTCTGGG	CCCAGCTGAG	TACAGCCCTG	ACTAGGGACA	240
10	ACCGCCACTT	CTATAACCAG	AACTTCTGCC	GGGGACCTAC	AGCAGAATAG	CTCAGAGCAA	300
	CAGCTGGAAA	CAGCTGCATC	TTCATGTCTG	GTTCCCGAGT	TGCTCTGCCT	GCCTTGCTCT	360
	GTCTTACTAC	ACCGCTATTT	GGCAAGTGCG	CAATATATGC	TACCAAGCCA	CCAGGCCAC	420
15	GGAGCAAAGG	TTGGCTGTAA	AGGGTAGTTT	TCTTCCCATG	CATCTTTCAG	GAGAGTGAAG	480
	ATAGACACCA	AACCCAC					497
20							
25							
30							
35							
40							
45							
50							
55							
60							
65							