



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
  
ESPAÑA

(11) Número de publicación: **2 289 610**

(51) Int. Cl.:

**C12N 15/12** (2006.01)

**C07K 14/72** (2006.01)

(12)

## TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Número de solicitud europea: **05003040 .2**

(86) Fecha de presentación : **13.10.1999**

(87) Número de publicación de la solicitud: **1584683**

(87) Fecha de publicación de la solicitud: **12.10.2005**

(54) Título: **Receptor acoplado a proteína G humano huérfano RUP3.**

(30) Prioridad: 20.11.1998 US 109213 P  
16.02.1999 US 120416 P  
26.02.1999 US 121852 P  
12.03.1999 US 123946 P  
12.03.1999 US 123949 P  
28.05.1999 US 136436 P  
28.05.1999 US 136437 P  
28.05.1999 US 136439 P  
28.05.1999 US 136567 P  
28.05.1999 US 137127 P  
28.05.1999 US 137131 P  
29.06.1999 US 141448 P  
29.09.1999 US 156653 P  
29.09.1999 US 156633 P  
29.09.1999 US 156555 P  
29.09.1999 US 156634 P  
01.10.1999 US 157280 P  
01.10.1999 US 157294 P  
01.10.1999 US 157281 P  
01.10.1999 US 157293 P  
01.10.1999 US 157282 P  
12.10.1999 US 417044  
12.10.1999 US 416760

(45) Fecha de publicación de la mención BOPI:  
**01.02.2008**

(73) Titular/es: **Arena Pharmaceuticals, Inc.**  
6166 Nancy Ridge Drive  
San Diego, California 92121, US

(72) Inventor/es: **Chen, Ruoping y**  
**Leonard, James N.**

(45) Fecha de la publicación del folleto de la patente:  
**01.02.2008**

(74) Agente: **Zea Checa, Bernabé**

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

Receptor acoplado a proteína G humano huérfano RUP3.

### 5 Campo de la invención

La invención descrita en este documento de patente se relaciona con los receptores transmembrana, y más particularmente con los receptores acoplados a proteína G humanos, huérfanos, endógenos, ("GPCRs", G protein-coupled receptors). La presente invención se relaciona sólo con Rup3; se tratan otros GPCRs solamente de referencia.

### 10 Estado de la técnica anterior

Aunque existen varias clases de receptores en los humanos, con diferencia la más abundante y terapéuticamente relevante es la representada por los receptores acoplados a proteína G (GPCR o GPCRs). Se estima que hay unos 15 100.000 genes en el genoma humano, y de estos, aproximadamente el 2%, o 2.000 genes, se estima que codifican para GPCRs. Los receptores, incluyendo los GPCRs, para los cuales ha sido identificado su ligando endógeno son referidos como receptores "conocidos", mientras que los receptores para los cuales no se ha identificado el ligando endógeno son referidos como receptores "huérfanos". Los GPCRs representan una importante área para el desarrollo 20 de productos farmacéuticos: a partir de aproximadamente 20 de los 100 GPCRs conocidos se han desarrollado el 60% de todos los fármacos de prescripción. Esta distinción no es sólo semántica, particularmente en el caso de los GPCRs. Así, los GPCRs huérfanos son a la industria farmacéutica lo que fue el oro para California a finales del siglo XIX—una oportunidad de crecimiento, expansión, mejora y desarrollo.

Los GPCRs comparten un motivo estructural común. Todos estos receptores tienen siete secuencias de entre 22 y 25 24 aminoácidos hidrofóbicos que forman siete hélices alfa, cada una de las cuales atraviesa la membrana (cada región transmembrana se identifica con un número, es decir, transmembrana-1 (TM-1), transmembrana-2 (TM-2), etc.). Las hélices transmembrana están unidas mediante cadenas de aminoácidos entre la transmembrana-2 y la transmembrana-3, la transmembrana-4 y la transmembrana-5, y la transmembrana-6 y la transmembrana-7 en el exterior, o lado "extracelular", de la membrana celular (llamadas regiones "extracelulares" 1, 2 y 3 (EC-1, EC-2 y EC-3), respectivamente). Las hélices transmembrana también están unidas por cadenas de aminoácidos entre la transmembrana-1 y la transmembrana-2, la transmembrana-3 y la transmembrana-4, y la transmembrana-5 y la transmembrana-6 en el interior, o lado "intracelular", de la membrana celular (llamadas regiones "intracelulares" 1, 2 y 3 (IC-1, IC-2 y IC-3) respectivamente). El "carboxi" ("C") terminal del receptor se localiza en el espacio intracelular en la célula, y el "amino" ("N") terminal del receptor se localiza en el espacio extracelular fuera de la célula.

35 Generalmente, cuando un ligando endógeno se une con el receptor (a menudo referido como "activación" del receptor), hay un cambio en la conformación de la región intracelular que permite el acoplamiento entre la región intracelular y una "proteína-G" intracelular. Se ha descrito que los GPCRs son "promiscuos" respecto a las proteínas G, es decir, que un GPCR puede interactuar con más de una proteína G. Ver, Kenakin, T., 43 *Life Sciences* 1095 40 (1988). Aunque existen otras proteínas G, actualmente, Gq, Gs, Gi y Go son proteínas G que han sido identificadas. El acoplamiento de GPCR activado por ligando endógeno con la proteína G, inicia un proceso en cascada de señalización (llamado "transducción de señal"). Bajo condiciones normales, la transducción de señal resulta en último término en la activación celular o la inhibición celular. Se piensa que el bucle IC-3, así como el carboxi terminal del receptor interactúan con la proteína G.

45 Bajo condiciones fisiológicas, los GPCRs existen en la membrana celular en equilibrio entre 2 conformaciones diferentes: un estado "inactivo" y un estado "activo". Un receptor en un estado inactivo es incapaz de enlazar el circuito de transducción de señal intracelular para producir una respuesta biológica. Cambiando la conformación del receptor al estado activo se permite el enlace al circuito de transducción (vía la proteína G) y se produce una respuesta biológica. Un receptor puede estabilizarse en un estado activo mediante un ligando endógeno o un compuesto tal como un fármaco.

### Resumen de la invención

55 La presente invención proporciona un método de rastreo de compuestos candidatos para identificar un agente farmacéutico para una enfermedad o estado de desorden relacionado con el páncreas, el método comprendiendo:

proporcionar una célula hospedadora eucariota que comprende un receptor acoplado a proteína G que es una versión activa independiente de ligando de un receptor que tiene la SEQ ID NO: 8, donde el receptor se acopla a una proteína G; y

60 rastrear compuestos candidatos contra dicho receptor acoplado a proteína G.

### Breve descripción de los dibujos

65 Las Figuras 1A y 1B proporcionan "tablas" de referencia para ciertos dot-blots proporcionados aquí (ver también, Figuras 2A y 2B, respectivamente).

# ES 2 289 610 T3

Las Figuras 2A y 2B proporcionan reproducciones de los resultados de ciertos análisis de dot-blot resultantes de hCHN3 y hCHN8, respectivamente (*ver también*, Figuras 1A y 1B, respectivamente).

La Figura 3 proporciona una reproducción de los resultados del análisis de hRUP3 por RT-PCR.

5

La Figura 4 proporciona una reproducción de los resultados del análisis de hRUP4 por RT-PCR.

La Figura 5 proporciona una reproducción de los resultados del análisis de hRUP6 por RT-PCR.

## 10 Descripción detallada

La literatura científica que se ha desarrollado en torno a receptores ha adoptado un número de términos para referirse a ligandos con diferentes efectos sobre los receptores. Para la claridad y consistencia, se usarán las siguientes definiciones a lo largo de este documento de patente. En el caso de que estas definiciones entren en conflicto con otras 15 definiciones para estos términos, las siguientes definiciones prevalecerán:

Las *ABREVIATURAS DE AMINOÁCIDOS* usadas aquí se definen en la Tabla 1:

20 **TABLA 1**

ALANINA	ALA	A
ARGININA	ARG	R
ASPARAGINA	ASN	N
ÁCIDO ASPÁRTICO	ASP	D
CISTEÍNA	CYS	C
ÁCIDO GLUTÁMICO	GLU	E
GLUTAMINA	GLN	Q
GLICINA	GLY	G
HISTIDINA	HIS	H
ISOLEUCINA	ILE	I
LEUCINA	LEU	L
LISINA	LYS	K
METIONINA	MET	M
FENILALANINA	PHE	F
PROLINA	PRO	P
SERINA	SER	S
TREONINA	THR	T
TRIPTÓFANO	TRP	W
TIROSINA	TYR	Y
VALINA	VAL	V

60 *COMPOSICIÓN* significa un material que comprende al menos un componente.

ENDÓGENO significará un material que un mamífero produce naturalmente. Por ejemplo, y sin ser una limitación, ENDÓGENO en relación al término “receptor”, significará aquél que es producido naturalmente por un mamífero (por ejemplo, y sin ser una limitación, un humano) o un virus. En contraste, el término *NO ENDÓGENO* en este contexto 65 significará aquél que no es producido naturalmente por un mamífero (por ejemplo, y sin ser una limitación, un humano) o un virus.

5      *CÉLULA HOSPEDADORA* significará una célula capaz de tener incorporado un Plásmido y/o un Vector en su interior. En el caso de una Célula Hospedadora procariota, un Plásmido se replica típicamente como una molécula autónoma mientras la Célula Hospedadora se replica (generalmente, el Plásmido se aísla entonces para su introducción en una Célula Hospedadora eucariota); en el caso de una Célula Hospedadora eucariota, un Plásmido se integra en el ADN celular de la Célula Hospedadora de manera que cuando la Célula Hospedadora se replica, se replica el Plásmido. Para los propósitos de la invención descrita aquí, la Célula Hospedadora es eucariota, más preferiblemente, de mamífero, y más preferiblemente seleccionada del grupo que consiste en las células 293, 293T y COS-7.

10     *LIGANDO* significará una molécula de origen natural, endógena, específica para un receptor de origen natural y endógeno.

15     *RECEPTOR NO HUÉRFANO* significará una molécula de origen natural, endógena, específica para un ligando de origen natural y endógeno, en el que la unión de un ligando a un receptor activa un circuito de señalización intracelular.

20     *RECEPTOR HUÉRFANO* significará un receptor endógeno para el cual el ligando endógeno específico para ese receptor no ha sido identificado o no es conocido.

25     *PLÁSMIDO* significará la combinación de un Vector y ADNc. Generalmente, un Plásmido es introducido en una Célula Hospedadora con el propósito de la replicación y/o la expresión del ADNc como una proteína.

30     *VECTOR*, en referencia a ADNc, significará un ADN circular capaz de incorporar al menos un ADNc y capaz de incorporarse en una Célula Hospedadora.

El orden de las siguientes secciones está pensado para una mejor eficiencia en la descripción y no se pretende, ni debe interpretarse, como una limitación en la descripción o en las reivindicaciones que le siguen.

35     A. *Identificación de GPCRs humanos*

Los esfuerzos del proyecto Genoma Humano han llevado a la identificación de una pléthora de información respecto a secuencias de ácido nucleico localizadas en el genoma humano; en este esfuerzo se ha dado el caso de que la información de secuencia genética se ha hecho disponible sin una comprensión o reconocimiento de si alguna secuencia genómica en particular contiene o podría contener información de marco abierto de lectura que traduzca proteínas humanas o no. Varios métodos de identificación de secuencias de ácido nucleico dentro del genoma humano están dentro del ámbito de conocimiento de aquéllos expertos en la materia. Por ejemplo, y sin suponer una limitación, una variedad de GPRCs, descritos aquí, fueron descubiertos revisando la base de datos GenBank™, mientras que otros GPRCs fueron descubiertos usando una secuencia de ácido nucleico de un GPRC, previamente secuenciado, para llevar a cabo una búsqueda BLAST™ en la base de datos EST. La Tabla A, a continuación, lista los GPRCs huérfanos endógenos, descritos con un GPCR homólogo respectivo de GPCR:

45     (Tabla pasa a página siguiente)

50

55

60

65

TABLA A

5 GPRCs Humanos Huérfanos Descritos	10 Número de Acceso Identificado	15 Marco Abierto de Lectura (Pares de Bases)	20 Porcentaje de Homología al GPCR Designado	25 Referencia al GPCR Homólogo (Nº de Acceso)
<b>hARE-3</b>	AL033379	1.260 pb	52.3% LPA-R	U92642
<b>hARE-4</b>	AC006087	1.119 pb	36% P2Y5	AF000546
<b>hARE-5</b>	AC006255	1.104 pb	32% <i>Oryzias latipes</i>	D43633
<b>hGPR27</b>	AA775870	1.128 pb		
<b>hARE-1</b>	A1090920	999 pb	43% KIAA0001	D13626
<b>hARE-2</b>	AA359504	1.122 pb	53% GPR27	
<b>hPPR1</b>	H67224	1.053 pb	39% EBI1	L31581
<b>hG2A</b>	AA754702	1.113 pb	31% GPR4	L36148
<b>hRUP3</b>	AL035423	1.005 pb	30% <i>Drosophila melanogaster</i>	2133653
<b>hRUP4</b>	AI307658	1.296 pb	32% pNPGPR 28% y 29 % Pez Zebra Ya y Yb, respectivamente	NP_004876 AAC41276 y AAB94616
<b>hRUP5</b>	AC005849	1.413 bp	25% DEZ 23% FMLPR	Q99788 P21462
<b>hRUP6</b>	AC005871	1.245 bp	48% GPR66	NP_006047
<b>hRUP7</b>	AC007922	1.173 pb	43% H3R	AF140538
<b>hCHN3</b>	EST 36581	1.113 pb	53% GPR27	
<b>hCHN4</b>	AA804531	1.077 pb	32% trombina	4503637
<b>hCHN6</b>	EST 2134670	1.503 pb	36% edg-1	NP_001391
<b>hCHN8</b>	EST 764455	1.029 pb	47% KIAA0001	D13626
<b>hCHN9</b>	EST 1541536	1.077 pb	41% LTB4R	NM_000752
<b>hCHN10</b>	EST 1365839	1.055 pb	35% P2Y	NM_002563

La homología del receptor es útil en términos de conseguir una apreciación del papel de los receptores descritos en el cuerpo humano. Adicionalmente, esta homología puede proporcionar ideas de posibles ligando(s) endógenos que puedan ser activadores naturales para los GPRCs huérfanos descritos.

#### B. Rastreo de Receptores

Las técnicas se han vuelto más fácilmente accesibles durante los últimos años para la identificación de ligandos endógenos (esto, en principio, con el propósito de proporcionar una manera de llevar a cabo ensayos de unión de receptor que requieren un ligando endógeno del receptor) porque el estudio tradicional de receptores siempre se ha basado en el supuesto *a priori* (basado históricamente) de que el ligando endógeno debe ser identificado en primer lugar antes de que el descubrimiento pueda continuar buscando antagonistas y otras moléculas que puedan afectar al

receptor. Incluso en casos donde un antagonista pudiera ser conocido previamente, la búsqueda se extendería inmediatamente a buscar el ligando endógeno. Este modo de pensar ha perdurado en la investigación de receptores incluso después del descubrimiento de receptores activados constitutivamente. Lo que no había sido reconocido hasta ahora es que es el estado activado del receptor el más útil para el descubrimiento de agonistas, agonistas parciales y agonistas inversos del receptor. Para aquellas enfermedades que resultan de un receptor demasiado activado o de un receptor infraactivado, lo deseado en un fármaco terapéutico es un compuesto que actúe reduciendo el estado activo del receptor o aumentando la actividad del receptor, respectivamente, no necesariamente un fármaco que sea un antagonista del ligando endógeno. Esto es debido a que un compuesto que reduce o aumenta la actividad del estado activo del receptor no necesita unirse en el mismo sitio que el ligando endógeno. Así, como enseña un método de esta invención, cualquier búsqueda de compuestos terapéuticos debería comenzar rastreando compuestos contra el estado activo independiente de ligando.

Como es conocido en el estado de la técnica, los GPCRs pueden ser “activos” en su estado endógeno incluso sin la unión del ligando endógeno del receptor. Tales receptores naturalmente activos pueden ser probados para identificación directa (es decir, sin la necesidad del ligando endógeno del receptor) de, en particular, agonistas inversos. Alternativamente, el receptor puede ser “activado” vía, por ejemplo, mutación del receptor para establecer una versión no endógena del receptor que será activa en ausencia del ligando endógeno del receptor.

Rastreando compuestos candidatos contra una versión endógena o no endógena activada constitutivamente de los GPCRs humanos huérfanos descritos aquí, se pueden proporcionar para la identificación directa de compuestos candidatos que actúan en este receptor de la superficie celular, sin necesitar el uso del ligando endógeno del receptor. Mediante la determinación de las áreas del cuerpo donde se expresa y/o se sobreexpresa la versión endógena de los GPCRs humanos descritos aquí, es posible determinar estados de enfermedad/desorden relacionados que estén asociados con la expresión y/o sobreexpresión del receptor; en este documento de patente se describe una aproximación a ello.

En relación a la creación de una mutación que pueda probar la activación constitutiva de los GPCRs humanos huérfanos descritos aquí, se basa en la distancia del residuo de prolina donde se presume que está localizada dentro del TM6 del GPCR típicamente cercano a la interfaz TM6/IC3 (tal residuo de prolina parece estar bastante conservado). Mutando el residuo de aminoácido localizado a 16 residuos de este residuo (presumiblemente localizado en la región IC3 del receptor) a, preferiblemente, un residuo de lisina, se puede obtener tal activación. Otros residuos de aminoácido pueden ser útiles en la mutación de esta posición para conseguir este objetivo.

#### C. Identificación y/o Selección de la Enfermedad/Desorden

Preferiblemente, la secuencia de ADN del GPCR humano huérfano se puede usar para hacer una sonda para (a) análisis dot-blot contra ARNm tisular, y/o (b) identificación por RT-PCR de la expresión del receptor en muestras de tejido. La presencia de un receptor en un tejido, o en un tejido enfermo, o la presencia del receptor en concentraciones elevadas en tejido enfermo comparado con un tejido normal, puede ser preferiblemente utilizada para identificar una correlación con un régimen de tratamiento, incluyendo pero no limitado a, una enfermedad asociada a esa enfermedad. Mediante esta técnica los receptores pueden igualmente ser localizados en regiones de órganos. Basándose en las funciones conocidas de los tejidos específicos donde el receptor está localizado, se puede deducir el rol funcional putativo del receptor.

#### D. Rastreo de Compuestos Candidatos

##### 1. Técnicas de ensayo genéricas para el rastreo de GPCR

Cuando un receptor de proteína G se vuelve constitutivamente activo (es decir, activo en ausencia de unión con ligando endógeno), se une a una proteína G (por ejemplo, Gq, Gs, Gi, Go) y estimula la unión de GTP a la proteína G. Entonces la proteína G actúa como una GTPasa e hidroliza lentamente el GTP a GDP, por lo cual el receptor, bajo condiciones normales, queda desactivado. Sin embargo, los receptores activados constitutivamente continúan cambiando GDP a GTP. Un análogo no hidrolizable de GTP, [<sup>35</sup>S]GTPγS, puede usarse para monitorizar uniones mejoradas a membranas que expresan receptores activados constitutivamente. Se ha descrito que [<sup>35</sup>S]GTPγS puede ser usado para monitorizar acoplamiento de proteína G a membranas en ausencia y presencia de ligando. Un ejemplo de esta monitorización, entre otros ejemplos bien conocidos y disponibles para los expertos en la materia, fue descrito por Traynor y Nahorski en 1995. El uso preferido de este sistema de ensayo es para el rastreo inicial de compuestos candidatos ya que el sistema es genéricamente aplicable a todos los receptores acoplados a proteína G, cual sea la proteína G en particular que interacciona con el dominio intracelular del receptor.

##### 2. Técnicas de ensayo específicas para el rastreo de GPCR

Una vez los compuestos candidatos están identificados usando el ensayo “genérico” de receptores acoplados a proteína G (es decir, un ensayo para seleccionar compuestos que son agonistas, agonistas parciales, o agonistas inversos), es preferible rastrear más para confirmar que los compuestos han interactuado en el sitio del receptor. Por ejemplo, un compuesto identificado mediante el ensayo “genérico” podría no unirse con el receptor, pero en cambio podría simplemente “desacoplar” la proteína G del dominio intracelular.

## ES 2 289 610 T3

### a. *Gs y Gi*

Gs estimula la enzima adenilil ciclase. Gi (y Go), al contrario, inhiben esta enzima. La adenilil ciclase cataliza la conversión de ATP en AMPc; así, los GPCRs activados constitutivamente que acoplan la proteína Gs están asociados a niveles celulares incrementados de AMPc. Por otra parte, GPCRs activados constitutivamente que acoplan la proteína Gi (o Go) están asociados a niveles celulares reducidos de AMPc. Ver, en general, “Indirect Mechanisms of Synaptic Transmission,” Chpt. 8, *From Neuron To Brain* (3<sup>a</sup> Ed.) Nichols, J.G. et al eds. Sinauer Associates, Inc. (1992). De esta manera, se pueden utilizar ensayos que detectan AMPc para determinar si un compuesto candidato es, por ejemplo, un agonista inverso al receptor (es decir, un compuesto así disminuiría los niveles de AMPc). Se pueden utilizar varias aproximaciones para medir AMPc conocidas en el estado de la técnica; una aproximación más preferida radica en el uso de anticuerpos antiAMPc en un formato basado en ELISA. Otro tipo de ensayo que puede ser utilizado es un ensayo de sistema de detección de segundo mensajero de la célula entera. Los promotores en los genes dirigen la expresión de las proteínas que un gen en particular codifica. El AMP cíclico dirige la expresión génica promoviendo la unión de una proteína de unión a ADN sensible a AMPc o un factor de transcripción (CREB) que se une al promotor en sitios específicos llamados elementos de respuesta a AMPc, y lleva a la expresión del gen. Se pueden construir sistemas de detección que tengan un promotor que contenga múltiples elementos de respuesta a AMPc antes del gen de detección, p. ej.,  $\beta$ -galactosidasa o luciferasa. Así, un receptor activado constitutivamente unido a Gs causa la acumulación de AMPc, el cual activa el gen y la expresión de la proteína de detección. De esta manera, la proteína de detección como la  $\beta$ -galactosidasa o la luciferasa pueden ser detectadas usando ensayos bioquímicos estándares (Chen et al. 1995).

### b. *Go y Gq*

Gq y Go están asociadas a la activación de la enzima fosfolipasa C, que, a su vez, hidroliza al fosfolípido PIP<sub>2</sub>, liberando dos mensajeros intracelulares: diacioglicerol (DAG) e inistol 1,4,5-trifosfato (IP<sub>3</sub>). La acumulación creciente de IP<sub>3</sub> está asociada a la activación de los receptores asociados a Gq y Go. Ver, en general, “Indirect Mechanisms of Synaptic Transmission,” Chpt. 8, *From Neuron To Brain* (3<sup>a</sup> Ed.) Nichols, J.G. et al eds. Sinauer Associates, inc. (1992). Los ensayos que detectan la acumulación de IP<sub>3</sub> pueden ser utilizados para determinar si un compuesto candidato es, p. ej., un agonista inverso de un receptor asociado a Gq o Go (es decir, un compuesto así disminuiría los niveles de IP<sub>3</sub>). Los receptores asociados a Gq también pueden ser examinados utilizando un ensayo de detección AP1 en el cual la fosfolipasa C dependiente de Gq provoca la activación de genes que contienen elementos AP1; así, los receptores asociados a Gq activados provocarán un aumento en la expresión de dichos genes, por lo que los agonistas inversos provocarán un descenso de tal expresión, y los agonistas provocarán un aumento de tal expresión. Se encuentran disponibles comercialmente ensayos para tal detección.

### 35 3. Proteína de Fusión GPCR

El uso de un GPCR endógeno huérfano activado constitutivamente, o de un GPCR no endógeno huérfano activado constitutivamente, para el rastreo de compuestos candidatos para la identificación directa de agonistas inversos, agonistas y agonistas parciales tiene la dificultad única de que, por definición, el receptor sea activo incluso en ausencia de unión con un ligando endógeno. Así, a menudo es útil utilizar una aproximación que mejore la señal obtenida por el receptor activado. Una aproximación preferida es el uso de una Proteína de Fusión GPCR.

En general, una vez se determina que un GPCR está o ha estado constitutivamente activado, utilizando las técnicas de ensayo descritas anteriormente (así como otras), es posible determinar la proteína G predominante que se acopla con el GPCR endógeno. El acoplamiento de la proteína G al GPCR proporciona un circuito de señalización que puede ser analizado. Debido que es más preferido que el rastreo se lleve a cabo usando un sistema de expresión mamífero, se espera que tal sistema tenga proteína G endógena. De esta manera, por definición, en un sistema así, el GPCR huérfano activado constitutivamente dará señal continuamente. En relación a esto, es preferido que esta señal sea mejorada, de manera que en presencia de, p. ej., un agonista inverso al receptor, es más probable que permita diferenciar más fácilmente, particularmente en el contexto de rastreo, entre el receptor cuando está en contacto con el agonista inverso.

Se pretende que la Proteína de Fusión GPCR mejore la eficacia del acoplamiento de la proteína G con el GPCR. La Proteína de Fusión GPCR se prefiere para rastreos con un GPCR no endógeno activado constitutivamente, ya que esa aproximación incrementa la señal que se utiliza más preferentemente en este tipo de técnicas de rastreo, aunque la Proteína de Fusión GPCR también puede ser (y preferiblemente lo es) utilizada con un GPCR endógeno activado constitutivamente. Esto es importante para dar una relación “señal-ruido” significativa.

Tal relación significativa es preferible para el rastreo de los compuestos candidatos como se describe aquí.

La realización de una construcción útil para la expresión de una Proteína de Fusión GPCR está dentro del ámbito del experto en la materia. Los vectores de expresión y sistemas disponibles comercialmente ofrecen una variedad de aproximaciones que se pueden ajustar a las necesidades particulares de un investigador. El criterio de importancia para tal construcción de una Proteínas de Fusión GPCR es que la secuencia del GPCR y la secuencia de la proteína G sean ambas de marco interno (preferentemente, la secuencia para el GPCR está más arriba (upstream) de la secuencia de la proteína G) y que el codón de “terminación” del GPCR debe ser eliminado o sustituido de tal manera que tras la expresión del GPCR, la proteína G también pueda ser expresada. El GPCR puede estar enlazado directamente a la proteína G, o pueden haber residuos espaciadores entre los dos (preferiblemente, no más de unos 12, aunque este

# ES 2 289 610 T3

número puede ser fácilmente determinado por el experto en la materia). Tenemos una preferencia (basada en la conveniencia) del uso de un espaciador en algunos de aquellos sitios de restricción que no son usados efectivamente y que bajo expresión se convierten en espaciadores. Más preferiblemente, la proteína G que se acopla al GPCR habrá sido identificada previamente a la creación de la construcción de la Proteína de Fusión GPCR. Ya que sólo unas pocas 5 proteínas G han sido identificadas, es preferible que una construcción que comprende la secuencia de la proteína G (es decir, una construcción de proteína G universal) sea válida para la inserción de una secuencia de un GPCR endógeno; esto proporciona eficiencia en el contexto de rastreos a gran escala de varios GPCRs diferentes con diferentes secuencias.

## 10 E. Otra Utilidad

Un uso preferido de los GPCRs humanos huérfanos descritos aquí puede ser para la identificación directa de compuestos candidatos como agonistas inversos, agonistas, o agonistas parciales (preferiblemente para uso como agentes farmacéuticos). Estas versiones de GPCRs humanos también pueden ser utilizadas en investigación. Por ejemplo, sistemas 15 *in vitro* e *in vivo* que incorporen GPCRs pueden utilizarse tanto para aclarar más y comprender los papeles que estos receptores juegan en la condición humana, sana o patológica, como para entender el papel de la activación constitutiva como se aplica en la comprensión de la cascada de señalización. El valor de los GPCRs humanos huérfanos es que su utilidad como herramienta de investigación se puede mejorar determinando la localización(es) de esos receptores 20 en el cuerpo, los GPCRs se pueden usar para entender el papel de estos receptores en el cuerpo humano antes de que se identifique el ligando endógeno. Otros usos de los receptores descritos serán evidentes para los expertos en la materia basándose en, *inter alia*, el análisis de este documento de patente.

## Ejemplos

25 Los siguientes ejemplos se presentan con el propósito de aclarar, y no limitar, la presente invención y para proporcionar información de referencia. Aunque se describen secuencias específicas de ácido nucleico y aminoácido, se considera al experto en la materia con la capacidad de hacer pequeñas modificaciones a estas secuencias mientras consiga los mismos resultados u otros sustancialmente similares a los proporcionados a continuación. Excepto cuando 30 se especifica lo contrario, todas las secuencias de ácido nucleico para los GPCRs humanos huérfanos endógenos han sido secuenciadas y verificadas. En relación a receptores equivalentes, el experto en la materia apreciará fácilmente que pueden llevarse a cabo sustituciones conservativas a las secuencias descritas con el objetivo de obtener receptores funcionalmente equivalentes.

### Ejemplo 1

#### 35 GPCRs humanos endógenos

##### 1. Identificación de GPCRs Humanos

40 Varios de los GPCRs humanos endógenos descritos fueron identificados en base a la revisión de la información de la base de datos GenBank. Buscando en la base de datos, fueron identificados los siguientes clones de ADNc según se describe a continuación:

45	<b>GPCRs Humanos Huérfanos Descritos</b>	<b>Número de Acceso</b>	<b>Secuencia Completa de ADN (Pares de Bases)</b>	<b>Marco Abierto de Lectura (Pares de Bases)</b>	<b>Ácido Nucleico SEQ.ID. NO.</b>	<b>Aminoácido SEQ.ID.NO.</b>
50	<b>hARE-3</b>	AL033379	111.389 pb	1.260 pb	1	2
55	<b>hARE-4</b>	AC006087	226.925 pb	1.119 pb	3	4
60	<b>hARE-5</b>	AC006255	127.605 pb	1.104 pb	5	6
65	<b>hRUP3</b>	AL035423	140.094 pb	1.005 pb	7	8
	<b>hRUP5</b>	AC005849	169.144 pb	1.413 pb	9	10
	<b>hRUP6</b>	AC005871	218.807 pb	1.245 pb	11	12
	<b>hRUP7</b>	AC007922	158.858 pb	1.173 pb	13	14

# ES 2 289 610 T3

Otros GPCRs humanos endógenos fueron identificados mediante una búsqueda BLAST de la base de datos EST (dbest), utilizando los siguientes clones EST como secuencias de búsqueda. Fueron identificados los siguientes clones EST y usados a continuación como sonda para rastrear una librería genómica humana.

	<b>GPRCs Humanos Huérfanos Descritos</b>	<b>Búsqueda (Secuencia)</b>	<b>Clon EST/ No. Acceso Identificado</b>	<b>Marco Abierto de Lectura (Pares de Bases)</b>	<b>Ácido Nucleico SEQ.ID.NO.</b>	<b>Aminoácido SEQ.ID.NO.</b>
5	<b>hGPR27</b>	Ratón GPCR27	AA775870	1.125 pb	15	16
10	<b>hARE-1</b>	TDAG	1689643 AI090920	999 pb	17	18
15	<b>hARE-2</b>	GPCR27	68530 AA359504	1.122 pb	19	20
20	<b>hPPR1</b>	Bovino PPRI	238667 H67224	1.053 pb	21	22
25	<b>hG2A</b>	Ratón 1179426	ver ejemplo 2(a), a continuación	1.113 pb	23	24
30	<b>hCHN3</b>	N.A.	EST 36581 (longitud total)	1.113 pb	25	26
35	<b>hCHN4</b>	TDAG	1184934 AA804531	1.077 pb	27	28
40	<b>hCHN6</b>	N.A.	EST 2134670 (longitud total)	1.503 pb	29	30
45	<b>hCHN8</b>	KIAA0001	EST 764455	1.029 pb	31	32
50	<b>hCHN9</b>	1365839	EST 1541536	1.077 pb	33	34
55	<b>hCHN10</b>	Ratón EST 1365839	Humano 1365839	1.005 pb	35	36
60	<b>hRUP4</b>	N.A.	AI307658	1.296 pb	37	38
<i>N.A. = "no applicable"</i>						

# ES 2 289 610 T3

## 2. Clonación de Longitud Total

### a. hG2A (Seq. Id. Nos. 23 y 24)

5 Se usó el clon EST de ratón 1179426 para obtener un clon genómico humano que contuviera las tres secuencias codificantes de aminoácidos hG2A. El extremo 5' de esta secuencia codificante se obtuvo mediante 5'RACE<sup>TM</sup>, y el molde para PCR fue ADNc de bazo humano Marathon-ready<sup>TM</sup> de Clontech. El G2A humano descrito se amplificó mediante PCR utilizando, durante el primer y segundo ciclo de PCR, los cebadores específicos para ADNc de G2A como se muestran en las SEQ.ID.NO.:39 y SEQ.ID.NO.:40: 5'-CTGTGTACAGCAGTCGCAGAGTG-3' (SEQ.ID.NO.:39; 1<sup>er</sup> ciclo PCR) 5'-GAGTGCCAGGCAGAGCAGGTAGAC-3' (SEQ.ID.NO.:40; 2<sup>do</sup> ciclo PCR). La PCR se llevó a cabo usando el kit Advantage<sup>TM</sup> GC Polymerase (Clontech; se seguirán las instrucciones de fabricación), a 94°C durante 30 seg. seguido de 5 ciclos a 94°C durante 5 seg. y 72°C durante 4 min.; y 30 ciclos a 94° durante 5 seg. y 70° durante 4 min. Se purificó un fragmento de aproximadamente 1,3 kb mediante gel de agarosa, se digirió con Hind III y Xba I y se clonó en el vector de expresión pRC/CMV2 (Invitrogen). El inserto clonado se secuenció mediante el kit T7 Sequenase<sup>TM</sup> (USB Amersham; se seguirán las instrucciones del fabricante) y la secuencia se comparó con la secuencia presentada. Se detectará la expresión de G2A humano sondando un dot-blot de ARN (Clontech; se seguirán las instrucciones del fabricante) con el fragmento marcado con P<sup>32</sup>.

### b. hCHN9 (Seq. Id. Nos. 33 y 34)

20 La secuenciación del clon EST 1541536 indicó que hCHN9 es un clon parcial de ADNc que sólo tiene un codón de iniciación; es decir, el codón de terminación se perdió. Cuando se utilizó hCHN9 para lanzarlo contra la base de datos (nr), la secuencia 3' de hCHN9 fue 100% homóloga a la región 5' no traducida del ADNc del receptor luecotrieno B4, 25 que contenía un codón de terminación en el marco con secuencia codificante de hCHN9. Para determinar si la región 5' no traducida del ADNc de LTB4R era la secuencia 3' de hCHN9, se realizó un PCR usando cebadores basados en la secuencia 5' que flanquea el codón de iniciación encontrado en hCHN9 y la secuencia 3' alrededor del codón de terminación encontrado en la región 5' no traducida de LTB4R. La secuencia del cebador 5' fue la siguiente:

30 5'-CCCGAATTCCCTGCTTGCTCCCAGCTTGGCCC-3' (SEQ.ID.NO.:41; sentido) y  
5'-TGTGGATCCTGCTGTCAAAGGTCCCATTCCGG-3' (SEQ.ID.NO.:42; antisentido).

35 La PCR se realizó usando ADNc de timo como molde y polimerasa rTth (Perkin Elmer) con el sistema tampón proporcionado por el fabricante, 0,25 μM de cada cebador, y 0,2 mM de cada uno de los 4 nucleótidos. Las condiciones de ciclado fueron 30 ciclos a 94°C durante 1 min., 65°C durante 1 min. y 72°C durante 1 min. y 10 seg. De la PCR se obtuvo un fragmento de 1,1 kb coherente con el tamaño esperado. Este fragmento de PCR se subclonó en pCMV (*ver*, SEQ.ID.NO.:33).

40

### c. hRUP 4 (Seq. Id. Nos. 37 y 38)

45 Se clonó la longitud total de hRUP4 mediante RT-PCR con ADNc de cerebro humano (Clontech) como molde:

50 5'-TCACAATGCTAGGTGTGGTC-3' (SEQ.ID.NO.:43; sentido) y  
5'-TGCATAGACAATGGGATTACAG-3' (SEQ.ID.NO.:44; antisentido).

55 La PCR se realizó usando TaqPlus<sup>TM</sup> Precision<sup>TM</sup> polimerasa (Stratagene; se seguirán las instrucciones de fabricación) con los siguientes ciclos: 94°C durante 2 min.; 94°C 30 seg.; 55°C durante 30 seg., 72°C durante 45 seg., y 72°C durante 10 min. Los ciclos del 2 al 4 se repitieron 30 veces.

60 Los productos de PCR se separaron en un gel de agarosa al 1% y se aisló un fragmento de PCR de 500 pb que fue clonado en el vector pCRII-TOPO (Invitrogen) y secuenciado mediante el kit T7 DNA Sequenase<sup>TM</sup> (Amsham) y los cebadores SP6/T7 (Stratagene). El análisis de la secuencia reveló que el fragmento de PCR era de hecho una forma de AI307658 empalmada alternativamente que tenía un marco abierto de lectura continuo con similitud a otros GPRCs. La secuencia completa de este fragmento de PCR era la siguiente:

5'-TCACAATGCTAGGTGTGGCTGGCTGGCAGTCATCGTAG  
65 GATCACCCATGTGGCACGTGCAACAACTTGAGATCAAATATGAC

ES 2 289 610 T3

TTCCTATATGAAAAGGAACACATCTGCTGCTTAGAAGAGTGGACC  
AGCCCTGTGCACCAGAAGATCTACACCACCTTCATCCTTGTCACTC  
CTCTTCCTCCTGCCTCTTATGGTGATGCTTATTCTGTACGTAAAAA  
TTGGTTATGAACCTTGGATAAAAGAAAAGAGTTGGGATGGTTCA  
GTGCTTCGAACATTGAAACTATTGAAAGAAAATGTCCAAAATAGCCAGG  
AAGAAGAAACGAGCTGTCATTATGATGGTGACAGTGGTGGCTCT  
CTTGCTGTGCTGGCACCATTCATGTTGTCCATATGATGAT  
TGAATACAGTAATTTGAAAAGGAATATGATGATGTCACAATCAA  
GATGATTTTGCTATCGTGCAAATTATTGGATTTCCAACCTCCATC  
TGTAATCCCATTGTCTATGCA-3' (SEQ.ID.NO.:45).

Basándose en la secuencia anterior, se utilizaron como cebadores dos conjuntos de oligonucleótidos sentido, con las siguientes secuencias:

5'-CTGCTTAGAACAGTGGACCAG-3' (SEQ.ID.NO.:46; oligo 1)

5'-CTGTGCACCAAGATCTACAC-3' (SEQ ID NO.47; oligo 2)

y dos conjuntos de oligonucleótidos antisentido, con las secuencias siguientes:

5'-CAAGGATGAAAGTGTTGAGA-3' (SEQ ID NO:48; oligo 3)

5'-GTGTAGATCTTCTGGTGCACAGG-3' (SEQ ID NO:49; oligo 4)

ra la PCR 3'- y 5'- RACE con ADNc de cerebro humano Marathon-Ready™ (Clontech, Cat# 7400-1) como molde, según las instrucciones del fabricante. Los fragmentos de ADN generados por la RACE PCR se clonaron en el vector pRII-TOPO™ (Invitrogen) y se secuenciaron utilizando los cebadores SP6/T7 (Stratagene) y algunos cebadores internos. El producto 3' RACE contenía una cola polí(A) y un marco abierto de lectura completo finalizando en un dón de terminación TAA. El producto 5' RACE contenía un final 5' incompleto; es decir, el codón de iniciación CG no estaba presente.

Basándose en la nueva secuencia 5', el oligo 3 y el siguiente cebador:

5'-GCAATGCAGGTCTAGTGAGC-3' (SEQ.ID.NO.:50; oligo 5)

fueron utilizados para un segundo proceso 5' RACE PCR y los productos de PCR se analizaron como antes. Un tercer proceso de 5' RACE PCR se llevó a cabo usando los cebadores antisentido:

5'-TGGAGCATGGTACGGGAATGCAGAAG-3' (SEQ.ID.NO.:51; oligo 6) y,

5'-GTGATGAGCAGGTCACTGAGCGCCAAG-3' (SEQ.ID.NO.:52; oligo7).

La secuencia de los productos 5' RACE PCR reveló la presencia del codón de iniciación ATG, y un proceso más de 5' RACE PCR no generó más secuencia 5'. La secuencia 5' completada se confirmó por RT-PCR utilizando el cebador sentido 5'-GCAATGCAGGCCTTAACATTAC-3' (SEQ.ID.NO.: 53; oligo 8) y oligo 4 como cebadores y el análisis de secuencia del producto de PCR de 650 pb generado a partir de moldes de ADNc de corazón y cerebro humano (Clontech, Cat# 7404-1). La secuencia 3' completada fue confirmada mediante RT-PCR utilizando oligo 2 y el siguiente cebador antisentido:

5'-TTGGGTTACAACTGAGGGCA-3' (SEQ.ID.NO.: 54; oligo 9)

y el análisis de secuencia del producto de PCR de 670 pb generado a partir de moldes de ADNc de cerebro y corazón humano (Clontech, Cat# 7404-1).

# ES 2 289 610 T3

## d. hRUP5 (*Seq. Id. Nos. 9 y 10*)

La longitud total de hRUP5 se clonó mediante RT-PCR usando un cebador sentido más arriba, upstream, de ATG, el codón de iniciación (SEQ.ID.NO.: 55), y un cebador antisentido que contenía TCA como codón de terminación (SEQ.ID.NO.:56), el cual tenía las siguientes secuencias:

5'-ACTCCGTGTCCAGCAGGACTCTG-3' (SEQ.ID.NO.:55)

10 5'-TGCCTGTTCCCTGGACCCTCACGTG-3' (SEQ.ID.NO.:56)

y un ADNc de leucocito periférico humano (Clontech) como molde. Para la amplificación se usó Advantage cDNA polimerasa (Clontech) en una reacción de 50 µL mediante los siguientes ciclos, repitiendo 30 veces desde el paso 2 al 4: 94°C durante 30 seg.; 94°C durante 15 seg.; 69° durante 40 seg.; 72°C durante 3 min.; 72°C durante 6 min. Se aisló un fragmento de PCR de 1,4 kb, se clonó con el vector pCRII-TOPO™ (Invitrogen) y se secuenció completamente utilizando el kit T7 ADN Sequenase™ (Amsham). Ver, SEQ.ID.NO.:9

## e. hRUP6 (*Seq. Id. Nos. 11 y 12*)

20 La longitud total de hRUP6 se clonó mediante RT-PCR utilizando los cebadores:

5'-CAGGCCTTGGATTTAATGTCAGGGATGG-3' (SEQ.ID.NO.: 57) y

25 5'-GGAGAGTCAGCTCTGAAAGAATTCAAGG-3' (SEQ.ID.NO.: 58);

y ADNc de timo humano Marathon-Ready™ (Clontech) como molde. Para la amplificación se usó Advantage cDNA polimerasa (Clontech, de acuerdo con las instrucciones del fabricante) en una reacción de 50 µl con los siguientes ciclos: 94°C durante 30 seg.; 94°C durante 5 seg.; 66°C durante 40 seg.; 72°C durante 2,5 seg. y 72°C durante 7 min. Los ciclos del 2 hasta el 4 se repitieron 30 veces. Se aisló un fragmento de PCR de 1,3 kb y se clonó en el vector pCRII-TOPO™ (Invitrogen) y se secuenció completamente (ver, SEQ.ID.NO.: 11) utilizando el kit ABI Big Dye Terminator™ (P.E. Biosystem).

## f. hRUP7 (*Seq. Id. Nos. 13 y 14*)

35 La longitud total de hRUP7 se clonó mediante RT-PCR usando los cebadores:

5'-TGATGTGATGCCAGATACTAATAGCAC-3' (SEQ.IDNO.: 59; sentido) y

40 5'-CCTGATTCAATTAGGTGAGATTGAGAC-3' (SEQ.ID.NO.:60; antisentido)

y ADNc de leucocito periférico humano (Clontech) como molde. Para la amplificación se usó Advantage™ cDNA polimerasa (Clontech) en una reacción de 50 µl con los siguientes ciclos, repitiendo 30 veces desde el paso 2 al 4: 94°C durante 2 min.; 94°C durante 15 seg.; 60°C durante 20 seg.; 72°C durante 2 min.; 72°C durante 10 min. Se aisló un fragmento de PCR de 1,25 kb y se clonó en el vector pCRII-TOPO™ (Invitrogen) y se secuenció completamente utilizando el kit ABI Big Dye Terminator™ (P.E. Biosystem). Ver, SEQ.ID.NO.:13.

## g. hARE-5 (*Seq. Id. Nos. 5 y 6*)

50 La longitud total de hARE-5 se clonó mediante PCR utilizando los cebadores específicos para hARE5 5'-CAGCG CAGGGTGAAGCCTGAGAGC-3' SEQ.ID.NO.:69 (sentido, 5' del codón de iniciación ATG) y 5'-GGCACCTGCT GTGACCTGTGCAGG-3' SEQ.ID.NO.:70 (antisentido, 3' del codón de terminación TGA) y ADN genómico humano como molde. Para la amplificación se usó TaqPlus Precision™ DNA polimerasa (Stratagene) con el siguientes ciclos, con los pasos de 2 a 4 repetidos 35 veces: 96°C, 2 minutos; 96°C, 20 segundos; 58°C, 30 segundos; 72°C, 2 minutos; y 72°C, 10 minutos.

60 Se aisló un fragmento de PCR de 1,1 kb del tamaño predicho y se clonó en el vector pCRII-TOPO™ (Invitrogen) y se secuenció completamente (SEQ.ID.NO.:5) utilizando el kit T7 DNA Sequenase™ (Amsham).

## h. hARE-4 (*Seq. Id. Nos.: 3 y 4*)

65 La longitud total de hARE-4 se clonó mediante PCR utilizando los cebadores específicos de hARE-4 5'-CTGGTGT GCTCCATGGCATCCC-3' SEQ.ID.NO.:67 (sentido, 5' del codón de iniciación ATG) y 5'-GTAAGCCTCCA GAACGAGAGG-3' SEQ.ID.NO.:68 (antisentido, 3' del codón de terminación TGA) y ADN genómico humano como molde. Para la amplificación se usó Taq DNA polimerasa (Stratagene) y 5% DMSO mediante los siguientes ciclos, con los pasos de 2 al paso 3 repetidos 35 veces: 94°C, 3 minutos; 94°C, 30 segundos; 59°C, 2 minutos; 72°C, 10 minutos.

## ES 2 289 610 T3

Se aisló un fragmento PCR de 1,12 kb del tamaño predicho y se clonó en el vector pCRII-TOPO™ (Invitrogen) y se secuenció completamente (SEQ.ID.NO.:3) utilizando el kit T7 DNA Sequenase™ (Amsham).

### i. hARE-3 (*Sq. Id. Nos.: 1 y 2*)

La longitud total de hARE-3 se clonó mediante PCR utilizando los cebadores específicos de hARE-3 5'-gatcaagcttCCATCCTACTGAAACCATTGGTC-3' SEQ.ID.NO.:65 (sentido, los nucleótidos en minúscula representan Hind III saliente, **ATG** como codón de iniciación) y 5'-gatcagatctCAGTTCCAATATTCACACCCACCGTC-3' SEQ.ID.NO.:66 (antisentido, los nucleótidos en minúscula representan Xba I saliente, **TCA** como codón de terminación) y ADN genómico humano como molde. Para la amplificación se usó ADN polimerasa TaqPlus Precision™ con los siguientes ciclos, los pasos del 2 al 4 repetidos 35 veces: 94°C, 3 minutos; 94°C, 1 minuto; 55°C, 1 minuto; 72°C, 2 minutos; 72°C, 10 minutos.

Se aisló un fragmento PCR de 1,3 kb del tamaño predicho, se digirió con Hind III y Xba I, se clonó en el vector pRC/CMV2 (Invitrogen) en los sitios Hind III y Xba I y se secuenció completamente (SEQ.ID.NO.:1) utilizando el kit T7 DNA Sequenase™ (Amsham).

### j. hRUP3 (*Sq. Id. Nos.:7 y 8*)

La longitud total de hRUP3 se clonó mediante PCR utilizando los cebadores específicos de hRUP3 5'-GTCCTGC CACTTCGAGACATGG-3' SEQ.ID.NO.:71 (sentido, **ATG** como codón de iniciación) y 5'-GAAACTTCTGCCC TTACCGTC-3' SEQ.ID.NO.:72 (antisentido, 3' del codón de terminación TAA) y ADN genómico humano como molde. Para la amplificación se usó TaqPlus Precision™ DNA polimerasa (Stratagene) mediante los siguientes ciclos, los pasos del 2 al 4 repetidos 35 veces: 94°C, 3 minutos; 94°C, 1 minuto; 58°C, 1 minuto; 72°C, 2 minutos; 72°C, 10 minutos.

Se aisló un fragmento PCR de 1,0 kb del tamaño predicho y se clonó en el vector pCRII-TOPO™ (Invitrogen) y se secuenció completamente (SEQ.ID.NO.:7) utilizando el kit T7 DNA Sequenase™ (Amsham).

## 30 Ejemplo 2

### *Expresión de receptores*

Aunque está disponible una variedad de células para la expresión de proteínas, es más preferible que se utilicen 35 células mamíferas. La razón principal para ello es práctica, es decir, la utilización de, p.ej., células de levadura para la expresión de un GPCR, cuando es posible, introduce en el protocolo una célula no mamífera la cual puede no (de hecho, en el caso de levadura, no lo hace) incluir el acoplamiento de receptor, el mecanismo genético y las rutas de secreción que han evolucionado en sistemas mamíferos - así, los resultados obtenidos en células no mamíferas, respecto al uso potencial, no son tan preferidos como aquéllos obtenidos a partir células mamíferas. De las células mamíferas, 40 las células COS-7, 293 y 293T son particularmente preferidas, aunque la célula mamífera específica utilizada puede ser predicha en base a las necesidades particulares del artesano. El procedimiento general para la expresión de los GPCRs descritos es como sigue.

En el día uno, se cultivaron en placa 1X10<sup>7</sup> células 293T por placa de 150 mm. En el día dos, se prepararán dos 45 tubos de reacción (las proporciones a seguir para cada tubo son por placa): el tubo A se preparará mezclando 20 µg de ADN (p. ej., vector pCMV; vector pCMV con ADNc de receptor, etc.) en 1,2 ml de DMEM libre de suero (Irvine Scientific, Irvine, CA); el tubo B se preparará mezclando 120 µl de lipofectamine (Gibco BRL) en 1,2 ml de DMEM libre de suero. Los tubos A y B son mezclados mediante inversiones (varias veces), seguido de incubación a temperatura ambiente durante 30-45 min. La mezcla puede ser referida como la "mezcla de transfección". Se lavan las 50 células cultivadas 293T con 1XPBS, seguido por la adición de 10 ml de DMEM libre de suero. Se añadirán entonces a las células 2,4 ml de la mezcla de transfección, seguido de incubación durante 4 h a 37°C/5% de CO<sub>2</sub>. Entonces, se eliminó por aspiración la mezcla de transfección, seguido de la adición de 25 ml de DMEM/10% de Suero Fetal Bovino. Entonces, las células serán incubadas a 37°C/5% de CO<sub>2</sub>. Despues de 72 h de incubación, las células entonces pueden ser recolectadas y utilizadas para análisis.

## 55 Ejemplo 3

### *Distribución tisular de los GPCRs humanos descritos*

60 Se pueden utilizar varias aproximaciones para la determinación de la distribución tisular de los GPCRs aquí descritos.

#### 1. Análisis Dot-Blot

65 Usando un formato de dot-blot para tejidos humanos disponible comercialmente, se sondaron GPCRs huérfanos endógenos para una determinación de las áreas donde se localizan tales receptores. Los fragmentos de ADNc de los GPCRs del Ejemplo 1 (marcados radioactivamente) fueron (o pueden ser) utilizados como sonda: la sonda marcada radioactivamente fue o (puede ser) generada utilizando el ADNc completo del receptor (escindido del vector) usando

# ES 2 289 610 T3

un kit Prime-It II™ Random Primer Labelling (Stratagene, #300385); según las instrucciones del fabricante. Se hibridó un RNA Master Blot™ humano (Clontech, #7770-1) con la sonda marcada radioactivamente de GPCR humano endógeno y se lavó bajo condiciones severas según instrucciones del fabricante. El blot se expuso al Kodak BioMax™ Autoradiography film durante la noche a -80°C. Los resultados están resumidos para varios receptores en la Tabla 5 B y C (ver Figuras 1 A y 1 B para una tabla identificando los varios tejidos y sus localizaciones, respectivamente). Se proporcionan dot-blots ejemplares en la Figura 2A y 2B para resultados derivados usando hCHN3 y hCHN8, respectivamente.

TABLA B

<b>Distribución tisular</b>	
<b>(niveles más elevados, relativos a otros tejidos en el dot-blot)</b>	
10	hGPCR27 Cerebro fetal, putamen, glándula pituitaria, caudate nucleus
15	hARE-1 Bazo, leucocitos periféricos, bazo fetal
20	hPPR1 Glándula pituitaria, corazón, glándula salival, intestino delgado, testículos
25	hRUP3 Páncreas
30	hCHN3 Cerebro fetal, putamen, cortex occipital
35	hCHN9 Páncreas, intestino delgado, hígado
40	hCHN 10 Riñón, tiroides

TABLA C

<b>Distribución tisular</b>	
<b>(niveles más elevados, relativos a otros tejidos en el dot-blot)</b>	
45	hARE-3 Cerebelo izquierdo, cerebelo derecho, testículos, accumbens
50	hGPCR3 Corpus collusum, caudate nucleus, hígado, corazón, septum interventricular
55	hARE-2 Cerebelo izquierdo, cerebelo derecho, Substantia nigra
60	hCHN8 Cerebelo izquierdo, cerebelo derecho, hígado, pulmón
65	

# ES 2 289 610 T3

## 2. RT-PCR

### a. hRUP3

5 Para averiguar la distribución tisular de ARNm hRUP3, se realizó una RT-PCR usando cebadores específicos para hRUP3 y paneles de ADNc de múltiples tejidos humanos (MTC, Clontech) como moldes. Se utilizó Taq DNA polimerasa (Stratagene) para la reacción de PCR, utilizando los siguientes ciclos de reacción en una reacción de 40  $\mu$ l : 94°C durante 2 min.; 94°C durante 15 seg.; 55°C durante 30 seg.; 72°C durante 1 min.; 72°C durante 10 min. Los cebadores fueron como sigue:

10 5'-GACAGGTACCTGCCATCAAG-3' (SEQ.ID.NO.:61; sentido)

15 5'-CTGCACAATGCCAGTGATAAGG-3' (SEQ.ID.NO.:62; antisentido)

15 Se cargaron 20  $\mu$ l de la reacción en un gel de agarosa al 1%; los resultados se presentan en la Figura 3.

20 Como se soporta por los datos de la Figura 3, de los 16 tejidos humanos del panel de ADNc utilizado (cerebro, colon, corazón, riñón, ovario, páncreas, placenta, próstata, esqueleto, intestino delgado, bazo, testículos, leucocito de timo, e hígado) sólo en el páncreas se evidencia una banda simple hRUP3. Análisis comparativos adicionales de la secuencia de proteínas de hRUP3 con otros GPCRs sugieren que hRUP3 está relacionado con GPCRs que tienen moléculas pequeñas como ligandos endógenos así que se predijo que el ligando endógeno para hRUP3 es una molécula pequeña.

25 b. hRUP4

Se realizó una RT-PCR usando los oligo 8 y 4 de hRUP4 como cebadores y los paneles de ADNc de múltiples tejidos humanos (MTC, Clontech) como moldes. Para la amplificación se usó Taq DNA polimerasa (Stratagene) en una reacción de 40  $\mu$ l mediante los siguientes ciclos: 94°C durante 30 segundos, 94°C durante 10 segundos, 55°C durante 30 segundos, 72°C durante 2 minutos, y 72°C durante 5 minutos con los ciclos del 2 al 4 repetidos 30 veces.

30 Para analizar los productos RT-PCR se cargaron 20  $\mu$ l de la reacción en un gel de agarosa al 1%, y se encontró ARNm de hRUP4 expresado en muchos tejidos humanos, con la expresión más fuerte en corazón y riñón. (ver, Figura 4). Para confirmar la autenticidad de los fragmentos de PCR, se usó como sonda para el análisis Southern Blot un fragmento de 300 pb derivado del extremo 5' de hRUP4. La sonda se marcó con  $^{32}$ P-dCTP usando el kit Prime-It II<sup>TM</sup> Random Primer Labeling (Stratagene) y se purificó utilizando las microcolumnas ProbeQuant<sup>TM</sup> G-50 (Amersham). La hibridación se hizo durante la noche a 42°C después de 12 horas de prehibridación. Finalmente, se lavó el blot a 65°C con 0,1 x SCC. El Southern blot confirmó los fragmentos de PCR como hRUP4.

40 c. hRUP5

Se realizó una RT-PCR usando los siguientes cebadores específicos para hRUP5:

45 5'-CTGACTTCTTGTTCCTGGCAGCAGCGG-3' (SEQ.ID.NO.:63; sentido)

50 5'-AGACCAGCCAGGGCACGCTGAAGAGTG-3' (SEQ.ID.NO.:64; antisentido)

y los paneles de ADNc de múltiples tejidos humanos (MTC, Clontech) como moldes. Para la amplificación se usó Taq DNA polimerasa (Stratagene) en una reacción de 40  $\mu$ l mediante los siguientes ciclos: 94°C durante 30 segundos, 94°C durante 10 segundos, 62°C durante 1,5 minutos, 72°C durante 5 minutos, y con los ciclos del 2 al 3 repetidos 30 veces. Se cargaron 20  $\mu$ l de la reacción en un gel de agarosa al 1,5% para analizar los productos de RT-PCR, y el ARNm hRUP5 se encontró expresado sólo en los leucocitos de sangre periférica (datos no facilitados).

55 d. hRUP6

Se aplicó RT-PCR para confirmar la expresión y para determinar la distribución tisular de hRUP6. Los oligonucleótidos usados, basados en un alineamiento de los segmentos AC005871 y GPR66, tenían las siguientes secuencias:

60 5'-CCAACACCAGCATCCATGGCATCAAG-3' (SEQ.ID.NO.:73; sentido),

5'-GGAGAGTCAGCTCTGAAAGAATTCAAGG-3' (SEQ.ID.NO.:74; antisentido).

65 Se utilizaron los paneles de ADNc de múltiples tejidos humanos (MTC, Clontech) como moldes. La PCR se realizó utilizando TaqPlus Precision<sup>TM</sup> polimerasa (Stratagene; se seguirán las instrucciones de fabricación) en una reacción de 40  $\mu$ l mediante los siguientes ciclos: 94°C durante 30 segundos, 94°C durante 5 segundos, 66°C durante 40 segundos, 72°C durante 2,5 minutos, y 72°C durante 7 min. Los ciclos del 2 al 4 se repitieron 30 veces.

Se cargaron 20  $\mu$ l de la reacción en un gel de agarosa al 1,2% para analizar los productos de RT-PCR, y un fragmento específico de ADN de 760 pb que representaba hRUP6 se expresó predominantemente en el timo y con menos expresión en el corazón, riñón, pulmón, próstata, intestino delgado y testículos. (ver, Figura 5).

5 Aunque está disponible una variedad de Vectores para aquéllos en la materia, con el objetivo de la utilización para GPCRs humanos, tanto endógenos como no endógenos, es más preferido que el Vector utilizado sea pCMV. Este vector se depositó con la Colección de Cultivos Tipo Americana (American Type to Culture Collection, ATCC) el 13 de Octubre de 1998 (10801 University Blvd., Manassas, VA 20110-2209 USA) bajo las condiciones del Tratado de  
10 Budapest para el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en Materia de Patentes. El ADN fue probado por la ATCC y determinado el serlo. La ATCC ha asignado el siguiente número de depósito a pCMV: ATCC #203351.

#### Referencias citadas en la descripción

15 La lista de referencias citadas por el solicitante es sólo para la comodidad del lector. No forma parte del documento de patente europea. Aunque se ha puesto atención en compilar las referencias, no se puede excluir que haya errores u omisiones y la EPO deniega toda responsabilidad respecto a esta cuestión.

#### 20 Documentos de patente citados en la descripción

- EP 05003040 A [0060]
- US 60137131 B [0060]
- EP 99972682 A [0060]
- US 60141448 B [0060]
- 25 • US 9923687 W[0060]
- US 60156653 B [0060]
- US 60109213 B [0060]
- US 60156333 B [0060]
- 30 • US 60120416 B [0060]
- US 60156555 B [0060]
- US 60121852 B [0060]
- US 60156634 B [0060]
- 35 • US 60123946 B [0060]
- US 60157280 B [0060]
- US 60123949 B [0060]
- US 60157294 B [0060]
- US 60136436 B [0060]
- 40 • US 60157281 B [0060]
- US 60136437 B [0060]
- US 60157293 B [0060]
- US 60136439 B [0060]
- US 60157282 B [0060]
- 45 • US 60136567 B [0060]
- US 09417044 B [0060]
- US 60137127 B [0060]
- US 09416760 B [0060]

#### Literatura de no patente citada en la descripción

- 50 • Indirect Mechanism of Synaptic Transmission. From Neuron To Brain. Sinauer Associates, Inc, 1992 [0019]  
[0020]

55

60

65

# ES 2 289 610 T3

## REIVINDICACIONES

1. Método de rastreo de compuestos candidatos para identificar un agente farmacéutico para una enfermedad o 5 estado de desorden relacionado con el páncreas, el método comprendiendo:
  - proporcionar una célula hospedadora eucariota que comprende un receptor acoplado a proteína G que es una versión activa independiente de ligando de un receptor que tiene la SEQ ID NO: 8, donde el receptor se acopla a una proteína G; y
  - 10 rastrear compuestos candidatos contra dicho receptor acoplado a proteína G.
2. Método de la reivindicación 1, donde la célula hospedadora eucariota es una célula hospedadora de mamífero.
3. Método de la reivindicación 1, donde la célula hospedadora eucariota es una célula de levadura.
- 15 4. Método de la reivindicación 1 o la reivindicación 2, donde dicho rastreo comprende detectar AMPc.
5. Método de la reivindicación 4, donde dicha detección de AMPc comprende ELISA utilizando un anticuerpo 20 anti-AMPc.
6. Método de la reivindicación 4, donde la célula comprende un sistema reportero que comprende múltiples elementos que responden a AMPc operativamente unidos a un gen reportero.
- 25 7. Método de cualquiera de las reivindicaciones anteriores, donde el método comprende identificar un agonista del receptor acoplado a proteína G.
8. Método de cualquiera de las reivindicaciones 1 a 6, donde el método comprende identificar un agonista parcial del receptor acoplado a proteína G.
- 30 9. Método de cualquiera de las reivindicaciones 1 a 6, donde el método comprende identificar un agonista inverso del receptor acoplado a proteína G.
10. Método según cualquiera de las reivindicaciones 7 a 9, donde el método además comprende confirmar que el 35 compuesto candidato se une al receptor.
11. Método según cualquiera de las reivindicaciones 7 a 10, donde el método además comprende formular dicho agonista, agonista parcial o agonista inverso como un fármaco.
- 40 12. Método de cualquiera de las reivindicaciones anteriores, donde el receptor es un receptor endógeno que tiene la SEQ ID NO: 8.
13. Método de cualquiera de las reivindicaciones 1 a 11, donde dicho receptor es un receptor no endógeno que 45 tiene una mutación localizada 16 residuos de aminoácido N-terminal desde el residuo conservado de prolina dentro del dominio TM6 de la SEQ ID NO: 8.

50

55

60

65

1	2	3	4	5	6	7	8
A	Amigdala	Caudate Nucleous	Cerebelo	Córtex Cerebral	Córtex Frontal	Hipocampo	Bulbo Raquídeo
B	Córtex Occipital	Putamen	Substantia Nigra	Córtex Temporal	Tálamo	Núcleo Accumbens	
C	Corazón	Aorta	Músculo esquelético	Colon	Vejiga	Útero	Medula espinal
D	Testículos	Ovario	Páncreas	Pituitaria	Glándula adrenal	Próstata	Estómago
E	Riñon	Hígado	Intestino Delgado		Tiroídes	Glándula Salival	Glándula Mamaria
F	Apéndice	Pulmón	Tráquea			Leucocito Periférico	Ganglio Linfático
G	Cerebro	Corazón Fetal	Riñón Fetal	Hígado	Bazo Fetal	Timo Fetal	Pulmón Fetal
H							

Fig. 1A

	1	2	3	4	5	6	7	8	9	10	11	12
A	Cerebro Izquierdo	Substantia Nigra	Corazón	Esófago	Colon Transversal	Riñón	Pulmón	Hígado	Leucemia HL-60	Cerebro Fetal		
B	Cortex Cerebral Derecho	Núcleo Accumbens	Aorta	Estómago	Colon Descendente	Músculo Esquelético	Placenta	Páncreas	HeLa S3	Corazón Fetal		
C	Cortex Frontal	Cuerpo Calloso	Tálamo	Aurícula Izquierda	Duodeno	Recto	Bazo	Vejiga	Glándula Adrenal	Leucemia K562	Riñón Fetal	
D	Lóbulo Parietal	Amígdala	Glándula Pituitaria	Aurícula Derecha	Enterocito					Leucemia MOLT-4	Hígado Fetal	
E	Cortex Occipital	Claudete Nucleous	Médula espinal	Ventriculo Izquierdo	ileón					Glándula Salival	Linfoma Raji de Burkitt	Bazo Fetal
F	Cortex Temporal	Hipocampo		Ventriculo Derecho	Ciego del intestino grueso					Glándula Mamaria	Linfoma Daudi de Burkitt	Timo Fetal
G	Circunvolución Paracentral del Cortex Cerebral	Bulbo Raquídeo		Septum Inter-ventricular	Apéndice					Mérida Ósea	Ovario	Adeno-carcinoma colorectal
H											Tráquea	Carcinoma Pulmonar A549

Fig. 1B

Fig. 2A

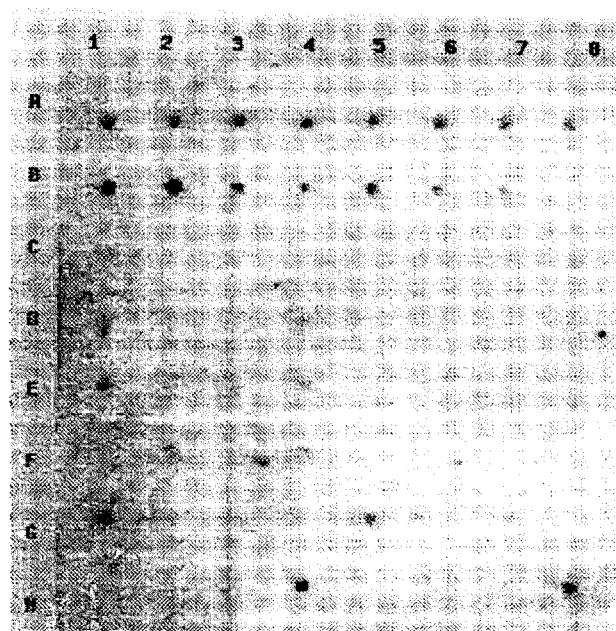
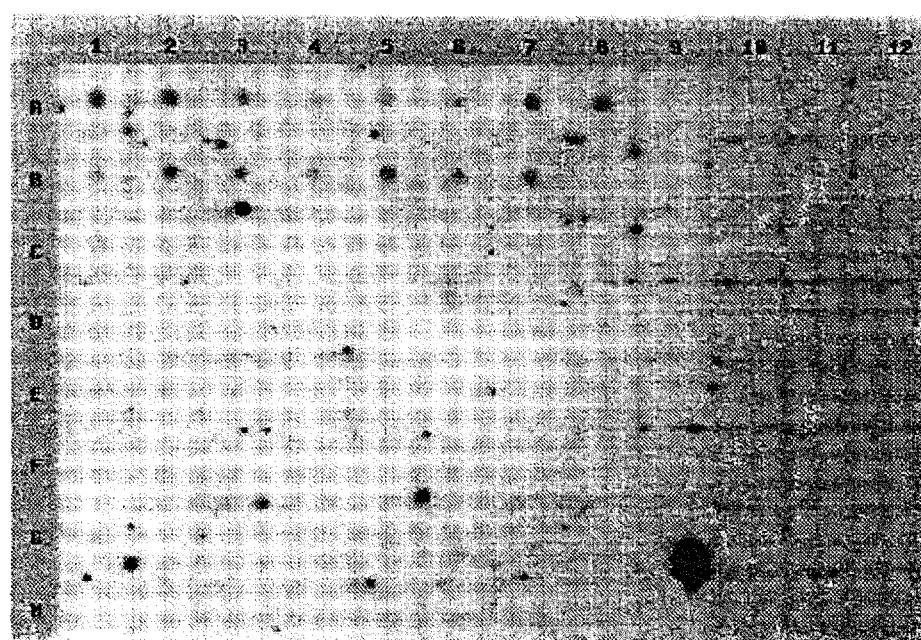


Fig. 2B



ES 2 289 610 T3

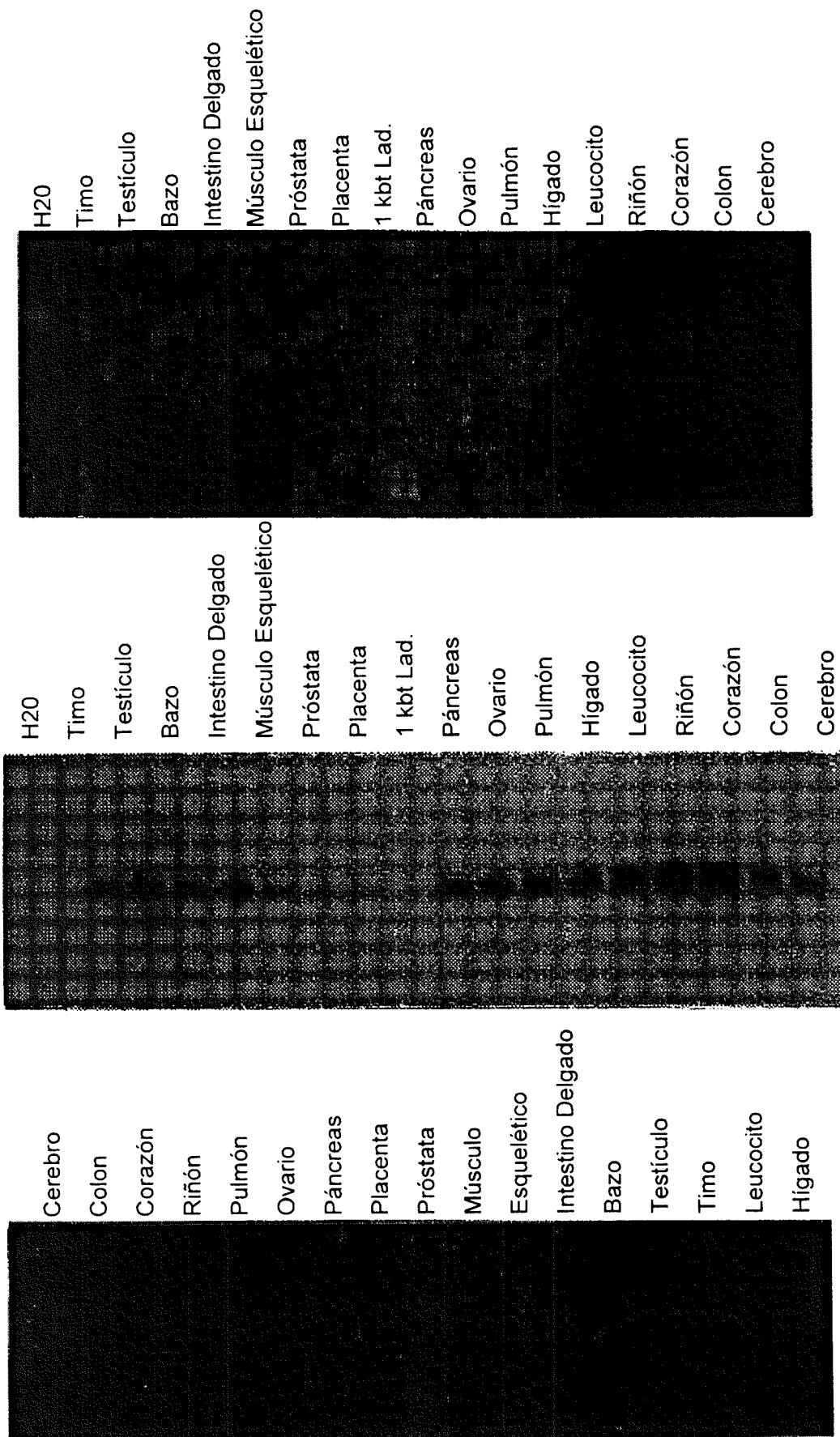


Fig. 3

Fig. 4

Fig. 5

**LISTA DE SECUENCIAS**

**(1) INFORMACIÓN GENERAL**

- 5           (i) SOLICITANTE: Chen, Ruoping  
               Dang, Huong T,  
               Liaw, Chen W.  
               Lin, I-Lin
- 10          (ii) TITULO DE LA INVENCIÓN: Receptores acoplados a Proteína G humanos huérfanos
- (iii) NÚMERO DE SECUENCIAS: 74
- 15          (iv) DIRECCIÓN DE CORRESPONDENCIA:  
             (A) DESTINATARIO: Arena Pharmaceuticals, Inc.  
             (B) DIRECCIÓN: 6166 Nancy Ridge Drive  
             (C) CIUDAD: San Diego  
             (D) ESTADO: California  
             (E) PAÍS: Estados Unidos de América  
             (F) (ZIP): 92121
- 20          (v) FORMATO LEGIBLE POR ORDENADOR:  
             (A) TIPO DE MEDIO: disquette  
             (B) (B) ORDENADOR: IBM PC compatible  
             (C) SISTEMA OPERATIVO: PC-DOS/MS-DOS  
             (D) PROGRAMA: PatentIn Release #1.0, Versión #1.30
- 25          (vi) DATOS DE LA SOLICITUD ACTUAL:  
             (A) NÚMERO DE LA SOLICITUD: US  
             (B) FECHA DE REGISTRO:  
             (C) CLASIFICACIÓN:
- 30          (viii) INFORMACIÓN DEL ABOGADO/AGENTE:  
             (A) NOMBRE: Burgoon, Richard P.  
             (B) NÚMERO DE REGISTRO: 34.787
- 35          (ix) INFORMACIÓN PARA TELECOMUNICACIÓN:  
             (A) TELÉFONO: (858) 453-7200  
             (B) TELEFAX: (858) 453-7210

**(2) INFORMACIÓN PARA LA SEC. NO ID: 1:**

- 50          (i) CARACTERÍSTICAS DE LA SECUENCIA:  
             (A) LONGITUD: 1260 pares de base  
             (B) TIPO: ácido nucléico  
             (C) TIPO DE HEBRA: sencilla  
             (D) TOPOLOGÍA: lineal
- 55          (ii) TIPO DE MOLÉCULA: DNA (genomic)
- 60          (iii) DESCRIPCIÓN DE LA SECUENCIA: SEQ NO ID: 1:

ES 2 289 610 T3

A T G C T T C T C G G C A G T G T G A C T G C G T T C C A T A C C G G G A C A T C A A C A C A C A A C A T T T G T C 60

5 G T G T A T G A A A A C A C T A C A T G A A T A T T A C A C T C C C T C C A C C A T T C C A G C A T C T G A C C T C 120

10 A G T C C A T T G C T T A G A T A T A G T T T G A A J C C A T G G C T C C C A C T G G T T T G A G G T T C C T T G A C C 180

G T G A A T A G T A C A G C T G G C C C A C A C A C C A G C A G C A T T A A G A G C C T A A A C T T G C C T C T T 240

C A G A T C A C C C T T T C T C T A T A A T G A T A T T C A T T C T G T T G T G C T T T T C T G G G A A C T T G 300

15 G T T G T T G C C C T A G G T T T A C C A A A A G C T G C C A T G A O G T C T G C A A T T A A C A T C T C C P T 360

G C C A G C C T A G C T T T G C A G A C A T G G C T T G C A G T G C T G A A C A T G C C C T T T C C C T G G T A 420

20 A C T A T T C T T A C T A C C G A T G G A T T T T T G G A A A T T C T T C T G A G G G T A T C T G T A T G T T T 480

T T C T G G T T A T T T G T G A T A G A A G G A G T A G C C A T C C T G C T C A C T A T G C A T A G A T A G G T T C 540

25 C T T A T T A T A G T C C A G A G G C A G G A T A A C C C A T A T A G G C T A A G G T A A G G T C T G A 600

G T T T C T T G G G C A A C T T C C T T T G T G A T G C T T T C C G T A G G A A A C C C G A C C T G 660

30 C A G A T A C C T T C C C G A G C T C C C A G T G T G T G T T G G G T A C A C A C C A A T C C A G G T A C C A G 720

G C T T A T G T G A T T T C T C T A T T T C T T C C T C A T A C C T T C C T G G T A A T A C T G T A C 780

35 T C A T T I A T G G G C A T A C T C A A C A C C T T C G G C A C A A T G C T T G A G G A T O C A T A C T A C C C T 840

G A A G G T A T A T G C C T A G G C C A G C A A A C T G G G T C T C A T G A T C T G C A G A G C T T T C 900

40 C A G A T G A G C A T T G A C A T O G G C T T A A A A C A C G T G C T C A C C A T T T T G A T T C T C T T T 960

G C T G T C T T C A T T G C T G C T G G C C C A T T C C A C C A T T A C A G C T T G G C 1020

45 A A G C A T T T T A C T A T G C A C A A T T T T T G A G A T T A G C A C C T G G C T A C T G T G G C 1080

T A C C T C A A G T C T G A A T C C G C T G A T C T A C T A C T G G A G G A T A A G A T A G A T 1140

50 G C T T G C T G G A C A T G A T G C C T A A G T C C T T C A G T T T T G C C G A G C T C C C T G G T C A C A C A 1200

A A G G G A C C G G A T A C T G C T C A G T G T C T A T G T G G G G A A C A T C G G A C G G T G G T G A 1260

55

(3) INFORMACIÓN PARA LA SEC. NO ID: 2:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 60 (A) LONGITUD: 1260 pares de base  
(B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal

65 (ii) TIPO DE MOLÉCULA: DNA (genomic)

ES 2 289 610 T3

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

5	Met Val Phe Ser Ala Val Leu Thr Ala Phe His Thr Gly Thr Ser Asn 1                   5                   10                   15
10	Thr Thr Phe Val Val Tyr Glu Asn Thr Tyr Met Asn Ile Thr Leu Pro 20                   25                   30
15	Pro Pro Phe Gln His Pro Asp Leu Ser Pro Leu Leu Arg Tyr Ser Phe 35                   40                   45
20	Glu Thr Met Ala Pro Thr Gly Leu Ser Ser Leu Thr Val Asn Ser Thr 50                   55                   60
25	Ala Val Pro Thr Thr Pro Ala Ala Phe Lys Ser Leu Asn Leu Pro Leu 65                   70                   75                   80
30	Gln Ile Thr Leu Ser Ala Ile Met Ile Phe Ile Leu Phe Val Ser Phe 85                   90                   95
35	Leu Gly Asn Leu Val Val Cys Leu Met Val Tyr Gln Lys Ala Ala Met 100               105               110
40	Arg Ser Ala Ile Asn Ile Leu Leu Ala Ser Leu Ala Phe Ala Asp Met 115               120               125
45	Leu Leu Ala Val Leu Asn Met Pro Phe Ala Leu Val Thr Ile Leu Thr 130               135               140
50	Thr Arg Trp Ile Phe Gly Lys Phe Phe Cys Arg Val Ser Ala Met Phe 145               150               155               160
55	Phe Trp Leu Phe Val Ile Glu Gly Val Ala Ile Leu Leu Ile Ser 165               170               175
60	Ile Asp Arg Phe Leu Ile Ile Val Gln Arg Gln Asp Lys Leu Asn Pro 180               185               190
65	Tyr Arg Ala Lys Val Leu Ile Ala Val Ser Trp Ala Thr Ser Phe Cys 195               200               205
70	Val Ala Phe Pro Leu Ala Val Gly Asn Pro Asp Leu Gln Ile Pro Ser 210               215               220
75	Arg Ala Pro Gln Cys Val Phe Gly Tyr Thr Thr Asn Pro Gly Tyr Gln 225               230               235               240
80	Ile Tyr Val Ile Leu Ile Ser Leu Ile Ser Phe Phe Ile Pro Phe Leu 255               260               265               270
85	Ala Val Phe Ile Val Cys Trp Ala Pro Phe Thr Thr Tyr Ser Leu Val 285               290               295               300
90	Ala Thr Phe Ser Lys His Phe Tyr Tyr Gln His Asn Phe Phe Glu Ile 310               315               320               325
95	Ser Thr Trp Leu Leu Trp Leu Cys Tyr Leu Lys Ser Ala Leu Asn Pro 335               340               345               350
100	Ile Ile Tyr Tyr Trp Arg Ile Lys Lys Phe His Asp Ala Cys Leu Asp 365               370               375               380
105	Met Met Pro Lys Ser Phe Lys Phe Leu Pro Gln Leu Pro Gly His Thr 395               400               405               410
110	Lys Arg Arg Ile Arg Pro Ser Ala Val Tyr Val Cys Gly Glu His Arg 415               420               425               430

(4) INFORMACIÓN PARA LA SEC. NO ID: 3:

65 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 1119 pares de base

# ES 2 289 610 T3

- (B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal

5 (ii) TIPO DE MOLÉCULA: DNA (genomic)

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

10 ATGTTAGCCA ACAGCTCCTC AACCAACAGT TCTGTTCTCC CCGTCCTGA CTACCGACCT 60  
ACCCACCGCC TGCACTTGTT GGTCTACAGC TTGGTGCTTG CTGCCCGGCT CCCCCCTCAAC 120  
15 GCGCTAGCCC TCTGCGCTT CCTGCCTGCG CTGGCGGTGC ACTCGGTTGT GACCGTGTAC 180  
ATGTTAACCG TGGCGGCCAG CGACCTCTTC TTCACCCCTCT CGCTCCCCGT TCTGTTCTCC 240  
TACTACGCAC TGCAACCTTG GCCCTTCCCC GACCTCTGT CCCAGACGAC GGCGCCATC 300  
20 TTCCAGATGA ACATGTACGG CAGCTGCATC TTCTGATGC TCATCAACGT GGACCGCTAC 360  
GCCGCCATCG TGCACCCGCT GCGACTGCGC CACCTUCOOGC GGCCCCGGGT GGCGCGGTG 420  
25 CTCTGCCTGG CGCTCTGGGC GCTCATCTTG GTGTTTGGCG TGCCCCGCCG CGCGGTGAC 480  
AGGCCCTCGC GTTGCCTGTA CGGGGACCTC GAGGTGCGCC TATGCTTGA GAGCTTCAGC 540  
GACGAGCTGT GGAAAGGCAG GCTGCTGCC CGCTGCTGCG TGGCCGAGGC GCTGGGCTTC 600  
30 CTGCTGCCCG TGGCGCGGT GGTCTACTCG TCGGGCCGAG TCTTCTGGAC CCTGGCGCGC 660  
CCCGACGCCA CGCGAGCCA CGCGCGCGG AAGACCGCTGC GCGTCTCTCT GCGTAACCTC 720  
35 GTCATCTTCC TCTGCTGCTT CGTGCCTAC AACAGCACGC TGGCGGTCTA CGGGCTGCTG 780  
CGAGCGACGC TGGTGGCGGC CAGCGTGCCT GCGGCGCGATC GCGTGCCTCG GGTGCTGATG 840  
40 GTGATGGTGC TGCTGGCCGG CGCCAACCTGC GTGCTGGACC CGCTGGTATA CTACTTTAGC 900  
GCCGAGGCGT TCCGCAACAC CCTGCGCGGC CTGGGCACTC CGCACCGGGC CAGGACCTCG 960  
GCCACCAACG GGACCGCGGC GGCCTGCGCG CAATCGAAA GGTCCGGCGT CACCAACGAC 1020  
45 GCCACCAAGGC CGGATGCCGC CAGTCAGGCG CTGCTCCGAC CCTCCGACTC CCACCTCTG 1080  
TCTTCTTCA CACAGTGTCC CGAGGATTCC GCCCTCTGA 1119

50 (5) INFORMACIÓN PARA LA SEC. NO ID: 4:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 372 aminoácidos  
(B) TIPO: aminoácido  
55 (C) TIPO DE HEBRA:  
(D) TOPOLOGÍA: no relevante

60 (ii) TIPO DE MOLÉCULA: proteína

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

# ES 2 289 610 T3

	Met Leu Ala Asn Ser Ser Ser Thr Asn Ser Ser Val Leu Pro Cys Pro			
1	5	10	15	
	Asp Tyr Arg Pro Thr His Arg Leu His Leu Val Val Tyr Ser Leu Val			
5	20	25	30	
	Leu Ala Ala Gly Leu Pro Leu Asn Ala Leu Ala Leu Trp Val Phe Leu			
	35	40	45	
10	Arg Ala Leu Arg Val His Ser Val Val Ser Val Tyr Met Cys Asn Leu			
	50	55	60	
	Ala Ala Ser Asp Leu Leu Phe Thr Leu Ser Leu Pro Val Arg Leu Ser			
15	65	70	75	80
	Tyr Tyr Ala Leu His His Trp Pro Phe Pro Asp Leu Leu Cys Gln Thr			
	85	90	95	
	Thr Gly Ala Ile Phe Gln Met Asn Met Tyr Gly Ser Cys Ile Phe Leu			
	100	105	110	
20	Met Leu Ile Asn Val Asp Arg Tyr Ala Ala Ile Val His Pro Leu Arg			
	115	120	125	
	Leu Arg His Leu Arg Arg Pro Arg Val Ala Arg Leu Leu Cys Leu Gly			
25	130	135	140	
	Val Trp Ala Leu Ile Leu Val Phe Ala Val Pro Ala Ala Arg Val His			
	145	150	155	160
	Arg Pro Ser Arg Cys Arg Tyr Arg Asp Leu Glu Val Arg Leu Cys Phe			
30	165	170	175	
	Glu Ser Phe Ser Asp Glu Leu Trp Lys Gly Arg Leu Leu Pro Leu Val			
	180	185	190	
	Leu Leu Ala Glu Ala Leu Gly Phe Leu Leu Pro Leu Ala Ala Val Val			
35	195	200	205	
	Tyr Ser Ser Gly Arg Val Phe Trp Thr Leu Ala Arg Pro Asp Ala Thr			
	210	215	220	
	Gln Ser Gln Arg Arg Lys Thr Val Arg Leu Leu Leu Ala Asn Leu			
40	225	230	235	240
	Val Ile Phe Leu Leu Cys Phe Val Pro Tyr Asn Ser Thr Leu Ala Val			
	245	250	255	
	Tyr Gly Leu Leu Arg Ser Lys Leu Val Ala Ala Ser Val Pro Ala Arg			
45	260	265	270	
	Asp Arg Val Arg Gly Val Leu Met Val Met Val Leu Leu Ala Gly Ala			
	275	280	285	
50	Asn Cys Val Leu Asp Pro Leu Val Tyr Tyr Phe Ser Ala Glu Gly Phe			
	290	295	300	
	Arg Asn Thr Leu Arg Gly Leu Gly Thr Pro His Arg Ala Arg Thr Ser			
55	305	310	315	320
	Ala Thr Asn Gly Thr Arg Ala Ala Leu Ala Gln Ser Glu Arg Ser Ala			
	325	330	335	
60	Val Thr Thr Asp Ala Thr Arg Pro Asp Ala Ala Ser Gln Gly Leu Leu			
	340	345	350	
	Arg Pro Ser Asp Ser His Ser Leu Ser Ser Phe Thr Gln Cys Pro Gln			
65	355	360	365	
	Asp Ser Ala Leu			
	370			

# ES 2 289 610 T3

## (6) INFORMACIÓN PARA LA SEC. NO ID: 5:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- 5 (A) LONGITUD: 1107 base de pares
- (B) TIPO:ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

### 10 (ii)TIPO DE MOLÉCULA: DNA (genomic)

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

15 ATGGCCAACT CCACAGGGCT GAAAGCTCA GAAAGTGGAG GCTCGTTGGG GTTGATCCTG 60  
GCAGACTCTCG TGGAGGTGGG GGCACTCTCG GGCAACGGCG CGCTGCTGGT CGTGGTGCTG 120  
20 CGCACGCCCG GACTCCCGCA CGCGCTCTAC CTGGGCCACC TGTGGCTCGT GGACCTGCTG 180  
GCGGCCGCGT CCATCATGCC GCTGGCCCTG CTGGCCGACG CGCCGCCCGG GCTGGGCCGC 240  
25 GTGGGCTCG GCCCCCGGCC ATGCCGGGCC GCTCGCTTCC TCTCCGCCGC TCTGCTGCCG 300  
GCCTGCACCC TCAGGGCTGAC CGCACTTGGC CTGGCACCGT ACCGCCTCAT CGTGCACCCG 360  
30 CTGGGCCAG GCTCGGGGCCG GCGGCCCTGCTG CTGGTCTCA CGCGCGTGTG GGCGGGGGCG 420  
GGACTCTGG GCGCGCTCTC CCTGCTCGGC CGCGGCCCGG CACCGCCCG TGCTCTCT 480  
35 CCTGCTCG TCTGGCTCG GGGCGCTCGG CGCTTCCCGG CGCTCTGGC CCTGCTGCCG 540  
TTGCGCTGC CGGCCCTCTG GCTGCTCGGC GCCTACGGGG GCACTCTTCGT GGTCGGCGGT 600  
CGCGCTGCCG TGAGGGCCCC AGGGGGGGGG CGGGGGTCCC GACTCGCTC CGACTCTCTG 660  
GATAGCCGCC TTTCCATCTT GCGGCCCTCTC CGGCCCTGCC TGCCCCGGGG CAAGGGCGCC 720  
40 CTGGCCCCAG CGCTGGCGGT GGGCAATTTC GCAAGCTACT CGCTGCCCTTA TGGCTGCACG 780  
TGCCTGGCGC CGCGACGCCG GGCGCGGGAA GCGGAAGCG CGTCACCTG GGTCGCTAC 840  
45 TCGGCCCTCG CGGCTCACCC CTTCCTGTAC GGGCTGCTGC AGCGCCCCGT CGCGTTGGCA 900  
CTGGGCCGCC TCTCTCGCCG TCACTGCTT GGACCTGTGC GGCGCTGCCAC TCCGCTAGCC 960  
TOGCACCCGC GGGCACTCTT GCAATGCCCTC CAGAGACCCC CAGAGGGCCC TGCCGTAGGC 1020  
50 CCTTCTGAGG CTCCAGAAC ACGCCCGAG TTGGCAGGAG GGCGGAGGCC CGCATACCAAG 1080  
GGGCCACCTG AGAGTTCTCT CTCCCTGA 1107

## (7) INFORMACIÓN PARA LA SEC. NO ID: 6:

### (i) CARACTERÍSTICAS DE LA SECUENCIA:

- 55 (A) LONGITUD: 368 aminoácidos
- (B) TIPO:aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

### 60 (ii) TIPO DE MOLÉCULA: proteína

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

# ES 2 289 610 T3

	Met Ala Asn Ser Thr Gly Leu Asn Ala Ser Glu Val Ala Gly Ser Leu			
1	5	10	15	
5	Gly Leu Ile Leu Ala Ala Val Val Glu Val Gly Ala Leu Leu Gly Asn			
	20	25	30	
10	Gly Ala Leu Leu Val Val Val Leu Arg Thr Pro Gly Leu Arg Asp Ala			
	35	40	45	
15	Leu Tyr Leu Ala His Leu Cys Val Val Asp Leu Leu Ala Ala Ala Ser			
	50	55	60	
20	Ile Met Pro Leu Gly Leu Leu Ala Ala Pro Pro Pro Gly Leu Gly Arg			
	65	70	75	80
25	Val Arg Leu Gly Pro Ala Pro Cys Arg Ala Ala Arg Phe Leu Ser Ala			
	85	90	95	
30	Ala Leu Leu Pro Ala Cys Thr Leu Gly Val Ala Ala Leu Gly Leu Ala			
	100	105	110	
35	Arg Tyr Arg Leu Ile Val His Pro Leu Arg Pro Gly Ser Arg Pro Pro			
	115	120	125	
40	Pro Val Leu Val Leu Thr Ala Val Trp Ala Ala Ala Gly Leu Leu Gly			
	130	135	140	
45	Ala Leu Ser Leu Leu Gly Pro Pro Pro Ala Pro Pro Pro Ala Pro Ala			
	145	150	155	160
50	Arg Cys Ser Val Leu Ala Gly Gly Leu Gly Pro Phe Arg Pro Leu Trp			
	165	170	175	
55	Ala Leu Leu Ala Phe Ala Leu Pro Ala Leu Leu Leu Gly Ala Tyr			
	180	185	190	
60	Gly Gly Ile Phe Val Val Ala Arg Arg Ala Ala Leu Arg Pro Pro Arg			
	195	200	205	
65	Pro Ala Arg Gly Ser Arg Leu Arg Ser Asp Ser Leu Asp Ser Arg Leu			
	210	215	220	
70	Ser Ile Leu Pro Pro Leu Arg Pro Arg Leu Pro Gly Gly Lys Ala Ala			
	225	230	235	240
75	Leu Ala Pro Ala Leu Ala Val Gly Gln Phe Ala Ala Cys Trp Leu Pro			
	245	250	255	
80	Tyr Gly Cys Ala Cys Leu Ala Pro Ala Ala Arg Ala Ala Glu Ala Glu			
	260	265	270	
85	Ala Ala Val Thr Trp Val Ala Tyr Ser Ala Phe Ala Ala His Pro Phe			
	275	280	285	
90	Leu Tyr Gly Leu Leu Gln Arg Pro Val Arg Leu Ala Leu Gly Arg Leu			
	290	295	300	
95	Ser Arg Arg Ala Leu Pro Gly Pro Val Arg Ala Cys Thr Pro Gln Ala			
	305	310	315	320
100	Trp His Pro Arg Ala Leu Leu Gln Cys Leu Gln Arg Pro Pro Glu Gly			
	325	330	335	
105	Pro Ala Val Gly Pro Ser Glu Ala Pro Glu Gln Thr Pro Glu Leu Ala			
	340	345	350	
110	Gly Gly Arg Ser Pro Ala Tyr Gln Gly Pro Pro Glu Ser Ser Leu Ser			
	355	360	365	

# ES 2 289 610 T3

(8) INFORMACIÓN PARA LA SEC. NO ID: 7:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1008 pares de bases
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: DNA (genomic)

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

ATGGAATCAT CTTTCTCATT TGGAGTGATC CTTAGCTGTCC TGGCTCCCT CATCATTGCT 60  
ACTAACACAC TAGTGGCTGT GGCTGTGCTG CTGGTGTGCC ACAAGAAATGA TGGTGTCACT 120  
CTCTGCTTCA CCTTGAATCT GGCTGTGGCT GACACCTTGA TTGGTGTGTC CATCTCTGGC 180  
CTACTCACAG ACCAGCTCTC CAGCCCTTCT CGGCCCCACAC AGAAGACCCCT GTGCAGCCTG 240  
COGATGGCAT TTGTCACTTC CTCCCGCACT GCCTCTGTCC TCACGGTCAT GCTGATCACC 300  
TTTGACAGGT ACCTTGCCAT CAAGCAGCCC TCCGGCTACT TGAAGATCAT GAGTGGGTTTC 360  
GTGGCCGGGG CCTGCGATTCG CGGGCTGTGG TTAGTGTCTT ACCTCAATTGG CTTCCCTCCCCA 420  
CTGGAAATCC CGATGTTCCA GCAGACTGCC TACAAAGGCG AGTGCAGCTT CCTTGCTGTA 480  
TTTCACCCCTC ACTTCGTGCT GACCCCTCTCC TGGGTTGGCT TCTTCCAGC CATGCTCTC 540  
TTTGCTCTCT TCTACTGCGA CATGCTCAAG ATGGCTCCA TGACACGCCA GCAGATTOGA 600  
AAGATOGAAC ATGGCAGGAGC CATGGCTGGA GGTATCGAT CCCCACGGAC TCCCAGCGAC 660  
TTCAAAAGCTC TCCGTACTGT GTCGTGTTCTC ATTGGGAGCT TTGCTCTATC CTGGACCCCC 720  
TTCCCTTATCA CTGGCATTGT CCAGGTGGCC TGCCAGGAGT GTCACCTCTA CCTAGTGTG 780  
GAACGGTACC TGTGGCTGCT CGGCGTGGGC AACTCCCTGC TCAACCCACT CATCTATGCC 840  
TATGGCAGA AGGAGGTOCG ACTUGCAGCTC TACCACATGG CCCTAGGAGT GAAGAAGGTG 900  
CTCACCTCAT TCCCTCTCTT TCTCTCGGCC AGGAATTGTC GCCCAGGAGG GCGGAGGAA 960  
AGTTCCCTGTC ACATCGTCAC TATCTCCAGC TCAAGAGTTG ATGGCTAA 1008

(9) INFORMACIÓN PARA LA SEC. NO ID: 8:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1008 pares de bases
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: DNA (genomic)

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

# ES 2 289 610 T3

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

(10) INFORMACIÓN PARA LA SEC. NO ID: 9:

- 65           (i) CARACTERÍSTICAS DE LA SECUENCIA:  
               (A) LONGITUD: 1413 pares de bases  
               (B) TIPO: ácido nucléico

ES 2 289 610 T3

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

5 (ii) TIPO DE MOLÉCULA: DNA (genomic)

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

ATGGACACTA CCATGGAAGC TGACCTGGGT GCCACTGGCC ACAGGCCCCG CACAGAGCTT 60  
10 GATGATGAGG ACTCTCTACCC CCAAGGTGGC TGGGACACGG TCTTCCTGGT GGCCCTGCTG 120  
15 CTCTTGGGC TGCCAGCAA TGGGTTGATG GCGTGGCTGG CGCGCTCCA GGCCCCGAT 180  
19 GGAGCTGGCA CGCGCTCTGGC GCTGCTCTTG CTCAGGCTGG CGCGCTCTGA CTCTTGTTC 240  
20 CTGGCAGCAG CGGCGCTTCCA GATCCTAGAG ATCCGGCATG GGGGACACTG GCCGCTGGGG 300  
ACAGCTGGCT GCGCGCTCTA CTACTTCTTA TGGGGCGTGT CCTACTCTTC CGGCGCTTTC 360  
CTGCTGGCGG CGCTCAGCCT CGACCGCTGC CTGCTGGCGC TGTGCCACA CTGGTACCCCT 420  
45 000CACCGGC CAACTCCGCT GCGCGCTCTGG GTCTCGCCG GTGTCCTGGT GCTGGCCACA 480  
CTCTTCAGCG TGGCTGGCT GGTCTTCCCC GAGGCTGCGG TCTGGTGGTA CGACCTGGTC 540  
ATCTCTGGAG ACTTCTGGGA CAGCGAGGAG CTGTOGCTGA GGTGCTGGGG 600  
50 GGCTTCCTGC CTTCCTCTCT GCTGCTCGTC TGCCACGTC TCACCCAGGC CACAGCCTGT 660  
CGCACCTGCC ACCGCCAACA GCAGCCCCGA GCCTGGGGGG GCTTCGGGGG TGTGGCCAGG 720  
ACCATTCTGT CAGCTTATGT GGTCTTGAGG CTGCGCTACC AGCTGGCCCA GCTGCTCTAC 780  
CTGGCTTCC TGTGGACGT CTACTCTGGC TACCTGCTCT GGGAGGCGCT GGTCTACTCC 840  
40 GACTACCTGA TCTTACTCAA CAGCTGCCCTC AGCCCCCTCC TCTGCCCTAT GGCCAGTCCC 900  
GACCTCCCGA CCTGCTGCG CTCCGTCCTCG CGGCAGCTCT CTGGGAGGAG 960  
55 CGGCTGGGGCA GCTTCAGGCC CACTBAGCCA CAGACCCAGC TAGATTCTGA GGGTCCAACTT 1020  
CTGCCAGAGC CGATGCCAGA CGCCCACTCA CAGATGGAGC CTGTGGCCCA CCCTCAOGTG 1080  
AACCCACAC TCCAGGCCAG ATCGGATCCC ACAGCTCAGC CACAGCTGAA CCCTACGGCC 1140  
CAGCCACAGT CGGATCCCAC AGCCCCAGCA CAGCTGAAAC TCATGGCCCA GCCACAGTCA 1200  
GATTCTGTGG CCCAGCCACA GGCAGACACT AACGTCCAGA CCCTGCAACC TGCTGCCAGT 1260  
60 TCTGTGCCCA GTCCCTGTA TGAAGCTTCC CCAACCCCAT CCTCGCATCC TACCCCAAGGG 1320  
GCCCTGAGG ACCCAGCCAC ACCTCTTGCC TCTGAAGGAG AAAGCCCCAG CAGCACCCCG 1380  
CCAGAGGCCGG CCCCCGGCGC AGGCCCCACG TGA 1413

65 (11) INFORMACIÓN PARA LA SEC. NO ID: 10:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

# ES 2 289 610 T3

(A) LONGITUD: 468 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE HEBRA:

5 (D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

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

Met Asp Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly His Arg Pro  
1 5 10 15

Arg Thr Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly Gly Trp Asp  
15 20 25 30

Thr Val Phe Leu Val Ala Leu Leu Leu Gly Leu Pro Ala Asn Gly  
20 25 30 35

Leu Met Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr

25 50 55 60

Arg Leu Ala Ieu Leu Leu Ser Leu Ala Leu Ser Asp Phe Leu Phe  
30 65 70 75 80

Leu Ala Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His Gly Gly His  
85 90 95

Trp Pro Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe Leu Trp Gly  
35 100 105 110

Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu Ser Leu Asp  
40 115 120 125

Arg Cys Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly His Arg Pro  
130 135 140

Val Arg Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val Leu Ala Thr  
45 145 150 155 160

Leu Phe Ser Val Pro Trp Leu Val Phe Pro Glu Ala Ala Val Trp Trp  
50 165 170 175

55

60

65

# ES 2 289 610 T3

Tyr Asp Leu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu Glu Leu Ser  
 180 185 190  
 Leu Arg Met Leu Glu Val Val Gly Gly Phe Leu Pro Phe Leu Leu Leu  
 195 200 205  
 Leu Val Cys His Val Leu Thr Gln Ala Thr Arg Thr Cys His Arg Gln  
 210 215 220  
 Gln Gln Pro Ala Ala Cys Arg Gly Phe Ala Arg Val Ala Arg Thr Ile  
 225 230 235 240  
 Leu Ser Ala Tyr Val Val Leu Arg Leu Pro Tyr Gln Leu Ala Gln Leu  
 245 250 255  
 Leu Tyr Leu Ala Phe Leu Trp Asp Val Tyr Ser Gly Tyr Leu Leu Trp  
 260 265 270  
 Glu Ala Leu Val Tyr Ser Asp Tyr Leu Ile Leu Leu Asn Ser Cys Leu  
 275 280 285  
 Ser Pro Phe Leu Cys Leu Met Ala Ser Ala Asp Leu Arg Thr Leu Ile  
 290 295 300  
 Arg Ser Val Leu Ser Ser Phe Ala Ala Ala Leu Cys Glu Glu Arg Pro.  
 305 310 315 320  
 Gly Ser Phe Thr Pro Thr Glu Pro Gln Thr Gln Leu Asp Ser Glu Gly  
 325 330 335  
 Pro Thr Leu Pro Glu Pro Met Ala Glu Ala Gln Ser Gln Met Asp Pro  
 340 345 350  
 Val Ala Gln Pro Gln Val Asn Pro Thr Leu Gln Pro Arg Ser Asp Pro  
 355 360 365  
 Thr Ala Gln Pro Gln Leu Asn Pro Thr Ala Gln Pro Gln Ser Asp Pro  
 370 375 380  
 Thr Ala Gln Pro Gln Leu Asn Leu Met Ala Gln Pro Gln Ser Asp Ser  
 385 390 395 400  
 Val Ala Gln Pro Gln Ala Asp Thr Asn Val Gln Thr Pro Ala Pro Ala  
 405 410 415  
 Ala Ser Ser Val Pro Ser Pro Cys Asp Glu Ala Ser Pro Thr Pro Ser  
 420 425 430  
 Ser His Pro Thr Pro Gly Ala Leu Glu Asp Pro Ala Thr Pro Pro Ala  
 435 440 445  
 Ser Glu Gly Glu Ser Pro Ser Ser Thr Pro Pro Glu Ala Ala Pro Gly  
 450 455 460  
 Ala Gly Pro Thr  
 465

(12) INFORMACIÓN PARA LA SEC. NO ID: 11:

- 55 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
 (A) LONGITUD: 1248 pares de bases  
 (B) TIPO: ácido nucléico  
 (C) TIPO DE HEBRA: sencilla  
 (D) TOPOLOGÍA: lineal
- 60 (ii) TIPO DE MOLÉCULA: DNA (genomic)  
 (ii) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 11:

65

# ES 2 289 610 T3

ATGTCAGGGA TGAAAAGT TCAGAAATGCT TCCCTGGATCT ACCAGCAGAA ACTAGAAGAT 60  
CCATTCAGAA AACACCTGAA CAGCACCGAG GAGTATCTGG CCTTCCTCTG CGGACCTCG 120  
CGCAAGCCACT TCTTCCTCCC CTGTGCTGTC GTGTATGTCAC CAATTGTTGT GGTOGGGGTC 180  
ATTOGCAATG TCCCTGGATG CCTGCTGATT CTGCAGCACC AGGCTATGAA GACCCCCACC 240  
AACTACTACC TCTTCAGGCT AGCGATCTCTG GACCTCTGG TCTCTCTCTG TCGAATGCC 300  
CTGGAGGTCT ATGAGATGTC AGCGAACTAC CCTTTCCTGT TCGGGGGCGT AGGCTGCTAC 360  
TTCAGAACGG CCCTCTTGA GACCGTGTAC TCGGCTCCA TCTTCAGCAT CACCACGGTC 420  
AOGCTGGAGC GCTACGTTGC CATCCATACG CGCTTCCCGA CCAAACTGCA GACACCCGG 480  
CGCCGGGACCC TCAAGGATCTC CGCGATGTC TCGGGCTCTG CGCTACTCTT CTCCCTUCCC 540  
AACACCAAGCA TCCATOGCAT CAAGTTCAC TACTTCCTCA ATGGTCCCT AGTCACAGGT 600  
TOGGCCACCT GTACGGTCAT CAAGCCATG TGGATCTACA ATTTCATCAT CTAGGTCACC 660  
TCCTTCTAT TCTACCTCTT CCCCATGACT GTCATGTCG TCTCTACTA CCTCATOGCA 720  
CTCAGACTAA AGAAGACAA ATCTCTTGAG GCAGATGAG GGAATGCAA TATTCAAGA 780  
CCCTOCAGAA AATCAGTCAA CAAGATGCTG TTGTCTTGG TCTTAGTGT TCTATCTGT 840  
TGGGGGGCGT TCCACATTGA CGGACTCTTC TICAGCTTGT TGGAGGAGTG GAGTQAATCC 900  
CTGGCTGCTG TGTCAACCT CTCATGATG TGTCAGGTG TCTTCCTCTA CCTGAGCTCA 960  
GCTGTCAACC CCATTATCTA TAACCTACTG TCTCGGGCT TCCAGGGAGC ATTCAGAAAT1020  
GTGATCTCTT CTTCACCAA ACAGTGGAC TCCCAGCATG ACCCACAGTT GCCACCTGCC1080  
CAGCGGAAAC TCTTCCTGAC AGAATGCCAC TTGTGAGC TGACCTGAGA TATAAGTCCC1140  
CAATTCCAT GTCACTTCATC CATGACAAAC TCTCACCTCC CAACAGCCCT CTCTAGTCAA1200  
CAGATGTCAA GAACAAACTA TCAAGCTTC CACTTTAACCA AAACCTGA 1248

35 (13) INFORMACIÓN PARA LA SEC. NO ID: 12:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 415 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

40 45 (ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 12:

50

55

60

65

# ES 2 289 610 T3

	Met Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln	
	1 5 10 15	
	Lys Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr	
5	20 25 30	
	Leu Ala Phe Leu Cys Gly Pro Arg Arg Ser His Phe Phe Leu Pro Val	
	35 40 45	
	Ser Val Val Tyr Val Pro Ile Phe Val Val Gly Val Ile Gly Asn Val	
10	50 55 60	
	Leu Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr	
	65 70 75 80	
	Asn Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu	
15	85 90 95	
	Leu Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe	
	100 105 110	
	Leu Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr	
	115 120 125	
20	Val Cys Phe Ala Ser Ile Leu Ser Ile Thr Thr Val Ser Val Glu Arg	
	130 135 140	
	Tyr Val Ala Ile Leu His Pro Phe Arg Ala Lys Leu Gln Ser Thr Arg	
	145 150 155 160	
	Arg Arg Ala Leu Arg Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu	
25	165 170 175	
	Phe Ser Leu Pro Asn Thr Ser Ile His Gly Ile Lys Phe His Tyr Phe	
	180 185 190	
	Pro Asn Gly Ser Leu Val Pro Gly Ser Ala Thr Cys Thr Val Ile Lys	
30	195 200 205	
	Pro Met Trp Ile Tyr Asn Phe Ile Ile Gln Val Thr Ser Phe Leu Phe	
	210 215 220	
	Tyr Leu Leu Pro Met Thr Val Ile Ser Val Leu Tyr Tyr Leu Met Ala	
	225 230 235 240	
35	Leu Arg Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala	
	245 250 255	
	Asn Ile Gln Arg Pro Cys Arg Lys Ser Val Asn Lys Met Leu Phe Val	
40	260 265 270	
	Leu Val Leu Val Phe Ala Ile Cys Trp Ala Pro Phe His Ile Asp Arg	
	275 280 285	
45	Leu Phe Phe Ser Phe Val Glu Glu Trp Ser Glu Ser Leu Ala Ala Val	
	290 295 300	
	Phe Asn Leu Val His Val Val Ser Gly Val Phe Phe Tyr Leu Ser Ser	
50	305 310 315 320	
	Ala Val Asn Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala	
	325 330 335	
	Ala Phe Gln Asn Val Ile Ser Ser Phe His Lys Gln Trp His Ser Gln	
55	340 345 350	
	His Asp Pro Gln Leu Pro Pro Ala Gln Arg Asn Ile Phe Leu Thr Glu	
	355 360 365	
	Cys His Phe Val Glu Leu Thr Glu Asp Ile Gly Pro Gln Phe Pro Cys	
60	370 375 380	
	Gln Ser Ser Met His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu	
	385 390 395 400	
65	Gln Met Ser Arg Thr Asn Tyr Gln Ser Phe His Phe Asn Lys Thr	
	405 410 415	

# ES 2 289 610 T3

(14) INFORMACIÓN PARA LA SEC. NO ID: 13:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

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

(ii) TIPO DE MOLÉCULA: DNA (genomic)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 13:

ATGCCAGATA CTAATAGCAC AATCAATTAA TCACTAAGCA CTCGTGTTAC TTTAGCATTT 60  
TTTATGTCCT TAGTAGCTTT TGCTATAATG CTAGGAAATG CTTTGGTCAT TTTAGCTTT 120  
GTGCTGGACA AAAACCTTAG ACATCGAAAGT AGTTTATTTT TTCTTAACTT GGCCATCTCT 180  
GACTTCTTTCG TGGGTGTGAT CTCCATTCCCT TTGTACATCC CTCACACGCT GTTCGAATGG 240  
GATTTTGAA AGGAAATCTG TGTATTTTGG CTCACTACTG ACTATCTGTT ATOTACAGCA 300  
TCTGTATATA ACATTGTCCT CATCAAGCTAT GATCGATACC TGTCACTCTC AATGCTGTG 360  
TCTTATAGAA CTCAACATAC TGGGGTCTTG AAGATTGTTA CTCTGATGGT GGCCGTTTGG 420  
GTGCTGGCCT TCTTAGTGAA TGGGCCAATG ATTCTAGTTT CAGAGCTTG GAAAGATQAA 480  
GOTAGTGAAT GTGAACCTGG ATTTTTTTTCG GAATGGTACA TCCCTGGCAT CACATCATTC 540  
TTGQAATTGG TGATCCCAGT CATCTTAGTC GCTTATTTCAC ACATGAATAT TTATTGGAGC 600  
CTGTOGAAGC GTGATCATCT CAGTAGGTGC CAAAGCCATC CTGGACTGAC TGTCTGCTCT 660  
TCCAACATCT GTGGACACTC ATTCTAGAGGT AGACTATCTT CAAGGAGATC TCTTTCTGCA 720  
TCQACAGAAAG TTCTCTGATCT CTTCATTCA GAGAGACAGA GGAGAANGAG TAGTCTCATG 780  
TTTTCTCAA GAACCAAGAT GAATAGCAAT ACAATTGCTT CCAAAATGCG TTCTTCTCC 840  
CAATCAGATT CTGTAGCTCT TCACCAAGG GAACATOTGG AACTGCTTAG AGCCAGGAGA 900  
TTAGCCAGT CACTGGCAT TCTCTTAAGG GTTTTTGCTG TTTGCTGGGC TCCATATTCT 960  
CTGTTTCAAA TTGTCCTTTC ATTITATTCCT TCACCAACAG GTCTAAATC AGTTTGTAT 1020  
AGAATTGCAAT TTTOGCTTCA GTGGTTCAAT TCCCTTTGTCA ATCCCTTTT GTATCCATTG 1080  
TGTCAAGC GCTTCAAAA GGCTTCTTG AAAATATTTT GTATAAAA GCAACCTCTA 1140  
CCATCACAAAC ACAGTCGGTC AGTATCTCT TAA 1173

(15) INFORMACIÓN PARA LA SEC. NO ID: 14:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 390 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 14:

# ES 2 289 610 T3

	Met Pro Asp Thr Asn Ser Thr Ile Asn Leu Ser Leu Ser Thr Arg Val			
1	5	10	15	
	Thr Leu Ala Phe Phe Met Ser Leu Val Ala Phe Ala Ile Met Leu Gly			
5	20	25	30	
	Asn Ala Leu Val Ile Leu Ala Phe Val Val Asp Lys Asn Leu Arg His			
	35	40	45	
10	Arg Ser Ser Tyr Phe Phe Leu Asn Leu Ala Ile Ser Asp Phe Phe Val			
	50	55	60	
	Gly Val Ile Ser Ile Pro Leu Tyr Ile Pro His Thr Leu Phe Glu Trp			
	65	70	75	80
15	Asp Phe Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp Tyr Leu			
	85	90	95	
20	Leu Cys Thr Ala Ser Val Tyr Asn Ile Val Leu Ile Ser Tyr Asp Arg			
	100	105	110	
	Tyr Leu Ser Val Ser Asn Ala Val Ser Tyr Arg Thr Gln His Thr Gly			
	115	120	125	
25	Val Leu Lys Ile Val Thr Leu Met Val Ala Val Trp Val Leu Ala Phe			
	130	135	140	
	Leu Val Asn Gly Pro Met Ile Leu Val Ser Glu Ser Trp Lys Asp Glu			
	145	150	155	160
30	Gly Ser Glu Cys Glu Pro Gly Phe Phe Ser Glu Trp Tyr Ile Leu Ala			
	165	170	175	
	Ile Thr Ser Phe Leu Glu Phe Val Ile Pro Val Ile Leu Val Ala Tyr			
	180	185	190	
35	Phe Asn Met Asn Ile Tyr Trp Ser Leu Trp Lys Arg Asp His Leu Ser			
	195	200	205	
40	Arg Cys Gln Ser His Pro Gly Leu Thr Ala Val Ser Ser Asn Ile Cys			
	210	215	220	
	Gly His Ser Phe Arg Gly Arg Leu Ser Ser Arg Arg Ser Leu Ser Ala			
	225	230	235	240
45	Ser Thr Glu Val Pro Ala Ser Phe His Ser Glu Arg Gln Arg Arg Lys			
	245	250	255	
	Ser Ser Leu Met Phe Ser Ser Arg Thr Lys Met Asn Ser Asn Thr Ile			
	260	265	270	
50	Ala Ser Lys Met Gly Ser Phe Ser Gln Ser Asp Ser Val Ala Leu His			
	275	280	285	
	Gln Arg Glu His Val Glu Leu Leu Arg Ala Arg Arg Leu Ala Lys Ser			
	290	295	300	
55	Leu Ala Ile Leu Leu Gly Val Phe Ala Val Cys Trp Ala Pro Tyr Ser			
	305	310	315	320
	Leu Phe Thr Ile Val Leu Ser Phe Tyr Ser Ser Ala Thr Gly Pro Lys			
	325	330	335	
60	Ser Val Trp Tyr Arg Ile Ala Phe Trp Leu Gln Trp Phe Asn Ser Phe			
	340	345	350	
	Val Asn Pro Leu Leu Tyr Pro Leu Cys His Lys Arg Phe Gln Lys Ala			
	355	360	365	
	Phe Leu Lys Ile Phe Cys Ile Lys Lys Gln Pro Leu Pro Ser Gln His			
	370	375	380	
65	Ser Arg Ser Val Ser Ser			
	385	390		

# ES 2 289 610 T3

(16) INFORMACIÓN PARA LA SEC. NO ID: 15:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1128 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: DNA (genomic)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 15:

ATGGCGAACG CAAACGAGCC 00GTGGCAAC GGGGAGCGGC AGACGGCCGC CCTGGGGCTC 60  
AAAGCTGGCCA CGCTCAAGCT GCTGCTGTGC ATGAGCCTAG CGGCGAACGT GCTGTTGGG 120  
CTGCTGATCG TGCGGGGGCG CAGCTTGCAAC CGGGCCCCGT ACTACCTGCT GCTGGACCTG 180  
TACCTGGCCG ACGGGCTGCG CGCGCTCGCC TGCGCTCCGG CGCGTATACT CGCGGGCGCG 240  
CGTCGAGGGG CGCGCGCGGG CGCGCGCGCG CGCGCGCTGG CGTCGAGACT GCTGGACCTG 300  
  
CTGGCCCGCGC TCTTCCTGCTT CCACGCCGCC TTCTCTGCTG TGCGCTGGCG CGTCACCCGC 360  
TACCTGGCCA TCGCGCACCA CGCGTCTCAT GCAGAGCGCC TGCGCGCTG CGCGTGCGCC 420  
GCCATGCTGG TGTGCGCCGC CTGGCGCTG CGCGTGGCGA CGGCCTTCGC CGCAGTGCTG 480  
GACCGCGGTG CGACACACGA GGACCGCGCG TGCGCCCTGG AGCACCGGCC CGACGGCGCC 540  
CCCGCGCGCG TGGGCTTCCT GCTGCTGCTG CGCGTGGTGG TGCGGCCAC CGACCTCGTC 600  
TACCTCCGCC TGCTCTTCCT CATCCACGAC CGCCGCAAGA TGCGGCCCG CGCGCTGGTG 660  
CCCGCGCGTC CGCCACGACT GACCTCCAC CGCCCGCGCG CGACCGGCCA CGCGGCCGCC 720  
AACTGGACGG CGGGCTTCGG CGCGCGGCC ACGGCGCCCG CGCTTGCTGG CATCGGCCCG 780  
GCAGGGCGCG CGCGCGGCCG CGCGCGCTC CTGCTGCTGG AAGAATTCAA GACGGAGAAG 840  
AGGCTTGTCA AGATGTTCTA CGCCGTCACG CTGCTCTTC TGCTCTCTG GGGGCGCTAC 900  
GTCGTGGCCA OCTACCTGCG GTGCTGCTGG CGGGCGCGCG CGGTCCCCCA CGCTTACCTG 960  
ACGGCGCTCCG TGTGGCTGAC CTTCGCGCAAG CGCGGCATCA ACCCGCTGCT GTGCTTCTC 1020  
TTCAACAGGG AGCTGAAGGA CTGCTTCAGG CGCCAGTTCC CCTGCTGGCA GAGCCCCCGG 1080  
ACCACCCAGG CGACCCATCC CTGGGACCTG AAAGGATTCG GTTTATGA 1128

(17) INFORMACIÓN PARA LA SEC. NO ID: 16:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 375 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 16:

# ES 2 289 610 T3

Met Ala Asn Ala Ser Glu Pro Gly Gly Ser Gly Gly Glu Ala Ala  
 1                   3                   10                   15  
 Ala Leu Gly Leu Lys Leu Ala Thr Leu Ser Leu Leu Leu Cys Val Ser  
 20                   25                   30  
 5 Leu Ala Gly Asn Val Leu Phe Ala Leu Leu Ile Val Arg Glu Arg Ser  
 16                   40                   45  
 Leu His Arg Ala Pro Tyr Tyr Leu Leu Leu Asp Leu Cys Leu Ala Asp  
 20                   25                   30  
 Gly Leu Arg Ala Leu Ala Cys Leu Pro Ala Val Met Leu Ala Ala Arg  
 45                   50                   75                   80  
 Arg Ala Ala Ala Ala Gly Ala Pro Pro Gly Ala Leu Gly Cys Lys  
 85                   90                   95  
 Leu Leu Ala Phe Leu Ala Ala Leu Phe Cys Phe His Ala Ala Phe Leu  
 100                  105                  110  
 10 Leu Leu Gly Val Gly Val Thr Arg Tyr Leu Ala Ile Ala His His Arg  
 115                  120                  125  
 Phe Tyr Ala Glu Arg Leu Ala Gly Trp Pro Cys Ala Ala Met Leu Val  
 130                  135                  140  
 20 Cys Ala Ala Trp Ala Leu Ala Leu Ala Ala Phe Pro Pro Val Leu  
 145                  150                  155                  160  
 Asp Gly Gly Asp Asp Glu Asp Ala Pro Cys Ala Leu Glu Gln Arg  
 165                  170                  175  
 Pro Asp Gly Ala Pro Gly Ala Leu Gly Phe Leu Leu Leu Leu Ala Val  
 180                  185                  190  
 25 Val Val Gly Ala Thr His Leu Val Tyr Leu Arg Leu Leu Phe Phe Ile  
 195                  200                  205  
 His Asp Arg Arg Lys Met Arg Pro Ala Arg Leu Val Pro Ala Val Ser  
 210                  215                  220  
 His Asp Trp Thr Phe His Gly Pro Gly Ala Thr Gly Gln Ala Ala Ala  
 225                  230                  235                  240  
 30 Asn Trp Thr Ala Gly Phe Gly Arg Gly Pro Thr Pro Pro Ala Leu Val  
 245                  250                  255  
 Gly Ile Arg Pro Ala Gly Pro Gly Arg Gly Ala Arg Arg Leu Leu Val  
 260                  265                  270  
 35 Leu Glu Glu Phe Lys Thr Glu Lys Arg Leu Cys Lys Met Phe Tyr Ala  
 275                  280                  285  
 Val Thr Leu Leu Phe Leu Leu Leu Trp Gly Pro Tyr Val Val Ala Ser  
 290                  295                  300  
 40 Tyr Leu Arg Val Leu Val Arg Pro Gly Ala Val Pro Gln Ala Tyr Leu  
 305                  310                  315                  320  
 Thr Ala Ser Val Trp Leu Thr Phe Ala Gln Ala Gly Ile Asn Pro Val  
 325                  330                  335  
 Val Cys Phe Leu Phe Asn Arg Glu Leu Arg Asp Cys Phe Arg Ala Gln  
 340                  345                  350  
 45 Phe Pro Cys Cys Gln Ser Pro Arg Thr Thr Gln Ala Thr His Pro Cys  
 355                  360                  365  
 Asp Leu Lys Gly Ile Gly Leu  
 370                  375

50 (18) INFORMACIÓN PARA LA SEC. NO ID: 17:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1002 pares de base

55 (B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

60 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 17:

ES 2 289 610 T3

ATGAAACACCA CAGTGATGCA AGGCTTCAAC AGATCTGAGC GGTGGCCCCAG AGACACTCGG 60  
ATAGTACAOCT TGGTATTCCC AGCCCTCTAC ACAGTGGTTT TCTTGACCGG CATCCTGCTG 120  
5 AATACTTTGG CTCTGTGGGT GTTGTTCAC ATCCCCAGCT CCTOCACCTT CATCATCTAC 180  
CTCAAAAACA CTTTGGTGGC CGACTTGATA ATGACACTCA TGCTTCCTT CAAAATCCTC 240  
10 TCTGACTCAC ACCTGGCACC CTGGCAGCTC AGAGCTTTG TGTGTGCTTT TTCTTCGGTG 300  
ATATTTTATG AGACCATGTA TGTGGGCATC GTGCTGTTAG GCCTCATAGC CTTTGACACAGA 360  
15 TTCTCTAAGA TCATCAQACC TTTGAGAAAT ATTTTCTAA AAAAACCTGT TTTTGCAAAA 420  
ACGGTCTCAA TCTTCATCTG GTTCTTTTG TTCTTCATCT CCCTGCCAAA TACGATCTG 480  
AQCAACAAAGG AAGCAACACC ATCGTCTGTG AAAAAGTGTG CTTCTTAAA GGGGCTCTG 540  
20 GGGCTGAAAT GGCATCAAAT GGTAAATAAC ATATGCCAGT TTATTTCTG GACTGTTTT 600  
ATCCTAATGC TTGTGTTTA TGTGGTTATT GCAAAAAAAG TATATGATTG TTATAGAAAG 660  
TCCAAAAGTA AGGACAGAA AACAACAAA AAGCTGGAG GCAAAAGTATT TGTGTCGTG 720  
25 GCTGTCTCT TTGTGTTTT TGTCCATT CATTGGCCA GAGTTCCATA TACTCACAGT 780  
CIAACCAACA ATAAGACTGA CTGTAGACTG CAAAATCAAC TGTGTTATTGC TAAAGAAACA 840  
30 ACTCTCTTTT TGGCAGCAAC TAACATTTGT ATGGATCCCT TAATATACAT ATTCTTATGT 900  
AAAAAATTCA CAGAARAGCT ACCATGTATG CAAGGGAGAA AGACCACAGC ATCAAAGCAA 960  
35 GAAAATCATA GCAGTCAGAC AGACAAACATA ACCTTAAGCT GA 1002

(19) INFORMACIÓN PARA LA SEC. NO ID: 18:

- 40 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 333 aminoácidos  
(B) TIPO: aminoácido  
45 (C) TIPO DE HEBRA:  
(D) TOPOLOGÍA: no relevante  
(ii) TIPO DE MOLÉCULA: proteína  
50 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 18:

55

60

65

# ES 2 289 610 T3

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

(20) INFORMACIÓN PARA LA SEC. NO ID: 19:

- 60       (i) CARACTERÍSTICAS DE LA SECUENCIA:  
          (A) LONGITUD: 1122 pares de base  
          (B) TIPO: ácido nucléico  
          (C) TIPO DE HEBRA: sencilla  
          (D) TOPOLOGÍA: lineal
- 65       (ii) TIPO DE MOLÉCULA: ADN (genomic)

# ES 2 289 610 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 19:

ATGGCCAAACA CTACCGGAGA GCCTGAGGGG GTGAOCCTGG CTCCTGCCCC ACCGTCCGCA 60  
5  
TCAGCTTATG TGAAGCTGGT ACTGCTGGG A CTGATTATGT GCGTGAGGCT GGC GGTAAC 120  
GCCATCCTGT CCTCTCTGGT GCTCAAGGAG CGTCCCTCTGC ACAAGGGCTCC TTACTACTTC 180  
10 CTGCTGGACC TGTGCCTGGC CGATGGATA CGCTCTOCCG TCTGCTTCCC CTTTGTGCTG 240  
GCTTCTGTGC GCCACGGCTC TTCATGGACC TTCAGTOCAC TCAGCTGCAA GATTGTGGCC 300  
15 TTATATGCCG TGCCTTTTG CTCATGGCG GCCTTCATOC TGTTCTGCAT CAAGCTCACCC 360  
CGCTACATGG CCATCGCCCA CCACCGCTTC TACGCCAAGC GCATGACACT CTGGACATGC 420  
GC GGCTGTCA TCTGCATGGC CTGGACCTG TCTGTTGCCA TGGCTTCCC ACCTGTCTT 480  
20 GACGTOGGCA CCTACRAAGTT TATTCGGGAG GAGGACCA GT GCATCTTGA GCATCGTAC 540  
TTCAAGGCCA ATGACACGCT GGGCTTCATG CTTATGTTGG CTGTGCTCAT CGCAGCTTAC 600  
CATGCTGTCT ACGGCAAGCT GCTCTCTTC GAGTATGTC ACCGCAAGAT GAAGCCAGTG 660  
25 CAGATGGTGC CAGCCATCAG CCAGAACTGG ACATTCATG GTCGGGGGGC CACCGGGCAG 720  
GCTGCTGCCA ACTGGATGGC CGGCTTGGC CGTGGGGCCCA TGCCACCAAC CCGCTGCGT 780  
30 ATCCGGCAGA ATGGGCATGC AGCCAGCCCG CGGCTACTGG GCATGGACCA GGTCAAGGGT 840  
AAAAAGCAGC TGGCCCGCAT GTTCTACGGG ATCACACTOC TCTTCTGCT CCTCTGGTCA 900  
35 CCCTACATGG TGGCTGCTA CTGGCGAGTG TTTTGAAAG CCTGTCGTGT GCCCCACCGC 960  
TACCTGGCCA CTGCTGTTTG GATGAGCTTC GCCCAGGCTG CCGTCACCC AATTGTCTGC 1020  
TTCCCTGCTCA ACAAGGACCT CAAGAAGTGC CTGACCAACTC ACGCCCCCTG CTGGGGCAC 1080  
40 :GGAGGCTGCCAG CGGCTCCAG AGAACCTAC TGTGTCATGT GA 1122

(21) INFORMACIÓN PARA LA SEC. NO ID: 20:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 373 aminoácidos
- (B) TIPO: aminoácido
- 50 (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

55 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 20:

60

65

# ES 2 289 610 T3

	Met Ala Asn Thr Thr Gly Glu Pro Glu Glu Val Ser Gly Ala Leu Ser	
1	5	10
	15	
5	Pro Pro-Ser Ala Ser Ala Tyr Val Lys Leu Val Leu Leu Gly Leu Ile	
	20	25
	30	
	Met Cys Val Ser Leu Ala Gly Asn Ala Ile Leu Ser Leu Leu Val Leu	
	35	40
	45	
10	Lys Glu Arg Ala Leu His Lys Ala Pro Tyr Tyr Phe Leu Leu Asp Leu	
	50	55
	60	
	Cys Leu Ala Asp Gly Ile Arg Ser Ala Val Cys Phe Pro Phe Val Leu	
	65	70
	75	
15	80	
	Ala Ser Val Arg His Gly Ser Ser Trp Thr Phe Ser Ala Leu Ser Cys	
	85	90
	95	
	Lys Ile Val Ala Phe Met Ala Val Leu Phe Cys Phe His Ala Ala Phe	
	100	105
	110	
20	Met Leu Phe Cys Ile Ser Val Thr Arg Tyr Met Ala Ile Ala His His	
	115	120
	125	
	Arg Phe Tyr Ala Lys Arg Met Thr Leu Trp Thr Cys Ala Ala Val Ile	
	130	135
	140	
25	Cys Met Ala Trp Thr Leu Ser Val Ala Met Ala Phe Pro Pro Val Phe	
	145	150
	155	
	160	
	Asp Val Gly Thr Tyr Lys Phe Ile Arg Glu Glu Asp Gln Cys Ile Phe	
	165	170
	175	
30	Glu His Arg Tyr Phe Lys Ala Asn Asp Thr Leu Gly Phe Met Leu Met	
	180	185
	190	
	Leu Ala Val Leu Met Ala Ala Thr His Ala Val Tyr Gly Lys Leu Leu	
	195	200
	205	
35	Leu Phe Glu Tyr Arg His Arg Lys Met Lys Pro Val Gln Met Val Pro	
	210	215
	220	
	Ala Ile Ser Gln Asn Trp Thr Phe His Gly Pro Gly Ala Thr Gly Gln	
	225	230
	235	
	240	
40	Ala Ala Ala Asn Trp Ile Ala Gly Phe Gly Arg Gly Pro Met Pro Pro	
	245	250
	255	
	Thr Leu Leu Gly Ile Arg Gln Asn Gly His Ala Ala Ser Arg Arg Leu	
	260	265
	270	
45	Leu Gly Met Asp Glu Val Lys Gly Glu Lys Gln Leu Gly Arg Met Phe	
	275	280
	285	
	Tyr Ala Ile Thr Leu Leu Phe Leu Leu Leu Trp Ser Pro Tyr Ile Val	
	290	295
	300	
50	Ala Cys Tyr Trp Arg Val Phe Val Lys Ala Cys Ala Val Pro His Arg	
	305	310
	315	
	320	
	Tyr Leu Ala Thr Ala Val Trp Met Ser Phe Ala Gln Ala Ala Val Asn	
	325	330
	335	
55	Pro Ile Val Cys Phe Leu Leu Asn Lys Asp Leu Lys Lys Cys Leu Thr	
	340	345
	350	
	Thr His Ala Pro Cys Trp Gly Thr Gly Gly Ala Pro Ala Pro Arg Glu	
	355	360
	365	
60	Pro Tyr Cys Val Met	
	370	

(22) INFORMACIÓN PARA LA SEC. NO ID: 21:

- 65           (i) CARACTERÍSTICAS DE LA SECUENCIA:  
              (A) LONGITUD: 1053 pares de base  
              (B) TIPO: ácido nucléico

# ES 2 289 610 T3

- (C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal  
(ii) TIPO DE MOLÉCULA: ADN (genomic)

- 5 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 21:

ATGGCTTTGG AACAGAACCA GTCACAGAT TATTATTATG AGGAAATGA AATGAATGCC 60  
ACTTATGACT ACAGTCATA TGAATTGATC TGTATCAAAG AAGATGTCAG AGAATTGCA 120  
10 AAAGTTTTCG TCCCTTATT CCTCACAATA GCTTTGTC 180  
ATGGTAGTGG CAATTATOC CTATTACAAG AACAGAGAA CCAAAACAGA TGTGTACATC 240  
CTGAATTGG CTGTAGCAGA TTCTACTCCTT CTATTCACTC TGCCCTTTTG GGCTGTAAAT 300  
15 GCAGTTCATG GGTGGGTTT AGGGAAATA ATGTGCAAAA TAACCTTCAGC CTTGTACACA 360  
CTAAACTTTC TCTCTGAAAT GCAGTTCTG GCTTGATCA GCATAGACAG ATATGTTGCA 420  
GTAACAAATG TCCCCAGCCA ATCAGGAGTGG GAAAAACCAT OCTGGATCAT CTGTTTCGT 480  
20 GTCTGGATGG CTGCCATCTT GCTGAGCATA CCCCAAGCTGG TTTTTTATAC AGTAAATGAC 540  
AATGCTAGGT GCATTCCCCTT TTTCCCCCGC TACCTAGGAA CATCAATGAA AGCATTGATT 600  
CAAATGCTAG AGATCTGCAT TGGATTTGTA GTACCCCTTC TTATTATGGG GGTGTOCTAC 660  
TTTATCACCG CAAGGACACT CATGAGGATG CCAAACATTA AAATATCTCG ACCCCTAAAA 720  
25 GTTCTGCTCA CGTCTTTAT AGTTTCATT CTCACTCAAC TGCCCTATAA CATTGTCAAG 780  
TTCTGCCGAG CCATAGACAT CATCTACTCC CTGATCACCA GCTGCAACAT GAGCAACCGC 840  
ATGGACATCG CCATCCAGT CACAGAAAGC ATTGCACTCT TTCACAGCTG CCTCAACCCA 900  
ATCCCTTATG TTTTTATGGG AGCATCTTC AAAAATCAGG TTATGAAAGT GGCCTAAGAA 960  
30 TATGGGTCTT GGAGAAGACA GAGACAAAGT GTGGAGGGT TTCCCTTTGA TTCTGAGGGT1020  
CCTACAGAGC CAACCACTAC TTTTACATT TAA 1093

- 35 (23) INFORMACIÓN PARA LA SEC. NO ID: 22:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 350 aminoácidos  
40 (B) TIPO: aminoácido  
(C) TIPO DE HEBRA:  
(D) TOPOLOGÍA: no relevante

- 45 (ii) TIPO DE MOLÉCULA: proteína

- (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 22:

50 Met Ala Leu Glu Gln Asn Gln Ser Thr Asp Tyr Tyr Tyr Glu Glu Asn  
1 5 10 15

55 Glu Met Asn Gly Thr Tyr Asp Tyr Ser Gln Tyr Glu Leu Ile Cys Ile  
20 25 30

Lys Glu Asp Val Arg Glu Phe Ala Lys Val Phe Leu Pro Val Phe Leu  
35 40 45

60 Thr Ile Ala Phe Val Ile Gly Leu Ala Gly Asn Ser Met Val Val Ala

# ES 2 289 610 T3

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

50 (24) INFORMACIÓN PARA LA SEC. NO ID: 23:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 55 (A) LONGITUD: 1116 pares de base  
 (B) TIPO: ácido nucléico  
 (C) TIPO DE HEBRA: sencilla  
 (D) TOPOLOGÍA: lineal

- 60 (ii) TIPO DE MOLÉCULA: ADN (genomic)  
 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 23:

# ES 2 289 610 T3

ATGCCAGGAA ACGCCACCCC AGTGACCACC ACTGCCCGT GGGCCTCCCT GGGCCTCTCC 60  
GCCAAGACT GTCAACAAACGT GTCCTTCGAA GAGAGCAGGA TAGTCTGGT CGTGGTGTAC 120  
5 AGCGCGGTGT GCACGCTGCG GGTGCCGCC AACTGCCCTGA CTGCGTGGCT GGCCTGCTG 180  
CAGGTACTGC AGGGCAACGT GCTGCCGCC TACCTGCTCT GCCTGGCACT CTGCGAACTG 240  
CTGTACACAG GCACGCTGCC ACTCTGGTC ATCTATATCC GCAACCAACCA CCGCTGGAC 300  
10 CTAGGCCCTGC TGGCCTCGAA GGTGACCGCC TACATCTCTCT TCTGCAACAT CTACGTCAGC 360  
ATCCTCTTCC TGTGCTGCAT CTCCCTGGAC CGCTTCGCG 420  
AGTCGGGGCC GCGGCCGCC GAGGACCOCC ATCCTCATCT CGCCCTGAT CTTCATCCTC 480  
15 GTCGGGATCG TTCACTACCC GGTTTCCAG ACAGGAAGACA AGGAGACCTG CTTTGACATG 540  
CTGCGAGATGG ACAGCAAGAT TGCCGGGTAC TACTACGCCA GGTTGACCGT TGGCTTTGCC 600  
ATCCTCTCTCT CCATCATOGC CTTCACCAAC CACCGQATTI TCAGGAGCAT CAAGCAGAGC 660  
20 ATGGGCTTAA CGCGTGGCCCA GAAGGCCAAG GTGAAGCACT CGGCCATCGC GGTCGTTGTC 720  
ATCTTCCTAG TCTGCTTCGC CGCGTACCAAC CTGGTTCTCC TCGTCAAAGC CGCTGCTTT 780  
TCCTACTACA GAGGAGACAG GAACGOCATG TCCGGCTTGG AGGAAAGGCT GTACACAGCC 840  
25 TCTGTGCGTGT TTCTGTGCGT GTCCACGGTG AACCGCGTGG CTGACCCCAT TATCTACGTG 900  
CTGGCCACGG ACCATTCGG ACCAGGAAGTG TCCAGAAATCC ATAAGGGGTG GAAAGACTGG 960  
TCCATGAAGA CAGACGTCAC CAGGCTCACC CACAGCAAGG ACACCGAGGA GCTGCACTCG 1020  
30 CGCGTGGCCC TTGAGACCA CTACACCTTC TCCAGGCCCG TGCACCCACC AGGGTCACCA 1080  
TGCCCTGCMA AGAGGCTGAT TGAGGAGTCC TOCTGA 1116

35 (25) INFORMACIÓN PARA LA SEC. NO ID: 24:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 371 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

45 (ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 24:

50

55

60

65

# ES 2 289 610 T3

	Met Pro Gly Asn Ala Thr Pro Val Thr Thr Ala Pro Trp Ala Ser	
	1 5 10 15	
5	Leu Gly Leu Ser Ala Lys Thr Cys Asn Asn Val Ser Phe Glu Glu Ser	
	20 25 30	
	Arg Ile Val Leu Val Val Tyr Ser Ala Val Cys Thr Leu Gly Val	
	35 40 45	
10	Pro Ala Asn Cys Leu Thr Ala Trp Leu Ala Leu Leu Gln Val Leu Gln	
	50 55 60	
	Gly Asn Val Leu Ala Val Tyr Leu Leu Cys Leu Ala Leu Cys Glu Leu	
	65 70 75 80	
15	Leu Tyr Thr Gly Thr Leu Pro Leu Trp Val Ile Tyr Ile Arg Asn Gln	
	85 90 95	
	His Arg Trp Thr Leu Gly Leu Leu Ala Ser Lys Val Thr Ala Tyr Ile	
	100 105 110	
20	Phe Phe Cys Asn Ile Tyr Val Ser Ile Leu Phe Leu Cys Cys Ile Ser	
	115 120 125	
	Cys Asp Arg Phe Val Ala Val Val Tyr Ala Leu Glu Ser Arg Gly Arg	
	130 135 140	
25	Arg Arg Arg Arg Thr Ala Ile Leu Ile Ser Ala Cys Ile Phe Ile Leu	
	145 150 155 160	
	Val Gly Ile Val His Tyr Pro Val Phe Gln Thr Glu Asp Lys Glu Thr	
	165 170 175	
30	Cys Phe Asp Met Leu Gln Met Asp Ser Arg Ile Ala Gly Tyr Tyr Tyr	
	180 185 190	
	Ala Arg Phe Thr Val Gly Phe Ala Ile Pro Leu Ser Ile Ile Ala Phe	
	195 200 205	
35	Thr Asn His Arg Ile Phe Arg Ser Ile Lys Gln Ser Met Gly Leu Ser	
	210 215 220	
	Ala Ala Gln Lys Ala Lys Val Lys His Ser Ala Ile Ala Val Val Val	
	225 230 235 240	
40	Ile Phe Leu Val Cys Phe Ala Pro Tyr His Leu Val Leu Leu Val Lys	
	245 250 255	
45	Ala Ala Ala Phe Ser Tyr Tyr Arg Gly Asp Arg Asn Ala Met Cys Gly	
	260 265 270	
	Leu Glu Glu Arg Leu Tyr Thr Ala Ser Val Val Phe Leu Cys Leu Ser	
	275 280 285	
50	Thr Val Asn Gly Val Ala Asp Pro Ile Ile Tyr Val Leu Ala Thr Asp	
	290 295 300	
	His Ser Arg Gln Glu Val Ser Arg Ile His Lys Gly Trp Lys Glu Trp	
55	305 310 315 320	
	Ser Met Lys Thr Asp Val Thr Arg Leu Thr His Ser Arg Asp Thr Glu	
	325 330 335	
60	Glu Leu Gln Ser Pro Val Ala Leu Ala Asp His Tyr Thr Phe Ser Arg	
	340 345 350	
65	Pro Val His Pro Pro Gly Ser Pro Cys Pro Ala Lys Arg Leu Ile Glu	
	355 360 365	

# ES 2 289 610 T3

(26) INFORMACIÓN PARA LA SEC. NO ID: 25:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1113 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 25:

ATGGCGAACT ATAGCCATGC AGCTGACAAAC ATTTTGCAAAT ATCTCTGCC TCTAACAGCC 60  
TTTCTGAAC TGACTTCCTT CGGTTTCATA ATAGGACTCA CGGTGGTGGG CAAACCTCTG 120  
ATCTCCATTG TGCTAGTGAA AGATAAGACC TTGCTAGAG CACCTTACTA CTTCCTGTTG 180  
GATCTTTGGT GTTCAGATAT CCTCAGATCT GCAATTGGTT TCCCATTGTT GTTCAGCTCT 240  
GTCAAAAATG GCTCTACCTG GACTTATGGG ACTCTGACTT GCAAAGTGAT TGCCCTTCTG 300  
GGGGTTTTGT CCTGTTTCCA CACTGTTTC ATGCTTTCT GCATCAGTGT CACCAAGATAC 360  
TTAGCTATCG CCCATCACCG CTTCTATACA AAGAGGCTGA CCTTTTGGAC GTGTCTGGCT 420  
GTGATCTGTA TGGTGTGGAC TCTGCTGTG GCCATGGCAT TTOCCCCGGT TTTAGACCTG 480  
GGCACTTACT CATTCAATTAG GGAGGAAGAT CAATGCACCT TCCAAACACCG CTCCCTCAAG 540  
GCTAATGATT CCTTAGGATT TATGCTCTT CTTGCTCTCA TCCCTCTAGC CACACAGCTT 600  
GTCTACCTCA AGCTGATATT TTTCGTOCAC GATCGAAGAA AAATGAAGCC AGTCCAGTTT 660  
GTAGCAGCAAG TCAGCCAGAA CTGGACTTTT CATGGCTCTG GAQCCAGTGG CCAGGGAGCT 720  
GCCAATTGGC TAGCAGGATT TGGAAAGGGT CCCACACCCAC CCACCTTGCT GGGCATCAGG 780  
CAAAATGCA ACACCACAGG CAGAAGAAGG CTATTGGCT TAQACGAGTT CAAAATGGAG 840  
AAAGAATCA GCAGAAATGTT CTATATAATG ACTTTTCTGT TTCTAACCTT GTGGGGCCCC 900  
TACCTGGTGG CCTTGTATTG GAGAGTTTTT GCAAGAGGGC CTGTAGTACC AGGGGGATT 960  
CTAACAGCTG CTGCTCGAT GAGTTTGCC CAAGCAOGAA TCAATCCTTT TGTCTGCATT 1020  
TTCTCAAACA CGGAGCTGAG GCGCTGTTTC ACCACAAACCC TTCTTACTG CAGAAAATCC 1080  
AGGTTACCAA GGGAACTTA CTGTGTATA TGA 1113

(27) INFORMACIÓN PARA LA SEC. NO ID: 26:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 370 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 26:

# ES 2 289 610 T3

	Met Ala Asn Tyr Ser His Ala Ala Asp Asn Ile Leu Gln Asn Leu Ser			
1	5	10	15	
	Pro Leu Thr Ala Phe Leu Lys Leu Thr Ser Leu Gly Phe Ile Ile Gly			
5	20	25	30	
	Val Ser Val Val Gly Asn Leu Leu Ile Ser Ile Leu Leu Val Lys Asp			
10	35	40	45	
	Lys Thr Leu His Arg Ala Pro Tyr Tyr Phe Leu Leu Asp Leu Cys Cys			
15	50	55	60	
	Ser Asp Ile Leu Arg Ser Ala Ile Cys Phe Pro Phe Val Phe Asn Ser			
20	65	70	75	80
	Val Lys Asn Gly Ser Thr Trp Thr Tyr Gly Thr Leu Thr Cys Lys Val			
25	85	90	95	
	Ile Ala Phe Leu Gly Val Leu Ser Cys Phe His Thr Ala Phe Met Leu			
30	100	105	110	
	Phe Cys Ile Ser Val Thr Arg Tyr Leu Ala Ile Ala His His Arg Phe			
35	115	120	125	
	Tyr Thr Lys Arg Leu Thr Phe Trp Thr Cys Leu Ala Val Ile Cys Met			
40	130	135	140	
	Val Trp Thr Leu Ser Val Ala Met Ala Phe Pro Pro Val Leu Asp Val			
45	145	150	155	160
	Gly Thr Tyr Ser Phe Ile Arg Glu Glu Asp Gln Cys Thr Phe Gln His			
50	165	170	175	
	Arg Ser Phe Arg Ala Asn Asp Ser Leu Gly Phe Met Leu Leu Ala			
55	180	185	190	
	Leu Ile Leu Leu Ala Thr Gln Leu Val Tyr Leu Lys Leu Ile Phe Phe			
60	195	200	205	
	Val His Asp Arg Arg Lys Met Lys Pro Val Gln Phe Val Ala Ala Val			
65	210	215	220	
	Ser Gln Asn Trp Thr Phe His Gly Pro Gly Ala Ser Gly Gln Ala Ala			
70	225	230	235	240
	Ala Asn Trp Leu Ala Gly Phe Gly Arg Gly Pro Thr Pro Pro Thr Leu			
75	245	250	255	
	Leu Gly Ile Arg Gln Asn Ala Asn Thr Thr Gly Arg Arg Arg Leu Leu			
80	260	265	270	
	Val Leu Asp Glu Phe Lys Met Glu Lys Arg Ile Ser Arg Met Phe Tyr			
85	275	280	285	
	Ile Met Thr Phe Leu Phe Leu Thr Leu Trp Gly Pro Tyr Leu Val Ala			
90	290	295	300	
	Cys Tyr Trp Arg Val Phe Ala Arg Gly Pro Val Val Pro Gly Gly Phe			
95	305	310	315	320
	Leu Thr Ala Ala Val Trp Met Ser Phe Ala Gln Ala Gly Ile Asn Pro			
100	325	330	335	
	Phe Val Cys Ile Phe Ser Asn Arg Glu Leu Arg Arg Cys Phe Ser Thr			
105	340	345	350	
	Thr Leu Leu Tyr Cys Arg Lys Ser Arg Leu Pro Arg Glu Pro Tyr Cys			
110	355	360	365	
	Val Ile			
115	370			

# ES 2 289 610 T3

(28) INFORMACIÓN PARA LA SEC. NO ID: 27:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1080 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 27:

ATGCAGGTCC CBAACAGCAC CGGCCGGAC AACGGGACGC TGCAGATGCT CGGGAAACCGG 60  
GGGATCGCGG TGGCCCTOCC CGTGGTGTAC TCCTGGTGG CGGGGGTCAG CATCCCCGGC 120  
AACCTCTTCT CTCNGTGGGT GCTGTGCCGG CGCATGGGOC CGAGATCCCC GTGGTTCATC 180  
TTCATGATCA ACCTGAGGT CACGGACCTG ATGCTGGCA GCCTGTTGCC TTTCCAATC 240  
TACTACCATT GCAACGCCA CCACTGGGT ATOGGGGTGC TCTTTCAA CGTGGTGACC 300  
GTCGGCTTTT ACGCAAACAT GATTCAGC ATCTCAACCA TGACCTGTAT CAGCGTGGAG 360  
CGCTTCTCGG GGGTCTGTA CCCGCTCAGC TCCAAGCGT GGCGCCGCCG TCGTTACCG 420  
GTGGCCCGGT GTGCAGGGAC CTGGCTGCTG CTCTGACCG CCTGTTGCC GCTGGGGCGC 480  
ACCGATCTCA CCTAACCGGT GCAACGCCCTG GACATCACTA CCTGCTTGA CGTCTCAAG 540  
TGGACGATGC TCCCCAGGCT GACCATGTGG CCTGGTGTCC TCTTCACCAT CTTCATCTG 600  
CTGTTCTCA TCCCCTCTGT GATCACCGTG CCTGTTACA CGGCCACCAT CCTCAAGCTG 660  
TTGCGCACGG AGGAGGGCGCA CGGCCGGGAG CAGCGGAGGC CGGGGGTGG OCTGGCCGCG 720  
GTGGTCTTGC TGGCCTTGT CACCTCTTC GCCCCCCAAC AACTCTGTCT CCTGGCGCAC 780  
ATCGTGAGCC CCTGTTCTA CGGCAAGAQC TACTACCACG TGTACAAAGCT CACGCTGTGT 840  
CTCAGCTGCC TCAACAATG TCTGGACCCG TTGTTTATT ACTTTGGTC CGGGAAATT 900  
CAGCTGCC CGGGGGATA TTTGGGCTGC CGCCGGGTGC CCAAGAACAC CCTGGACACG 960  
CGGGGGGAGA GCTCTTCTC CGCCAGGACC ACGTCCGTGC GCTCOGAGGC CGGTGCGCAC 1020  
CCTGAAGGAGA TCGAGGGAGC CACCAAGGCC CGCCCTCCAGA GGCAGGAGAG TGTGTTCTGA 1080

(29) INFORMACIÓN PARA LA SEC. NO ID: 28:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 359 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

# ES 2 289 610 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 28:

```

Met Gln Val Pro Asn Ser Thr Gly Pro Asp Asn Ala Thr Leu Gln Met
1           5          10          15
5
Ieu Arg Asn Pro Ala Ile Ala Val Ala Leu Pro Val Val Tyr Ser Leu
20          25          30
10
Val Ala Ala Val Ser Ile Pro Gly Asn Leu Phe Ser Leu Trp Val Leu
35          40          45
15
Cys Arg Arg Met Gly Pro Arg Ser Pro Ser Val Ile Phe Met Ile Asn
50          55          60
Leu Ser Val Thr Asp Leu Met Leu Ala Ser Val Leu Pro Phe Gln Ile
65          70          75          80
15
Tyr Tyr His Cys Asn Arg His His Trp Val Phe Gly Val Leu Leu Cys
85          90          95
20
Asn Val Val Thr Val Ala Phe Tyr Ala Asn Met Tyr Ser Ser Ile Leu
100         105         110
20
Thr Met Thr Cys Ile Ser Val Glu Arg Phe Leu Gly Val Leu Tyr Pro
115         120         125
25
Leu Ser Ser Lys Arg Trp Arg Arg Arg Tyr Ala Val Ala Ala Cys
130         135         140
25
Ala Gly Thr Trp Leu Leu Leu Thr Ala Leu Cys Pro Leu Ala Arg
145         150         155         160
30
Thr Asp Leu Thr Tyr Pro Val His Ala Leu Gly Ile Ile Thr Cys Phe
165         170         175
30
Asp Val Leu Lys Trp Thr Met Leu Pro Ser Val Ala Met Trp Ala Val
180         185         190
35
Phe Leu Phe Thr Ile Phe Ile Leu Leu Phe Leu Ile Pro Phe Val Ile
195         200         205
35
Thr Val Ala Cys Tyr Thr Ala Thr Ile Leu Lys Leu Leu Arg Thr Glu
210         215         220
40
Glu Ala His Gly Arg Glu Gln Arg Arg Ala Val Gly Leu Ala Ala
225         230         235         240
40
Val Val Leu Leu Ala Phe Val Thr Cys Phe Ala Pro Asn Asn Phe Val
245         250         255
45
Leu Leu Ala His Ile Val Ser Arg Leu Phe Tyr Gly Lys Ser Tyr Tyr
260         265         270
45
His Val Tyr Lys Leu Thr Leu Cys Leu Ser Cys Leu Asn Asn Cys Leu
275         280         285
50
Asp Pro Phe Val Tyr Tyr Phe Ala Ser Arg Glu Phe Gln Leu Arg Leu
290         295         300
50
Arg Glu Tyr Leu Gly Cys Arg Arg Val Pro Arg Asp Thr Leu Asp Thr
305         310         315         320
55
Arg Arg Glu Ser Leu Phe Ser Ala Arg Thr Thr Ser Val Arg Ser Glu
325         330         335
55
Ala Gly Ala His Pro Glu Gly Met Glu Gly Ala Thr Arg Pro Gly Leu
340         345         350

```

60 (30) INFORMACIÓN PARA LA SEC. NO ID: 29:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1503 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla

# ES 2 289 610 T3

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 29:

ATGGAGCGTC CCTGGGAGGA CAGCCCAGGC CCAGGAGGGGG CAGCTGAGGG CTGGCTGTG 60  
CCAGTCGCCG CCAGGAGGGGG CTGGGGTGCC GCGGGAGTG GCACAGGGTG GCAGCCATGG 120  
OCTGAGTGCC CGGGACCAA GGGGAGGGGG CAACTGCTGG CGACGGCCGG CCCTTTCGGT 180  
CGCTGGCCG CCCCCCTGCC TGCCAGCTCC AGCCCCGCCCC CGGGAGGGC GTCCGCTCAC 240  
TCGGTTCAAG CGAGCGCGAC TGCGGGTACG GCAGGACAG GCGCGAGACC TTGGGGCGCG 300  
CGGCCCATGG AGTGGGGCT GCTGGGGCCG GCGGGGTGA GCGAGGTCAAT CGTCCCTGCAT 360  
TACAATACA CGCGCAAGCT CGGGGGTGCC AGCTACCAGC CGGGTGCCCG CCTGGCGGCC 420  
GACCCCGTGG TGTGCGTGGC GGTGTGCGCC TTCACTGTGC TAGAGAACTT AGCCGTGTTG 480  
TTGGTGTCTCG GACGCCACCC CGCGCTTCCAC GCTCCCATGT TCTGCTCTG GGGCAGGCTC 540  
ACGTGGTGGG ATCTGGTGGC AGGCAGCGCC TACGCCGCCA ACATCTACT GTGGGGGCCG 600  
CTCACGCTGA AACTGTCCCC CGCGCTCTGG TTOGACAGGG AGGGAGGCCT CTTCGTGGCA 660  
CTCACTGGGT CGGTGCTGAG CCTCTGGCC ATCGCGCTGG AGCGCAGCCT CACCATGGCG 720  
CGCAGGGGGC CGCGCCCCGT CTCCAGTGGG GGGCGCACAC TGGCGATGGC AGCCGGCGCC 780  
TGGGGCGTGT CGGTGCTCTG CGGGCTCTG CCAGGGCTGG GCTGGATTG CCTGGTGTGG 840  
CTGGACGCTT GCTCCACTGT CTGGCCGCTC TACGCCAAGG CCTACGCTCT CTTCCTGCGT 900  
CTCGCCCTCG TGGGCATCCP GGCCCGGATC TGTOCACTCT ACGCGCOCAT CTACTGCCAG 960  
GTACCGGCCA ACACCGGGCG CCTGCCGCCA CGGGCCCCGA CTGGGGGGAC CACCTCGACC1020  
CGGGCGCGTC GCAAGCGCG CGCTCTGGCC TTGCTGCGCA CGCTCAQCGT GGTGCTCTG1080  
GCCTTGTGG CATGTTGGGG CGCCCTCTTC CTGCTGCTGT TGCTGAGGT GGCGTGGCCG1140  
GGCGCGACCT GTCTGTACT CCTGCAGGCC GATCCCTTCC TGGGACTGGC CATGGCCAAC1200  
TCACCTCTGA ACCCCATCAT CTACACGCTC ACCAACCGCG ACCTGCGGCCA CGCGCTCTGG1260  
CGCCTGGTCT GCTCGGGACG CCACCTCTGC GGCAGAGACC CGAGTGGCTC CGAGCAAGTCG1320  
CGGAAGGGCGG CTGAGGCTTC CGGGGGCTG CGGGCGCTGCC TCCCCCGGG CCTTGATGGG1380  
AGCTTCAGCG GCTCGGGACG CTCATCGCCC CAGCGCGACG GGCTGGACAC CAGCGGCTCC1440  
ACAGGCAGCC CGGGTGACCC CACAGCCGCC CGGACTCTGG TATCAGAACG GGCTGCAAGAC1500  
TGA 1503

(31) INFORMACIÓN PARA LA SEC. NO ID: 30:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 500 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE HEBRA:

(D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

# ES 2 289 610 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 30:

	Met Glu Arg Pro Trp Glu Asp Ser Pro Gly Pro Glu Gly Ala Ala Glu			
1	5	10		
5	Gly Ser Pro Val Pro Val Ala Ala Gly Ala Arg Ser Gly Ala Ala Ala			
	20	25	30	
10	Ser Gly Thr Gly Trp Gln Pro Trp Ala Glu Cys Pro Gly Pro Lys Gly			
	35	40	45	
15	Arg Gly Gln Ieu Leu Ala Thr Ala Gly Pro Leu Arg Arg Trp Pro Ala			
	50	55	60	
20	Pro Ser Pro Ala Ser Ser Ser Pro Ala Pro Gly Ala Ala Ser Ala His			
	65	70	75	80
25	Ser Val Gln Gly Ser Ala Thr Ala Gly Gly Ala Arg Pro Gly Arg Arg			
	85	90	95	
30	Pro Trp Gly Ala Arg Pro Met Glu Ser Gly Leu Leu Arg Pro Ala Pro			
	100	105	110	
35	Val Ser Glu Val Ile Val Leu His Tyr Asn Tyr Thr Gly Lys Leu Arg			
	115	120	125	
40	Gly Ala Ser Tyr Gln Pro Gly Ala Gly Leu Arg Ala Asp Ala Val Val			
	130	135	140	
45	Cys Leu Ala Val Cys Ala Phe Ile Val Leu Glu Asn Leu Ala Val Leu			
	145	150	155	160
50	Leu Val Leu Gly Arg His Pro Arg Phe His Ala Pro Met Phe Leu Leu			
	165	170	175	
55	Leu Gly Ser Leu Thr Leu Ser Asp Leu Leu Ala Gly Ala Ala Tyr Ala			
	180	185	190	
60	Ala Asn Ile Leu Leu Ser Gly Pro Leu Thr Leu Lys Leu Ser Pro Ala			
	195	200	205	
65	Leu Trp Phe Ala Arg Glu Gly Gly Val Phe Val Ala Leu Thr Ala Ser			
	210	215	220	
70	Val Leu Ser Leu Leu Ala Ile Ala Leu Glu Arg Ser Leu Thr Met Ala			
	225	230	235	240
75	Arg Arg Gly Pro Ala Pro Val Ser Ser Arg Gly Arg Thr Leu Ala Met			
	245	250	255	
80	Ala Ala Ala Ala Trp Gly Val Ser Leu Leu Leu Gly Leu Leu Pro Ala			
	260	265	270	
85	Leu Gly Trp Asn Cys Leu Gly Arg Leu Asp Ala Cys Ser Thr Val Leu			
	275	280	285	
90	Pro Leu Tyr Ala Lys Ala Tyr Val Leu Phe Cys Val Leu Ala Phe Val			
	290	295	300	
95	Gly Ile Leu Ala Ala Ile Cys Ala Leu Tyr Ala Arg Ile Tyr Cys Gln			
	305	310	315	320
100	Val Arg Ala Asn Ala Arg Arg Leu Pro Ala Arg Pro Gly Thr Ala Gly			
	325	330	335	
105	Thr Thr Ser Thr Arg Ala Arg Arg Lys Pro Arg Ser Leu Ala Leu Leu			
	340	345	350	
110	Arg Thr Leu Ser Val Val Leu Leu Ala Phe Val Ala Cys Trp Gly Pro			
	355	360	365	
115	Leu Phe Leu Leu Leu Leu Asp Val Ala Cys Pro Ala Arg Thr Cys			

# ES 2 289 610 T3

370                    375                    380

5 Pro Val Leu Leu Gln Ala Asp Pro Phe Leu Gly Leu Ala Met Ala Asn  
385                    390                    395                    400

Ser Leu Leu Asn Pro Ile Ile Tyr Thr Leu Thr Asn Arg Asp Leu Arg  
405                    410                    415

10 His Ala Leu Leu Arg Leu Val Cys Cys Gly Arg His Ser Cys Gly Arg  
420                    425                    430

Asp Pro Ser Gly Ser Gln Gln Ser Ala Ser Ala Ala Glu Ala Ser Gly  
435                    440                    445

15 Gly Leu Arg Arg Cys Leu Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly  
450                    455                    460

Ser Glu Arg Ser Ser Pro Gln Arg Asp Gly Leu Asp Thr Ser Gly Ser  
465                    470                    475                    480

20 Thr Gly Ser Pro Gly Ala Pro Thr Ala Ala Arg Thr Leu Val Ser Glu  
485                    490                    495

Pro Ala Ala Asp  
25                    500

(32) INFORMACIÓN PARA LA SEC. NO ID: 31:

30 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1029 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

35 (ii) TIPO DE MOLÉCULA: ADN (genomic)

40 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 31:

ATGCAAGCCG TCGACAATCT CACCTCTCG CCTGGGAACA CCAGTCGTG CACCAAGAGAC 60  
TACAAATCA CCCAOGCTCT CTTCCTCACTG CTCTACACTG TCCCTGTTTTT TGTTGGACTT 120  
ATCACAAATG GCCTGGCGAT GAGGATTTTC TTTCAAATCC GGAGTAATTC AAACTTATT 180  
ATTTTCTTA AGAACACAGT CATTCTGAT CTTCATCA TTCTGACTTT TCCATTCAA 240  
50 ATTCTTAGTG ATGCCAAACT GGGAACAGGA CCACTGAGAA CTTTTGTTG TCAAGMTACC 300  
TCCGTATAT TTATTTAC AATGTATATC AGTATTTCAT TCCTGGACT GATAACTATC 360  
55 GATCGCTACC AGAAGACCAAC CAGGCCATTT AAAACATCCA ACCCCCCAAA TCTCTTGGG 420  
GCTAAGATTC TCTCTGTTGT CATCTGGCA TTCATGTTCT TACTCTTTT GCCTAACATG 480

60

65

ES 2 289 610 T3

ATTCTGACCA ACAGGCAGCC GAGAGACAAG AATGTGAAGA AATGCTCTT CCTAAATCA 540  
 GAGTTGGTC TAGTCTGGCA TGAAATAGTA AAATTACATCT GTCAAGTCAT TTCTGGATT 600  
 5 AATTCTTAA TIGTTATTGT ATGTTATACA CTCATTACAA AAGAAGTGTA CGGGTCATAC 660  
 GTAAGAACGA GGGGTGAGG TAAAGTCCCC AGGAAAAAGG TGAACGTCAA AGTTTCATT 720  
 10 ATCATTOCTG TATTCTTAT TTGTTTGTT CCTTTCCATT TTGCCCCAT TCCCTACACC 780  
 CTGAGCCAAA CCCGGGATGT CTTTGACTGC ACTGCTGAAA ATACTCTGTT CTATGTGAAA 840  
 15 GAGAGCACTC TGCGGTTAAC TTCCCTAAAT GCATGCCGG ATCCGTTCAT CTATTTTTC 900  
 CTTTGCAAGT CCTTCAGAAA TTCCCTGATA AGTATGCTGA AGTGCCCAA TTCTOCARCA 960  
 20 TCTCTGTCCC AGGACAATAG GAAAAAGAA CAGGATGGTG GTGACCCAAA TGAAGAGACT1020  
 CCAATGTAA 1029

25 (33) INFORMACIÓN PARA LA SEC. NO ID: 32:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 342 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

35 (ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 32:

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

ES 2 289 610 T3

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

(34) INFORMACIÓN PARA LA SEC. NO ID: 33:

45 (j) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1077 pares de base
  - (B) TIPO: ácido nucléico
  - (C) TIPO DE HEBRA: sencilla
  - (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

55 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 33:

60

65

ES 2 289 610 T3

ATGTCGGTCT GCTACCGTCC CCCAGGGGAC GAGACACTGC TGAGCTGGAA GACTTCGGGG 60  
GCCACAGGCA CAGCCTTCCT GCTGCTGGCG GCGCTGCTGG CGCTGGCTGG CAACGGGTTC 120  
5 GTGGTGTGGA GCTTGGCGGG CTGGCGGGCT GCACGGGGGC GACCGCTGGC GGCCACGTT 180  
GTGCTACACC TGGCAGCTGGC CGACGGCGCG GTGCTGCTGC TCACGGCGCT CTTTGTTGCC 240  
10 TTCCCTGACCC GGCAGGGCTG GCGCGTGGGC CAGGCAGGCT GCAAGGGCGT GTACTACGTG 300  
TCCGCGCTCA GCATGTACGC CAGCGTGTG CTCACGGGCC TGTCTGGCT GCAGCGCTGC 360  
15 CTCGCACTCA CCCGGCCCCP CCTGGCGGCCT CGGCTGCGCA GCGCGACCCG GCGCCGGCGC 420  
CTACTGCTGG CGTCTGGCT GGCAGGGCTG TTGCTGGCG TCCCGGCCGC CATCTACCGC 480  
CACCTGTTGA CGCACCGCGT ATGCCAGCTG TGCCACCCGT CGCCGATCCA CGCCGGCGCC 540  
20 CACCTGAGCC TGGAGACTCT GACCGCTTTC GTGCTTCCTT TCGGGACTGAT GCTCGGGCTGC 600  
TACACCGTGA CGCTGOCACG GCTGGGGGGC GCGCGCTGGG GCTCCGGCGC GCACGGGGCG 660  
CGGGTGGGCC GGCTGGTGAG CGCCATCGTG CTGCGCTTG GCTTGCTCTG GCGCCCGTAC 720  
25 CACCGAGTCA ACCTTCTGCA GGCAGTCGCA GCGCTGGCTC CACCGGAAGG GCGCTGGCG 780  
AAGCTGGCGG QACCCGGCCA GCGCGCGCGA GCGGAACTA CGGCCTGGC CTTCTTCAGT 840  
TCTAGCGTCA ACCCGGTGCT CTACGTCCTC ACGGCTGGAG ATCTGCTGCC CGGGCAGGT 900  
30 CCGCGTTTCC TCACCGCGCT CTTCGAAGGC TCTGGGGAGG CCCGAAGGGGG CGGGCGCTCT 960  
AGGGAAAGGGG CCATGGAGCT CGGAACCTACC CCTCAGCTGA AAGTGGTGGG GCAAGGGCCGC1020  
GGCAATGGAG ACCCGGGGGG TGCGATGGAG AAGGACGGTC CGGAATGGGA CCTTTGA 1077  
35

(35) INFORMACIÓN PARA LA SEC. NO ID: 34:

40 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 358 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

50 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 34:

Met Ser Val Cys Tyr Arg Pro Pro Gly Asn Glu Thr Leu Leu Ser Trp  
1 5 10 15

55 Lys Thr Ser Arg Ala Thr Gly Thr Ala Phe Leu Leu Ala Ala Leu

60

65

# ES 2 289 610 T3

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

(36) INFORMACIÓN PARA LA SEC. NO ID: 35:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 65 (A) LONGITUD: 1005 pares de base  
 (B) TIPO: ácido nucléico

# ES 2 289 610 T3

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

5 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 35:

10 ATGCTGGGA TCATGGCATG GAATGCAGT TGCAAAACT GGCTGGCAGC AGAGGCTGCC 60  
CTGGAAAAGT ACTACCTTC CATTCTTAT GGGATTGAGT TCGTTGTGGG AGTCCTTGGA 120  
AATACCAATTG TTGTTTACGG CTACATCTTC TCTCTGAAGA ACTGGAACAG CAATAATATT 180  
15 TATCTCTTAA ACCTCTCTGT CTCTGACTTA CCTTTCTGT GCACCCCTCCC CATGCTGATA 240  
AGGAGTTATG CCAATGGAAA CTGGATATAT GGAGACGTCG TCTGCTAAAG CAACGGATAT 300  
20 GTGCTTCATG CCAACCTCTA TACCAAGCATT CTCTTTCTCA CTTTTATCAG CATAAGATCGA 360  
TACTTGATAA TTAAGTATCC TTCCGGAGAA CACCTCTCTOC AAAAGAAAGA GTTTOCTATT 420  
TTAATCTCCT TGGCCATTTC GGTTTTAATA ACCTTAGAGT TACTACCCAT ACTTCCCCCTT 480  
25 ATAAATCCTG TTATAACTGA CAATGGCACC ACCTGTAATG ATTTTGCAAG TTCTGGAGAC 540  
CCCAACTACA ACCTCATTTA CACCATGAGT CTAACACTGT TGGGTTCCCT TATTCCCTTT 600  
TTTGTGATGT GTTCTTTTA TTACAAGATT GCTCTCTCC TAAAGCAGAG GAATAGGCAG 660  
30 GTTGCTACTG CTCTGCCCCCT TGAAAAGCCT CTCAACTTGG TCATCATGCC AGTGGTAATC 720  
TTCTCTGTGC TTTTACACC CTATCACGTC ATCGGAATG TGAGGATGCC TTCACCCCTG 780  
GGGACTTGGG ACCAGTATCA GTGCACTCAG GTCGTCACTA ACTCCTTTA CATTGTGACA 840  
35 CGGCTTGGG CCTTTCTGAA CAGTGTCACTA AACCCGTCT TCTATTTCT TTTGGAGAT 900  
CACTTCAGGG ACATGCTGAT GAATCAACTG AGACACAACT TCAAAATCCCT TACATCCTTT 960  
40 ACCAGATGGG CTCATGAACT CCTACTTTCA TTCAAGAGAA AGTGA 1005

(37) INFORMACIÓN PARA LA SEC. NO ID: 36:

45 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 334 aminoácidos

(B) TIPO: aminoácido

(C) TIPO DE HEBRA:

50 (D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

55 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 36:

60

65

# ES 2 289 610 T3

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

# ES 2 289 610 T3

(38) INFORMACIÓN PARA LA SEC. NO ID: 37:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 5 (A) LONGITUD: 1296 pares de base  
(B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
10 (D) TOPOLOGÍA: lineal

10 (ii) TIPO DE MOLÉCULA: ADN (genomic)

15 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 37:

ATGCAGGGCG ITAACATTAC CCCGGACGAG TTCTCTCGGC TGCTGGGGGA CCACAACCTG 60  
ACGCCGGAGC AGTTCATCGC TCTGTACCGG CTGCGACCGC TCGTCTACAC CCCAGAGCTG 120  
CCGGGACCGG CCAAACCTGGC CCTCTGTCTC ACCGGGCTGC TCATCTCGC CCTGGCGCTC 180  
TTTGGCAATG CTCTGGTGTG CTACGGTGTG ACCGGCAGCA AGGCCATGCG CACCOTCACC 240  
AACATCTTA TCTGCTCTT GGCCTCTAGT GACCTGCTCA TCACCTCTT CTGCATTCCC 300  
GTCACCATGC TCCAGAACAT TTCCGACAAAC TGGCTGGGG GTGCTTTCAT TTGCAAGATG 360  
GTGCCATTG TCGAGTCTAC CGCTGTGTG ACAGAAATGC TCACTATGAC CTGCATTGCT 420  
GTGAAAGGC ACCAAGGGACT TGTGCATCTT TTAAATGA ACTGGCAATA CACCAACCGA 480  
AGGGCTTCA CAATGCTAGG TGTGGCTGG CTGCTGGCAG TCATGGTAGG ATCACCCATG 540  
TGGCACOTGC AACAACTTGA GATCAAATAT GACTTCTAT ATGAAAAGGA ACACATCTGC 600  
TGCTTAGAAG AGTGGACCAAG CCCTGTGAC CAGAAGATCT ACACCACTT CATCTTGTC 660  
ATCCTCTTCC TCCCTCCCTCT TATGGTGTG CTTATTCGT ACAGTAAAAT TGGTTATGAA 720  
CTTGATGAA AGAAAAGAGT TGGGGATGGT TCACTGCTTC GAACTATTCA TGGAAAAGAA 780  
ATGTCACAAA TAGCCAGGAA GAAGAAACGA GCTGTCATTA TGATGGTGAC AGTGGTGGCT 840  
CTCTTGTG TGCTGGGGC ACCATTCCAT GTGTCATTA TGATGGTGA ATACAGTAAT 900  
TTTGAAAAGG AATATGATGA TGTACAAATC AAGATGATTG TTGCTATGTT GCAAATTATT 960  
GGATTTCCA ACTCCATCTG TAATCCCATT GTCTATGCA TTATGAATGA AAACCTCAAA 1020  
AAAAATGTTT TGTCTGCAGT TTGTTATTGC ATAGTAAATA AAACCTCTC TCCAGCACAA 1080  
AGGCATGGAA ATTCAAGGAAT TACAATGATG CGGAAGAAAG CAAAGTTTC CCTCAGAGAG 1140  
AATCCAGTGG AGGAAACCAA AGGAGAAGCA TTCACTGATG GCAACATTGA AGTCAAATTG 1200  
TGTGAACAGA CAGAGGAGAA QAAAAGCTC AAACGACATC TTGCTCTCTT TAGGCTGAA 1260  
CTGGCTGAGA ATTCTCTTT AGACAGTGGG CATTAA

1296

# ES 2 289 610 T3

(39) INFORMACIÓN PARA LA SEC. NO ID: 38:

5 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 431 aminoácidos
- (B) TIPO: aminoácido
- (C) TIPO DE HEBRA:
- (D) TOPOLOGÍA: no relevante

10 (ii) TIPO DE MOLÉCULA: proteína

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 38:

15 Met Gln Ala Leu Asn Ile Thr Pro Glu Gln Phe Ser Arg Leu Leu Arg  
1 5 10 15

20 Asp His Asn Leu Thr Arg Glu Gln Phe Ile Ala Leu Tyr Arg Leu Arg  
20 25 30

25 Pro Leu Val Tyr Thr Pro Glu Leu Pro Gly Arg Ala Lys Leu Ala Leu  
35 40 45

Val Leu Thr Gly Val Leu Ile Phe Ala Leu Ala Leu Phe Gly Asn Ala  
50 55 60

30 Leu Val Phe Tyr Val Val Thr Arg Ser Lys Ala Met Arg Thr Val Thr  
65 70 75 80

Asn Ile Phe Ile Cys Ser Leu Ala Leu Ser Asp Leu Leu Ile Thr Phe  
85 90 95

35 Phe Cys Ile Pro Val Thr Met Leu Gln Asn Ile Ser Asp Asn Trp Leu

40

45

50

55

60

65

ES 2 289 610 T3

(40) INFORMACIÓN PARA LA SEC. NO ID: 39:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24 pares de base

(B) TIPO: ácido nucléico

ES 2 289 610 T3

	(C) TIPO DE HEBRA: sencilla (D) TOPOLOGÍA: lineal	
5	(ii) TIPO DE MOLÉCULA: ADN (genomic) (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 39:	
	CTGTGTACAG CAGTCGCAG AGTG	24
10	(41) INFORMACIÓN PARA LA SEC. NO ID: 40:  (i) CARACTERÍSTICAS DE LA SECUENCIA:  (A) LONGITUD: 24 pares de base (B) TIPO: ácido nucléico (C) TIPO DE HEBRA: sencilla (D) TOPOLOGÍA: lineal	
15	(ii) TIPO DE MOLÉCULA: ADN (genomic) (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 40:	
20	GAGTGCCAGG CAGAGCAGGT AGAC	24
25	(42) INFORMACIÓN PARA LA SEC. NO ID: 41:  (i) CARACTERÍSTICAS DE LA SECUENCIA:  (A) LONGITUD: 31 pares de base (B) TIPO: ácido nucléico (C) TIPO DE HEBRA: sencilla (D) TOPOLOGÍA: lineal	
30	(ii) TIPO DE MOLÉCULA: ADN (genomic) (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 41:	
35	CCCGAATTCC TGCTTGCTCC CAGCTTGGCC C	31
40	(43) INFORMACIÓN PARA LA SEC. NO ID: 42:  (i) CARACTERÍSTICAS DE LA SECUENCIA:  (A) LONGITUD: 32 pares de base (B) TIPO: ácido nucléico (C) TIPO DE HEBRA: sencilla (D) TOPOLOGÍA: lineal	
45	(ii) TIPO DE MOLÉCULA: ADN (genomic) (iv) ANTISENTIDO: SI (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 42:	
50	TGTGGATCCT GCTGTCAAAG GTCCCATTCC GG	32
55	(44) INFORMACIÓN PARA LA SEC. NO ID: 43:  (i) CARACTERÍSTICAS DE LA SECUENCIA:  (A) LONGITUD: 20 pares de base (B) TIPO: ácido nucléico	
60		
65		

# ES 2 289 610 T3

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

5 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: NO

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 43:

10 TCACAATGCT AGGTGTGGTC

20

(45) INFORMACIÓN PARA LA SEC. NO ID: 44:

15 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 22 pares de base

(B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

20 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

25 (iv) ANTISENTIDO: SI

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 44:

30 TGCATAGACA ATGGGATTAC AG

22

(46) INFORMACIÓN PARA LA SEC. NO ID: 45:

35 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 511 pares de base

(B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

40 (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 45:

45 TCACAATGCT AGGTGTGGTC TGGCTGGTGG CAGTCATCGT AGGATCACCC ATGTOGCACG 60

TGCAACAACT TGAGATCAA TATGACTTCC TATATGAAAA GGAACACATC TGCTGCTTAG 120

AAGAGTGGAC CAGCCCTGTG CACCAGAAGA TCTACACCAAC CTTCATCCTT GTCATCTCT 180

TCCTCTGCC TCTTATGGTG ATGCTTATTC TGTACGTAAR ATTGGTTATG AACTTTGGAT 240

AAAGAAAAGA GTTGGGGATG GTTCAOTGCT TCGAACTATT CATGGAAAAG AAATGTCCAA 300

AAATAGCCAGG AAGAAGAAC GAGCTGTCAAT TATGATGGTG ACAGTGGTGG CTCTCTTGC 360

TGTGTGCTGG GCACCATTCCT ATGTTGTCCA TATGATGATT GAATACAGTA ATTGTTGAAAAA 420

GGAATATGAT GATGTCACAA TCAAGATGAT TTTTGCTATC GTGCAAATTA TTGGATTTTC 480

60 CAACCTCCATC TGTAAATCCCC TTGTCTATGC A 511

(47) INFORMACIÓN PARA LA SEC. NO ID: 46:

65 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 21 pares de base

ES 2 289 610 T3

- (B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal
- 5 (ii) TIPO DE MOLÉCULA: ADN (genomic)  
(iv) ANTISENTIDO: SI  
10 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 46:  
  
CTGCTTAGAA GAGTGGACCA G 21
- (48) INFORMACIÓN PARA LA SEC. NO ID: 47:  
15 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 22 pares de base  
20 (B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal  
25 (ii) TIPO DE MOLÉCULA: ADN (genomic)  
(iv) ANTISENTIDO: NO  
30 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 47:  
  
CTGTGCACCA GAAGATCTAC AC 22
- (49) INFORMACIÓN PARA LA SEC. NO ID: 48:  
35 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 21 pares de base  
(B) TIPO: ácido nucléico  
40 (C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal  
45 (ii) TIPO DE MOLÉCULA: ADN (genomic)  
(iv) ANTISENTIDO: SI  
50 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 48:  
  
CAAGGATGAA GGTGGTGTAG A 21
- (50) INFORMACIÓN PARA LA SEC. NO ID: 49:  
55 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 23 pares de base  
(B) TIPO: ácido nucléico  
60 (C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal  
65 (ii) TIPO DE MOLÉCULA: ADN (genomic)  
(iv) ANTISENTIDO: SI  
(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 49:  
  
GTGTAGATCT TCTGGTGCAC AGG 23

ES 2 289 610 T3

(51) INFORMACIÓN PARA LA SEC. NO ID: 50:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
5 (A) LONGITUD: 21 pares de base  
(B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal  
10 (ii) TIPO DE MOLÉCULA: ADN (genomic)  
(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 50:

15 GCAATGCAGG TCATAGTGAG C

21

(52) INFORMACIÓN PARA LA SEC. NO ID: 51:

- 20 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 27 pares de base  
(B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
25 (D) TOPOLOGÍA: lineal  
(ii) TIPO DE MOLÉCULA: ADN (genomic)  
(iii) HIPOTÉTICO: SI  
30 (iv) ANTISENTIDO: SI  
(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 51:

35 TGGAGCATGG TGACGGAAAT GCAGAAG

27

(53) INFORMACIÓN PARA LA SEC. NO ID: 52:

- 40 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 27 pares de base  
(B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
45 (D) TOPOLOGÍA: lineal  
(ii) TIPO DE MOLÉCULA: ADN (genomic)  
(iv) ANTISENTIDO: SI  
50 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 52:

GTGATGAGCA GGTCACTGAG CGCCAAG

27

55 (54) INFORMACIÓN PARA LA SEC. NO ID: 53:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
60 (A) LONGITUD: 23 pares de base  
(B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal  
65 (ii) TIPO DE MOLÉCULA: ADN (genomic)  
(iv) ANTISENTIDO: NO

ES 2 289 610 T3

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 53:

GCAATGCAGG CGCTTAACAT TAC

23

5 (55) INFORMACIÓN PARA LA SEC. NO ID: 54:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 10 (A) LONGITUD: 22 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

15 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: SI

20 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 54:

TTGGGTTACA ATCTGAAGGG CA

22

(56) INFORMACIÓN PARA LA SEC. NO ID: 55:

25 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 23 pares de base
- (B) TIPO: ácido nucléico
- 30 (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

35 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: NO

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 55:

40 ACTCCGTGTC CAGCAGGACT CTG

23

(57) INFORMACIÓN PARA LA SEC. NO ID: 56:

45 (i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 24 pares de base
- (B) TIPO: ácido nucléico
- 50 (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: SI

55 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 56:

TGCGTGTTC TGGACCCTCA CGTG

24

60 (58) INFORMACIÓN PARA LA SEC. NO ID: 57:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- 65 (A) LONGITUD: 29 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla

ES 2 289 610 T3

(D) TOPOLOGÍA: lineal

5 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: NO

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 57:

CAGGCCTTGG ATTTAATGT CAGGGATGG

29

10 (59) INFORMACIÓN PARA LA SEC. NO ID: 58:

15 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 pares de base

(B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

20 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: SI

25 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 58:

GGAGAGTCAG CTCTGAAAGA ATTCAAG

27

30 (60) INFORMACIÓN PARA LA SEC. NO ID: 59:

35 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 pares de base

(B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

40 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: NO

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 59:

45 TGATGTGATG CCAGATACTA ATAGCAC

27

(61) INFORMACIÓN PARA LA SEC. NO ID: 60:

50 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 pares de base

(B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

55 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: SI

60 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 60:

CCTGATTCTT TAGGTGAGA TTGAGAC

27

65 (62) INFORMACIÓN PARA LA SEC. NO ID: 61:

ES 2 289 610 T3

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 22 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: NO

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 61:

15 GACAGGTACC TTGCCATCAA G

21

(63) INFORMACIÓN PARA LA SEC. NO ID: 62:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 22 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: SI

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 62:

CTGCACAATG CCAGTGATAA GG

22

35 (64) INFORMACIÓN PARA LA SEC. NO ID: 63:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 27 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: NO

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 63:

50 CTGACTTCTT GTTCCTGGCA GCAGCGG

27

(65) INFORMACIÓN PARA LA SEC. NO ID: 64:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 27 pares de base
- (B) TIPO: ácido nucléico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: SI

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 64:

ES 2 289 610 T3

AGACCAGCCA GGGCACGCTG AAGAGTG

27

(66) INFORMACIÓN PARA LA SEC. NO ID: 65:

- 5 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 32 pares de base  
(B) TIPO: ácido nucléico  
10 (C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal  
(ii) TIPO DE MOLÉCULA: ADN (genomic)  
15 (iv) ANTISENTIDO: NO  
(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 65:

20 GATCAAGCTT CCATCCTACT GAAACCATGG TC

32

(67) INFORMACIÓN PARA LA SEC. NO ID: 66:

- 25 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 35 pares de base  
(B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
30 (D) TOPOLOGÍA: lineal  
(ii) TIPO DE MOLÉCULA: ADN (genomic)  
(iv) ANTISENTIDO: SI  
35 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 66:

GATCAGATCT CAGTTCCAAT ATTCACACCA CCGTC

35

40 (68) INFORMACIÓN PARA LA SEC. NO ID: 67:

- (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 22 pares de base  
45 (B) TIPO: ácido nucléico  
(C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal  
(ii) TIPO DE MOLÉCULA: ADN (genomic)  
(iv) ANTISENTIDO: NO  
50 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 67:

55 CTGGTGTGCT CCATGGCATC CC

22

(69) INFORMACIÓN PARA LA SEC. NO ID: 68:

- 60 (i) CARACTERÍSTICAS DE LA SECUENCIA:  
(A) LONGITUD: 22 pares de base  
(B) TIPO: ácido nucléico  
65 (C) TIPO DE HEBRA: sencilla  
(D) TOPOLOGÍA: lineal

ES 2 289 610 T3

	(ii) TIPO DE MOLÉCULA: ADN (genomic)	
5	(iv) ANTISENTIDO: SI	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 68:	
	GTAAGCCTCC CAGAACGAGA GG	22
10	(70) INFORMACIÓN PARA LA SEC. NO ID: 69:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
15	(A) LONGITUD: 24 pares de base	
	(B) TIPO: ácido nucléico	
20	(C) TIPO DE HEBRA: sencilla	
	(D) TOPOLOGÍA: lineal	
	(ii) TIPO DE MOLÉCULA: ADN (genomic)	
25	(iv) ANTISENTIDO: NO	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 69:	
	CAGCGCAGGG TGAAGCCTGA GAGC	24
30	(71) INFORMACIÓN PARA LA SEC. NO ID: 70:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
35	(A) LONGITUD: 24 pares de base	
	(B) TIPO: ácido nucléico	
40	(C) TIPO DE HEBRA: sencilla	
	(D) TOPOLOGÍA: lineal	
	(ii) TIPO DE MOLÉCULA: ADN (genomic)	
45	(iv) ANTISENTIDO: SI	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 70:	
	GGCACCTGCT GTGACCTGTG CAGG	24
50	(72) INFORMACIÓN PARA LA SEC. NO ID: 71:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
55	(A) LONGITUD: 22 pares de base	
	(B) TIPO: ácido nucléico	
60	(C) TIPO DE HEBRA: sencilla	
	(D) TOPOLOGÍA: lineal	
	(ii) TIPO DE MOLÉCULA: ADN (genomic)	
	(iv) ANTISENTIDO: NO	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 71:	
	GTCCTGCCAC TTCGAGACAT GG	22
65	(73) INFORMACIÓN PARA LA SEC. NO ID: 72:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
	(A) LONGITUD: 23 pares de base	

ES 2 289 610 T3

(B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

5 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: SI

10 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 72:

GAACTTCTC TGCCCTTACC GTC

23

15 (74) INFORMACIÓN PARA LA SEC. NO ID: 73:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 26 pares de base

(B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

20 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: NO

25 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 73:

30 CCAACACCAG CATCCATGGC ATCAAG

26

(75) INFORMACIÓN PARA LA SEC. NO ID: 74:

35 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 pares de base

(B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

40 (ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: SI

45 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 74:

GGAGAGTCAG CTCTGAAAGA ATTCAAG

27

50

**LISTADO DE SECUENCIAS**

55 <110> Arena Pharmaceuticals, Inc.

<120> Human Orphan G Protein-Coupled Receptors

60 <130> JEC/FP6282123

<140> EP 05003040.2

<140> 1999 10 13

65 <150> EP 99972682.1

<151> 1999 10 13

# ES 2 289 610 T3

<150> PCT/US99/23687  
<151> 1999 10 13

5 <150> 601109,213  
<151> 1998 11 20

10 <150> 60/120,416  
<151> 1999 02 16

15 <150> 60/121,852  
<151> 1999 02 26  
<150> 60/123,946  
<151> 1999 03 12

20 <150> 60/123,949  
<151> 1999 03 12

25 <150> 60/136,436  
<151> 1999 05 28

30 <150> 60/136,437  
<151> 1999 05 28  
<150> 60/136,439  
<151> 1999 05 28

35 <150> 60/136,567  
<151> 1999 05 28

40 <150> 60/137,127  
<151> 1999 05 28

45 <150> 60/137,131  
<151> 1999 05 28  
<150> 601141,448  
<151> 1999 06 29

50 <150> 60/156,653  
<151> 1999 09 29

55 <150> 60/156,333  
<151> 1999 09 29  
<150> 60/156,555  
<151> 1999 09 29

60 <150> 60/156,634  
<151> 1999 09 29

65 <150> 60/157,280  
<151> 1999 10 01

ES 2 289 610 T3

## ES 2 289 610 T3

&lt;400&gt; 2

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

55

60

65

ES 2 289 610 T3

Ala Leu Arg Ile His Ser Tyr Pro Glu Gly Ile Cys Leu Ser Gln Ala  
275 280 285

5 Ser Lys Leu Gly Leu Met Ser Leu Gln Arg Pro Phe Gln Met Ser Ile  
290 295 300

Asp Met Gly Phe Lys Thr Arg Ala Phe Thr Thr Ile Leu Ile Leu Phe  
305 310 315 320

10 Ala Val Phe Ile Val Cys Trp Ala Pro Phe Thr Thr Tyr Ser Leu Val  
325 330 335

Ala Thr Phe Ser Lys His Phe Tyr Tyr Gln His Asn Phe Phe Glu Ile  
340 345 350

15 Ser Thr Trp Leu Leu Trp Leu Cys Tyr Leu Lys Ser Ala Leu Asn Pro  
353 360 365

20 Leu Ile Tyr Tyr Trp Arg Ile Lys Lys Phe His Asp Ala Cys Leu Asp  
           370                         375                         380

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

25 Lys Arg Arg Ile Arg Pro Ser Ala Val Tyr Val Cys Gly Glu His Arg  
405 410 415

Thr Val Val

The Val Val

<210> 3

<211> 1119

<212> DNA

35 <213> *Homo sapiens*

<400> 3

atgttagcca	acagctccctc	ttaccaaacagt	tctgttctcc	cgtgtcttgc	ctaccggacct	60
accccacggcc	tgcacttgggt	ggtctacaggc	ttgggtctgg	ctggccggggt	ccccctcaac	120
gcgcctagccc	tctgggttctt	cctgcgcgcgc	ctgcgcgtgc	actcggtgggt	gagcgtgtac	180
atgtgttaacc	tggggggccagg	cgacactgttc	ttcacccctt	cgctgtccccgt	tctgttcttcc	240
tactacgcac	tgcacccactg	gccttttttttt	gaccctctgt	gcccaggacac	ggggggccatc	300
tcccgatgt	acatgttaegg	tagetgtcate	tteetgtatgc	tcatacaatgt	ggacccgttac	360
ggccgcattcg	tgcaccccgct	gcgactgcgc	cacctggggc	ggccccccgt	ggggggggctg	420
ctctgttctgg	gggtgttgggc	gtctatcttg	gtgttgggg	tgggggggg	eeeggtgtcac	480
aggcccttcgc	gttggcgatca	ccggggaccc	gagggtggcgc	tatgttgcga	gagtttccage	540
gacgagctgt	ggaaaggcgag	gctgtgtccc	ctcggtgtgc	ttggccggggc	gtctgggttcc	600
ctgtgtccccc	tgggggggggt	ggtctactcg	tcgggggggg	tcttttggac	gttggggcgc	660
cccgacggca	cgcagagccca	ggggcgccgg	aagacccgtgc	gcctcttgcgt	ggctaaaccc	720
gtcatettcc	tgttgtgttt	cggtggccatcc	aaacagacacgc	ttgggggttcta	cgggctgtcg	780
cgggagcaagg	tgggtggcggt	cggcggtgtt	gcggcgatcc	gggtgtcgccgg	gggtgtgtgt	840
gtgttgtgtgc	tgttgtgggg	cgccaaactgc	gtgttgtggacc	cgcttgtgtat	ctacttttgc	900
gcccgagggt	tccgcacacac	cctgcgcgcgc	ctgggcacttc	cgccatccgggc	caggacccgt	960
gccaccaacg	ggacgcggggc	ggcgctcgcc	caatccggaaa	ggtccggccgt	caccacccgac	1020
gccaccaggc	cggtgtccgc	tgttgtgggg	ctgttgtggac	cctccgacttc	ccactctctg	1080
tcttcttca	cacagtgtcc	ccaggattee	geeeetgtga			1119

<210> 4

60 <211> 372

<212> PTR

<213> *Homo sapiens*

## ES 2 289 610 T3

&lt;400&gt; 4

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

60

65

ES 2 289 610 T3

Asn Cys Val Leu Asp Pro Leu Val Tyr Tyr Phe Ser Ala Glu Gly Phe  
290 295 300

5 Arg Asn Thr Leu Arg Gly Leu Gly Thr Pro His Arg Ala Arg Thr Ser  
           305                 310                 315                 320

Ala Thr Asn Gly Thr Arg Ala Ala Leu Ala Gin Ser Glu Arg Ser Ala  
325 330 335

10 Val Thr Thr Asp Ala Thr Arg Pro Asp Ala Ala Ser Gln Gly Leu Leu  
                   340           345           350

Arg Pro Ser Asp Ser His Ser Leu Ser Ser Phe Thr Gln Cys Pro Gln  
355 360 365

Asp Ser Ala Leu  
370

20 <210> 5

<211> 1107

<212> DNA

<213> *Homo sapiens*

25

<400> 5

<210> 6

<211> 368

<212> PTR

<213> *Homo sapiens*

55 <400> 6

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

60 Gly Leu Ile Leu Ala Ala Val Val Glu Val Gly Ala Leu Leu Gly Asn  
60 20 25 30

## ES 2 289 610 T3

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

# ES 2 289 610 T3

<210> 7

<211> 1008

<212> DNA

5 <213> *Homo sapiens*

<400> 7

```

10      sttggaaatcat ctttcattt tggatgtatc ctttgtgtttt tggccctccctt catcattgtt 60
       actsaacacac tagtggtgtt ggctgtgtg ctgttgtatcc acsaaatgtt tgggtgtcaat 120
       ctctgtttaa ctttgaatctt ggctgtggctt gacacccttga ttgggtgtggc catctctggc 180
       ctactcacatc accagcttcc cagcccttctt cggccccacac agaagaccctt gtgcagccctg 240
       cggatggcat ttgtcaatcc ctcggcagctt ccctctgtttc tcacggatcat gctgtatcacc 300
       tttgacaggat accttggccat caagcagccc ttccgttact tgaagatcat gagtgggttc 360
       gtggccgggg cctgcatttc cgggtgtgg tttagtgtttt acettatggg ttteetccca 420
       ctcggaaatcc ccatgttcca gcaactgtcc tacaaaaggge agtgeagettt tttgtgttta 480
       tttcaaccctt acttcgtgtt gaccccttcc tgcgttggct ttttccage catgttccctt 540
       tttgttcttctt tttactgtccat catgttcaag atgttccca tgcacagccca gcaatccca 600
       aagatggaaatc atgeaggagc catggcttggg ggttatgtat ccccaacggac tcccacggcac 660
       ttcaaaatgtt tccgttctt gtcgttctt attggggatgt ttgttataatc ctggacccccc 720
       ttcctttatca ctggcatgtt gcaagggtggcc tgccaggagt gtcaccccttta cctagtgtctg 780
       gaacggtaacc tttgttggctt cggcggtggcc aactcccttc tcaacccact catctatgcc 840
       ttttggcaga aggagggttggc actgeagatc taccacatgg ccctaggagt gaagaagggtg 900
       ctcaacctcat ttttcccttt ttttctggcc aggaattgtt gccccagagag gcccaggggaa 960
       agtttctgttccatc ttttcccttgc ttttctccaggc tcagatgttgc atggctaa 1008

```

<210> 8

30 <211> 335

<212> PTR

<213> *Homo sapiens*

35 <400> 8

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

60

65

# ES 2 289 610 T3

	Tyr Leu Lys Ile Met Ser Gly Phe Val Ala Gly Ala Cys Ile Ala Gly			
	115	120	125	
5	Leu Trp Leu Val Ser Tyr Leu Ile Gly Phe Leu Pro Leu Gly Ile Pro			
	130	135	140	
	Met Phe Gln Gln Thr Ala Tyr Lys Gly Gln Cys Ser Phe Phe Ala Val			
10	145	150	155	160
	Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro			
	165	170	175	
15	Ala Met Leu Leu Phe Val Phe Phe Tyr Cys Asp Met Leu Lys Ile Ala			
	180	185	190	
	Ser Met His Ser Gln Gln Ile Arg Lys Met Glu His Ala Gly Ala Met			
20	195	200	205	
	Ala Gly Gly Tyr Arg Ser Pro Arg Thr Pro Ser Asp Phe Lys Ala Leu			
	210	215	220	
25	Arg Thr Val Ser Val Leu Ile Gly Ser Phe Ala Leu Ser Trp Thr Pro			
	225	230	235	240
	Phe Leu Ile Thr Gly Ile Val Gln Val Ala Cys Gln Glu Cys His Leu			
	245	250	255	
30	Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser			
	260	265	270	
	Leu Leu Asn Pro Leu Ile Tyr Ala Tyr Trp Gln Lys Glu Val Arg Leu			
35	275	280	285	
	Gln Leu Tyr His Met Ala Leu Gly Val Lys Val Leu Thr Ser Phe			
	290	295	300	
40	Leu Leu Phe Leu Ser Ala Arg Asn Cys Gly Pro Glu Arg Pro Arg Glu			
	305	310	315	320
	Ser Ser Cys His Ile Val Thr Ile Ser Ser Ser Glu Phe Asp Gly			
45	325	330	335	
	<210> 9			
	<211> 1413			
	<212> DNA			
50	<213> <i>Homo sapiens</i>			
	<400> 9			
55	atggacacta ccatggaaagc tgacctgggt gccactggcc acaggccccg cacagagctt 60 gatgtatggg actccctaaaa ccmaagggtggc tgggacacgg ttttttttgtt ggccctgttg 120 ctcccttgggc tgccagccaa tgggtttagt gctgtggctgg cccgcctccca ggccctggcat 180 ggagctggca cgcgcttggc qctgtcttcg ctcaagcttgg ccctctctga ttctttgttc 240 ctggcagcag cggecttcca gatccctagag atccggcatg ggggacactg gcccctgggg 300 acagctgcct gcccgtttta ctacttccca tggggcgtgtt ctacttctc cggcccttttc 360 ctgtctggccg ccctcagectt cggccgttgc ctgtctggcgc tggcccaaca ctgttaccct 420 gggcacccgcc cagtccggctt gcccctctgg gtctgtgcgg ggtttttgggtt gttggccaca 480 ccctttcagecg tgccctggctt ggtttttttttt gaggctgtccg ttctgggttata cggccctggtc 540			
60				

ES 2 289 610 T3

```

atctgcctgg acttctggga cagcgaggag ctgtcgctga ggatgtcgaa ggctctgggg 600
ggcttcctgc cttttctctt gctgtcttgtc tggccacgtgc tcaccctaggc cacagccctgt 660
cgcacctggc accgccaaca gcagccccca gcttgcgggg gcttcgtccccg tggccagg 720
accattctgt cagccatgtt ggccctgagg ctgccttacc agetggccca gctgtcttac 780
ctggccctcc tggggacgt ctactctggc tacctgtctt gggaggccct ggtctactcc 840
gacttactga ttctactcaz cagetgcctc agcccttcc tctgtccat ggccagtgcc 900
gacctccggc ccctgtcgcc cccctgtgtc tggcccttgc cggcagctt ctggcgaggag 960
cggtccgggcgatccacgtt cacttgaggca cagacccaggc tagatttctga gggccaaact 1020
ctggccaggc cgtatggcaga gggccaggta cagatggat ctgtgggcca gcttcaggcg 1080
aaaccccttccat tccggccacg atcggtttcc acgttgttgc cccatggggcc 1140
cagccacatgtt cggatccccac agccccggccaa cagctgttgcacc tggccaggatc 1200
qattctgttg cccatggccaca ggcagacact aacgttccaga cccctgtcacc tgctggccat 1260
tttgtgtccca gtccctgtgt tgaagcttcc caaacccccat cccatgttgcacc taaaaagggg 1320
ggcccttgagg acccttagccat accttctggc tctgtgggg gaaagccccat cagcaccctgg 1380
ccatggggggggccatggccgc agggcccttacgt tga 1413

```

<210> 10

<211> 468

20 <212> PTR

<213> *Homo sapiens*

<400> 10

25

Met Asp Thr Thr Met Glu Ala Asp Leu Gly Ala Thr Gly His Arg Pro  
1 5 10 15

30 Arg Thr Glu Leu Asp Asp Glu Asp Ser Tyr Pro Gln Gly Gly Trp Asp  
                   20                 25                                 30

Thr Val Phe Leu Val Ala Leu Leu Leu Leu Gly Leu Pro Ala Asn Gly  
35 40 45

35 Leu Met Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr

Arg Leu Ala Leu Leu Leu Ser Leu Ala Leu Ser Asp Phe Leu Phe

Leu Ala Ala Ala Ala Phe Gln Ile Leu Glu Ile Arg His Gly Gly His

45 Trp Pro Leu Gly Thr Ala Ala Cys Arg Phe Tyr Tyr Phe Leu Trp Gly

Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu Ser Leu Asp  
15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30

50 Arg Cys Leu Leu Ala Leu Cys Pro His Trp Tyr Pro Gly His Arg Pro  
51 52 53 54 55 56 57 58 59 60 61 62 63 64

Val Arg Leu Pro Leu Trp Val Cys Ala Gly Val Trp Val Leu Ala Thr  
145 150 155 160

Leu Phe Ser Val Pro Trp Leu Val Phe Pro Glu Ala Ala Val Trp Trp  
165 170 175

60 Tyr Asp Leu Val Ile Cys Leu Asp Phe Trp Asp Ser Glu Glu Leu Ser  
180 185 190

# ES 2 289 610 T3

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

<210> 11

<211> 1248

60 <212> DNA

<213> *Homo sapiens*

# ES 2 289 610 T3

<400> 11

```

5      atgtcaggqa tggaaaaact tcagaatqct tcctggatct accagcagaa actagaagat 60
       ccattccaga aacacctgaa cagcacccgag gaggatctgg ctttcttcgt cggacatcg 120
       cgcaagccact tctteetccc cgtgtctgtg gtgtatgtc caattttgt ggtgggggtc 180
       attggcaatg tcctgggtgtg ctgggtgtt ctgeageace aggetatgaa gacgccccacc 240
       aactactacc tcttagectt ggccggatctt gaccccttgg tcctgtccctt cggaaatgccc 300
       ctggagggtct atgagatgtg ggcacactac cttttcttgt tcgggcccgt gggctgtac 360
       ttcaagacggg ccccttttga gacegtgtgc ttgcctcca tcctcagcat caccacccgtc 420
       agcgtggacg gctacgtggc catctacac cccgtccggc ccaactgca gggcaccggg 480
       cgccggggcc tcaggatctt cggcatagtc tggggctttt cctgtctttt cccctgccc 540
       aacacccatgca tccatggcat caagtccac tacttccca atgggtccctt gttcccaagg 600
       tcggccacctt gtacggtcat caagccatgt tggatctaca atttcatcat ccaggtaacc 660
       tccttcttat tcttccctt cccatgact gtcatcaigt tccttacta cctcatggca 720
       ccacatcaa agaaagccaa strictinga gcaatgaaa gyaatgcaaa tattcaaaa 780
       cccctgcayaa aatcaatgtcaaa caatgtctg tttgtcttgg tctttagtgtt tqctatctgt 840
       tqggccccgyt tccacattga ccgactcttc ttcaqctttg tqqaggagtg qagtgaatcc 900
       ctggctgtctg tgatcaatctt cgtccatgtg gtgtcagggtg tcttcttcta cctgagctca 960
       gctgtcaacc ccaatataca taactactg tctcggccgt tccaggcagc attccagaat 1020
       gtgatotctt ctcccacaa acgtggcacttccatgtt accacatgtt gcccacgtgtcc 1080
       cagcggaaada ttttctgtac agaatggcactt tttgtggagc tgaccgaaga tatagggtcc 1140
       tatttccat gtcaatgttcatgcaatc tcttcccttcc caatgttccctt ctctatgtca 1200
       cagatgtcas caatcaatca tcaaaatgttcc cacttttaca aaacctgt 1248

```

25 <210> 12

<211> 415

<212> PTR

<213> *Homo sapiens*

30 <400> 12

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

# ES 2 289 610 T3

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

<210> 13

<211> 1173

55 <212> DNA

<213> *Homo sapiens*

<400> 13

```
atggccagata cttatagcgc acatcaatcta tcactaaaggca ctctgtgttac ttttagccatcc 60
tttatgtctt tagtagcttc tgctataatg ctggaaatg ctttggatc tttagcttt 120
```

ES 2 289 610 T3

20 <210> 14

<211> 300

<211> 590

25 <213> *Homo sapiens*

<400> 14

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

ES 2 289 610 T3

Ile Thr Ser Phe Leu Glu Phe Val Ile Pro Val Ile Leu Val Ala Tyr  
 180 185 190  
 5 Phe Asn Met Asn Ile Tyr Trp Ser Leu Trp Lys Arg Asp His Leu Ser  
 195 200 205  
 Arg Cys Gln Ser His Pro Gly Leu Thr Ala Val Ser Ser Asn Ile Cys  
 210 215 220  
 10 Gly His Ser Phe Arg Gly Arg Leu Ser Ser Arg Arg Ser Leu Ser Ala  
 225 230 235 240  
 15 Ser Thr Glu Val Pro Ala Ser Phe His Ser Glu Arg Gln Arg Arg Lys  
 245 250 255  
 Ser Ser Leu Met Phe Ser Ser Arg Thr Lys Met Asn Ser Asn Thr Ile  
 260 265 270  
 20 Ala Ser Lys Met Gly Ser Phe Ser Gln Ser Asp Ser Val Ala Leu His  
 275 280 285  
 Gln Arg Glu His Val Glu Leu Leu Arg Ala Arg Arg Leu Ala Lys Ser  
 290 295 300  
 25 Leu Ala Ile Leu Leu Gly Val Phe Ala Val Cys Trp Ala Pro Tyr Ser  
 305 310 315 320  
 30 Leu Phe Thr Ile Val Leu Ser Phe Tyr Ser Ser Ala Thr Gly Pro Lys  
 325 330 335  
 Ser Val Trp Tyr Arg Ile Ala Phe Trp Leu Gln Trp Phe Asn Ser Phe  
 340 345 350  
 35 Val Asn Pro Leu Leu Tyr Pro Leu Cys His Lys Arg Phe Gln Lys Ala  
 355 360 365  
 Phe Leu Lys Ile Phe Cys Ile Lys Lys Gln Pro Leu Pro Ser Gln His  
 370 375 380  
 40 Ser Arg Ser Val Ser Ser  
 385 390  
 45 <210> 15  
 <211> 1128  
 <212> DNA  
 <213> *Homo sapiens*  
 50 <400> 15  
 55 atggcgaacg cgagcgagcc ggggtggcagc ggcggcggcg aggcggccgc cctggggcetc 60  
 aagctggcca cgctcagctt gctgttgtgc gtgagccatg cggccaaacgt gtcgttgtcg 120  
 ctgtctgtatcg tgcgggagcg cagccgtcac cgcgcggccgt actactgtctgt gtcgacatcg 180  
 tgcctggccg acgggtctcg cgcgcgtcgcc tgcctccccgg ccgtcatgtctgt ggccgggggggg 240  
 cggcggggggg cggcggggggg ggcggccggcg ggcgcgtcg gtcgtcaagct gtcgcgtttt 300  
 ctggcccgcc ttttgtttt ccaatggcc ttttgtgtgc tgggcgtggg cgtcacccggc 360  
 tacctggcca tgcgcgtccca ccgcatttat gcaagtgccgc tggccggctg gcccgtgcgc 420  
 gtcgtgtcg tgcgtgcgcgc ctggggatgt ggcgtqgcgc qggcccttccc gccaggatgtcg 480  
 gacgggggtg gcaacgcaga ggacgcgcgc tgcgccttgg agcagggggcc cggacggcgcc 540  
 cccqqqcgccqtc tqqqgtttctt gtcgtgtcg gccgtggggq tggggccac qcacctcgtc 600

# ES 2 289 610 T3

taccctccggcc tgcgtttttt cattccacgac cggcggcaaga tgccggccccgc ggcgcctggtg 660  
 cccggccgttc gccaacgactg gacccctccac ggccccggggcg ccacccggccaa ggccggccggcc 720  
 aactggacgg cgggttttgg ccggggggcc \*cggccggcc eggttgtggg catccggccc 780  
 gcaggggccgg gceggggcgc gcggggccctc ctgtgtttgg aaaaatccaa gacggagaag 840  
 aggtgtgtca agatgtttta cgccgttcacq ctgtttttcc tgctcccttg gggggccctac 900  
 gtcgtggccca gctacccctggg ggtccctggtg cggccggccgg cccgtccccca ggcctacctg 960  
 accggccctccg tttgggtgtggc cttccggcggcgg wccggccatca accccccylcyl wccggccatca 1020  
 tccaaacaggg agctgagggg ctgtttcagg gccccagttcc cctgtgtccaa gageccccgg 1080  
 accacccaggc cgaccatcc ctggggatcg aaaggccatgg ttatgtaa 1128

10

<210> 16

<211> 375

15 <212> PTR

<213> *Homo sapiens*

<400> 16

20

Met	Ala	Asn	Ala	Ser	Glu	Pro	Cly	Gly	Ser	Gly	Gly	Gly	Glu	Ala	Ala
1										10					15

25

Ala	Leu	Cly	Leu	Lys	Leu	Ala	Thr	Leu	Ser	Leu	Leu	Leu	Cys	Val	Ser
														20	30

Leu	Ala	Gly	Asn	Val	Leu	Phe	Ala	Leu	Leu	Ile	Val	Arg	Glu	Arg	Ser
												35	40	45	

30

Leu	His	Arg	Ala	Pro	Tyr	Tyr	Leu	Leu	Leu	Asp	Leu	Cys	Leu	Ala	Asp
												55	60		

35

Gly	Leu	Arg	Ala	Leu	Ala	Cys	Leu	Pro	Ala	Val	Met	Leu	Ala	Ala	Arg
											65	70	75	80	

40

Arg	Ala	Ala	Ala	Ala	Ala	Gly	Ala	Pro	Pro	Gly	Ala	Leu	Gly	Cys	Lys
											85	90	95		

Leu	Leu	Ala	Ile	Leu	Ala	Ala	Leu	Phe	Cys	Phe	His	His	Ala	Phe	Leu
											100	105	110		

Leu	Leu	Gly	Val	Gly	Val	Thr	Arg	Tyr	Leu	Ala	Ile	Ala	His	His	Arg
											115	120	125		

45

Phe	Tyr	Ala	Glu	Arg	Leu	Ala	Gly	Trp	Pro	Cys	Ala	Ala	Met	Leu	Val
											130	135	140		

Cys	Ala	Ala	Trp	Ala	Leu	Ala	Leu	Ala	Ala	Phe	Pro	Pro	Val	Leu	
											145	150	155	160	

50

Asp	Gly	Gly	Asp	Asp	Glu	Asp	Ala	Pro	Cys	Ala	Leu	Glu	Gln	Arg	
											165	170	175		

55

Pro	Asp	Gly	Ala	Pro	Gly	Ala	Leu	Gly	Phe	Leu	Leu	Leu	Ala	Val	
											180	185	190		

Val	Val	Gly	Ala	Thr	His	Leu	Val	Tyr	Leu	Arg	Leu	Leu	Phe	Phe	Ile
											195	200	205		

60

His	Asp	Arg	Arg	Lys	Met	Arg	Pro	Ala	Arg	Leu	Val	Pro	Ala	Val	Ser
											210	215	220		

65

# ES 2 289 610 T3

	His Asp Trp Thr Phe His Gly Pro Gly Ala Thr Gly Gln Ala Ala Ala			
225	230	235	240	
5	Asn Trp Thr Ala Gly Phe Gly Arg Gly Pro Thr Pro Pro Ala Leu Val			
	245	250	255	
	Gly Ile Arg Pro Ala Gly Pro Gly Arg Gly Ala Arg Arg Leu Leu Val			
	260	265	270	
10	Leu Glu Glu Phe Lys Thr Glu Lys Arg Leu Cys Lys Met Phe Tyr Ala			
	275	280	285	
	Val Thr Leu Leu Phe Leu Leu Trp Gly Pro Tyr Val Val Ala Ser			
15	290	295	300	
	Tyr Leu Arg Val Leu Val Arg Pro Gly Ala Val Pro Gln Ala Tyr Leu			
	305	310	315	320
20	Thr Ala Ser Val Trp Leu Thr Phe Ala Gln Ala Gly Ile Asn Pro Val			
	325	330	335	
	Val Cys Phe Leu Phe Asn Arg Glu Leu Arg Asp Cys Phe Arg Ala Gln			
	340	345	350	
25	Phe Pro Cys Cys Gin Ser Pro Arg Thr Thr Gln Ala Thr His Pro Cys			
	355	360	365	
	Asp Leu Lys Gly Ile Gly Leu			
	370	375		
30	<210> 17			
	<211> 1002			
	<212> DNA			
35	<213> <i>Homo sapiens</i>			
	<400> 17			
40	atgaaacacca cagtgtatgca aggcttcaac agatctgagc ggtgcggccagg agacactcggt 60 atagtacayc tggtatccc aecctctac acagtggttt tcttgaccgg catcctgtq 120 aatactttgg ctctgtgggt gttgttac atccccagct cttccacccat catcatctac 180 ctaaaaaaaca ctttgggtggc cgacttgata atgacactca tggccatccc caaaatccct 240 tcgtactcac acctggcacc ctggcagetc agagcttttg tggtgttgtt ttcttcgggt 300 atattttatgt agaccatgtt tggccatcc tggctgttag ggctcatagc ctttgacaga 360 ttcttcggaa tcatacgttcc tttggaaaat atttttctaa aaaaacatgt ttttgcaaaa 420 acggctctcaa ttttcatctg ttcttcttgc ttcttcatctt ccctgcacaaa taatgttttg 480 agcaacaagg aagcaacacc atcgtctgtg aaaaatgtgtg ttctttaaa gggccctctg 540 gggtgtttttt ggcataaatat ggttaataaac atatgccat ttttttctg gactgtttt 600 atccaaatgtc ttgtgtttta tgggtttttt gcaaaaaaaaaat tataatgttcc ttatagaaaag 660 tccaaaaatgtt aggcacagaaa aaccaacaaa aagctggaaag gcaaaatgtt tggctgtgtq 720 gttgtttttt tttgtgtttt tggccatccat tttttgtttt gggccatccat tactcacatgt 780 caaaaaaaaca ataaatgtt tggtagactg caaaaatcaac tggccatccat taaagaaaaca 840 actctttttt tggcagcaac tttttttttt atggatccat tttttttttt tttttttttt 900 aaaaatccatca cagaaaaatgtt accatgtatg caagggagaa agaccacacgc atcaagccaa 960 aaaaatccatca cagaaaaatgtt accatgtatg caagggagaa agaccacacgc atcaagccaa 1002			
45				
50				
55				
	<210> 18			
	<211> 333			
60	<212> PTR			
	<213> <i>Homo sapiens</i>			

## ES 2 289 610 T3

&lt;400&gt; 18

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

&lt;210&gt; 19

&lt;211&gt; 1122

ES 2 289 610 T3

<212> DNA

<213> *Homo sapiens*

5 <400> 19

<210> 20

<211> 373

30 <212> PTR

<213> *Homo sapiens*

<400> 20

35

Met Ala Asn Thr Thr Gly Glu Pro Glu Glu Val Ser Gly Ala Leu Ser  
1 5 10 15

40 Pro Pro Ser Ala Ser Ala Tyr Val Lys Leu Val Leu Leu Gly Leu Ile  
20 25 30

Met Cys Val Ser Leu Ala Gly Asn Ala Ile Leu Ser Leu Leu Val Leu  
35 40 45

45 Lys Glu Arg Ala Leu His Lys Ala Pro Tyr Tyr Phe Leu Leu Asp Leu  
50 55 60

Cys Leu Ala Asp Gly Ile Arg Ser Ala Val Cys Phe Pro Phe Val Leu  
65 70 75 80

50 Ala Ser Val Arg His Gly Ser Ser Trp Thr Phe Ser Ala Leu Ser Cys  
55 60 65

55

60

65

# ES 2 289 610 T3

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

<210> 21

<211> 1053

<212> DNA

<213> *Homo sapiens*

ES 2 289 610 T3

<400> 21

	atggctttgg aacagaacca gtcacagat tattattatg aggaaaaatga aatgaatggc 60
5	acttatgact acagtcaata tgaattgtac tgatcaag aagatgtcag aeaatttgc 120
	aaaggccccc tccccgtatt ctcacaata gctttcgta ttggacttgc aggcaatcc 180
	atggtagtgq caatitatac ctatcacaaq aaacagagaa ccaaaacaga ttgtacate 240
	atggatggg ctgtggccgg tttttttttt tttttttttt tgggtttttt ygcgttttt 300
	qcagttcatq gggtgggtttt agggaaaaata atgtgcaaaa taacttcage cttgtacaca 360
10	ctaaatcttq tctctgqaat gcagtttctg gcttgcatac goatagacac atatggggca 420
	qtaactaatq tcccccagcc atcaggatq ggaaaaccat getggatcat ctgtttctgt 480
	gtctggatgg ctggccatctt gtcgagcata cccccatgtt ttttttatac agtaaatgac 540
	aatgttaggt gcattcccat tttccccccgc tacctaggaa catcaatgaa agcattgtatt 600
	caaattgttag agatctgcat tggatttgtt gtaccctttt tttttatggg ggtgtgtac 660
	tttatacagg caaggacact catgaagatq ccaaaacatta aatatactcg acccctaataa 720
15	gtctgtcata cagtctttat agttttcatc gtcaactcaat tgccttataa cattgtcaan 780
	tctgtccatc ccatacgtatc catctactcc ctgtatccca gtcacaaat gagoaaacgc 840
	atggacatcg ccatacgtatc cacaagaaatc atggcactt ttcacacgtq ctcacaaatcca 900
	atccctttatggg agatcttttcc aaaaactatcg ttatgaaatgt ggcacaaagaaa 960
	ttalgggttttgg gagaagaca gagacaaatgt gtcggaggat ttccttttgc ttctgtgggt 1020
	catacaatgc caaccaatgtac ttttagcatt taa 1053

20

<210> 22

<211> 350

25 <212> PTR

<213> *Homo sapiens*

<400> 22

30

# ES 2 289 610 T3

	165	170	175
	Thr Val Asn Asp Asn Ala Arg Cys Ile Pro Ile Phe Pro Arg Tyr Leu 180	185	190
5	Gly Thr Ser Met Lys Ala Leu Ile Gln Met Leu Glu Ile Cys Ile Gly 195	200	205
10	Phe Val Val Pro Phe Leu Ile Met Gly Val Cys Tyr Phe Ile Thr Ala 210	215	220
	Arg Thr Leu Met Lys Met Pro Asn Ile Lys Ile Ser Arg Pro Leu Lys 225	230	235
15	Val Leu Leu Thr Val Val Ile Val Phe Ile Val Thr Gln Leu Pro Tyr 245	250	255
20	Asn Ile Val Lys Phe Cys Arg Ala Ile Asp Ile Ile Tyr Ser Leu Ile 260	265	270
	Thr Ser Cys Asn Met Ser Lys Arg Met Asp Ile Ala Ile Gln Val Thr 275	280	285
25	Glu Ser Ile Ala Leu Phe His Ser Cys Leu Asn Pro Ile Leu Tyr Val 290	295	300
	Phe Met Gly Ala Ser Phe Lys Asn Tyr Val Met Lys Val Ala Lys Lys 305	310	315
30	Tyr Gly Ser Trp Arg Arg Gln Arg Gln Ser Val Glu Glu Phe Pro Phe 325	330	335
	Asp Ser Glu Gly Pro Thr Glu Pro Thr Ser Thr Phe Ser Ile 340	345	350
35	<210> 23		
	<211> 1116		
40	<212> DNA		
	<213> Homo sapiens		
	<400> 23		
45	atgccaggaa acgcacaaaa agtgaccacc actgccccgt gggccctccc ggcccttcctc 60 qcccaagacct gcaacaaacgt gtcccttqaa gagagcaggaa tggctctgggt cgtgggtgtac 120 aqcqccqgtgt gcacgcgtgg gggtccggcc aactgcctgt ctgcgtggct ggcgcgtgtc 180 cgggtactqc aqggcaacgt gctggccgtt tacctgtctt gcctggcact ctgcgtactg 240 ctqtacacag qcacgcgtgcc accttgggtt atctataatcc gcaaccagca ccgcgtggacc 300 ctaggccatgc tggccctcgaa ggtgacceggcc tacatetttt tctgcacat ctacgtcagc 360 atcccttttcc tggctgtcat ctccgtcgac cgcttcgtgg ccgggggtgtac cgcgcgtggag 420 aglcggggcc gcccggccgg gaggaccggcc atccctcatct ccgcgtgtcat ctteatccccc 480 gtcqqqatcg ttcaacttcccc ggtgttcccg aqggaaqaca aqggaaqactg cttagacatg 540 ctgcagatgg acagcaggat tgccgggtac tactacggca ggttccatcg tggctttggcc 600 atccctttttcc ccatcatcgc ctccatccaac caccggattt tcaggagcat caagcagagc 660 atggggtttaa gcgctggccca gaaaggccaaag gtgaagcaat cggccatcgc qgtgggtgtc 720 atcttccatcg tctgttttcgc cccgttaccac ctggttttcc tggccaaatgc cgctgccttt 780 tccatccatca yaggagacag gaacggccatg tgccggcttgg aggaaaggatc gtacacagcc 840 tctgtgtgtt ttclglgtct gtcacqgtg aqcgqgtgg ctgacccat tatctacgtg 900 ctggccacgg accatccccg ccaagaagtg tcccaaatcc ataaagggtg qaaaaggtgg 960 tccatgaaga caqacgtcac cgggttccacc cacagcaggaa acaccggggg gctgcgttcg 1020		
55	ccccgtggccc ttgcagacca ctacacccccc tccaggccccg tgcacccacc agggtcacca 1080 tgcctgtcaa agaggctgtat tgaggagttcc tgctgt 1116		
60			
65			

# ES 2 289 610 T3

<210> 24  
 <211> 371  
 <212> PTR  
 5 <213> *Homo sapiens*  
 <400> 24

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

60

65

ES 2 289 610 T3

<210> 26  
55 <211> 370  
<212> PTR  
<213> *Homo sapiens*

60 <400> 26

Met Ala Asn Tyr Ser His Ala Ala Asp Asn Ile Leu Gln Asn Leu Ser  
1 5 10 14

# ES 2 289 610 T3

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

ES 2 289 610 T3

<210> 27

<211> 1080

<212> DNA

5 <213> *Homo sapiens*

<400> 27

10	atgcayggcc	cgaaacagcac	cggccccggac	aacgcgcacgc	tgcagatgtc	gcggaaaccgg	60
	qcqatcgccg	ggcccttgtcc	cgtgtgtlac	tgcgtgggtt	ggcggttcag	cattttggggc	120
	aaaccttttct	ctctgtgggt	gtctgtccgg	ccatggggc	ccagatcccc	gtcggttcata	180
	ttcatgtatca	acctgaggcte	cacggacactg	atgttgttcca	gggtgttgee	tttccaaatc	240
	tactaccatt	gcaaccggca	ccactgggtt	ttcggttgta	tgcgttgcac	cggtgtgace	300
15	gtggccctttt	acgaaacat	gtatccatgc	atcccttacca	tgcacgttat	cagegtggag	360
	ccctttccctgg	gggttctgtt	cccgttcagc	tccaaaggctt	ggcgccggccg	tgcgttacgcg	420
	gtggccggcggt	gtgcaggggac	ctggctgttg	ctccctgacccg	ccctgtgecc	gttggccggcc	480
	accgtatctca	cccttccctgg	geacggccctg	ggcatcata	ctgttttgcg	cgttctcaag	540
	tqqacgtatqc	tcccccgggt	ggccatgtgg	ggcggtttcc	ttttccat	cttcatcttg	600
	ctgtttctca	tcccccgttgc	gtttccatgg	gtttgttaca	ggccacat	cttcataatgt	660
	ttgcgcacgg	aggaggccga	cggccggggag	cacggagge	gcgggggtggg	tctggccgcg	720
	gtgttttgc	tgggtttgt	caatgtttgt	ggccuuauaa	attttgttgc	tctggccgcac	780
	atcggtggcc	gcctgtttca	cggttttttt	tactaccacg	tgttacatgt	ccgttgcgt	840
	ccatgtttcc	tcaacaactg	tctggaccccg	tttggttatt	actttgcgtc	ccggggaaatc	900
	ccatgtgcgc	tgcgggaaata	tttgggttgc	cgccgggttgc	ccaaqagacac	cttggacacac	960
	ccatgtgcgc	gccttttttc	cgccaggacc	aatgtttgtc	gttccggagge	cggttgcgcac	1020
	ccatgtgcgc	tgcggggggc	tttttttttt	ggccatccaga	ggccatccaga	tgcgttgcgtt	1080

30 <210> 28

<211> 359

<212> PTR

<213> *Homo sapiens*

35

<400> 28

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

55

60

65

# ES 2 289 610 T3

	Tyr Tyr His Cys Asn Arg His His Trp Val Phe Gly Val Leu Leu Cys			
	85	90	95	
5	Asn Val Val Thr Val Ala Phe Tyr Ala Asn Met Tyr Ser Ser Ile Leu			
	100	105	110	
	Thr Met Thr Cys Ile Ser Val Glu Arg Phe Leu Gly Val Leu Tyr Pro			
	115	120	125	
10	Leu Ser Ser Lys Arg Trp Arg Arg Arg Arg Tyr Ala Val Ala Ala Cys			
	130	135	140	
	Ala Gly Thr Trp Leu Leu Leu Thr Ala Leu Cys Pro Leu Ala Arg			
15	145	150	155	160
	Thr Asp Leu Thr Tyr Pro Val His Ala Leu Gly Ile Ile Thr Cys Phe			
	165	170	175	
20	Asp Val Leu Lys Trp Thr Met Leu Pro Ser Val Ala Met Trp Ala Val			
	180	185	190	
	Phe Leu Phe Thr Ile Phe Ile Leu Leu Phe Leu Ile Pro Phe Val Ile			
	195	200	205	
25	Thr Val Ala Cys Tyr Thr Ala Thr Ile Leu Lys Leu Leu Arg Thr Glu			
	210	215	220	
	Glu Ala His Gly Arg Glu Gln Arg Arg Arg Ala Val Gly Leu Ala Ala			
30	225	230	235	240
	Val Val Leu Leu Ala Phe Val Thr Cys Phe Ala Pro Asn Asn Phe Val			
	245	250	255	
35	Leu Leu Ala His Ile Val Ser Arg Leu Phe Tyr Gly Lys Ser Tyr Tyr			
	260	265	270	
	His Val Tyr Lys Leu Thr Leu Cys Leu Ser Cys Leu Asn Asn Cys Leu			
	275	280	285	
40	Asp Pro Phe Val Tyr Tyr Phe Ala Ser Arg Glu Phe Gln Leu Arg Leu			
	290	295	300	
	Arg Glu Tyr Leu Gly Cys Arg Arg Val Pro Arg Asp Thr Leu Asp Thr			
45	305	310	315	320
	Arg Arg Glu Ser Leu Phe Ser Ala Arg Thr Thr Ser Val Arg Ser Glu			
	325	330	335	
50	Ala Gly Ala His Pro Glu Gly Met Glu Gly Ala Thr Arg Pro Gly Leu			
	340	345	350	
	Gin Arg Gin Glu Ser Val Phe			
	355			
55	<210> 29			
	<211> 1503			
	<212> DNA			
60	<213> <i>Homo sapiens</i>			

# ES 2 289 610 T3

<400> 29

```

5      atggagcqtc cctgggaggga cagccctagge cgggggggggg cagctggggg ctgcgcgtq 60
      ccagtgcggc cggggggggcg ctccgggtgcc qcggcgagtg gcacaggctq gcagccatgg 120
      gctqsgtgcg cgggaccggaa ggggagggggg caactgtgg cgaccggggg ccccttgggt 180
      cgctggcccg cccccctcgcc tgccagttcc agccccggcc cggggggggc gtccggcteac 240
      ccgggtcaag gcaugccggac tggggggggggc ycaucyacauu tttttttgggg 300
      cggccatgg agtcgggggt getggggggc ggccgggtqa gegagggtat cgttcgtat 360
      tacaactaca cggccaaqct cggcggtqcg agtaccacgc cgggtqccgg cctgcgcqcc 420
      gacqccqtcg tttgtgcgtgc qgtgtqcgcc ttcatcggtc tagagaaatc agccqgtgtcg 480
      ttgtgtgcgtg gacgcccaccc gcttcgttccac gctccatgt ttttgcgttccggcc 540
      acgttgttcgg atctgtgttgc agggccggcc taacccggcca acatctact gtcggggggc 600
      ctcaactgtg aactgtccccc cgggttcgtgg ttcgttcgggggg agggaggggt ctgtgtggc 660
      ctcaactgtgtc cgggtgtgtgg ccttcgtggcc atcgccgtgg agcgccgttcc cccatggcg 720
      cccacaaaaaccc cccaccccccgt ctccatgggg gggccggccggc tggggatggc agccggggcc 780
      tgggggggtgt cgttcgttccatgggttccggcc tcaacgtgtgg gctggatgttccgtgtggc 840
      cgggggggtgt cgttcgttccatgggttccggcc tcaacgtgtgg gctggatgttccgtgtggc 900
      ctccgttcgg tggggccatcc gggccggccgt tttgtgcactt acggccgtat ctactggccag 960
      gtaacggccca atggccggccg cctggggggca cggccggggggc ctggggggac cccctggggcc 1020
      cggggccggccg gcaacggccggc ctccgttccatgggttccggcc ttcgttcggggggc cccctggggcc 1080
      gtcgttgggg catgttgggg ccccttcgttcc ttcgttcggggc ttcgttcggggc cccctggggcc 1140
      tcgttcggccat gtcgttgggg ccccttcgttcc ttcgttcggggc ttcgttcggggc cccctggggcc 1200
      ttcgttcggccat gtcgttgggg ccccttcgttcc ttcgttcggggc ttcgttcggggc cccctggggcc 1260
      cggccgtggcc ttcgttcgggg ccccttcgttcc ttcgttcggggc ttcgttcggggc cccctggggcc 1320
      gcaacggccggg ctggggccgttcc ttcgttcggggc ttcgttcggggc ttcgttcggggc cccctggggcc 1380
      agcttcacccg gtcgttcggccg ctccatggggcc ttcgttcggggcc ttcgttcggggcc 1440
      acaacggccggcc cccgttcggcc ttcgttcggggcc ttcgttcggggcc ttcgttcggggcc 1500
      tga                                         1503

```

30 <210> 30

<211> 500

<212> PTR

<213> *Homo sapiens*

35

<400> 30

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

65

# ES 2 289 610 T3

	Gly Ala Ser Tyr Gln Pro Gly Ala Gly Leu Arg Ala Asp Ala Val Val			
	130	135	140	
5	Cys Leu Ala Val Cys Ala Phe Ile Val Leu Glu Asn Leu Ala Val Leu			
	145	150	155	160
	Leu Val Leu Gly Arg His Pro Arg Phe His Ala Pro Met Phe Leu Leu			
	165	170	175	
10	Leu Gly Ser Leu Thr Leu Ser Asp Leu Leu Ala Gly Ala Ala Tyr Ala			
	180	185	190	
	Ala Asn Ile Leu Leu Ser Gly Pro Leu Thr Leu Lys Leu Ser Pro Ala			
	195	200	205	
15	Leu Trp Phe Ala Arg Glu Gly Gly Val Phe Val Ala Leu Thr Ala Ser			
	210	215	220	
20	Val Leu Ser Leu Leu Ala Ile Ala Leu Glu Arg Ser Leu Thr Met Ala			
	225	230	235	240
	Arg Arg Gly Pro Ala Pro Val Ser Ser Arg Gly Arg Thr Leu Ala Met			
	245	250	255	
25	Ala Ala Ala Ala Trp Gly Val Ser Leu Leu Leu Gly Leu Leu Pro Ala			
	260	265	270	
	Leu Gly Trp Asn Cys Leu Gly Arg Leu Asp Ala Cys Ser Thr Val Leu			
	275	280	285	
30	Pro Leu Tyr Ala Lys Ala Tyr Val Leu Phe Cys Val Leu Ala Phe Val			
	290	295	300	
	Gly Ile Leu Ala Ala Ile Cys Ala Leu Tyr Ala Arg Ile Tyr Cys Gln			
	305	310	315	320
35	Val Arg Ala Asn Ala Arg Arg Leu Pro Ala Arg Pro Gly Thr Ala Gly			
	325	330	335	
40	Thr Thr Ser Thr Arg Ala Arg Arg Lys Pro Arg Ser Leu Ala Leu Leu			
	340	345	350	
	Arg Thr Leu Ser Val Val Leu Leu Ala Phe Val Ala Cys Trp Gly Pro			
	355	360	365	
45	Leu Phe Leu Leu Leu Leu Asp Val Ala Cys Pro Ala Arg Thr Cys			
	370	375	380	
	Pro Val Leu Leu Gln Ala Asp Pro Phe Leu Gly Leu Ala Met Ala Asn			
	385	390	395	400
50	Ser Leu Leu Asn Pro Ile Ile Tyr Thr Leu Thr Asn Arg Asp Leu Arg			
	405	410	415	
	His Ala Leu Leu Arg Leu Val Cys Cys Gly Arg His Ser Cys Gly Arg			
	420	425	430	
55	Asp Pro Ser Gly Ser Gln Gln Ser Ala Ser Ala Ala Glu Ala Ser Gly			
60				

ES 2 289 610 T3

435                    440                    445

Gly Leu Arg Arg Cys Leu Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly  
5                    450                    455                    460                    465

Ser Glu Arg Ser Ser Pro Gln Arg Asp Gly Leu Asp Thr Ser Gly Ser  
10                    470                    475                    480                    485

Thr Gly Ser Pro Gly Ala Pro Thr Ala Ala Arg Thr Leu Val Ser Glu  
15                    490                    495

Pro Ala Ala Asp  
20                    500

<210> 31  
<211> 1029  
<212> DNA  
20 <213> *Homo sapiens*  
<400> 31

25                    atgcggcccg tcgacaaatct caccctctgcg cctgggaaca ccagtctgtg ccccaagagac 60  
tacaaaatca cccaggrrctt ctteccactg ctclacactg ttcgtttttt tttttggactt 120  
atcacaaatg gcttggcgat gaggatttc ttccaaatcc ggatcaaatac aaactttttt 180  
atccccctta agaacacagt cattctgtat ttcttcatgs ttctgactttt tccatccaaa 240  
atccccatgt atgccaaact gggAACAGGA ccactgagaa cttttgtgtg tcaagtttacc 300  
tccgtccatat ttatccac aatgtatatac agtatttcat ttctggact gataactatac 360  
gtatcgctacc agaagaccac caggccatcc aaaaacatcca accccaaaaa ttcttgggg 420  
gtttttttttt tctctgttgt catctggccat ttcatgtttt tttttttttt gcttccatgt 480  
atcttgacca acaggcagcc gagggccaaag aatgtgtaaag aatgtttttt cttttatca 540  
gatgtttttt tagtttggca tggaaatgtat aatccatgtt gtcaagtcat ttcttggatt 600  
atccctttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 660  
gtttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 720  
35                    gttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 780  
ttttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 840  
ttttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 900  
40                    tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 960  
ttttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1020  
ttttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt tttttttttt 1029

<210> 32  
<211> 343  
<212> PTR  
<213> *Homo sapiens*  
50 <400> 32

Met Gln Ala Val Asp Asn Leu Thr Ser Ala Pro Gly Asn Thr Ser Leu  
55                    1                    5                    10                    15

Cys Thr Arg Asp Tyr Lys Ile Thr Gln Val Leu Phe Pro Leu Leu Tyr  
60                    20                    25                    30

Thr Val Leu Phe Phe Val Gly Leu Ile Thr Asn Gly Leu Ala Met Arg  
65                    35                    40                    45

Ile Phe Phe Gln Ile Arg Ser Lys Ser Asn Phe Ile Ile Phe Leu Lys  
70                    50                    55                    60

# ES 2 289 610 T3

	Asn Thr Val Ile Ser Asp Leu Leu Met Ile Leu Thr Phe Pro Phe Lys			
	65	70	75	80
5	Ile Leu Ser Asp Ala Lys Leu Gly Thr Gly Pro Leu Arg Thr Phe Val			
	85	90	95	
	Cys Gln Val Thr Ser Val Ile Phe Tyr Phe Thr Met Tyr Ile Ser Ile			
	100	105	110	
10	Ser Phe Leu Gly Leu Ile Thr Ile Asp Arg Tyr Gln Lys Thr Thr Arg			
	115	120	125	
15	Pro Phe Lys Thr Ser Asn Pro Lys Asn Leu Leu Gly Ala Lys Ile Leu			
	130	135	140	
	Ser Val Val Ile Trp Ala Phe Met Phe Leu Leu Ser Leu Pro Asn Met			
	145	150	155	160
20	Ile Leu Thr Asn Arg Gln Pro Arg Asp Lys Asn Val Lys Lys Cys Ser			
	165	170	175	
	Phe Leu Lys Ser Glu Phe Gly Leu Val Trp His Glu Ile Val Asn Tyr			
	180	185	190	
25	Ile Cys Gln Val Ile Phe Trp Ile Asn Phe Leu Ile Val Ile Val Cys			
	195	200	205	
30	Tyr Thr Leu Ile Thr Lys Glu Leu Tyr Arg Ser Tyr Val Arg Thr Arg			
	210	215	220	
	Gly Val Gly Lys Val Pro Arg Lys Lys Val Asn Val Lys Val Phe Ile			
	225	230	235	240
35	Ile Ile Ala Val Phe Phe Ile Cys Phe Val Pro Phe His Phe Ala Arg			
	245	250	255	
	Ile Pro Tyr Thr Leu Ser Gln Thr Arg Asp Val Phe Asp Cys Thr Ala			
	260	265	270	
40	Glu Asn Thr Leu Phe Tyr Val Lys Glu Ser Thr Leu Trp Leu Thr Ser			
	275	280	285	
45	Leu Asn Ala Cys Leu Asp Pro Phe Ile Tyr Phe Phe Leu Cys Lys Ser			
	290	295	300	
	Phe Arg Asn Ser Leu Ile Ser Met Leu Lys Cys Pro Asn Ser Ala Thr			
	305	310	315	320
50	Ser Leu Ser Gln Asp Asn Arg Lys Lys Glu Gln Asp Gly Gly Asp Pro			
	325	330	335	
	Asn Glu Glu Thr Pro Met			
	340			
55	<210> 33			
	<211> 1077			
60	<212> DNA			
	<213> <i>Homo sapiens</i>			

# ES 2 289 610 T3

<400> 33

```

5      atgtcggtct qctaccgtcc cccaggaaac gagacactgc tgagctggaa gacttcgccc 60
       gcccacggca cagccttccct gctgtctggc gctgtgtggc gctgtgtggc caacggcttc 120
       gtgggtgtggc gcttgccggg ctggccggct gcaegggggc gaccgtggc ggccacgccc 180
       gatgtgtcccc tggcgctggc cgacggggcg gttgtgtggc tcacggccgt ctgtgtggcc 240
       ttttgtggcc ggcaggccctg gctgtgtggc tggcgccgt gcaaggccgt gtaatacily 300
       tgcgcgttca gcatgtacgc cagcggtgtg ctacccggcc tgcgtgtggc gcaacggcc 360
       ctgcgtgtca cccggccctt ctggccggct gggtgtggc gcccggccctt ggcggccggc 420
       ctgtgtgtgg cgggtgtggc ggcggccctg ctgtgtggcc tccggccggc ggttatccgc 480
       cacctgtggc gggaccgggt atggccatgt tggccatccgt cgccgggtca cgcggccggc 540
       cacctgagcc tggagatct gacccgttcc gttgtgtgtat gtcgggtgtc 600
       tacagcggtga cgctggccacg gctggggggc gcccggtggg gtcggggcg gacggggggc 660
       cgnntggggc ggtgtgtggag cgccatcggt ctgtgtgtggc gttgtgtgtc ggcggccctac 720
       cacgcgtgtca acccttctea ggcgggtcgca gcaatggcgtt caccqqaadq gaccttggcg 780
       aacgtggggg gageccggcc ggcggccggc gggggaaata tggccgtggc ttgttcaagt 840
       tctagcggtca accccgggtgtt ctacgtgttc accggccggag atctgtgtcc ccqqgcaggt 900
       ccccggttcc tcaacggccgtt ctccgtggcc tctggggagg cccggggggg cgcccggtgt 960
       agggaaayggc ccstggagct cccaaactacc ctctaggtgtgaa aqgtgtgtggg gcaaggccgc 1020
       ggcataatggag accccggggggg tgggtgtggag auggaaatggc cggaaatggga cttttga 1077

```

<210> 34

25 <211> 358

<212> PTR

<213> *Homo sapiens*

30 <400> 34

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

65

# ES 2 289 610 T3

	165	170	175
	His Ala Ala Ala His Leu Ser Leu Glu Thr Leu Thr Ala Phe Val Leu		
5	180	185	190
	Pro Phe Gly Leu Met Leu Gly Cys Tyr Ser Val Thr Leu Ala Arg Leu		
	195	200	205
	Arg Gly Ala Arg Trp Gly Ser Gly Arg His Gly Ala Arg Val Gly Arg		
10	210	215	220
	Leu Val Ser Ala Ile Val Leu Ala Phe Gly Leu Leu Trp Ala Pro Tyr		
	225	230	235
			240
15	His Ala Val Asn Leu Leu Gln Ala Val Ala Ala Leu Ala Pro Pro Glu		
	245	250	255
	Gly Ala Leu Ala Lys Leu Gly Gly Ala Gly Gln Ala Ala Arg Ala Gly		
	260	265	270
20	Thr Thr Ala Leu Ala Phe Phe Ser Ser Val Asn Pro Val Leu Tyr		
	275	280	285
	Val Phe Thr Ala Gly Asp Leu Leu Pro Arg Ala Gly Pro Arg Phe Leu		
	290	295	300
25	Thr Arg Leu Phe Glu Gly Ser Gly Glu Ala Arg Gly Gly Arg Ser		
	305	310	315
			320
	Arg Glu Gly Thr Met Glu Leu Arg Thr Thr Pro Gln Leu Lys Val Val		
	325	330	335
30	Gly Gln Gly Arg Gly Asn Gly Asp Pro Gly Gly Gly Met Glu Lys Asp		
	340	345	350
	Cys Phe Glu Trp Asp Leu		
35	355		
	<210> 35		
	<211> 1005		
40	<212> DNA		
	<213> <i>Homo sapiens</i>		
	<400> 35		
45	atgctgggya tcatggcatg gaatgcact tgcaaaaact ggctggcagc agaggtgcc 60 ctggaaaaagt actacccttc catttttat gggattgagt tcgttgtggg agtcccttgg 120 aaataccatttgc ttgttttacgg ctacatcttc tctctgaaga actggaaacag cagtaatattt 180 tatttttttttta accctctctgt ctctgactta gctttttctgt gcacccccc catgttgata 240 aggagttatg ccaatggaaa ctggatataat ggeqacgtgc tctgcataaag caaccqataat 300 qigtttcatg ccaaccctta taccaggatt ctcttttctca cttttatccatg catagatcg 360 tacttgataaa ttaaqtatcc ttccggagaa cacccttctgc aaaaqasaga gtttgetatt 420 ttaatcccttgc ttggccatrtg ggtttttagta accttagagt tactaccatc acttccccctt 480 ataaaatccctg ttataacttgc caatqgacc acctgtaaatg attttgcacag ttetggagac 540 cccaatctaca acccttattta cagcatgtgt cttaacactgt tggggttctt tatttcttctt 600 tttqgtatgt gtttcttttttta ttaaagatt gtctcttttttcaaggagag gaataggcag 660 gttqgtactg ctctgccccct tgaaaagccct cttaacttgg tctatcgccgc agtggtaatc 720 ttctctgtgc tttttacacc ctatcacgtc atgcggaaatg tgaggatgc ttcaacgcctq 780 ggaggttggaa agcagtatca gtgcacttgc gtgcateatca actcccttta cattqgtata 840		
50			
55			
60	cgcccttttgg cttttatgtaa cagtgatcactt aacctgttct tttttttttt tttttttttt 900 cacttcagggg acatgtgtat gaatcaactg agacacaaact tcaatccctt tacatccctt 960 agcagatggg ctcatgaact ccatacttca tccagagaaa agtga 1005		
65	<210> 36		
	<211> 334		
	<212> PTR		

# ES 2 289 610 T3

<213> *Homo sapiens*

<400> 36

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

ES 2 289 610 T3

<210> 37

<211> 1296

<212> DNA

5 <213> *Homo sapiens*

<400> 37

<210> 38

35 <211> 431

<212> PTR

<213> *Homo sapiens*

40 <400> 38

Met Gln Ala Leu Asn Ile Thr Pro Glu Gln Phe Ser Arg Leu Leu Arg  
1 5 10 15

45 Asp His Asn Leu Thr Arg Glu Gln Phe Ile Ala Leu Tyr Arg Leu Arg  
20 25 30

Pro Leu Val Tyr Thr Pro Glu Leu Pro Gly Arg Ala Lys Leu Ala Leu

55

60

65

# ES 2 289 610 T3

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

60

65

ES 2 289 610 T3

Asn Lys Thr Phe Ser Pro Ala Gin Arg His Gly Asn Ser Gly Ile Thr  
 355 360 365  
 Met Met Arg Lys Lys Ala Lys Phe Ser Leu Arg Glu Asn Pro Val Glu  
 370 375 380  
 Glu Thr Lys Gly Glu Ala Phe Ser Asp Gly Asn Ile Gln Val Lys Leu  
 385 390 395 400  
 Cys Glu Cln Thr Glu Glu Lys Lys Leu Lys Arg His Leu Ala Leu  
 405 410 415  
 Phe Arg Ser Glu Leu Ala Glu Asn Ser Pro Leu Asp Ser Gly His  
 420 425 430  
 15  
 <210> 39  
 <211> 24  
 20 <212> DNA  
 <213> *Homo sapiens*  
 <400> 39  
 25 ctgtgtacag cagttcgcag agtg  
 <210> 40  
 30 <211> 24  
 <212> DNA  
 <213> *Homo sapiens*  
 35 <400> 40  
 gagtgccagg cagaggcagg agac  
 40 <210> 41  
 <211> 31  
 <212> DNA  
 <213> *Homo sapiens*  
 45 <400> 41  
 cccgaattcc tgcttgctcc cagcttggcc c  
 50 <210> 42  
 <211> 32  
 <212> DNA  
 55 <213> *Homo sapiens*  
 <400> 42  
 60 tgtggatct gctgtcaaag gtcccatcc gg  
 <210> 43  
 <211> 20  
 65 <212> DNA  
 <213> *Homo sapiens*



# ES 2 289 610 T3

<210> 49		
<211> 23		
<212> DNA		
5 <213> <i>Homo sapiens</i>		
<400> 49		
10           gtgttagatct tccgggtgcac agg		23
<210> 50		
<211> 21		
15 <212> DNA		
<213> <i>Homo sapiens</i>		
<400> 50		
20           gcaatgcagg tcatagtgag c		21
<210> 51		
25 <211> 27		
<212> DNA		
<213> <i>Homo sapiens</i>		
30 <400> 51		
tggagcatgg tgacggaaat gcagaag		27
35 <210> 52		
<211> 27		
<212> DNA		
40 <213> <i>Homo sapiens</i>		
<400> 52		
45           gtgatgagca ggtcactgag cgccaag		27
<210> 53		
<211> 23		
50 <212> DNA		
<213> <i>Homo sapiens</i>		
<400> 53		
55           gcaatgcagg cgcttaacat tac		23
<210> 54		
60 <211> 23		
<212> DNA		
<213> <i>Homo sapiens</i>		
65 <400> 54		
ttgggttaca acctgaagg ca		22

# ES 2 289 610 T3

<210> 55		
<211> 23		
<212> DNA		
5 <213> <i>Homo sapiens</i>		
<400> 55		
10               actccgtgtc cagcaggact ctg		23
<210> 56		
<211> 24		
15 <212> DNA		
<213> <i>Homo sapiens</i>		
<400> 56		
20               tgcgtttcc tggaccctca cgta		24
<210> 57		
25 <211> 29		
<212> DNA		
<213> <i>Homo sapiens</i>		
30 <400> 57		
caggccttgg attttaatgt cagggatgg		
		29
35 <210> 58		
<211> 27		
<212> DNA		
40 <213> <i>Homo sapiens</i>		
<400> 58		
45               ggagagtcag ctctgaaaga attcagg		27
<210> 59		
<211> 27		
50 <212> DNA		
<213> <i>Homo sapiens</i>		
<400> 59		
55               tgatgtatcg ccagatacta atagcac		27
<210> 60		
60 <211> 27		
<212> DNA		
<213> <i>Homo sapiens</i>		
65 <400> 60		
cctgattcat ttagtgtgaga ttgagac		
		27

## ES 2 289 610 T3

<210> 61  
<211> 21  
<212> DNA  
5 <213> *Homo sapiens*  
  
<400> 61  
  
10 gacaggtacc ttgccatcaa g 21  
  
<210> 62  
<211> 22  
15 <212> DNA  
<213> *Homo sapiens*  
  
<400> 66  
20 ctgcacaatg ccagtataa gg 22  
  
<210> 63  
25 <211> 27  
<212> DNA  
<213> *Homo sapiens*  
  
30 <400> 63  
  
ctgacttcct gttcctggca gcagcgg 27  
  
35 <210> 64  
<211> 27  
<212> DNA  
40 <213> *Homo sapiens*  
  
<400> 64  
  
45 agaccagcca gggcacgctg aagagt 27  
  
<210> 65  
<211> 32  
50 <212> DNA  
<213> *Homo sapiens*  
  
<400> 65  
  
55 gatcaagctt ccatcctact gaaaccatgg tc 32  
  
<210> 66  
60 <211> 35  
<212> DNA  
<213> *Homo sapiens*  
  
65 <400> 66  
  
gatcagatct cagttccaat attcacacca ccgtc 35

ES 2 289 610 T3

<210> 67  
 <211> 22  
 <212> DNA  
 5 <213> *Homo sapiens*  
 <400> 67  
 10 ctggtgtgct ccatggcata cc 22  
 <210> 68  
 <211> 22  
 15 <212> DNA  
 <213> *Homo sapiens*  
 <400> 68  
 20 gtaaggctcc cagaacgaga gg 22  
 <210> 69  
 25 <211> 24  
 <212> DNA  
 <213> *Homo sapiens*  
 30 <400> 69  
 cagcgcaggg tgaaggctga gagc 24  
 35 <210> 70  
 <211> 24  
 <212> DNA  
 40 <213> *Homo sapiens*  
 <400> 70  
 45 ggcacacctgct gtgaccctgtc cagg 24  
 <210> 71  
 <211> 22  
 50 <212> DNA  
 <213> *Homo sapiens*  
 <400> 71  
 55 gtcctgccac ttcgagacat gg 22  
 <210> 72  
 <211> 23  
 60 <212> DNA  
 <213> *Homo sapiens*  
 65 <400> 72  
 gaaaacttctc tgcccttacc gtc 23

ES 2 289 610 T3

<210> 73  
<211> 26  
<212> DNA  
5 <213> *Homo sapiens*  
  
<400> 73  
  
10 ccaacaccag catccatggc atcaag 26  
  
<210> 74  
<211> 27  
15 <212> DNA  
<213> *Homo sapiens*  
  
<400> 74  
20 ggagagtcag ctctgaaaga attcagg 27  
  
25  
  
30  
  
35  
  
40  
  
45  
  
50  
  
55  
  
60  
  
65