



19



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

45 Fecha de la publicación del folleto de la patente:
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.

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 de productos farmacéuticos: a partir de aproximadamente 20 de los 100 GPCRs conocidos se han desarrollado el 60% 20 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 30 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.

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.

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

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).

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.

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.

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 definiciones para estos términos, las siguientes definiciones prevalecerán:

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

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

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 significará aquél que no es producido naturalmente por un mamífero (por ejemplo, y sin ser una limitación, un humano) o un virus.

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.

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

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.

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

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.

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.

A. Identificación de GPCRs humanos

Los esfuerzos del proyecto Genoma Humano han llevado a la identificación de una plétora 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 GPCRs, descritos aquí, fueron descubiertos revisando la base de datos GenBankTM, mientras que otros GPCRs fueron descubiertos usando una secuencia de ácido nucleico de un GPCR, previamente secuenciado, para llevar a cabo una búsqueda BLASTTM en la base de datos EST. La Tabla A, a continuación, lista los GPCRs huérfanos endógenos, descritos con un GPCR homólogo respectivo de GPCR:

(Tabla pasa a página siguiente)

TABLA A

GPRCs Humanos Huérfanos Descritos	Número de Acceso Identificado	Marco Abierto de Lectura (Pares de Bases)	Porcentaje de Homología al GPCR Designado	Referencia al GPCR Homólogo (N° de Acceso)
hARE-3	AL033379	1.260 pb	52.3% LPA-R	U92642
hARE-4	AC006087	1.119 pb	36% P2Y5	AF000546
hARE-5	AC006255	1.104 pb	32% <i>Oryzias latipes</i>	D43633
hGPR27	AA775870	1.128 pb		
hARE-1	A1090920	999 pb	43% KIAA0001	D13626
hARE-2	AA359504	1.122 pb	53% GPR27	
hPPR1	H67224	1.053 pb	39% EBI1	L31581
hG2A	AA754702	1.113 pb	31% GPR4	L36148
hRUP3	AL035423	1.005 pb	30% <i>Drosophila melanogaster</i>	2133653
hRUP4	AI307658	1.296 pb	32% pNPGPR 28% y 29 % <i>Pez Zebra</i> Ya y Yb, <i>respectivamente</i>	NP_004876 AAC41276 y AAB94616
hRUP5	AC005849	1.413 bp	25% DEZ 23% FMLPR	Q99788 P21462
hRUP6	AC005871	1.245 bp	48% GPR66	NP_006047
hRUP7	AC007922	1.173 pb	43% H3R	AF140538
hCHN3	EST 36581	1.113 pb	53% GPR27	
hCHN4	AA804531	1.077 pb	32% trombina	4503637
hCHN6	EST 2134670	1.503 pb	36% edg-1	NP_001391
hCHN8	EST 764455	1.029 pb	47% KIAA0001	D13626
hCHN9	EST 1541536	1.077 pb	41% LTB4R	NM_000752
hCHN10	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 in-
 5 fraactivado, 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
 10 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 identi-
 15 ficación 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
 20 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
 25 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).
 30 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.

35 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
 40 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.

45 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
 50 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, [³⁵S]GTPγS, puede usarse para monitorizar uniones
 55 mejoradas a membranas que expresan receptores activados constitutivamente. Se ha descrito que [³⁵S]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.

60 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 pro-
 65 teína 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.

a. *Gs* y *Gi*

Gs estimula la enzima adenilil ciclasa. Gi (y Go), al contrario, inhiben esta enzima. La adenilil ciclasa 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* (3ª 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 descendería 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., β -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 β -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 PIP₂, liberando dos mensajeros intracelulares: diacilglicerol (DAG) e inositol 1,4,5-trifosfato (IP₃). La acumulación creciente de IP₃ 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* (3ª Ed.) Nichols, J.G. et al eds. Sinauer Associates, inc. (1992). Los ensayos que detectan la acumulación de IP₃ 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 IP₃). 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.

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 Proteína 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 esta 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

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 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 *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 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

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 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

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:

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

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

2. Clonación de Longitud Total

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

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™, y el molde para PCR fue ADNc de bazo humano Marathon-ready™ 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'-CTGTGTACAGCAGTTCGCAGAGTG-3' (SEQ.ID.NO.:39; 1º ciclo PCR) 5'-GAGTGCCAGGCAGAGCAGGTAGAC-3' (SEQ.ID.NO.:40; 2º ciclo PCR). La PCR se llevó a cabo usando el kit Advantage™ 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°C durante 5 seg. y 70°C 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™ (USB Amersham; se seguirán las instrucciones del fabricante) y la secuencia se comparó con la secuencia presentada. Se detectará la expresión de *G2A* humano sondeando un dot-blot de ARN (Clontech; se seguirán las instrucciones del fabricante) con el fragmento marcado con P³².

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

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, 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:

5'-CCCGAATTCCTGCTTGCTCCAGCTTGGCCC-3' (SEQ.ID.NO.:41; sentido) y

5'-TGTGGATCCTGCTGTCAAAGGTCCCATTCCGG-3' (SEQ.ID.NO.:42; antisentido).

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* a continuación) y se secuenció (*ver*, SEQ.ID.NO.:33).

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

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

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

5'-TGCATAGACAATGGGATTACAG-3' (SEQ.ID.NO.:44; antisentido).

La PCR se realizó usando TaqPlus™ Precision™ 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.

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™ (Amsham) y los cebadores SP6/T7 (Stratagene). El análisis de la secuencia reveló que el fragmento de PCR era de hecho una forma de A1307658 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'-TCACAATGCTAGGTGTGGTCTGGCTGGTGGCAGTCATCGTAG
GATCACCCATGTGGCACGTGCAACAACCTTGAGATCAAATATGAC

TTCCTATATGAAAAGGAACACATCTGCTGCTTAGAAGAGTGGACC
 AGCCCTGTGCACCAGAAGATCTACACCACCTTCATCCTTGTCA
 5 CTCTTCCTCCTGCCTCTTATGGTGATGCTTATTCTGTACGTAAAA
 TTGGTTATGAACTTTGGATAAAGAAAAGAGTTGGGGATGGTTCA
 GTGCTTCGAACTATTCATGGAAGAAATGTCCAAAATAGCCAGG
 10 AAGAAGAAACGAGCTGTCATTATGATGGTGACAGTGGTGGCTCT
 CTTTGCTGTGTGCTGGGCACCATTCATGTTGTCCATATGATGAT
 TGAATACAGTAATTTTGAAGGAATATGATGATGTCACAATCAA
 15 GATGATTTTTGCTATCGTGCAAATTATTGGATTTTCCAACTCCATC
 TGTAAATCCCATTGTCTATGCA-3' (SEQ.ID.NO.:45).

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

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

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

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

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

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

35 para 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 pCRII-TOPO™ (Invitrogen) y se secuenciaron utilizando los cebadores SP6/T7 (Stratagene) y algunos cebadores internos. El producto 3' RACE contenía una cola poli(A) y un marco abierto de lectura completo finalizando en un
 40 codón de terminación TAA. El producto 5' RACE contenía un final 5' incompleto; es decir, el codón de iniciación ATG no estaba presente.

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

45 5'-GCAATGCAGGTCATAGTGAGC-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:

50 5'-TGGAGCATGGTGACGGGAATGCAGAAG-3' (SEQ.ID.NO.:51; oligo 6) y,

5'-GTGATGAGCAGGTCAGTCTGAGCGCCAAG-3' (SEQ.ID.NO.:52; oligo 7).

55 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'-GCAATGCAGGCGCTTAACATTAC-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
 60 el siguiente cebador antisentido:

5'-TTGGGTTACAATCTGAAGGGCA-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'-TGCGTGTTCTGACCCTCACGTG-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°C 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'-CAGGCCTTGGATTTTAATGTCAGGGATGG-3' (SEQ.ID.NO.: 57) y

25 5'-GGAGAGTCAGCTCTGAAAGAATTCAGG-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'-CCTGATTCATTTAGGTGAGATTGAGAC-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.

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)

60 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'-GTAAGCCTCCCA 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* (Seq. Id. Nos.: 1 y 2)

La longitud total de *hARE-3* se clonó mediante PCR utilizando los cebadores específicos de *hARE-3* 5'-gat-caagcttCCATCCTACTGAAACCATGGTC-3' SEQ.ID.NO.:65 (sentido, los nucleótidos en minúscula representan Hind III saliente, ATG como codón de iniciación) y 5'-gatcagatctCAGTTCCAATATTCACACCACCGTC-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* (Seq. 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'-GAAACTTCTCTGCCC 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).

Ejemplo 2

Expresión de receptores

Aunque está disponible una variedad de células para la expresión de proteínas, es más preferible que se utilicen 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, 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 1X10⁷ células 293T por placa de 150 mm. En el día dos, se prepararán dos 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 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 CO₂. 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 CO₂. Después de 72 h de incubación, las células entonces pueden ser recolectadas y utilizadas para análisis.

Ejemplo 3

Distribución tisular de los GPCRS humanos descritos

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

1. *Análisis Dot-Blot*

Usando un formato de dot-blot para tejidos humanos disponible comercialmente, se sondearon 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

un kit Prime-It II™ Random Primer Labelling (Stratagene, #300385); según las instrucciones del fabricante. Se hibridizó 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 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

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

TABLA C

GPCR HUÉRFANO	Distribución tisular (niveles más elevados, relativos a otros tejidos en el dot-blot)
hARE-3	Cerebelo izquierdo, cerebelo derecho, testículos, accumbens
hGPCR3	Corpus collusum, caudate nucleus, hígado, corazón, septum interventricular
hARE-2	Cerebelo izquierdo, cerebelo derecho, Substantia
hCHN8	Cerebelo izquierdo, cerebelo derecho, hígado, pulmón

2. RT-PCR

a. *hRUP3*

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 μ 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:

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

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

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

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.

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 μ 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.

Para analizar los productos RT-PCR se cargaron 20 μ 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™ Random Primer Labeling (Stratagene) y se purificó utilizando las microcolumnas ProbeQuant™ 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*.

c. *hRUP5*

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

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

5'-AGACCAGCCAGGGCAGCTGAAGAGTG-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 μ 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 μ 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).

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:

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

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

Se utilizaron los paneles de ADNc de múltiples tejidos humanos (MTC, Clontech) como moldes. La PCR se realizó utilizando TaqPlus Precision™ polimerasa (Stratagene; se seguirán las instrucciones de fabricación) en una reacción de 40 μ 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.

ES 2 289 610 T3

Se cargaron 20 μ 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 Budapest para el Reconocimiento Internacional del Depósito de Microorganismos a los Fines del Procedimiento en
10 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] |
| • US 60123946 B [0060] | • US 60157280 B [0060] |
| 35 • US 60123949 B [0060] | • US 60157294 B [0060] |
| • US 60136436 B [0060] | • US 60157281 B [0060] |
| 40 • US 60136437 B [0060] | • US 60157293 B [0060] |
| • US 60136439 B [0060] | • US 60157282 B [0060] |
| • US 60136567 B [0060] | • US 09417044 B [0060] |
| 45 • 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

REIVINDICACIONES

1. 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

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.

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 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.

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.

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 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.

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 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.

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

Fig. 1A

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

Fig. 1B

Fig. 2A

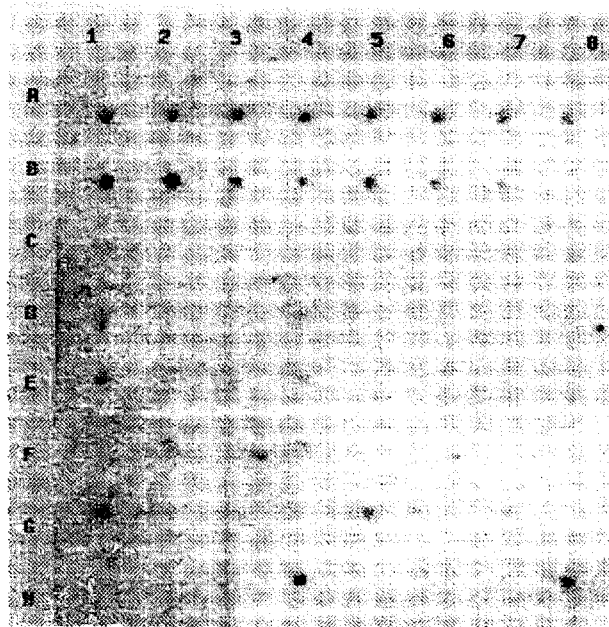
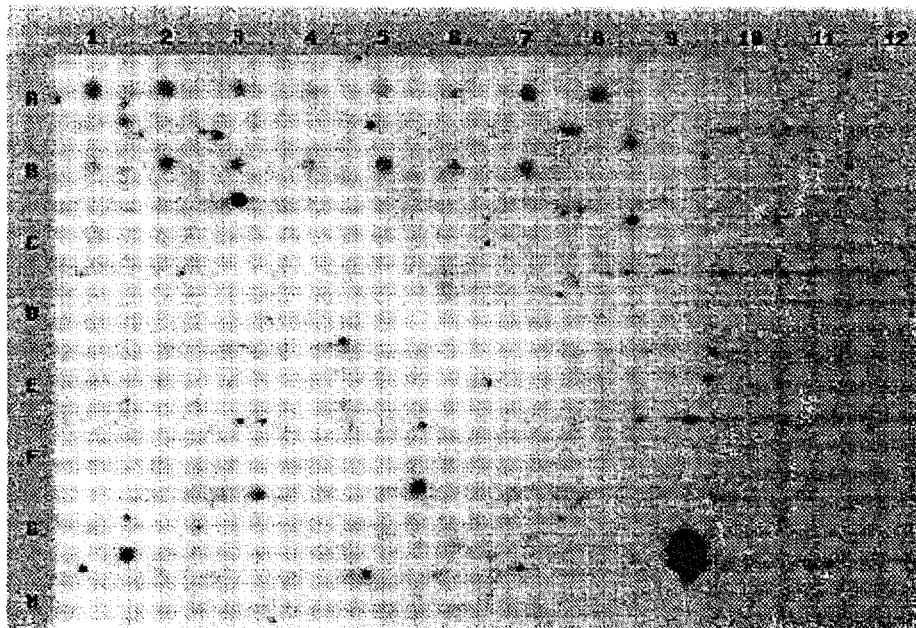


Fig. 2B



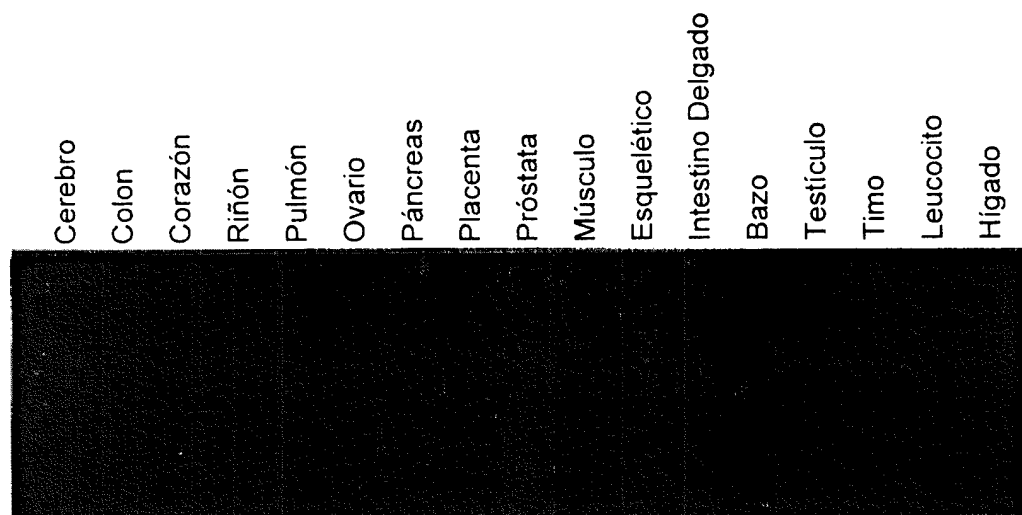


Fig. 3

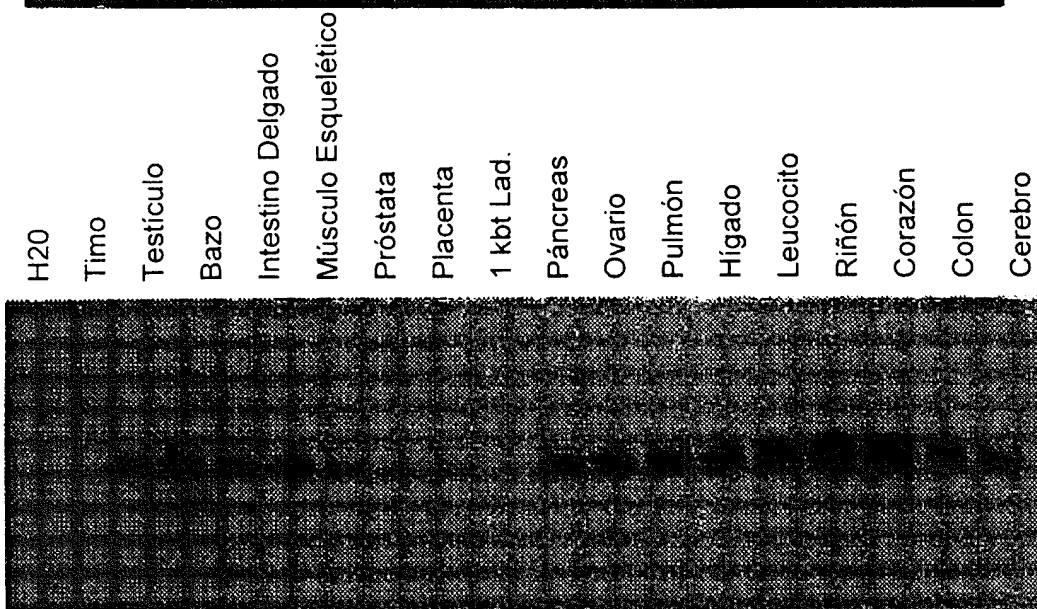


Fig. 4

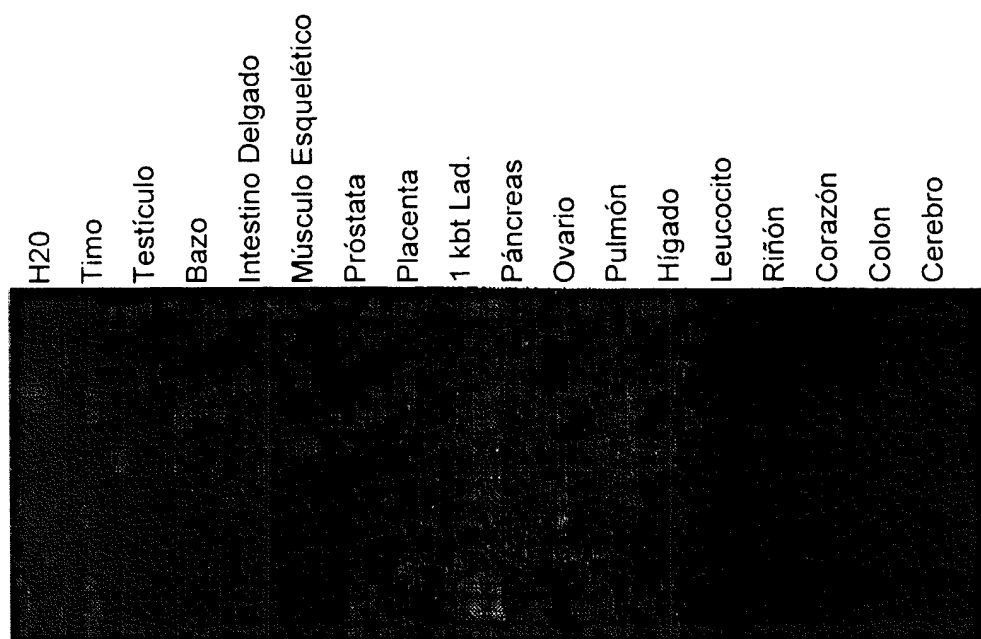


Fig. 5

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 INVENCION: Receptores acoplados a Proteína G humanos huérfanos
- (iii) NÚMERO DE SECUENCIAS: 74
- (iv) DIRECCIÓN DE CORRESPONDENCIA:
- 15 (A) DESTINATARIO: Arena Pharmaceuticals, Inc.
(B) DIRECCIÓN: 6166 Nancy Ridge Drive
(C) CIUDAD: San Diego
20 (D) ESTADO: California
(E) PAÍS: Estados Unidos de América
(F) (ZIP): 92121
- (v) FORMATO LEGIBLE POR ORDENADOR:
- 25 (A) TIPO DE MEDIO: disquette
(B) (B) ORDENADOR: IBM PC compatible
(C) SISTEMA OPERATIVO: PC-DOS/MS-DOS
30 (D) PROGRAMA: PatentIn Release #1.0, Versión #1.30
- (vi) DATOS DE LA SOLICITUD ACTUAL:
- (A) NÚMERO DE LA SOLICITUD: US
35 (B) FECHA DE REGISTRO:
(C) CLASIFICACIÓN:
- (viii) INFORMACIÓN DEL ABOGADO/AGENTE:
- 40 (A) NOMBRE: Burgoon, Richard P.
(B) NÚMERO DE REGISTRO: 34.787
- (ix) INFORMACIÓN PARA TELECOMUNICACIÓN:
- 45 (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
55 (C) TIPO DE HEBRA: sencilla
(D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: DNA (genomic)
- 60 (iii) DESCRIPCIÓN DE LA SECUENCIA: SEQ NO ID: 1:

ATGGTCTTCT CGGCAGTGT GACTGCGTTC CATACCGGGA CATCCAACAC AACATTGTTC 60

5 GTGTATGAAA ACACCTACAT GAATATTACA CTCCCTCCAC CATTCCAGCA TCCTGACCTC 120
 AGTCCATTGC TTAGATATAG TTTTGAAACC ATGGCTCCCA CTGGTTTGAG TTCCTTGACC 180
 10 GTGAATAGTA CAGCTGTGCC CACAACACCA GCAGCATTTA AGAGCCTAAA CTGCGCTCTT 240
 CAGATCACCC TTTCTGCTAT AATGATATTC ATTCTGTTTG TGTCTTTTCT TGGGAACCTG 300
 15 GTTGTTTGCC TCATGGTTTA CCAAAAAGCT GCCATGAAGT CTGCAATTAA CATCCTCCTT 360
 GCCAGCCTAG CTTTTCGAGA CATGTTGCTT GCAGTCTGA ACATGCCCTT TCCCCTGGTA 420
 20 ACTATTCTTA CTACCCGATG GATTTTGGG AAATCTTCT GTAGGGTATC TGCTATGTTT 480
 TTCTGTTTAT TTGTGATAGA AGGAGTAGCC ATCCTGCTCA TCATTAGCAT AGATAGGTTT 540
 25 CTTATTATAG TCCAGAGGCA GGATAAGCTA AACCCATATA GAGCTAAGGT TCTGATTGCA 600
 GTTCTTGGG CAACTTCCTT TTGTGTAGCT TTTCCTTTAG CCGTAGGAAA CCCCAGCCTG 660
 30 CAGATACCTT CCGAGCTCC CCAGTGTGTG TTGGGTACA CAACCAATCC AGGCTACCAG 720
 GCTTATGTGA TTTTGATTTC TCTATTCTT TTCTTCATAC CCTTCCTGGT AATACTGTAC 780
 35 TCATTATGG GCATACTCAA CACCCTTCGG CACAATGCCT TGAGGATCCA TAGCTACCCT 840
 GAAGGTATAT GCCTCAGCCA GGCAGCAAA CTGGGTCTCA TGAGTCTGCA GAGACCTTTC 900
 40 CAGATGAGCA TTGACATGG CTTTAAACA CGTGCTTCA CCACTATTTT GATTCTCTTT 960
 GCTGTCTTCA TTGTCTGCTG GGCCTCCTC ACCACTTACA GCCTTGTGGC
 AACATTCA0T1020
 45 AAGCACTTTT ACTATCAGCA CAACTTTTTT GAGATTAGCA CCTGGCTACT GTGGCTCTGC1080
 TACCTCAAGT CTGCATTGAA TCCGCTGATC TACTACTGGA GGATTAAGAA ATTCCATGAT1140
 50 GCTTGCCTGG ACATGATGCC TAAGTCTTC AAGTTTTTGC CGCAGCTCCC TGGTCACACA1200
 AAGCGACGGA TACGTCTAG TGCTGTCTAT GTGTGTGGGG AACATGGGAC GGTGTGTGTA1260

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

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

(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

Thr Thr Phe Val Val Tyr Glu Asn Thr Tyr Met Asn Ile Thr Leu Pro
20 25 30

10 Pro Pro Phe Gln His Pro Asp Leu Ser Pro Leu Leu Arg Tyr Ser Phe
35 40 45

15 Glu Thr Met Ala Pro Thr Gly Leu Ser Ser Leu Thr Val Asn Ser Thr
50 55 60

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

25 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
115 120 125

30 Leu Leu Ala Val Leu Asn Met Pro Phe Ala Leu Val Thr Ile Leu Thr
130 135 140

Thr Arg Trp Ile Phe Gly Lys Phe Phe Cys Arg Val Ser Ala Met Phe
145 150 155 160

35 Phe Trp Leu Phe Val Ile Glu Gly Val Ala Ile Leu Leu Ile Ile Ser
165 170 175

Ile Asp Arg Phe Leu Ile Ile Val Gln Arg Gln Asp Lys Leu Asn Pro
180 185 190

40 Tyr Arg Ala Lys Val Leu Ile Ala Val Ser Trp Ala Thr Ser Phe Cys
195 200 205

Val Ala Phe Pro Leu Ala Val Gly Asn Pro Asp Leu Gln Ile Pro Ser
210 215 220

Arg Ala Pro Gln Cys Val Phe Gly Tyr Thr Thr Asn Pro Gly Tyr Gln
225 230 235 240

45 Ala Tyr Val Ile Leu Ile Ser Leu Ile Ser Phe Phe Ile Pro Phe Leu
245 250 255 260

Ala Val Phe Ile Val Cys Trp Ala Pro Phe Thr Thr Tyr Ser Leu Val
265 270 275 280

50 Ala Thr Phe Ser Lys His Phe Tyr Tyr Gln His Asn Phe Phe Glu Ile
285 290 295 300

Ser Thr Trp Leu Leu Trp Leu Cys Tyr Leu Lys Ser Ala Leu Asn Pro
305 310 315 320

55 Leu Ile Tyr Tyr Trp Arg Ile Lys Lys Phe His Asp Ala Cys Leu Asp
325 330 335 340

Met Met Pro Lys Ser Phe Lys Phe Leu Pro Gln Leu Pro Gly His Thr
345 350 355 360

60 Lys Arg Arg Ile Arg Pro Ser Ala Val Tyr Val Cys Gly Glu His Arg
365 370 375 380 385 390 395 400 405 410 415

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1119 pares de base

ES 2 289 610 T3

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

(ii) TIPO DE MOLÉCULA: DNA (genomic)

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

```

ATGTTAGCCA ACAGCTCCTC AACCAACAAT TCTGTTCTCC CCGTCTCTGA CTACCGACCT  60
ACCCACCGCC TGCACCTTGGT GGTCTACAGC TTGGTGCTGG CTGCGGGGCT CCCCCCAAC 120
GCGCTAGCCC TCTGGGTCTT CCGCGCGCG CTGCGGTGC ACTCGGTGGT GAGCGGTAC 180
ATGTGTAACC TGGCGCCAG CGACCTGCTC TTCACCTCT CGCTGCCCCT TCGTCTCTCC 240
TACTACGCAC TGCACCACTG GCCCTTCCCC GACCTCTGT GCCAGACGAC GCGCGCCATC 300
TTCCAGATGA ACATGTACGG CAGCTGCATC TTCTGATGC TCATCAAGT GGACCGCTAC 360
GCCGCCATCG TGCACCCGCT GCGACTGCGC CAGCTGCGC GCGCCCGCT GCGCGGCTG 420
CTCTGCTGG GCGTGTGGG GCTCATCTG GTGTTTGGC TGCCCGCCG CCGGTGCAC 480
AGGCCCTCG GTTGGCGTA CCGGACCTC GAGGTGCGC TATGCTTGA GAGCTTCAG 540
GACGAGCTGT GGAAGGCAG GCTGCTGCC CTCGTGCTG TGGCGAGG CTTGGGCTT 600

CTGCTGCCCC TGGCGCGGT GGTCTACTG TCGGCGGAG TCTTCTGGAC GCTGGCGCG 660
CCCGACGCCA CCGAGGCCA GCGCGCGG AAGACCGTG CCTCTCTCT GGCTAACCTC 720
GTCTATCTCC TGTGTGCTT CCGGCTTAC AACAGCACG TGGCGGTCTA CCGGCTGCTG 780
CGAGGCAAGC TGGTGGCGC CAGCGTGCCT GCGCGCATC GCGTGGCGG GGTGCTGATG 840
GTGATGCTGC TGCTGGCGG CCGCAACTG GTGCTGGAC CGCTGGTOTA CTACTTTAGC 900
GCCGAGGGCT TCCCAACAC CCGCGCGCG CTGGGCACTC CGCACCGGC CAGGACCTG 960
GCCACCAACG GAGCGCGCG GCGCTGCG CAATCGAAG GGTCCGCCCT CACCACCGAC 1020
GCCACGAGG CGGATGCGG CAGTCAGGG CTGCTCGAC CCGCGACTC CCACTCTCTG 1080
TCTTCCTTCA CACAGTGTCC CCAGGATTCC GCCCTCTGA 1119

```

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

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

(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
20 25 30

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

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

Thr Gly Ala Ile Phe Gln 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

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

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

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

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

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
305 310 315 320

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

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

ES 2 289 610 T3

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1107 base de pares

(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: 5:

```

ATGCCCAACT CCACAGGGCT GAAGCCTCA GAAGTCGCAG GCTCGTTGGG GTTGATCCTG 60
GCAGCTGTTC TGGAGGTGGG GGCAGTCTG GGCACGGCG CGCTGCTGGT CGTGGTCTG 120
CGCAGCCCGG GACTGCGGA CGCGCTCTAC CTGGCGCACC TGTGCGTGT GACCTGCTG 180
GGGGCCGCT CCATCATGCC GCTGGGCTG CTGGCGCAC CGCGCCCGG GCTGGGCGC 240
GTGGCGCTGG GCGCCGCGC ATGCGCGGC GCTCGCTTC TCTCGCGC TCTGCTGCG 300
GCCTGCACGC TCGGGTGGC CGCACTTGG CTGGCAGCT ACCGCTCAT CGTGCACCG 360
CTGCGGCCAG GCTCGCGGC GCGGCTGTG CTGCTGCTCA CCGCGTGTG GCGCGCGCG 420
GGACTGCTGG GCGCGCTCT CCTGCTGGC CGCGCGCGG CACCGCCCC TGCTCTGCT 480
CGCTGCTCG TCCTGGCTGG GGGCCTCGG CCCTTCGCG CGCTCTGGG CCTGCTGCG 540
TTGCGCTGC CGCCCTCTT GCTGCTGCG GCCTACGGG GCATCTTCGT GGTGGCGCGT 600
CGCGTGCCC TGAGGCCCC ACGCGCGCG CGCGGTCCC GACTCGCTC GACTCTCTG 660
GATAGCCGC TTTCATCTT GCGCGGCTC CGGCTCGCC TGCCCGGGG CAAGCGCGC 720
CTGGCCCCAG CGCTGGCGT GGGCCAATT GCAGCCTGT GCTGCCTTA TGGCTGCGG 780
TGCTTGGCG CGCAGCGCG GGCGCGGAA GCGGAGCGG CTGTCACCT GGTCGCTAC 840
TCGGCTTCG CGGCTACCC CTCTCTGTAC GGGCTGCTG AGCGCCCGT GCGCTTGGCA 900
CTGGGCGGC TCTCTGCGG TGCACTGCT GGACCTGTG GGCGCTGAC TCGCAAGCC 960
TGGCACCGC GGGACTCTT GCAATGCTC CAGAGACCC CAGAGGGCC TCGGTAGGC 1020
CCTTCTGAG CTCCAGAACA GACCCCGAG TTGGCAGGAG GCGGAGCCC CGCATACCAG 1080
GGGCCACCT AGAGTTCTT CTCCTGA
1107

```

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 368 aminoácidos

(B) TIPO:aminoácido

(C) TIPO DE HEBRA:

(D) TOPOLOGÍA: no relevante

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

(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 CTGCTGTGCC TGGCCTCCCT CATCATTCCT 60
ACTAACACAC TAGTGGCTGT GGCTGTGCTG CTGTTGATCC ACAAGAATGA TGGTGTCAAT 120
CTCTGCTTCA CCTTGAATCT GGCTGTGGCT GACACCTTGA TTGTTGTGCG CATCTCTGGC 180
CTACTCACAG ACCAGCTCTC CAGCCCTTCT CGGCCACAC AGAAGACCCT GTGCAGCCTG 240
CGATGGCAT TTGCACTTC CTCCCAOCT GCCTCTGTCC TCACGGTCAT GCTGATCACC 300
TTTGACAGGT ACCTTGCCAT CAAGCAGCCC TTCCGCTACT TGAAGATCAT GAGTGGGTTT 360
GTGGCCGGGG CCTGCATTGC CGGCTGTGG TTAGTGTCTT ACCTCATTCG CTTCCTCCCA 420
CTCGGAATCC CCATGTTCCA GCAGACTGCC TACAAAGGCG AGTGACAGCT CTTCCTGTA 480
TTTCACCCCTC ACTTCGTGCT GACCCCTCTC TCGTTGGCT TCTTCCCAGC CATGCTCCTC 540
TTTGCTTCT TCTACTGCGA CATGCTCAAG ATTGCCTCCA TGCACAGCCA GCAGATTGGA 600
AAGATGGAAC ATGCAGGAGC CATGGCTGGA GTTTATCGAT CCCCACGGAC TCCCAGCGAC 660
TTCAAAGCTC TCCGTACTGT GTCTGTCTC ATTGGGAGCT TTGCTCTATC CTGGACCCCC 720
TTCTTATCA CTGGCATTGT GCAGGTGGCC TGCCAGGAGT GTCACCTCTA CCTAGTGTG 780
GAACGGTACC TGTGGCTGCT CGGCTGGGC AACTCCCTGC TCAACCCACT CATCTATGCC 840

TATTGGCAGA AGGAGGTGCG ACTGCAGCTC TACCACATGG CCCTAGGAGT GAAGAAGGTG 900
CTCACCTCAT TCCTCTCTT TCTCTCGGC AGGAATTGTG GCCCAGAGAG GCCCAGGAA 960
AGTTCCTGTC ACATCGTCAC TATCTCCAGC TCAGAGTTTG ATGGCTAA 1008

```

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1008 pares de bases

(B) TIPO: ácido nucleico

(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	
	Gln Leu Ser Ser Pro Ser Arg Pro Thr Gln Lys Thr Leu Cys Ser Leu	65	70	75	80
15	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	115	120	125	
25	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	145	150	155	160
30	Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro	165	170	175	
	Ala Met Leu Leu Phe Val Phe Phe Tyr Cys Asp Met Leu Lys Ile Ala	180	185	190	
35	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	225	230	235	240
45	Phe Leu Ile Thr Gly Ile Val Gln Val Ala Cys Gln Glu Cys His Leu	245	250	255	
	Tyr Leu Val Leu Glu Arg Tyr Leu Trp Leu Leu Gly Val Gly Asn Ser	260	265	270	
50	Leu Leu Asn Pro Leu Ile Tyr Ala Tyr Trp Gln Lys Glu Val Arg Leu	275	280	285	
	Gln Leu Tyr His Met Ala Leu Gly Val Lys Lys Val Leu Thr Ser Phe	290	295	300	
55	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

(ii) TIPO DE MOLÉCULA: DNA (genomic)

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

```

ATGGACACTA CCATGGAAGC TGACCTGGGT GCCACTGGCC ACAGGCCCCG CACAGAGCTT  60
GATGATGAGG ACTCCTACCC CCAAGGTGGC TGGGACACGG TCTTCCTGGT GGCCTGCTG 120
CTCCTTGGGC TGCCAGCCAA TGGGTTCATG GCGTGGCTGG CCGGCTCCCA GGCCCGGCAT 180
'GGAGCTGGCA CGCGTCTGGC GCTGCTCCTG CTCAGCCTGG CCTCTCTGA CTTCTTGTTC 240
CTGGCAGCAG CGGCCTTCCA GATCCTAGAG ATCCGGCATG GGGGACACTG GCCGCTGGGG 300
ACAGCTGCCT GCGGCTTCTA CTACTTCTTA TGGGGCGTGT CCTACTCCTC CGGCCTCTTC 360
CTGCTGGCCG CCTCAGCCT CGACCGCTGC CTGCTGGCGC TGTGCCACAA CTGGTACCCT 420
GGGCACCGCC CAATCCGCT GCGCCTCTGG GTCTGCCTCG GTGTCTGGGT GCTGGCCACA 480
CTCTTCAGCG TGCCCTGGCT GGTCTTCCCC GAGGCTGCCG TCTGGTGGTA CGACCTGGTC 540
ATCTGCTTGG ACTTCTGGGA CAGCGAGGAG CTGTGCTGTA GGATGCTGGA GGTCTCTGGG 600
GGCTTCTTGC CTTTCTCTCT GCTGCTCTGC TGCCAGTGC TCACCCAGGC CACAGCCTGT 660
CGCACCTGCC ACCGCCAACA GCAGCCCGCA GCCTGCCGGG GCTTCGCCCG TGTGGCCAAG 720
ACCATTCGTGT CAGCCTATGT GGTCTGAGG CTGCCCTACC AGCTGGCCCA GCTGCTCTAC 780
CTGGCCTTCC TGTGGGAGT CTACTCTGGC TACCTGCTCT GGGAGGCCCT GGTCTACTCC 840
GACTACCTGA TCTACTCAA CAGCTGCCTC AGCCCTTCC TCTGCTCAT GGCCAGTCC 900
GACCTCCGGA CCTGCTGCG CTCGCTGCTC TGTCTCTTG CGGCAGCTCT CTGCGAGGAG 960
CGGCCGGGCA GCTTCAGCC CACTGAGCCA CAGACCCAGC TAGATTCTGA GGTCCAACT1020
CTGCCAGAGC CGATGGCAGA GGGCCAGTCA CAGATGGATC CTGTGGCCCA GCCTCAGGTG1080
AACCCACAC TCCAGCCAG ATCGGATCCC ACAGCTCAGC CACAGCTGAA CCTACGGCC1140
CAGCCACAGT CGGATCCAC AGCCGAGCA CAGCTGAACC TCATGGCCCA GTCACAGTCA1200
GATTCTGTGG CCCAGCCACA GGCAGACACT AACGTCCAGA CCGCTGCACC TGCTGCCAGT1260
TCTGTGCCCA GTCCCTGTGA TGAAGCTTCC CCAACCCCAT CCTCGCATCC TACCCAGGG1320
GCCCTTGAGG ACCCAGCCAC ACCTCTTGGC TCTGAAGGAG AAAGCCCCAG CAGCACCCCG1380
CCAGAGGCGG CCCCAGGCGC AGGCCCCAGG TGA                                     1413

```

(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:

(D) TOPOLOGÍA: no relevante

(ii) TIPO DE MOLÉCULA: proteína

(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
                20           25           30

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

Leu Met Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr
                50           55           60

Arg Leu Ala Leu Leu Leu Ser Leu Ala Leu Ser Asp Phe Leu Phe
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
                100          105          110

Val Ser Tyr Ser Ser Gly Leu Phe Leu Leu Ala Ala Leu Ser Leu Asp
                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
145          150          155          160

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

```


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 Leu 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 Leu
 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:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

- (A) LONGITUD: 1248 pares de bases
- (B) TIPO: ácido nucleico
- (C) TIPO DE HEBRA: sencilla
- (D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: DNA (genomic)

(ii) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 11:

ES 2 289 610 T3

```

ATGTCAGGGA TGGAAAACT TCAGAATGCT TCTGGATCT ACCAGCAGAA ACTAGAAGAT 60
CCATTCCAGA AACACCTGAA CAGCACCGAG GAGTATCTGG CCTTCCTCTG CGGACCTCGG 120
CGCAGCCACT TCTTCCTCCC CATTGCTGTG GTGTATGTGC CAATTTTGTG GGTGGGGGTC 180
ATTGGCAATG TCTGGGTGTG CCTGGTGATT CTGCAGCACC AGGCTATGAA GACGCCCACC 240
AACTACTACC TCTTCAGCCT GCGGCTCTCT GACCTCTCTG TCTGCTCTCT TGGAAATGCC 300
CTGGAGGTCT ATGAGATGTG GCGCAACTAC CTTTCTTTGT TCGGGCCCGT GGGCTGCTAC 360
TTCAAAGAGG CCCTCTTTGA GACCGTGTGC TTGCGCTCCA TCTTCAGCAT CACCACCGTC 420
AGCGTGGAGC GCTACGTGGC CATCTTACAC CGTTTCCCGG CCAAACTGCA GAGCAGCCGG 480
CGCCGGGGCC TCAAGATCCT CGGCATGCTC TGGGGCTTCT CCGTCTCTTT CTCCTGCCCC 540

AACACCAGCA TCCATGGCAT CAAGTTCCAC TACTTCCCA ATGGTCCCT GGTCCAGGT 600
TGGGCCACCT GTACGGTCAT CAGGCCATG TGGATCTACA ATTTCATCAT CCAGGTCACC 660
TCTTCTCTAT TCTACCTCTT CCGCATGACT GTCATCAATG TCTCTACTA CCTCATGCA 720
CTCAGACTAA AGAAGACAA ATCTCTTGA GCGATGAA GGAATGCAA TATTCAAAGA 780
CCCTOCAGAA AATCAGTCAA CAGATGCTG TTTGTCTTG TCTTAGTGT TCTATCTGT 840
TGGCCCCCGT TCCACATTGA CCGACTCTTC TTCAGCTTGT TGGAGGAGTG GAGTGAATCC 900
CTGGCTGCTG TTTCAACCT COTCCATGTG GTGTCAGGTG TCTTCTCTA CCTGAGCTCA 960
GCTGTCAACC CCATTATCTA TAACCTACTG TCTGCGCGCT TCCAGGCAGC ATTCCAGAAT 1020
GTGATCTCTT CTTTCCACAA ACGTGGCAC TCCCAGCATG ACCCACAATT GCCACCTGCC 1080
CAGCGGAACA TCTTCTGAC AGAATGCCAC TTTGTGGAGC TGACCGAAGA TATAGGTCCC 1140
CAATTCCCAT GTCAGTCATC CATGCACAAC TCTCACCTCC CAACAGCCCT CTCTAGTGAA 1200
CAGATGTCAA GAACAAACTA TCAAGCTTC CACTTTAACA AAACCTGA 1248

```

(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

(ii) TIPO DE MOLÉCULA: proteína

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

ES 2 289 610 T3

Met Ser Gly Met Glu Lys Leu Gln Asn Ala Ser Trp Ile Tyr Gln Gln
 1 3 10 15
 Lys Leu Glu Asp Pro Phe Gln Lys His Leu Asn Ser Thr Glu Glu Tyr
 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
 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 Leu
 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
 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
 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
 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
 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
 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
 Phe Asn Leu Val His Val Val Ser Gly Val Phe Phe Tyr Leu Ser Ser
 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
 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
 370 375 380
 Gln Ser Ser Met His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu
 385 390 395 400
 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 nucleico

(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 AATCAATTTA TCACTAAGCA CTCGTGTTAC TTTAGCATTT 60
TTTATGTCCT TAGTAGCTTT TGCTATAATG CTAAGAAATG CTTTGCTCAT TTTAGCTTTT 120
GTGGTGGACA AAAACCTTAG ACATCGAAGT AGTTATTTTT TTCTTAACTT GGCCATCTCT 180
GACTTCTTTG TGGGTGTGAT CTCCATTCTT TTGTACATCC CTCACACGCT GTTCGAATGG 240
GATTTTGGAA AGGAAATCTG TGTATTTTGG CTCACTACTG ACTATCTGTT ATGTACAGCA 300
TCTGTATATA ACATTGTCTT CATCACTAT GATCGATACC TGTCACTCTC AAATCTCTGT 360
TCTTATAGAA CTCACATAC TGGGGTCTTG AAGATTGTTA CTCTGATGGT GGCCGTTTGG 420
GTGCTGGCCT TCTTAGTGAA TGGGCCAATG ATTCTAGTTT CAGAGTCTTG GAAAGATGAA 480
GGTAGTGAAT GTGAACCTGG ATTTTTTTTG GAAAGGTACA TCCTTGCCAT CACATCATTC 540
TTGGAATTCG TGATCCCAAT CATCTTAGTC GCTTATTTCA ACATGAATAT TTATTGGAGC 600
CTGTGGAAGC GTGATCATCT CAGTAGGTGC CAAAGCCATC CTGGACTGAC TGCTGTCTCT 660
TCCAACATCT GTGACACTC ATTCAAGGAT AGACTATCTT CAAGGAGATC TCTTTCTGCA 720
TCGACAGAAG TTCTGTATC CTTCATTCA GAGAGACAGA GGAGAANGAG TAGTCTCATG 780
TTTTCTCAA GAACCAAGAT GAATAGCAAT ACAATTGCTT CCAAAATGCG TTCCTTCTCC 840
CAATCAGATT CTGTAGCTCT TCACCAAAGG GAACATGTTG AACTGCTTAG AGCCAGGAGA 900
TTAGCCAAAT CACTGGCCAT TCTCTTAAGG GTTTTTGCTG TTTGCTGGGC TCCATATTCT 960
CTGTTCACAA TTGTCCTTTC ATTTTATTC TCAGCAACAG GTCCTAAATC AGTTTGGTAT1020
AGAATTGCAT TTTGGCTTCA GTGTTCAAT TCCTTTCTCA ATCCTCTTTT GTATCCATTG1080
TGTCAACAAG GCTTTCAAAA GGCTTTCTTG AAAATATTTT GTATAAAXX GCAACCTCTA1140
CCATCACAAC ACAGTCGGTC AGTATCTTCT 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
 20 25 30
 Asn Ala Leu Val Ile Leu Ala Phe Val Val Asp Lys Asn Leu Arg His
 35 40 45
 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
 Asp Phe Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp Tyr Leu
 85 90 95
 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
 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
 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
 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
 Gly His Ser Phe Arg Gly Arg Leu Ser Ser Arg Arg Ser Leu Ser Ala
 225 230 235 240
 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
 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
 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
 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
 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 nucleico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: DNA (genomic)

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

```

ATGCGGAACG CGAAGCAGCC GGGTGGCAGC GCGGGCGGCG AGGCGGCGC CCTGGGCGTC 60
AAGCTGGCCA CGCTCAGCCT GCTGCTGTGC GTGAGCCTAG CCGGCAACGT GCTGTTGCGG 120
CTGCTGATCG TGGGGAAGCG CAGCCTGCAC CGCGCCCCGT ACTACCTGCT GCTCGACCTG 180
TGCTTGCCCG ACGGGCTGCG CGCCTCGCC TGCTTCCCGG CGTCATGCT GCGCGCGCGG 240
CGTGGGCGCG CCGCGCGCGG GCGCGCGCGG GCGCGCGCTG GCTGCAAGCT GCTCGCCTTC 300

CTGCGCGCGC TCTTCTGCTT CCAGCGCGCC TTCCTGCTGC TGGGCGTGG CGTCACCGGC 360
TACCTGGCCA TCGCGACCA CGCTTCTAT GCAGAGCGCC TGGCGGCTG GCCGTGCGCC 420
GCCATGCTGG TGTGCGCGCG CTGGGCGCTG GCGCTGCGCG GGGCCTTCCG GCCAGTGCTG 480
GACCGCGGTG GCGACGACGA GGAAGCGCGG TCGCGCTGG AGCAGCGGCC CGACGCGGCC 540
CCCGCGCGCG TGGCTTCTT GCTGCTGCTG GCCGTGGTGG TGGGCGCCAC GCACCTCGTC 600
TACCTCGGCC TGCTTCTT CATCCAGAC CCGCGCAAGA TGGGCGCGC GCGCTGGTG 660
CCCGCGTCA GCGACGACTG GACCTTCCAC GCGCGGCGG CCACCGGCCA GCGCGCGGCC 720
AACTGGACGG CGGGCTTCGG CCGCGGCGCC AGCGCGCGCG CGCTTGTGG CATCGGCGCC 780
GCAGGCGCGG GCGCGCGCGC GCGCGCGCTC CTCGTGCTGG AAGAATTCAA GACGGAAGA 840
AGGCTGTGCA AGATGTTCTA CCGCGTCAGC CTGCTCTTCC TGCTCCTCTG GGGGCCCTAC 900
GTCGTGGCCA GTTACCTGCG GGTCTGTGTC CGGCCGCGG CGTCCCCCA GGCTTACCTG 960
ACCGCTCCG TGTGGCTGAC CTTCGCGCAG GCGGCGATCA ACCCGTGGT GTGCTTCTCT 1020
TTCAACAGGG AGCTGAGGGA CTGCTTCAAG GCGCAGTTCC CCTGCTGCCA GAGCCCCCG 1080
ACCACCCAGG CGACCCATCC CTGCGACCTG AAGGCAATG 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 Gly Glu Ala Ala
 1 5 10 15
 Ala Leu Gly Leu Lys Leu Ala Thr Leu Ser Leu Leu Leu Cys Val Ser
 20 25 30
 Leu Ala Gly Asn Val Leu Phe Ala Leu Leu Ile Val Arg Glu Arg Ser
 35 40 45
 Leu His Arg Ala Pro Tyr Tyr Leu Leu Leu Asp Leu Cys Leu Ala Asp
 50 55 60
 Gly Leu Arg Ala Leu Ala Cys Leu Pro Ala Val Met Leu Ala Ala Arg
 65 70 75 80
 Arg Ala 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
 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
 Cys Ala Ala Trp Ala Leu Ala Leu Ala Ala Phe Pro Pro Val Leu
 145 150 155 160
 Asp Gly 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
 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
 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
 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
 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
 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

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1002 pares de base

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

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

ES 2 289 610 T3

1 ATGAACACCA CAGTGATGCA AGGCTTCAAC AGATCTGAGC GGTGCCCCAG AGACACTCGG 60
 2 ATAGTACAGC TGGTATTCCC AGCCCTCTAC ACAGTGGTTT TCTTGACCGG CATCCTGCTG 120
 5 AATACTTTGG CTCTGTGGGT GTTGTTCAC ATCCCCAGCT CCTCCACCTT CATCATCTAC 180
 6 CTCAAAAACA CTTTGGTGGC OGACTTGATA ATGACACTCA TGCTTCCTTT CAAAATCCTC 240
 10 TCTGACTCAC ACCTGGCACC CTGGCAGCTC AGAGCTTTTG TGTGTCGTTT TTCTTCGGTG 300
 11 ATATTTTATG AGACCATGTA TGTGGGCATC GTGCTGTTAG GGCTCATAGC CTTTGACAGA 360
 12 TTCTCAAGA TCATCAGACC TTTGAGAAAT ATTTTCTAA AAAAACCTGT TTTTGCAAAA 420
 15 ACGGTCTCAA TCTTCATCTG GTTCTTTTGG TTCTTCATCT CCCTGCCAAA TACCATCTTG 480
 16 AGCAACAAGG AAGCAACACC ATCCTCTGTG AAAAAGTGTG CTTCTTAAA GGGGCTCTG 540
 20 GGGCTGAAAT GGCATCAAAT GGTAAATAAC ATATGCCAGT TTATTTTCTG GACTGTTTTT 600
 21 ATCTAATGC TTGTGTTTTA TGTGTTTATT GCAAAAAAAG TATATGATTC TTATAGAAAG 660
 22 TCCAAAAGTA AGACAGAAA AAACAACAAA AAGCTGGAAG GCAAAGTATT TGTGTCGCTG 720
 25 GCTGTCTTCT TTGTGTGTTT TGCTCCATTT CATTTTGCCA GAGTTCCATA TACTCACAGT 780
 26 CAAACCAACA ATAAGACTGA CTGTAGACTG CAAAATCAAC TGTTTATTGC TAAAGAAACA 840
 30 ACTCTCTTTT TGCCAGCAAC TAACATTGT ATGATCCCT TAATATACAT ATTCTTATGT 900
 31 AAAAAATCA CAGAAAAGCT ACCATGTATG CAAGGAGAA AGACCACAGC ATCAAGCCAA 960
 35 GAAAATCATA GCAGTCAGAC AGACAACATA ACCTTAAGCT GA 1002

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

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

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	Lys	Pro	Val	Phe	Ala	Lys	Thr	Val	Ser	Ile
				130				135					140			
25	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	
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	
45	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			
50	Ile	Cys	Met	Asp	Pro	Leu	Ile	Tyr	Ile	Phe	Leu	Cys	Lys	Lys	Phe	Thr
		290					295					300				
	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

65 (C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

ES 2 289 610 T3

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

5
10
15
20
25
30
35
40

```

ATGGCCAAACA CTACCGGAGA GCCTGAAGAG GTGAAGGCG CTCTGTCCC ACCGTCCGCA 60
TCAGCTTATG TGAAGCTGGT ACTGCTGGGA CTGATTATGT GCCTGAAGCT GCGCGGTAAC 120
GCCATCTTGT CCTCTCTGCT GCTCAAGGAG CGTCCCTGC ACAAGGCTCC TTACTACTTC 180
CTGCTGGACC TGTGCTGGC CGATGGCATA CGCTCTCCG TCTGCTTCC CTTTGTGCTG 240
GCTTCTGTGC GCCACGGCTC TTCATGGACC TTCAGTCCAC TCAGCTGCAA GATTGTGGCC 300
TTTATGGCCG TGTCTTTTG CTTCCATGG GCCTTCATOC TGTCTGCTAT CAGCGTCACC 360
CGCTACATGG CCATGGCCCA CCACCGCTTC TACGCCAAGC GCATGACACT CTGGACATGC 420
GCGGCTGTCA TCTGCATGG CTGGACCCTG TCTGTGGCCA TGGGCTTCC ACCTGTCTTT 480
GACGTGGGCA CCTACAAATT TATTCGGGAG GAGGACCAGT GCATCTTTGA GCATCGCTAC 540
TTCAAGGCCA ATGACACGCT GGGCTTCATG CTTATGTTGG CTGTGCTCAT GGCAGCTACC 600
CATGCTGTCT ACGGCAAGCT GCTCCTCTTC GAGTATGCTC ACCGCAAGAT GAAGCCAGTG 660
CAGATGGTGC CAGCCATCAG CCAGAACTGG ACATTCCATG GTCCCGGGGC CACCGGCCAG 720
GCTGCTGCCA ACTGGATGCG CCGCTTTGGC CGTGGGCCCA TGCCACCAAC CTTGCTGGGT 780
ATCCGGCAGA ATGGGCATGC AGCCAGCCGG CCGCTACTGG GCATGGACGA GGTCAAGGGT 840
GAAAAGCAGC TGGGCGGCAT GTTCTACGG ATCACACTGC TCTTTCTGCT CCTCTGGTCA 900
CCCTACATCG TGGCCTGCTA CTGGCGAGTG TTTGTGAAG CTTGTGCTGT GCCCCACCGC 960
TACCTGGCCA CTGCTGTTTG GATGAGCTTC GCCCAGGCTG CCGTCAACCC AATTGTCTGC1020
TTCTGTCTCA ACAAGGACCT CAAGAAGTGC CTGACCACTC ACGCCCCCTG CTGGGGCACA1080
GGAGGTGCCC CGGCTCCAG AGAACCTAC TGTGTATGT GA 1122

```

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 373 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: 20:

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
 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
 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
 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
 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
 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
 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
 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
 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
 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
 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
 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
 Pro Tyr Cys Val Met
 370

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1053 pares de base

(B) TIPO: ácido nucleico

ES 2 289 610 T3

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

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

```

ATGGCTTTGG AACAGAACCA GTCAACAGAT TATTATTATG AGGAAAATGA AATGAATGGC 60
ACTTATGACT ACAOTCAATA TGAATTGATC TGTATCAAAG AAGATGTCAG AGAATTTOCA 120
AAAGTTTTCC TCCCTGTATT CCTCACAATA GCTTTCTGTC TTGGACTTGC AAGCAATTCC 180
ATGGTAGTGG CAATTATGTC CTATTACAAG AAACAGAGAA CCAAAACAGA TGTGTACATC 240
CTGAATTTGG CTGTAGCAGA TTTACTCCTT CTATTCACCTC TGCCTTTTTG GGCTGTTAAT 300
GCAATTCATG GGTGGGTTTT AGGAAAATA ATGTGCAAAA TAACTTCAGC CTTGTACACA 360
CTAAACTTTG TCTCTGGAAT GCAGTTTCTG GCTTGATCA GCATAGACAG ATATGTGGCA 420
GTAACATAAT TCCCCAOCOA ATCAGGAGTG GAAAAACCAT GCTGGATCAT CTGTTTCTGT 480
GTCTGGATGG CTGCCATCTT GCTGAGCATA CCCCAGCTGG TTTTITATAC AGTAAATGAC 540
AATGCTAAGT GCATTCCCAT TTCCCCCGC TACCTAGGAA CATCAATGAA AGCATTGATT 600
CAAATGCTAG AGATCTGCAT TGGATTGTA GTACCTTTTC TTATTATGGG GGTGTCTTAC 660
TTTATCACCG CAAGGACACT CATGAAGATG CCAACATTA AAATATCTCG ACCCCTAAAA 720
GTTCTGCTCA CAGTCGTTAT AGTTTTCATT GTCACTCAAC TGCCTTATAA CATTGTCAAG 780
TTCTGCCGAG CCATAGACAT CATCTACTCC CTGATCACCA GCTGCAACAT GAGCAAAAGC 840
ATGGACATCG CCATCCAAAT CACAGAAAGC ATTGCACTCT TTCACAGCTG CCTCAACCCA 900
ATCCTTTATG TTTTATGGG AGCATCTTC AAAAACTACG TTATGAAGT GGCCAAGAAA 960
TATGGGTCTT GGAGAAGACA GAGACAAAGT GTGAGGAGT TTCCTTTTGA TTCTGAGGTT1020
CCTACAGAGC CAACCAATAC TTTTAGCATT TAA
1093

```

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 350 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: 22:

```

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

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

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

```

ES 2 289 610 T3

[illegible]

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1116 pares de base

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 23:

ES 2 289 610 T3

```

ATGCCAGGAA ACCCCACCCC AGTGACCACC ACTGCCCCGT GGGCCTCCCT GGGCCTCTCC 60
GCCAAGACCT GCAACAACGT GTCCCTCGAA GAGAGCAGGA TAGTCCTGGT CGTGGTGTAC 120
AGCGCGGTGT GCACGCTGGG GGTGCCGGCC AACTGCCCTGA CTGCCTGGCT GGCCTGCTG 180
CAGGTACTGC AGGCAACGT GCTGGCCGTC TACCTGCTCT GCCTGGCACT CTGCGAAGTG 240
CTGTACACAG GCACGCTGCC ACTCTGGGTC ATCTATATCC GCAACCAGCA CCGCTGGACC 300
CTAGGCCTGC TGGCCTCGAA GGTGACCGCC TACATCTTCT TCTGCAACAT CTACGTCAGC 360
ATCCTCTTCC TGTGCTGCAT CTCCTGGAC CGCTTCGTGG CCGTGGTGTG CGCGCTGGAG 420
AGTCGGGGCC GCGGCGCGCG GAGGACCGCC ATCCTCATCT CGGCTGCAT CTTCATCTTC 480
GTGGGATCG TTCACTACCC GGTGTTCCAG ACGGAAGACA AGGAGACCTG CTTTGACATG 540
CTGCAGATCG ACAACAGGAT TGCCGGGTAC TACTACGCCA GGTTCACCGT TGCTTTTGCC 600
ATCCCTCTCT CCATCATGCG CTTCACCAAC CACCGGATTT TCAGGAGCAT CAAOCAGAGC 660
ATGGGCTTAA GCGCTGCCCA GAAGGCCAAG GTGAAGCACT CGGCCATCGC GGTGGTTGTC 720
ATCTTCCTAG TCTGCTTCGC CCGTACCAC CTGGTTCTCC TCGTCAAAGC CGCTGCTTT 780
TCCTACTACA GAGGAGACAG GAACGCCATG TGCGGCTTGG AGGAAAGGCT GTACACAGCC 840
TCTGTGTTGT TTCTGTGCTT GTCCACGGTG AACGCGGTGG CTGACCCCAT TATCTACGTG 900
CTGGCCACGG ACCATTECCG CCAAGAGTG TCCAGAATCC ATAAGGGGTG GAAAGAGTGG 960
TCCATGAAGA CAGACGTCAC CAGGCTCACC CACAGCAGGG ACACCGAGGA GCTGCAATCG1020
CCCGTGGCCC TTGCAGACCA CTACACCTTC TCCAGGCCCC TGCACCCACC AGGGTCACCA1080
TGCCCTGCAA AGAGGCTGAT TGAGGAGTCC TGCTGA 1116

```

(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

(ii) TIPO DE MOLÉCULA: proteína

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

ES 2 289 610 T3

	Met	Pro	Gly	Asn	Ala	Thr	Pro	Val	Thr	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	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				
55	His	Ser	Arg	Gln	Glu	Val	Ser	Arg	Ile	His	Lys	Gly	Trp	Lys	Glu	Trp
	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 nucleico

(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 AGCTGACAAC ATTTTGCAAA ATCTCTCGCC TCTAACAGCC 60
TTTCTGAAC TGACTTCCTT GGGTTTCATA ATAGGAGTCA GCGTGGTGGG CAACCTCCTG 120
ATCTCCATT TGTAGTGAA AGATAAGACC TTGCATAGAG CACCTTACTA CTTCCTGTTG 180
GATCTTTGCT GTTCAGATAT CCTCAGATCT GCAATTTGTT TCCCATTTGT GTTCAACTCT 240
GTCAAAAATG GCTCTACCTG GACTTATGGG ACTCTGACTT GCAAAGTGAT TGCCTTCTG 300
GGGGTTTTGT CCTGTTTCCA CACTGCTTTC ATGCTCTTCT GCATCAGTGT CACCAGATAC 360
TTAGCTATCG CCCATCACCG CTTCATACA AAGAGGCTGA CCTTTGGAC GTGTCTGGCT 420
GTGATCTGTA TGGTGTGGAC TCTGTCTGTG GCCATGSCAT TTCCCCCGT TTTAGACGTG 480
GGCACTTACT CATTATTAG GGAGGAGAT CAATGCACCT TCCACACCG CTCCTTCAGG 540
GCTAATGATT CCTTAGGATT TATGCTGCTT CTGCTCTCA TCCTCCTAGC CACACAGCTT 600
GTCTACCTCA AGCTGATATT TTTGTCAC GATCGAAGAA AAATGAAGCC AGTCCAGTTT 660
GTAGCAGCAG TCAGCCAGAA CTGGACTTTT CATGGTCTG GAGCCAGTGG CCAGGCAGCT 720
GCCAATTGOC TAGCAGGATT TGGAAAGGCT CCCACACCAC CCACCTTGT GGGCATCAGG 780
CAAAATGCAA ACACCACAGG CAGAAGAAGG CTATTGGTCT TAGACGAGTT CAAAATGGAG 840
AAAAGAATCA GCAGAATGTT CTATATAATG ACTTTTCTGT TTCTAACCTT GTGGGGCCCC 900
TACCTGGTGG CCTTTATTG GAGAGTTTTT GCAAGAGGGC CTGTAGTACC AGGGGAGTTT 960
CTAACAGCTG CTGTCTGAT GAGTTTTGCC CAAGCAGGAA TCAATCCTTT TGTCTGCATT 1020
TTCTCAACA GGGAGCTGAG GCGCTGTTTC AGCACAACCC TTCTTTACTG 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
 20 25 30
 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
 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
 Ile Ala Phe Leu Gly Val Leu Ser Cys Phe His Thr Ala Phe Met Leu
 100 105 110
 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
 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
 Arg Ser Phe Arg Ala Asn Asp Ser Leu Gly Phe Met Leu Leu Leu Ala
 180 185 190
 Leu Ile Leu Leu Ala Thr Gln Leu Val Tyr Leu Lys Leu Ile Phe Phe
 195 200 205
 Val His Asp Arg Arg Lys Met Lys Pro Val Gln Phe Val Ala Ala Val
 210 215 220
 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
 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
 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
 305 310 315 320
 Leu Thr Ala Ala Val Trp Met Ser Phe Ala Gln Ala Gly Ile Asn Pro
 325 330 335
 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
 Val Ile
 370

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1080 pares de base

(B) TIPO: ácido nucleico

(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 AACGCGACGC TGCAGATGCT GCGGAACCCG 60
 GCGATCGCGG TGGCCCTGCC CGTGGTGTAC TCCTGGTGG CGGCGGTCAg CATCCCGGGC 120
 AACCTCTTCT CTCGTGGGT GCTGTCCGG CGCATGGGC CCAGATCCC GTCCGTCAATC 180
 TTCATGATCA AACTGAGCGT CACGGACCTG ATGCTGGCCA GCGTGTGCT TTTCCAAATC 240
 TACTACCATT GCAACGGCCA CCACTGGTA TTGGGGTGC TGCTTTGCA CGTGGTGACC 300
 GTGGCCTTTT ACGCAACAT GTATTCCAGC ATCTCACC A TGACCTGTAT CAGCGTGGAG 360
 CGCTTCCTGG GGGTCTGTA CCCGCTCAGC TCCAAGCGT GCGCCGCGG TCGTTACGGC 420
 GTGGCGCGT GTGCAGGAC CTGGCTGCTG CTCTGACCG CCCTGTGCC GCTGGCGCGC 480
 ACCGATCTCA CCTACCGGT GCACGCCCTG GGCATCATCA CCTGCTTGA CGTCTCAAG 540
 TGGACGATGC TCCCCAGCGT GGCATGTGG GCGGTGTTCC TCTCACCAT CTTCATCTG 600
 CTGTTCTCA TCCCGTTGCT GATCACCGTG GCTTGTACA CGGCCACCAT CCTCAAGCTG 660
 TTGCGACCG AGGAGGCGCA CGGCCGGAG CAGCGGAGC GCGCGTGGG OCTGGCCGCG 720
 GTGGTCTTG TGGCCTTGT CACCTGCTTC GCCCCAACA ACTTCGTGCT CCTGGCGCAC 780
 ATCGTGAGCC GCCTGTTCTA CGGCAAGAGC TACTACCAG TOTACAAGCT CACGCTGTGT 840
 CTCAGCTGCC TCAACAATG TCTGGACCG TTTGTTTAT ACTTTGCTC CCGGAATTTC 900
 CAGCTGCGCC TGGCGAATA TTGGGCTGC CGCCGGTGC CCAAGACAC CCTGGACAG 960
 CGCCGCGAGA GCCTCTTCTC GCGCAGGACC ACGTCCGTG GCTCGAGGC CGGTGCGCAC 1020
 CCTGAAGGGA TGGAGGGAGC CACCAGGCCC GGCCTCCAGA 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:

5	Met	Gln	Val	Pro	Asn	Ser	Thr	Gly	Pro	Asp	Asn	Ala	Thr	Leu	Gln	Met	1	5	10	15
	Leu	Arg	Asn	Pro	Ala	Ile	Ala	Val	Ala	Leu	Pro	Val	Val	Tyr	Ser	Leu	20	25	30	
	Val	Ala	Ala	Val	Ser	Ile	Pro	Gly	Asn	Leu	Phe	Ser	Leu	Trp	Val	Leu	35	40	45	
10	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	
	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	
	Leu	Ser	Ser	Lys	Arg	Trp	Arg	Arg	Arg	Arg	Tyr	Ala	Val	Ala	Ala	Cys	130	135	140	
25	Ala	Gly	Thr	Trp	Leu	Leu	Leu	Leu	Thr	Ala	Leu	Cys	Pro	Leu	Ala	Arg	145	150	155	160
	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	
	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	
	Glu	Ala	His	Gly	Arg	Glu	Gln	Arg	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	
	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	
	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
	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:

- 65 (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 CAGCCCAAGC CCGGAGGGG CAGCTGAAGG CTCGCCTGTG 60
CCAGTCGCCG CCGGGGCGCG CTCGGGTGCC GCGGCGAGTG GCACAGGCTG GCAGCCATGG 120
GCTGAGTGCC CGGGACCCAA GGGGAGGGG CAACTGCTGG CGACCGCCGG CCCTTTGCGT 180
CGCTGGCCCG CCCCCTGGCC TGCCAGCTCC AGCCCCGCCC CCGGAGCGGC GTCCGCTCAC 240
TCGGTTCAAG GCAGCGGAC TCGGGTGGC GCACGACCG GCGCGAAGC TTGGGGCGCG 300
CGGCCCATGG AGTCGGGGCT GCTGCGGCC GCGCCGTTGA GCGAGGTCAT CGTCCTGCAT 360
TACAACTACA CCGGCAAGCT CCGGGGTGG AGCTACCAGC CGGGTGCCCG CCTGCGCGCC 420
GACGCCGTGG TGTGCTGGC GGTGTGGCC TTCATCGTG TAGAGAATCT AGCCGTGTTG 480
TTGGTGCTCG GACGCCACC GCGCTTCAC GCTCCCATGT TCCTGCTCTT GGGCAGCCTC 540
ACGTTGTGG ATCTGCTGG AGGCGCGCC TACGCCGCA ACATCCTACT GTCGGGGCG 600
CTCACGCTGA AACTGTCCC CGCGCTCTG TTGCGACGG AGGGAAGCGT CTTCGTGGCA 660
CTCACTGCGT CCGTGCTGAG CCTCCTGGC ATCGCGCTGG AGCGCAGCCT CACCATGGCG 720
CGCAGGGGGC CGCGCCCGT CTCAGTGGG GGGCGCACG TGGCGATGG AGCCGCGCC 780
TGGGGCGTGT CCGTGTCTCT CGGGCTCCTG CCAGCGCTGG GCTGGAATTG CCTGGGTGGC 840
CTGAGCGTT GCTCCACTGT CTGCGGCTC TACGCCAAG CCTACGTCT CTCTGCGTG 900
CTCGCTTCG TGGGCATCT GCGCGGATC TGTGACTCT ACGCGCGAT CTACTGCCAG 960
GTACGCGCA ACGCGGCGC CTGCGGCA CGGCGCGGA CTGCGGGAC CACCTCGACC1020
CGGGCGGTC GCAAGCGCG CTCTGTGGC TTGCTGCGCA CGCTCAGCGT GGTGCTCCTG1080
GCGTTGTGG CATGTTGGG CCCCCTCTT CTGCTGCTGT TGCTCGACGT GCGGTGCCCG1140
GCGCGACCT GTCTGTACT CCTGCAGGC GATCCCTTC TGGGACTGG CATGGCCAAC1200
TCACTTCTGA ACCCATCAT CTACAGCTC ACCAACCAG ACCTGCGCA CGCGCTCCTG1260
CGCCTGGTCT GCTCGGACG CCACTCCTGC GGCAGAGACC CGAGTGGCTC CCAGCAGTCG1320
GCGAGCGCG CTGAGGCTT CGGGGGCTG CGCCGCTGCC TGCCCCGGG CCTTGATGGG1380
AGCTTCAGCG GCTCGGAGC CTCATCGCC CAGCGGACG GGCTGGACAC CAGCGGCTCC1440
ACAGGCAGC CCGGTGCACC CACAGCGCC CGGACTCTG TATCAGAAC GGTGCGAGAC1500
TGA

```

(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:

5	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	
10	Ser	Gly	Thr	Gly	Trp	Gln	Pro	Trp	Ala	Glu	Cys	Pro	Gly	Pro	Lys	Gly	35	40	45	
	Arg	Gly	Gln	Leu	Leu	Ala	Thr	Ala	Gly	Pro	Leu	Arg	Trp	Pro	Ala		50	55	60	
15	Pro	Ser	Pro	Ala	Ser	Ser	Ser	Pro	Ala	Pro	Gly	Ala	Ala	Ser	Ala	His	65	70	75	80
	Ser	Val	Gln	Gly	Ser	Ala	Thr	Ala	Gly	Gly	Ala	Arg	Pro	Gly	Arg	Arg	85	90	95	
20	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	
25	Gly	Ala	Ser	Tyr	Gln	Pro	Gly	Ala	Gly	Leu	Arg	Ala	Asp	Ala	Val	Val	130	135	140	
	Cys	Leu	Ala	Val	Cys	Ala	Phe	Ile	Val	Leu	Glu	Asn	Leu	Ala	Val	Leu	145	150	155	160
30	Leu	Val	Leu	Gly	Arg	His	Pro	Arg	Phe	His	Ala	Pro	Met	Phe	Leu	Leu	165	170	175	
	Leu	Gly	Ser	Leu	Thr	Leu	Ser	Asp	Leu	Leu	Ala	Gly	Ala	Ala	Tyr	Ala	180	185	190	
35	Ala	Asn	Ile	Leu	Leu	Ser	Gly	Pro	Leu	Thr	Leu	Lys	Leu	Ser	Pro	Ala	195	200	205	
	Leu	Trp	Phe	Ala	Arg	Glu	Gly	Gly	Val	Phe	Val	Ala	Leu	Thr	Ala	Ser	210	215	220	
40	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	
45	Ala	Ala	Ala	Ala	Trp	Gly	Val	Ser	Leu	Leu	Leu	Gly	Leu	Leu	Pro	Ala	260	265	270	
50	Leu	Gly	Trp	Asn	Cys	Leu	Gly	Arg	Leu	Asp	Ala	Cys	Ser	Thr	Val	Leu	275	280	285	
	Pro	Leu	Tyr	Ala	Lys	Ala	Tyr	Val	Leu	Phe	Cys	Val	Leu	Ala	Phe	Val	290	295	300	
55	Gly	Ile	Leu	Ala	Ala	Ile	Cys	Ala	Leu	Tyr	Ala	Arg	Ile	Tyr	Cys	Gln	305	310	315	320
	Val	Arg	Ala	Asn	Ala	Arg	Arg	Leu	Pro	Ala	Arg	Pro	Gly	Thr	Ala	Gly	325	330	335	
60	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	
65	Leu	Phe	Leu	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
20	Thr Gly Ser Pro Gly Ala Pro Thr Ala Ala Arg Thr Leu Val Ser Glu					
		485		490		495
25	Pro Ala Ala Asp					
		500				

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1029 pares de base

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

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

```

ATGCAAGCCG TCGACAATCT CACCTCTGCG CCTGGGAACA CCAGTCTGTG CACCAGAGAC 60
TACAAAATCA CCCAAGTCCT CTCCCACTG CTCTACACTG TCCTGTTTTT TGTGGACTT 120
ATCACAATG GCCTGGCGAT GAAGATTTTC TTCAAATCC GGAGTAAATC AAATTTTATT 180
ATTTTTCTTA AGAACACAGT CATTCTGAT CTTCTCATGA TTCTGACTTT TCCATTCAAA 240
ATTCTTAGTG ATGCCAACT GGGAAACAGG CCACTGAGAA CTTTGTGTG TCAAGTTACC 300
TCCGTCAAT TTTATTTCAC AATGTATATC AGTATTTTAT TCCTGGGACT GATAACTATC 360
GATCGCTACC AGAAGACCAC CAGGCCATTT AAAACATCCA ACCCCAAAAA TCTCTTGGGG 420
GCTAAGATTC TCTCTGTTGT CATCTGGGCA TTCATGTTCT TACTCTCTTT GCCTAACATG 480

```

ES 2 289 610 T3

ATTCTGACCA ACAGGCAGCC GAGAGACAAO AATGTGAAGA AATGCTCTTT CCTTAAATCA 540
 GAGTTCGGTC TAGTCTGGCA TGAATAGTA AATTACATCT GTCAAGTCAT TTTCTGGATT 600
 AATTTCTTAA TTGTTATTGT ATGTTATACA CTCATTACAA AAGAACTGTA CCGGTCATAC 660
 GTAAGAACGA GGGGTGTAGG TAAAGTCCCC AGGAAAAAGG TGAACGTCAA AGTTTTTCATT 720
 ATCATTGCTG TATTCTTTAT TTGTTTGTGTT CCTTTCATT TTGCCCGAAT TCCTTACACC 780
 CTGAGCCAAA CCCGGGATGT CTTTGACTGC ACTGCTGAAA ATACTCTGTT CTATGTGAAA 840
 GAGAGCACTC TGTGGTTAAC TTCCTTAAAT GCATGCCTGG ATCCGTTTAT CTATTTTTTC 900
 CTTTGCAAGT CCTTCAGAAA TTCCTTGATA AGTATGCTGA AGTGCCCCAA TTCTOCAACA 960
 TCTCTGTCCC AGGACAATAG GAAAAAGAA CAGGATGGTG GTGACCCAAA TGAAGAGACT1020
 CCAATGTAA 1029

(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

(ii) TIPO DE MOLÉCULA: proteína

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

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

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

- 45 (i) CARACTERÍSTICAS DE LA SECUENCIA:
- (A) LONGITUD: 1077 pares de base
 - (B) TIPO: ácido nucléico
 - 50 (C) TIPO DE HEBRA: sencilla
 - (D) TOPOLOGÍA: lineal
- (ii) TIPO DE MOLÉCULA: ADN (genomic)
- 55 (xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 33:

ES 2 289 610 T3

5
 10
 15
 20
 25
 30
 35

ATGTGGTCT GCTACGCTCC CCCAGGGAAC GAGACACTGC TGAGCTGGAA GACTTCGCGG 60
 GCCACAGGCA CAGCCTTCCT GCTGCTGGCG GCGCTGCTGG GGCTGCTGG CAACGGCTTC 120
 GTGGTGTGGA GCTTGGCGGG CTGGCGGCTT GCACGGGGGC GACCGCTGGC GGCCACGCTT 180
 GTGCTGCACC TGGCGCTGGC CGACGGGGCG GTGCTGCTGC TCACGCGCT CTTGTGGCC 240
 TTCTGACCC GGCAGGCTTG GCCGCTGGGC CAGGCGGGCT GCAAGGCGGT GTACTACGTG 300
 TCGCGCTCA GCATGTACGC CAGCGTGTG CTCACGGGC TGCTCAGCT GCAGCGCTGC 360
 CTGCGAGTCA CCGCCCCCTT CCTGGGCTT CGGCTGGCA GCCCGGCTT GGCCCGCGC 420
 CTGCTGCTGG CGTCTGGCT GGCGGCGCT TTGCTGCGCG TCCCGGCGC CTTCTACGC 480
 CACCTGTGGA GGCACGCGT ATGCCAGCTG TGCCACCGT CGCGGTCCA GCGCGCGCC 540
 CACCTGAGCC TGGAGACTCT GACCGTTTC GTGCTTCCTT TCGGGCTGAT GTCGGCTGC 600
 TACAGCGTGA CGCTGCAAG GCTGCGGGC GCCCGCTGG GCTCGGGCG GCACGGGGC 660
 CGGGTGGCC GGCTGGTGG CGCCATCGTG CTGCGCTTG GCTTGTCTG GGCCCCCTAC 720
 CACGCGTCA ACCTTCTGCA GCGGTGCGA GCGCTGGCT CACCGGAAGG GGCCTTGGC 780
 AAGCTGGCG GAGCGGCCA GCGGCGCGA GCGGGAATA CGGCTTGGC CTTCTTCACT 840
 TCTAGCGTCA ACCCGGTGCT CTACGTCTT ACCGCTGGAG ATCTGCTGCC CCGGGCAGGT 900
 CCGCGTTTC TCACGCGCT CTTGGAAGC TCTGGGAGG CCGAGGGG CGCGCGCTCT 960
 AGGGAAGGA CCATGGAGCT CCGAATACT CCTCAGCTGA AAGTGGTGG GCAGGGCGGC 1020
 GGCAATGGAG ACCCGGGGG TGCGATGGAG AAGGACGCT CGGAATGGA CTTTGA 1077

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

40 (i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 358 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: 34:

55
 60
 65

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

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
25	His 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 His 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
	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 Gly Arg Ser		
	305	310	315
50	Arg Glu Gly Thr Met Glu Leu Arg Thr Thr Pro Gln Leu Lys Val Val		
	325	330	335
55	Gly Gln Gly Arg Gly Asn Gly Asp Pro Gly Gly Gly Met Glu Lys Asp		
	340	345	350
60	Gly Pro Glu Trp Asp Leu		
	355		

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

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

ES 2 289 610 T3

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

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

```

ATGCTGGGGA TCATGGCATG GAATGCAACT TGCAAAAAC TGGCTGGCAGC AGAGGCTGCC 60
CTGGAAAAGT ACTACCTTTC CATTITTTAT GGGATTGAGT TCGTTGTGGG AGTCCTTGGA 120
AATACCATTG TTGTTTACGG CTACATCTTC TCTCTGAAGA ACTGGACAG CAOTAATATT 180
TATCTCTTTA ACCTCTCTGT CTCTGACTTA GCTTTTCTGT GCACCCCTCC CATGCTGATA 240
AGGAGTTATG CCAATGGAAA CTGGATATAT GGAGACGTGC TCTGCATAAG CAACCGATAT 300
GTGCTTCATG CCAACCTCTA TACCAGCATT CTCTTTCTCA CTTTTATCAG CATAGATCGA 360
TACTTGATAA TTAAGTATCC TTTCCGAGAA CACCTTCTGC AAAAGAAAGA GTTTOCTATT 420
TTAATCTCCT TGGCCATTTG GGTITTTAGTA ACCTTAGAGT TACTACCCAT ACTTCCCTTT 480
ATAAATCCTG TTATAACTGA CAATGGCACC ACCTGTAATG ATTTTGCAG TTCTGGAGAC 540
CCCAACTACA ACCTCATTTA CAGCATGTGT CTAACACTGT TGGGTTCTCT TATTCTCTTT 600
TTTGTGATGT GTTTCCTTTA TTACAAGATT GCTCTCTTCC TAAAGCAGAG GAATAGGCAG 660
GTTGCTACTG CTCTGCCCC TGAAGAGCCT CTCAACTTGG TCATCATGGC AGTGCTAATC 720
TTCTCTGTGC TTTTACACC CTATCACGTC ATGCGGAATG TGAGGATCGC TTCACGCCCTG 780
GGGAGTTGGA AGCAGTATCA GTGCACTCAG GTGCTCATCA ACTCCTTTTA CATTGTGACA 840
CGGCCCTTGG CCTTCTGAA CAGTGTCTAT AACCTGTCT TCTATTTTCT TTTOGGAQAT 900
CACTTCAGGG ACATGCTGAT GAATCAACTG AGACACAAC TCAAATCCCT TACATCCTTT 960
AGCAGATGGG CTCATGAACT CCTACTTTCA TTCAGAGAAA AGTGA 1005

```

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 334 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: 36:

ES 2 289 610 T3

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

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 1296 pares de base

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

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

```

ATGCAGGCGC TTAACATTAC CCCGAGCAG TTCTCTCGGC TGCTGCGGGA CCACAACCTG 60
ACGCGGAGC AGTTCATCGC TCTGTACCG CTGCGACCGC TCGTCTACAC CCCAGAGCTG 120
CCGGGACGCG CCAAGCTGGC CCTCGTCTC ACCGGCGTGC TCATCTTCGC CCTGGCGCTC 180
TTTGGCAATG CTCTGGTGT CTACGTGGTG ACCCGCAGCA AGGCCATGCG CACCTCACC 240
AACATCTTTA TCTGCTCCTT GGCCTCAGT GACCTGCTCA TCACCTTCTT CTGCATTCCC 300
GTCACCATGC TCCAGAACAT TTCCGACAAC TGGCTGGGGG GTGCTTTCAT TTGCAAGATG 360
GTGCCATTTG TCCAGTCTAC CGCTGTTGTG ACAGAAATGC TCACTATGAC CTGCATTGCT 420
GTGAAAGGC ACCAAGGACT TGTGCATCCT TTTAAATGA AGTGGCAATA CACCAACCGA 480
AGGCTTTCA CAATGCTAGG TGTGGTCTGG CTGTGGCAG TCATCGTAGG ATCACCCATG 540
TGCCACGTGC AACAACTTGA GATCAAATAT GACTTCCTAT ATGAAAAGGA ACACATCTGC 600
TGCTTAGAAG AGTGGACCAG CCCTGTGCAC CAGAAGATCT ACACCACCTT CATCCTTGTC 660
ATCCTCTTCC TCCTGCCTCT TATGGTGATG CTTATCTGT ACAGTAAAT TGGTTATGAA 720
CTTTGATAA AGAAAAGAGT TGGGATGGT TCAGTGCTTC GAACTATTCA TGGAAAAGAA 780
ATGTCCAAA TAGCCAGGAA GAAGAAACGA GCTGTCATTA TGATGGTGAC AGTGGTGGCT 840
CTCTTTGCTG TGTGCTGGGC ACCATTCCAT GTTGTCCATA TGATGATTGA ATACAGTAAT 900
TTTGAAAGG AATATGATGA TGTCAAAATC AAGATGATTT TTGCTATCGT GCAAATTATT 960
GGATTTTCCA ACTCCATCTG TAATCCCAT GTCTATGCAT TTATGAATGA AAACCTCAAA1020
AAAAATGTTT TGCTGCACT TTGTTATTGC ATAGTAAATA AAACCTTCTC TCCAGCACAA1080
AGGCATGGAA APTCAGGAAT TACAATGATG CGAAGAAAG CAAAGTTTTC CCTCAGAGAG1140
AATCCAGTGG AGGAAACCAA AGGAGAAGCA TTCAGTGATG GCAACATTGA AGTCAAATTG1200
TGTGAACAGA CAGAGGAGAA GAAAAGCTC AAACGACATC TTGCTCTCTT TAGGTCTGAA1260
CTGGCTGAGA ATTCTCCTTT AGACAGTGGG CATTAA

```

1296

ES 2 289 610 T3

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 431 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: 38:

```

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

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
35           40           45

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

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

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

```

ES 2 289 610 T3

	100	105	110
5	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		
10	Gln Gly Leu Val His Pro Phe Lys Met Lys Trp Gln Tyr Thr Asn Arg 145 150 155 160		
	Arg Ala Phe Thr Met Leu Gly Val Val Trp Leu Val Ala Val Ile Val 165 170 175		
15	Gly Ser Pro Met Trp His Val Gln Gln Leu Glu Ile Lys Tyr Asp Phe 180 185 190		
	Leu Tyr Glu Lys Glu His Ile Cys Cys Leu Glu Glu Trp Thr Ser Pro 195 200 205		
20	Val His Gln Lys Ile Tyr Thr Thr Phe Ile Leu Val Ile Leu Phe Leu 210 215 220		
	Leu Pro Leu Met Val Met Leu Ile Leu Tyr Ser Lys Ile Gly Tyr Glu 225 230 235 240		
25	Leu Trp Ile Lys Lys Arg Val Gly Asp Gly Ser Val Leu Arg Thr Ile 245 250 255		
	His Gly Lys Glu Met Ser Lys Ile Ala Arg Lys Lys Lys Arg Ala Val 260 265 270		
30	Ile Met Met Val Thr Val Val Ala Leu Phe Ala Val Cys Trp Ala Pro 275 280 285		
35	Phe His Val Val His Met Met Ile Glu Tyr Ser Asn Phe Glu Lys Glu 290 295 300		
	Tyr Asp Asp Val Thr Ile Lys Met Ile Phe Ala Ile Val Gln Ile Ile 305 310 315 320		
40	Gly Phe Ser Asn Ser Ile Cys Asn Pro Ile Val Tyr Ala Phe Met Asn 325 330 335		
	Glu Asn Phe Lys Lys Asn Val Leu Ser Ala Val Cys Tyr Cys Ile Val 340 345 350		
45	Asn Lys Thr Phe Ser Pro Ala Gln 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		
50	Glu Thr Lys Gly Glu Ala Phe Ser Asp Gly Asn Ile Glu Val Lys Leu 385 390 395 400		
55	Cys Glu Gln Thr Glu Glu Lys 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		
60			

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24 pares de base

(B) TIPO: ácido nucleico

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

(iv) ANTISENTIDO: NO

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

TCACAATGCT AGGTGTGGTC

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

(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: 44:

TGCATAGACA ATGGGATTAC AG

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 511 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: 45:

```
TCACAATGCT AGGTGTGGTC TGGCTGGTGG CAGTCATCGT AGGATCACCC ATGTGGCAGG 60
TGCAACAACT TGAGATCAAA TATGAATTCC TATATGAAAA GGAACACATC TGCTGCTTAG 120
AAGAGTGGAC CAGCCCTGTG CACCAGAAGA TCTACACCAC CTTCATCCTT GTCATCCTCT 180
TCCTCCTGCC TCTTATGGTG ATGCTTATTC TGTACGTAAA ATTGGTTATG AACTTTGGAT 240
AAAGAAAAGA GTTGGGGATG GTTCAATGCT TCGAACTATT CATGGAAAAG AAATGTCCAA 300
AATAGCCAGG AAGAAGAAAC GAGCTGTCAAT TATGATGGTG ACAGTGGTGG CTCTCTTTGC 360
TGTGTGCTGG GCACCATTCC ATGTTGTCCA TATGATGATT GAATACAGTA ATTTTGAAAA 420
GGAATATGAT GATGTCACAA TCAAGATGAT TTTTGCTATC GTGCAAAATTA TTGGATTTTC 480
CAACTCCATC TGTAATCCCA TTGTCTATGC A 511
```

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

(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
15	(48) INFORMACIÓN PARA LA SEC. NO ID: 47:	
	(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	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 47:	
30	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	
	(ii) TIPO DE MOLÉCULA: ADN (genomic)	
45	(iv) ANTISENTIDO: SI	
	(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 48:	
	CAAGGATGAA GGTGGTGTAG A	21
50	(50) INFORMACIÓN PARA LA SEC. NO ID: 49:	
	(i) CARACTERÍSTICAS DE LA SECUENCIA:	
55	(A) LONGITUD: 23 pares de base	
	(B) TIPO: ácido nucléico	
	(C) TIPO DE HEBRA: sencilla	
	(D) TOPOLOGÍA: lineal	
60	(ii) TIPO DE MOLÉCULA: ADN (genomic)	
	(iv) ANTISENTIDO: SI	
65	(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:

(A) LONGITUD: 21 pares de base

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

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

GCAATGCAGG TCATAGTGAG C

21

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 pares de base

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(iii) HIPOTÉTICO: SI

(iv) ANTISENTIDO: SI

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 51:

TGGAGCATGG TGACGGGAAT GCAGAAG

27

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 pares de base

(B) TIPO: ácido nucleico

(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: 52:

GTGATGAGCA GGTCAGTGAG CGCCAAG

27

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 23 pares de base

(B) TIPO: ácido nucleico

(C) TIPO DE HEBRA: sencilla

(D) TOPOLOGÍA: lineal

(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

25

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

(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:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

45

(A) LONGITUD: 24 pares de base

(B) TIPO: ácido nucléico

(C) TIPO DE HEBRA: sencilla

50

(D) TOPOLOGÍA: lineal

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: SI

55

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

TGCGTGTTCC 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

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: NO

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

CAGGCCTTGG ATTTTAATGT CAGGGATGG

29

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 pares de base

(B) TIPO: ácido nucleico

(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: 58:

GGAGAGTCAG CTCTGAAAGA ATTCAGG

27

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 pares de base

(B) TIPO: ácido nucleico

(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: 59:

TGATGTGATG CCAGATACTA ATAGCAC

27

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 27 pares de base

(B) TIPO: ácido nucleico

(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: 60:

CCTGATTCAT TTAGGTGAGA TTGAGAC

27

(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:

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

(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:

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:

GATCAAGCTT CCATCCTACT GAAACCATGG TC

32

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

- (i) CARACTERÍSTICAS DE LA SECUENCIA:
- 25 (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
- 50 (ii) TIPO DE MOLÉCULA: ADN (genomic)
- (iv) ANTISENTIDO: NO
- (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)

(iv) ANTISENTIDO: SI

(xi) DESCRIPCIÓN DE LA SECUENCIA: SEC. NO ID: 68:

GTAAGCCTCC CAGAACGAGA GG

22

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

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24 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: 69:

CAGCGCAGGG TGAAGCCTGA GAGC

24

(71) INFORMACIÓN PARA LA SEC. NO ID: 70:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 24 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: 70:

GGCACCTGCT GTGACCTGTG CAGG

24

(72) INFORMACIÓN PARA LA SEC. NO ID: 71:

(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: 71:

GTCCTGCCAC TTCGAGACAT GG

22

(73) INFORMACIÓN PARA LA SEC. NO ID: 72:

(i) CARACTERÍSTICAS DE LA SECUENCIA:

(A) LONGITUD: 23 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: 72:

GAACTTCTC TGCCCTTACC GTC

23

(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

(ii) TIPO DE MOLÉCULA: ADN (genomic)

(iv) ANTISENTIDO: NO

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

CCAACACCAG CATCCATGGC ATCAAG

26

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

(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: 74:

GGAGAGTCAG CTCTGAAAGA ATTCAGG

27

LISTADO DE SECUENCIAS

<110> Arena Pharmaceuticals, Inc.

<120> Human Orphan G Protein-Coupled Receptors

<130> JEC/FP6282123

<140> EP 05003040.2

<140> 1999 10 13

<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

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

<150> 60/121,852
<151> 1999 02 26

15 <150> 60/123,946
<151> 1999 03 12

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

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

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

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

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

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

<150> 60/137,131
<151> 1999 05 28

45 <150> 601141,448
<151> 1999 06 29

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

<150> 60/156, 333
55 <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

<150> 601157,294

<151> 1999 10 01

5 <150> 60/157,281

<151> 1999 10 01

<150> 60/157,293

10 <151> 1999 10 01

<150> 601157,282

<151> 1999 10 01

15

<150> 09/417,044

<151> 1999 10 12

20 <150> 091416,760

<151> 1999 10 12

<160> 74

25

<170> PatentIn Ver. 2.1

<210> 1

30 <211> 1260

<212> DNA

<213> *Homo sapiens*

35 <400> 1

```

atgggtcttct cggcagtggt gactgagttc cataccggga catccaacac aacatttgtc 60
gtgtatgaaa acacctacat gaalattaca ctccctccac cattccagca tectgacctc 120
agtccattgc ttagatatag ttttgaancc atggctccca ctggtttgag ttcccttgacc 180
gtgaatagta cagctgtgcc cacaacacca gcagcattta agagcctaaa cttgcctctt 240
cagatcacc cttctgctat aatgatattc attctgtttg tgtcttttct tgggaacttg 300
gttgtttgcc tcatgtttta ccaaaaagct gccatgaggt ctgcaattaa catcctcctt 360
gccagcctag cttttgcaga catgttgctt gcagtgtga acatgcccct tgccttggtt 420
actattccta ctaccgagtg gatcttttgg aaattcttct gtagggtatc tgetatgttt 480
ttctggttat ttgtgataga aggagttagc atcctgtcca tcattagcat agataagttc 540
cttattatag tccagaggca ggataagcta aaccatata gagctaagg tctgattgca 600
gtttctctgg caacttcctt ttgtgtagct tttcctttag ccgtaggaaa ccccgacctg 660
cagataacct cccgagctcc ccagtgtgtg ttgggttaca caaccaatcc aggcctaccg 720
gcttatgtga ttttgatttc tctcatttct ttcttcata ccttcttggt aetactgtac 780
tcatttatgg gcatactcaa cacccttcgg cacaatgcct tgaggatcca tagctaccct 840
gaaggtatat gcctcagcca ggccagcaaa ctgggtctca tgagtctgca gagacctttc 900
cagatgagca ttgacatggg ctttaaaaaca cgtgccttca ccactatctt gattctcttt 960
gctgtcttca ttgtctgtg ggccccattc accacttaca gccctgtggc aacatteagt 1020
aagcactttt actatcagca caacttttct gagattagca cctggctact gtggctctgc 1080
tacctcaagt ctgcattgaa tccgctgctc tactactgga ggattaagaa attccatgat 1140
gcttgccctg acatgatgcc taagtccctc aagtttttgc cgcagctccc tggtcacaca 1200

```

aagcagcggg tacgtcctag tgctgtctat gtgtgtgggg aacatcggac ggtgggtgtg 1260

60

<210> 2

<211> 419

<212> PTR

65 <213> *Homo sapiens*

ES 2 289 610 T3

<400> 2

5	Met	Val	Phe	Ser	Ala	Val	Leu	Thr	Ala	Phe	His	Thr	Gly	Thr	Ser	Asn	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	Gln	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	115	120	125	
25	Leu	Leu	Ala	Val	Leu	Asn	Met	Pro	Phe	Ala	Leu	Val	Thr	Ile	Leu	Thr	130	135	140	
	Thr	Arg	Trp	Ile	Phe	Gly	Lys	Phe	Phe	Cys	Arg	Val	Ser	Ala	Met	Phe	145	150	155	160
30	Phe	Trp	Leu	Phe	Val	Ile	Glu	Gly	Val	Ala	Ile	Leu	Leu	Ile	Ile	Ser	165	170	175	
	Ile	Asp	Arg	Phe	Leu	Ile	Ile	Val	Gln	Arg	Gln	Asp	Lys	Leu	Asn	Pro	180	185	190	
35	Tyr	Arg	Ala	Lys	Val	Leu	Ile	Ala	Val	Ser	Trp	Ala	Thr	Ser	Phe	Cys	195	200	205	
40	Val	Ala	Phe	Pro	Leu	Ala	Val	Gly	Asn	Pro	Asp	Leu	Gln	Ile	Pro	Ser	210	215	220	
	Arg	Ala	Pro	Gln	Cys	Val	Phe	Gly	Tyr	Thr	Thr	Asn	Pro	Gly	Tyr	Gln	225	230	235	240
45	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	

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
 355 360 365
 Leu Ile Tyr Tyr Trp Arg Ile Lys Lys Phe His Asp Ala Cys Leu Asp
 370 375 380
 20 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
 30 <210> 3
 <211> 1119
 <212> DNA
 35 <213> *Homo sapiens*
 <400> 3
 40 atgttagcca acagctcctc aaccaacagt tctgttctcc cgtgtcctga ctaccgaact 60
 acccaccgcc tgcacttggg ggtctacagc ttggtgctgg ctgccgggct cccctcaac 120
 gcgctagccc tctgggtcct cctgcgcgcg ctgcgcgtgc actcgggtgg gagcgtgtac 180
 atgtgtaacc tggcggccag cgaactgtct ttcacctct cgtgtcccg tctctctcc 240
 tactacgcac tgcaccactg gcccttcccc gacctcctgt gccagacgac gggcgccatc 300
 ttcacagatga acatgtacgg tagctgcate ttcctgatgc tcataaagct ggaccgctac 360
 45 gccgccatcg tgcaccgct gcgactgcgc cacctgcggc ggcccgcgt ggcgcggctg 420
 ctctgcctgg gcgtgtgggc gctcactctg gtgtttgcg tgcgcgcgc cgcgtgcac 480
 aggcctcgc gttgcccta ccgggacctc gaggtgcgc tatgcttca gacctcagc 540
 gacgagctgt ggaaggcag gctgctgccc ctgctgctgc tggcagaggc gctgggcttc 600
 ctgctgcccc tggcggcggg ggtctactcg tcgggcagag tcttctggac gctggcgcgc 660
 50 ccgacgcca cgcagagcca gcggcggcgg aagaccgtgc gctcctgct ggctaaccctc 720
 gtcattcttc tgcgtgtgct cgtgccctac aacagcacgc tggcgggtcta cggcgtgctg 780
 cggagcaagc tggtagcggc cagcgtgcct gccgcgac gccgtgcgcg ggtgtgtgat 840
 gtgatgggtg tgctggcgg cgcctaactc gtgctggacc cgttgggtga ctactttagc 900
 gccgagggct tccgcaaac cctgcgcggc ctgggcactc cgcaccgggc caggacctcg 960
 55 gccaccaacg ggcgcgggc ggcgtcgcg caatccgaaa ggtccgcctg caccaccgac 1020
 gccaccaggc cggatgccgc cagtcagggg ctgctccgac cctccgactc ccactctctg 1080
 tcttccctca cagatgtcc ccaggattcc gccctctga 1119
 60 <210> 4
 <211> 372
 <212> PTR
 <213> *Homo sapiens*
 65

ES 2 289 610 T3

<400> 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	Gln	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	
25	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
45	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	

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 Gln 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
 15 Asp Ser Ala Leu
 370

20 <210> 5
 <211> 1107
 <212> DNA
 <213> *Homo sapiens*
 25 <400> 5

atggccaact ccacagggtt gaacgcctca gaagtcgcag gctcgttggg gttgacatctg 60
 gcagctgtcg tggaggttgg ggcactgctg ggcaacggcg cgttgcctgt cgtggtgctg 120
 30 cgcacgcccg gactgogcga cgcctcttac ctggcgacac tgtgcgtcgt ggacctgctg 180
 ggggcgcctt ccatcatgcc gctgggcttg ctggcgacac cgcgcgcggg gctgggcgcg 240
 gtgcgccttg gccccgcgc atgcgcgcgc gctgccttcc tctcgcgcgc tctgctgctg 300
 gcttgcacgc tgggggttgg cgcacttggc ctggcagcgt accgcctcat cgtgcacccg 360
 ctgcggccag gctgcgcgc gccgcctgtg ctgcgtctca cgcgcgtgtg ggcgcgcggc 420
 35 ygaatgttg ggcgccttc cctgctcggc cgcgcgcgcg caccgccttc tgcctcctgt 480
 cgcgccttgg tcttggcttg gggcctcggc ccttccgcgc cgtcttgggc cctgcttggc 540
 ttcggtgttc cgcgccttc gctgctcggc gcctacggcg gcctcttctt ggttgcgttc 600
 cgcgccttgc tggggccccc acggcgcgcg cgcgggtccc gactcgcctc ggactctctg 660
 gatagccgcc ttcccatctt gcgcgcgcgc cggcctcgcg tgcctcgggg caaggcgggc 720
 40 ctggcctcag cgttggcgtt gggccaatct gcagcctgct ggcctgcctt tggctgcgcg 780
 tgccttgggc cgcgcgcgc ggcgcgcgaa gccgaagcgc ctgtcacttg ggtcgcctac 840
 tgggccttgc cggctcacc cttcctgtac gggctgtgct agcgcctcgt gcgccttggc 900
 ctgggcgcgc tctctgcgc tgcactgctt ggacctgtgc yggcctgcac tccgcaagcc 960
 tggcaccgcg gggcactctt gcaatgcctc cagagacccc cagagggccc tgccttaggc 1020
 45 ccttcttggg ctccagaaca gaccccgag ttggcaggag ggcggagccc cgcataccag 1080
 gggcaccctg agagttctct ctctga 1107

<210> 6
 <211> 368
 50 <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
 20 25 30

65

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
 10 Ile Met Pro Leu Gly Leu Leu Ala Ala Pro Pro Pro Gly Leu Gly Arg
 65 70 75 80
 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
 145 150 155 160
 25 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
 195 200 205
 35 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
 275 280 285
 50 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
 325 330 335
 60 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      atggaatcat ctttctcatt tggagtgatc cttggtgttc tggcctccct catcattgct 60
      actaacacac tagtggctgt ggctgtgctg ctgttgatcc accagaatga tgggtgtcagt 120
      ctctgcttca ccttgaatct ggctgtggct gacaccttga ttggtgtgqc catctctgqc 180
      ctactcacag accagctctc cagcccttct cggccacac agaagacct gtgcagcctg 240
      cggatggcat ttgtcacttc ctccgcagct gctctgtcc tcacggtcac gctgatcacc 300
15      ttgacagggt accttgccat caagcagccc ctccgctact tgaagatcat gagtgggttc 360
      gtggccgggg cctgcattgc cgggctgtgg ttagtgtctt acctcattgg cttcctccca 420
      ctccgaatcc ccatgttcca gcagaactgc taaaaagggc agtgcagctt ctttgcctga 480
      ttccaccctc acctcgtgct gacctctctc tgcgttggct tcttcccage catgctcttc 540
      ttgtcttctt tctactgcga catgctcaag attgcttcca tgcacagcca gcagattcga 600
20      aagatggaac atgcaggagc catggctgga ggttatcgat cccacggac tcccagcgac 660
      ttcaaaagtc tccgtactgt gtctgttctc attggtgagct ttgctctatc ctggaccccc 720
      tctcttatca ctggcattgt gcagggtggc tgcacggagt gtcacctcta cctagtgtctg 780
      gaacggtaac tgtggtgctc cggcgtgggc aactccctgc tcaaccact catctatgcc 840
      tattggcaga agggaggcgc actgcagctc taccacatgg ccctaggagt gaagaagggtg 900
25      ctcaactcat tctcctctt tctctcggcc aggaattgtg gccagagag gccagggaa 960
      agttcctgtc acatcgtcac tatctccagc tcagagtttg 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
40      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
50      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
55      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
 145 150 155 160
 10 Phe His Pro His Phe Val Leu Thr Leu Ser Cys Val Gly Phe Phe Pro
 165 170 175
 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
 20 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
 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
 275 280 285
 35 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
 Ser Ser Cys His Ile Val Thr Ile Ser Ser Ser Glu Phe Asp Gly
 325 330 335

<210> 9

<211> 1413

<212> DNA

50 <213> *Homo sapiens*

<400> 9

55 atggacacta ccatggaagc tgacctgggt gccactggcc acaggccccg cacagagett 60
 gatgatgagg actcctatccc ccaagggtggc tgggacacgg tcttccctggg ggccctgctg 120
 ctctctgggc tggcagccaa tgggttgatg gcgtggctgg ccggctccca ggcccggtat 180
 ggagctggca cgcgtctggc gctgctctcg ctacgctgg cctctctga ctctctgttc 240
 ctggtagcag cggccttcca gatcctagag atccggcatg ggggacactg gccgctgggg 300
 acagctgcct gccgcttcta ctacttcta tggggcgtgt cctactcttc cggcctcttc 360
 ctgctggccg ccttcagcct cgaccgctgc ctgctggcgc tgtgcccaca ctggtaccct 420
 gggcacccgc cagtcgcct gccctcttg gtctgcgcc gtgtctgggt gctggccaca 480
 ccttcagcgc tgcctggct ggtcttccc gaggtcgccg tctgggtgta cgaacctggc 540

65

ES 2 289 610 T3

```

atctgectgg acttctggga cagcgaggag ctgtcgetga ggatgctgga ggtccctggg 600
ggcttctctgc ctttctctct gctgctctgc tgcctacgtgc tcacccaggc cacagcctgt 660
cgcacctgccc accgccaaca gcagcccgca gcctgcccgg gcttcgcccg tgtggccagg 720
accattctgt cagcctatgt ggtcctgagg ctgccctacc agctggccca gctgctctac 780
ctggccttcc tgtgggacgt ctactctggc tacctgctct gggaggccct ggtctactcc 840
gactacctga tctactcaa cagctgcctc agccccttcc tctgctcat ggcagtgcc 900
gacctccgga ccttgcctgc ctcctgctc tcgtccttcc cggcagctcc ctgcgaggag 960
cggccgggga gtttcaagcc cactgagcca cagacccagc tagattctga gggccaaact 1020
ctgccagagc cgtatggcga ggcccagtca cagatggatc ctgtggccca gcctcaggct 1080
aaccccaaac tccagccacg atcggaatccc acagctcagc caccagctga ccttaccggc 1140
cagccacagt cggatcccaac agcccagcca cagctgaacc tcatggccca gccacagtc 1200
gattctctgt cccagccaca ggcagacact aacgtccaga cccctgcacc tgcctgccag 1260
tctgtgcccc gtccctgtga tgaagcttcc ccaaccccat cctcgcctcc taccacagg 1320
gcctctgagg acccagccac acctcctgcc tctgaaggag aaagccccag cagcaccacc 1380
ccagagggcg ccccggggcg agcccccacg tga 1413

```

<210> 10

<211> 468

<212> PTR

<213> *Homo sapiens*

<400> 10

```

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

Leu Met Ala Trp Leu Ala Gly Ser Gln Ala Arg His Gly Ala Gly Thr
 50          55          60

40 Arg Leu Ala Leu Leu Leu Leu Ser Leu Ala Leu Ser Asp Phe Leu Phe
 65          70          75          80

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

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

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

50 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
 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

```

65

ES 2 289 610 T3

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

<210> 11

<211> 1248

<212> DNA

<213> *Homo sapiens*

ES 2 289 610 T3

<400> 11

```

5      atgtcaggga tggaaaaact tcagaatgct tcttggatct accagcagaa actagaagat 60
      ccattccaga aacacctgaa cagcaccgag gagtatcttg ccttctctctg cggacctctg 120
      cgcagccact tcttctctcc cgtgtctgtg gtgtatgtgc caatttttgt ggtgggggtc 180
      attggcaatg tcttgggtgt cctgggtgatt ctgcagcacc aggctatgaa gacgcccacc 240
      aactactacc tcttcagcct ggcggctctt gacctctctg tcttgccctt cggaaacgcc 300
      ctggagggtct atgagatgtg gcgcaactac cctttcttgt tcggggcccg gggtctctac 360
10     ttcaagacgg ccttctttga gacgtgtgc ttgcctcca tcttcagcat caccaccgtc 420
      agcgtggagc gctacgtggc catctacac ccgttcgcg ccaactgca gacgcccgg 480
      cgcggggccc tcaggatcct cggcatcgtc tggggcttct ccgtgctctt ctccctgcc 540
      aacactagca tccatggcat caagtccac tacttccca atgggtccct ggtcccagg 600
      tcggccacct gtaagggtcat caagcccatg tggatctaca atttcatcat ccaggctacc 660
      tcttctctat tctacctctt cccatgact gtcactagt tcttctacta cctcatggca 720
15     ctccagctaa agaaagagca atctctgag gcagatgaa ggaatgcata tattcaaga 780
      ccttgcaaaa aatcagtcac caagatgctg ttgtcttttg tcttagtgtt tgcctctctt 840
      tgggcccctt tccacattga ccgactcttc ttcagctttg tggaggagtg gagtgaatcc 900
      ctgctctgtg tgttcaacct cgtccatgtg gtgtcaggtg tcttcttcta cctgagctca 960
      gctgtcacc ccattatcta caactactg tctgcgcgt tccaggcagc attccagaat 1020
20     gtgatctctt ctttccaca acagtggcac tccagcagc acccacagtt gccctctgcc 1080
      cagcggaaac tcttctgtac agaatgccac ttgtgtggag tgaccgaaga tataggtccc 1140
      caattcccat gtcagtcata catgcacac tcttctcttc caactagctct ctctagtga 1200
      cagatgtcaa gaacaaacta tcaagcttc cactttaaca aaacctga 1248

```

<210> 12

<211> 415

<212> PTR

<213> *Homo sapiens*

<400> 12

```

35     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
      20          25          30

40     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
      50          55          60

45     Leu Val Cys Leu Val Ile Leu Gln His Gln Ala Met Lys Thr Pro Thr
      65          70          75          80

50     Asn Tyr Tyr Leu Phe Ser Leu Ala Val Ser Asp Leu Leu Val Leu Leu
      85          90          95

      Leu Gly Met Pro Leu Glu Val Tyr Glu Met Trp Arg Asn Tyr Pro Phe
      100         105         110

55     Leu Phe Gly Pro Val Gly Cys Tyr Phe Lys Thr Ala Leu Phe Glu Thr
      115         120         125

      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

```

ES 2 289 610 T3

	145		150		155		160
5	Arg Arg Ala Leu Arg	Ile Leu Gly Ile Val Trp Gly Phe Ser Val Leu					
		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
	Leu Arg Leu Lys Lys Asp Lys Ser Leu Glu Ala Asp Glu Gly Asn Ala						
20		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						
30		290		295			300
	Phe Asn Leu Val His Val Val Ser Gly Val Phe Phe Tyr Leu Ser Ser						
		305		310			315
	Ala Val Asn Pro Ile Ile Tyr Asn Leu Leu Ser Arg Arg Phe Gln Ala						
35		325		330			335
	Ala Phe Gln Asn Val Ile Ser Ser Phe His Lys Gln Phe His Ser Gln						
		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						
45		370		375			380
	Gln Ser Ser Met His Asn Ser His Leu Pro Thr Ala Leu Ser Ser Glu						
		385		390			395
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

60

atgccagata ctaatagcac mactcattta tcaactaagca ctgcgtgttac tttagcattt 60
tttatgtctt taqtatcttt tgcataaatg ctggyaaatg ctitggctat tttagctttt 120

65

ES 2 289 610 T3

```

gtgggtggaca aaaaccttag acatcgaagc agttattttt ttcttaacct ggccatctct 180
gacttctcttg tgggtgtgat ctccattctt ttgtacatcc ctacacagct gttcgaatgg 240
gatttttgaa aggaatctg tgtatttttg ctcactactg actatctgtt atgtacagca 300
tcgtatata acattgtctt catcagctat gatcgatacc tgcagtcctt aaatgctctg 360
tcttatagaa ctcaacatac tgggtgtctg aagattgtta ctctgattgt ggccgttttg 420
gtgctggcct tcttagtgaa tgggccaatg attctagttt cagagtcttg gaaggatgaa 480
ggtagtgaat gtgaacctgg atttctctcg gaatgggaca tccctgccaat caccatcttc 540
ttggaattcg tgaacctcag catcttagtc gcttattcca acatgaatat ttattggagc 600
ctgtggaagc gtgattatct cagtaggctc caaagccatc ctggactgac tgcctctctt 660
tccaacatct gtgacatctt attcagaggt agactatctt caaggagatc tctttcttga 720
tcgacagaag ctctctgcat ctttcattca gagagacaga ggagaaagag tagtctcatg 780
tttctctcaa gaaccaagat gaatagcaat ccaatcgctt ccaaaatggg ttccttctcc 840
caatcagatt ctgtagctct tcaccaaaag gaacatgttg aactgcttag agccaggaga 900
ttagccaagt cactggccat tctcttaggg gtcttttctg tttgctgggc tccatattct 960
ctgttctcaa tcttctcttc atttctcttc tcaagcaacag gtcttaaatc agtttggat 1020
agaattgcat tttggtctca gtggttcant tcttctgcca atctctctt gtatccattg 1080
tctcacaagc gctttcaaaa ggcttctctg aaaaatctt gatataaaa qcaacctcta 1140
cctctcacaac acagtcgggc agtatctctt taa 1173

```

<210> 14

<211> 390

<212> PTR

<213> *Homo sapiens*

<400> 14

```

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
20 25 30

Asn Ala Leu Val Ile Leu Ala Phe Val Val Asp Lys Asn Leu Arg His
35 40 45

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

Asp Phe Gly Lys Glu Ile Cys Val Phe Trp Leu Thr Thr Asp Tyr Leu
85 90 95

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

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

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
 Ser Thr Glu Val Pro Ala Ser Phe His Ser Glu Arg Gln Arg Arg Lys
 15 245 250 255
 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
 Leu Phe Thr Ile Val Leu Ser Phe Tyr Ser Ser Ala Thr Gly Pro Lys
 30 325 330 335
 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 ggccggcgccg aggcggccgc cctgggcctc 60
 aagctggcca cgctcagcct gctgctgtgc gtgagcctag cgggcaacgt gctgttcgcg 120
 ctgctgatcg tgcgggagcg cagcctgcac cgcgcctcgt actacctgct gctcgacctg 180
 tgcctggccg acgggctgcg cgcgctcgcc tgcctcccg cgcctatgct ggcggcgccg 240
 cgtgcggcgg ccgcggcggg ggcgcgcgcg ggcgcgctgg gctgcaagct gctcgccctc 300
 ctggccgcgc tcttctgctt ccacgcgcgc ttcctgctgc tgggcgtggg cgtcaccgcg 360
 60 tacctggcca tgcgcacca ccgcttctat gcagagcgcc tggccggctg gccgtgccc 420
 gccatgctgg tgtgcgcgc ctgggcgctg gcgctggccg cggccttccc gccagtgetg 480
 gacggcggtg gcgacgacga ggacgcgcg tgcgccttg agcagcgcc cgacggcgcc 540
 cccggcgccg tgggcttccct gctgctgctg gccgtggctg tgggcgccac gcacctcgtc 600

65

ES 2 289 610 T3

```

taccctccgcc tgcctcttctt caccacagac cgcgcgaaga tgcggccccc gcgccttggtg 660
cccgccgtca gccacgactg gaccttcac gccccggggc ccaccggcca ggccggccgc 720
aactggacgg cgggcttcgg ccgcccggcc ccgcccgcgc cgtttgtggg caccggccc 780
gcaggggccg gccggggcgc gcgcgcctc ctgctgctgg aagaattcaa gacggagaag 840
aggtctgtca agatgttcta cgccttcac ctgctcttcc tgcctctctg ggggcctac 900
gtcgtggcca gctaccctgc ggtcctggtg cggcccgggc cgtccccca ggcctacctg 960
acggcctccg tctggctgac cttcgcgcag gccggcaca acccctcctc gctctctc 1020
ttcaacaggg agctgagggg ctgcttcagg gccagttcc cctgctgcca gagcccccgg 1080
accaccaggc cgaaccatcc ctgcgacctg aaaggcattg gtttatga 1128

```

<210> 16

<211> 375

<212> PTR

<213> *Homo sapiens*

<400> 16

```

Met Ala Asn Ala Ser Glu Pro Gly Gly Ser Gly Gly Gly Glu Ala Ala
 1              5              10              15

Ala Leu Gly Leu Lys Leu Ala Thr Leu Ser Leu Leu Leu Cys Val Ser
20              25              30

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

Leu His Arg Ala Pro Tyr Tyr Leu Leu Leu Asp Leu Cys Leu Ala Asp
50              55              60

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

Arg Ala 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 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

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 Ala Phe Pro Pro Val Leu
145             150             155             160

Asp Gly 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

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

```

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
 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
 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
 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
 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

30 <210> 17
 <211> 1002
 <212> DNA
 35 <213> *Homo sapiens*
 <400> 17

atgaacacca cagtgatgca aggcctcaac agatctgagc ggtgccccag agacactcgg 60
 atagtacagc tgggtattccc agccctctac acagtgggtt tcttgaccgg catcctgctg 120
 aatacttttg ctctgtgggt gtttgttcac atccccagct cctccacctt catcatctac 180
 ctcaaaaaca ctctgggtggc cgacttgata atgacactca tgettccttc caaaatcctc 240
 tctgaactcac acctggcacc ctggcagctc agagcttttg tgtgtcgttc ttcttcgggtg 300
 atatttlatg agaccatgta tgtgggcacc gtgctgttag ggctcatagc ctttgacaga 360
 ttcttcanga tcatcagacc ttigagaaat attttcttaa aaaaacctgt ttttgcaaaa 420
 acggtctcaa tcttcattct gttctttttg tcttcattct cctgccaata taagactctg 480
 agcaacaagg aagcaaacacc atcgtctgtg aaaaagtgtg cttctcttaa ggggcctctg 540
 gggctgaaat ggcatacaat ggtaataaac atatgccagt ttattttctg gactgttttt 600
 atcctaattg ttgtgtttta tgtggttact gcaaaaaaag tatatgattc ttatagaaag 660
 tccaaaagta aggacagaaa aaacaaacaa aagctggaag gcaaaagtatt tgtgtcgtg 720
 gctgtctctt ttgtgtgttt tgcctcattt catcttgcca gagttccata tactcacagt 780
 caaaccaaca ataagactga ctgtagactg caaatcaac tgtttattgc taaagaaaca 840
 actctctctt tggcagcaac taacatttgt atggatccct taatatatat attcttatgt 900
 aaaaaattca cagaaaagct acctatgatg caaggagaa agaccacagc atcaagccaa 960
 55 gaaaatcata gcagtcagac agacaaacata acctaggct ga 1002

<210> 18
 <211> 333
 60 <212> PTR
 <213> *Homo sapiens*

65

ES 2 289 610 T3

<400> 18

	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	
10			35					40					45				
	Val	His	Ile	Pro	Ser	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	
	Ser	Asp	Ser	His	Leu	Ala	Pro	Trp	Gln	Leu	Arg	Ala	Phe	Val	Cys	Arg	
				85						90					95		
20	Phe	Ser	Ser	Val	Ile	Phe	Tyr	Glu	Thr	Met	Tyr	Val	Gly	Ile	Val	Leu	
				100					105					110			
	Leu	Gly	Leu	Ile	Ala	Phe	Asp	Arg	Phe	Leu	Lys	Ile	Ile	Arg	Pro	Leu	
25			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			
	Gln	Phe	Ile	Phe	Trp	Thr	Val	Phe	Ile	Leu	Met	Leu	Val	Phe	Tyr	Val	
40			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	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	
55			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	
	Glu	Asn	His	Ser	Ser	Gln	Thr	Asp	Asn	Ile	Thr	Leu	Gly				
65				325						330							

<210> 19

<211> 1122

ES 2 289 610 T3

<212> DNA

<213> *Homo sapiens*

5 <400> 19

```

10      atgqccaaca ctaccggaga gcttgaggag gtgagcggcg ctctgtcccc accgtccgca 60
      tcagcttatg tgaagctggt accgtctggg ctgatttatgt gcgtgagcct ggccgggtaac 120
      gccatcttct cctgctggt gctcaaggag cgtgccctgc acaaggctcc ttactacttc 180
      ctgctggacc tgtgccctggc cgtatggcata cgtctctgccg tctgcttccc ctttctgtctg 240
      gcttctgtcg gccacggctc ttcattggacc ttcagtgcac tcagctgcaa gatttctggcc 300
      lltatggccg tgcctcttctg cttccatgcg gccttcatgc tgttctgcat cagcgtcaccc 360
      cgttatcatgg ccatcgccca ccuacgcttc tacgccaagc gcatgacact ctggacatgc 420
      gggctctgca tctgcatggc ctggaccctg tctgtggcca tggccttccc acctgtcttc 480
15      gacgtgggca cctacaagtc tcttcgggag gaggaccagt gcattcttga gcattcgctac 540
      ttcaggcca atgacacgct gggcttcatg cttaagtctg ctgtgctcat ggcagctacc 600
      catgctgtct acggcagct gctcctcttc gagtatctgc accgcaagat gaagcagtg 660
      cagatggtgc cagccatcag ccagaactgg acattccatg gtcccgggc caccggccag 720
      gctgctgcca actggtatgc cggcttctggc cgtgggcca tggcaccac cctgctgggt 780
20      atccggcaga atgggcatgc agccagccg cggctactgg gcattggacg ggtcaagggt 840
      gaaaagcagc tgggcgcgat gtctacgcg atcacactgc tcttctgtct cctctgtgca 900
      cctacatcgt tggctgcta ctgacgagt tttgtgaaag cctgtgctgt gccccaccgc 960
      tacctggcca ctgctgtttg gatgagcttc gccaggtcg ccgtcaacc aattgtctgc 1020
      tctctgctca acaaggacct caagaagtgc ctgaccttc acgcccctg ctggggcaca 1080
25      ggaggctgcc cggctccag agaaccctac tgtgtcatgt ga
      1122

```

<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
      85          90          95

```

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

<210> 21

60 <211> 1053

<212> DNA

<213> *Homo sapiens*

65

ES 2 289 610 T3

<400> 21

```

5      atggcttttg aacagaacca gtcaacagat ttttattatg aggaaatga aatgaatggc 60
      acctatgact acagtcaata tgaattgata tgtatcaaaag aagatgtcag agaattttgc 120
      aaagttttcc tccctgtatt cctcacaata gctttcgtca ttggacttgc aggcaattcc 180
      atggtagtgg caatttatgc ctattacaag aaacagagaa ccaaaacaga tgtgtacatc 240
      ctgaatttgg ctgtagcaga tttactcctt ctattcactc tgcctttttg ggctgttaat 300
      gcagttcatg ggtgggtttt agggaaaata atgtgcaaaa taacttcagc cttgtacaca 360
      ctanaattttg tctctggaat gcagtttctg gcttgcataa gcatagacag atatgtggca 420
10     gtaactaatg tccccagcca atcaggagtg ggaaaaccat gctggatcat ctgtttctgt 480
      gtctggatgg ctgccatctt gctgagcata cccagctgg ttttttatac agtaaatgac 540
      aatgtcaggt gcattcccat tttccccgc tacttaggaa catcaatgaa agcattgatt 600
      caaatgctag agatctgcat tggatttcta gtaccttctt ctattatggg ggtgtgctac 660
      tttatcacgg caaggacact catgaagatg ccaaacatta aaatatctcg acccctaana 720
15     gttctgctca cagtcttat agttttcatt gtcaattcaac tgccttataa cattgtcaan 780
      tttctgcagag ccatagacat catctactcc ctgatcacca gctgcaacat gagcaaacgc 840
      atygacaleg ccattccaagt cacagaaaagc attgcactct ttcacagctg cctcaaccca 900
      atccttttatg tttttatggg agcatctttt aaaaactacg ttatgaaagt ggccaagaaa 960
      tatgggtcct ggagaagaca gagacaaaat gtggaggagt ttccttttga ttctgagggt 1020
20     cctacagagc caaccagtac ttttagcatt taa 1053

```

<210> 22

<211> 350

25 <212> PTR

<213> *Homo sapiens*

<400> 22

```

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

```

65

ES 2 289 610 T3

						165																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
--	--	--	--	--	--	-----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

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	Thr	Ala	Pro	Tcp	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	Val	Tyr	Ser	Ala	Val	Cys	Thr	Leu	Gly	Val
			35					40					45			
	Pro	Ala	Asn	Cys	Leu	Thr	Ala	Trp	Leu	Ala	Leu	Leu	Gln	Val	Leu	Gln
20		50					55					60				
	Gly	Asn	Val	Leu	Ala	Val	Tyr	Leu	Leu	Cys	Leu	Ala	Leu	Cys	Glu	Leu
	65					70					75					80
	Leu	Tyr	Thr	Gly	Thr	Leu	Pro	Leu	Trp	Val	Ile	Tyr	Ile	Arg	Asn	Gln
25					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
	Val	Gly	Ile	Val	His	Tyr	Pro	Val	Phe	Gln	Thr	Glu	Asp	Lys	Glu	Thr
40					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

ES 2 289 610 T3

	260	265	270
	Leu Glu Glu Arg Leu Tyr Thr Ala Ser Val Val Phe Leu Cys Leu Ser		
	275	280	285
5	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		
10	305	310	315
	Ser Met Lys Thr Asp Val Thr Arg Leu Thr His Ser Arg Asp Thr Glu		
	325	330	335
	Glu Leu Gln Ser Pro Val Ala Leu Ala Asp His Tyr Thr Phe Ser Arg		
15	340	345	350
	Pro Val His Pro Pro Gly Ser Pro Cys Pro Ala Lys Arg Leu Ile Glu		
	355	360	365
20	Glu Ser Cys		
	370		

<210> 25

25 <211> 1113

<212> DNA

<213> *Homo sapiens*

30 <400> 25

	atggcgaaact atagccatgc agctgacaac attttgcaaa atctctcggc tctaacagcc 60
	tttctgaanac tgacttcctt ggggttcata ataggagtca ggggtggtgg caacctcctg 120
35	atctccattt tgctagttaa agataagacc ttgcatagag caccctacta ctctctgttg 180
	gactcttgcg gttcagatct cctcagatct gcaatttgtt tccattttgt gttcaactct 240
	gttcaaaagt gctctacctg gacttatggg actctgactt gcaaagtgat tgcctttctg 300
	ggggttctgt tcttcttcca cactgccttc atgctctctt gcatcagtgt caacagcttc 360
	ttagctatcg cccatcaccg ctcttatata aggaggctga ccttttggac gtgtctggct 420
40	gtgatctgta tgggtgtggac tctgtctgtg gccatggcat tccctccggg ttttagacgtg 480
	ggcacttact catccattag ggaggaagat caatgcacct tccacacacc ctccttcagg 540
	gctaattgat ccttaggatt tatgctgctt ctgtctctca tccctcctag caccacagct 600
	gtctacctca agctgatatt ttctgtccac gatcgaagaa aaatgaagcc agtccagttt 660
	gtagcagcag tcagccagaa ctggactttt catgggtctg gagccagtgg ccaggcagct 720
45	gtcaatttgg tagcaggatt tgggaaggggt cccacaccac ccaccttgct gggcatcagg 780
	caaaatgcaa acactacagg cagaagaagg ctattgggtc tagacgagtt caaaatggag 840
	aaaagaatca gcagaatgtt ctatataatg acttttctgt ttctaacctt gtggggcccc 900
	tacctggtgg cctgttattg gagagttttt gcaagagggc ctgtagtacc agggggattt 960
	ctaacagctg ctgtctggat gagtttttgc caagcaggaa tcaatccttt tgtctgcat 1020
50	ttctcaaaac gggagctgag gcgtgtttt agcacaacc cctcttactg cagaaaatcc 1080
	aggttaccac ggggaacctt ctgtgttata tga 1113

<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 15

65

ES 2 289 610 T3

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

ES 2 289 610 T3

<210> 27

<211> 1080

<212> DNA

5 <213> *Homo sapiens*

<400> 27

```

10      atgcaggtcc cgaacagcac cggcccggac aacgcgacgc tgcagatgct gcggaacccg 60
      gcgatcgccg tggccctgcc cgtggtgtac tcgctggtgg cggcgggtcag catcccgggc 120
      aacctcttct ctctgtgggt gctgtgccg cgcattgggc ccagatcccc gtcggtcata 180
      ttcattgatca acctgagcgt caccgacctg atgctggcca gctgtgtgcc ttcccaaatc 240
      tactaccatt gcaacccgcc ccaactgggtt ttccgggtgc tgcctttgcaa cgtggtgacc 300
15      gtggcctttt acgcaaacat gtattccagc atcctcacca tgacctgtat cagcgtggag 360
      cgttctctgg gggctctgta cccgctcagc tccaagcgcg ggcgcgcgcg tcgttacgcg 420
      gtggcgcgcg gtcgagggac ctggtgctg ctcccgaccg cctgtgtccc gctggcgcgc 480
      accgatctca cctaccgggt gcaagccctg ggcattatca cctgtctcga cgtcctcaag 540
      tggacgatgc tccccagcgt ggcctgttgg gcggtgttcc tcttcaccat cttcatcttg 600
20      ctgttctca tcccgttctg gatcaaccgt gcttgttaca cggccaccat cctcaagctg 660
      ttgcgacagg aggaggcgca cggcggggag caggggaggc gcgcggtggg cctggccgcg 720
      gtggctcttg tggcctttgt cacctgcttc gccuccaaca acttcgtgct cctggcgcac 780
      atcgtgagcc gctgttctt cggcaagagc tactaccacg tgtacaagct cagcctgtgt 840
      ctacgctgcc tcaacaactg tctggaccgg ttgtttatt actttgcgtc ccgggaattc 900
25      cagctgcgcc tgcgggaata ttgggctgc cgcgggtgc ccagagacac cctggacacg 960
      cgcgggagga gctctcttcc cgcaggacc acgtccgtgc gctccgaggc cgggtgcgac 1020
      cctgaacggg ttgaggggag caccagggcc ygcctccaga ggcgggagc tgcgtctctg 1080

```

30 <210> 28

<211> 359

<212> PTR

35 <213> *Homo sapiens*

<400> 28

```

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

      Leu Arg Asn Pro Ala Ile Ala Val Ala Leu Pro Val Val Tyr Ser Leu
          20              25              30

45      Val Ala Ala Val Ser Ile Pro Gly Asn Leu Phe Ser Leu Trp Val Leu
          35              40              45

      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

```

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	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			
	Gln	Arg	Gln	Glu	Ser	Val	Phe										
				355													

55
 <210> 29
 <211> 1503
 <212> DNA
 60
 <213> *Homo sapiens*

65

ES 2 289 610 T3

<400> 29

```

5  atggagcgtc cctgggagga cagcccagge ccggaggggg cagctgaggg ctgcctgtg 60
   ccagtcgccg ccggggcgcg ctccggtgcc gcggcgagtg gcacaggtg gcagccatgg 120
   gctgagtgcc cgggacccaa ggggaggggg caactgctgg cgaccgcggg ccccttgcgt 180
   cgtgggcccg cccctcgcc tgcagctcc agcccccggc ccggagcggc gtccgctcac 240
   tcggttcagg gcagcgcgac tgcgggtggc gcaagccag ggcgagacc tttyggcgcg 300
   cggcccatgg agtcgggggt gctggggcg gcgcgggtga gcgaggtcat cgtcctgcat 360
   tacaactaca ccggcaagct ccgcggtgcg agctaccagc cgggtgcggg cctgcgcgtc 420
10  gacgcctggg tgtgctggc ggtgtgcgcc ttcatcgtgc tagagaatct agccgtgtcg 480
   ttggtgctcg gacgccacc gcgcctccac gctcccatgt tctgctcct gggcagcctc 540
   acgltgtcg atctgctggc aggcgcggc tacgcgcgca acatctact gtcggggcg 600
   ctacagctga aactgtccc cgcgctctgg ttcgcacggg agggagggct ctctgtgga 660
   ctacagctg cgtgctgag cctcctggcc atcgcgtgg agcgcagcct caccatggcg 720
15  cccadqqqgc ccqcccccgt ctccagtcgg gggcgacgc tggcgatggc agccgcggcc 780
   tggggcggtg cgtgctcct ccggctcctg ccagcgctgg gctggaattg cctgggtcgc 840
   ctggagcctt gctccactgc cttgcgcgtc taagccaagg cctacgtgct ctctgtcggtg 900
   ctgcctctcg tgggcatctt ggcgcgcatc tgtgcatctt acgcgcgcat ctactgccag 960
   gtacgcgcga acgcgcggcg cctgcgggca cggcccgga ctgcggggac cactcgacc 1020
20  cgggcgcgtc gcaagcccg cttcttgcc ttgctgcga cgtcagcgt ggtgctcctg 1080
   gctctgtgg catgttggg cccctcttc ctgctgtgt tgcctgacgt ggcgtgccc 1140
   gcgcgcactt gctctgtact cctgcaggc gatcccttc tgggactggc catggccaac 1200
   tcaettctga accccatcat ctacacgctc accaaccggg acctgcgcca cgcgtcctg 1260
   cgcctgggtt gctgcggacg ccactcctgc ggcagagacc cgagtggctc ccagcagtcg 1320
   gcgagcgcg ctgaggttc cgggggctg cgcgcgtgcc tgcgccggg ccttgatggg 1380
25  agcttcaagc gctcggagcg ctcatcgccc cagcgcgacg ggttggacc cagcggctcc 1440
   acaggcagcc ccggtgcacc cacagccgce cggactctgg tatcagaacc ggtgcagac 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

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

   Pro Ser Pro Ala Ser 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

   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					
	Val	Leu	Ser	Leu	Leu	Ala	Ile	Ala	Leu	Glu	Arg	Ser	Leu	Thr	Met	Ala	
20	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			
	Thr	Thr	Ser	Thr	Arg	Ala	Arg	Arg	Lys	Pro	Arg	Ser	Leu	Ala	Leu	Leu	
40				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	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	

ES 2 289 610 T3

```

          435          440          445
Gly Leu Arg Arg Cys Leu Pro Pro Gly Leu Asp Gly Ser Phe Ser Gly
  450          455          460
5
Ser Glu Arg Ser Ser Pro Gln Arg Asp Gly Leu Asp Thr Ser Gly Ser
  465          470          475          480
10
Thr Gly Ser Pro Gly Ala Pro Thr Ala Ala Arg Thr Leu Val Ser Glu
          485          490          495
Pro Ala Ala Asp
          500
15
<210> 31
<211> 1029
<212> DNA
20
<213> Homo sapiens
<400> 31
25
atgcaagccg tcgacaatct cacctctgcg cctgggaaca ccagtcctgtg caccagagac 60
tacanaatca cccaggtcct ctteccactg etclacactg tccctgtttt tgttggactt 120
atcacaaatg gcctggcgat gaggattttc ttccaatcc ggagtaaatc aaactttatt 180
atttttctta agaacacagt cattctctgat ctctctatga ttctgacttt tccattcaaa 240
attctctagt atgccaaact gggaacagga ccactgagaa cttttgtgtg tcaagttacc 300
tcgcctatat ttattttcac aatgtatata agtattttcat tccctgggact gataaactatc 360
gatcgctacc agaagaccac caggccattt aaaacatcca accccaaaaa tctcttgggg 420
gctaagattc tctctgttgt catctgggca tccatgttct tactctcttt gccctaacatg 480
attctgacca acaggcagcc gagagacaag aatgtgaaga aatgctcttt ccttaaatca 540
gagctcgggc tagtctggca tgaatatgta aattacatct gtcaagtcct tttctggatt 600
aatctcttaa ctgttattgt atgttataca ctcatlacaa agaacctgta ccggtcctac 660
35
gttagaarya ggggtgtagg taaagtccc aggaaaaagg tgaacgtcaa agttttcatt 720
atcat tgcgt cttcttttat ttgttttgtt cttttccatt ttgcccqaat tctttacacc 780
ctgagcccaa cctgggatgt ctttgactgc actgctgaaa atactctgtt ctctgtgaaa 840
gagagcaatc tctgtttaa ttctttaa atgctgcttg atcgttcat ctattttttt 900
ctttgcaagt ccttcagaaa ttctttgata agtatgctga agtgccccaa ttctgcaaca 960
40
tctctgtccc aggaacaata gaaaaaagaa caggatgggt gtgacccaaa tgaagagact 1020
ccaatgtaa
1029
45
<210> 32
<211> 343
<212> PTR
<213> Homo sapiens
50
<400> 32
55
Met Gln Ala Val Asp Asn Leu Thr Ser Ala Pro Gly Asn Thr Ser Leu
  1          5          10          15
60
Cys Thr Arg Asp Tyr Lys Ile Thr Gln Val Leu Phe Pro Leu Tyr
          20          25          30
65
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
65

```

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	
	Pro Phe Lys Thr Ser Asn Pro Lys Asn Leu Leu Gly Ala Lys Ile Leu	130	135	140	
15	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	
	Tyr Thr Leu Ile Thr Lys Glu Leu Tyr Arg Ser Tyr Val Arg Thr Arg	210	215	220	
30	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	
	Leu Asn Ala Cys Leu Asp Pro Phe Ile Tyr Phe Phe Leu Cys Lys Ser	290	295	300	
45	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> *Homo sapiens*

65

ES 2 289 610 T3

<400> 33

```

5      atgtcgggtct gctaccgtcc cccagggaac gagacactgc tgagctggaa gacttcgcgg 60
      gccacaggca cagccttccct gctgctggcg gcgtgctgg ggctgctgg caacggcttc 120
      gtgggtgtgga gcttggcggg ctggcgggct gcacgggggc gacgctggc ggccacgctt 180
      ctgctgcacc tggcgctggc cgaaggcgcg gtgctgctgc tcacgctgct ctltgtggcc 240
      ttcctgaccc ggcaaggcctg gctgctgggc caggcgggct gcaaggcggt gtactacgty 300
      tgcgcgctca gcatgtacgc cagcgtgctg ctacgggccc tgcctagcct gcagcgctgc 360
10     ctgcgagtca cccgccccct cctggcgctc cggtgctgca gcccgccct ggccgcgcgc 420
      ctgctgctgg cggctctggct ggccgcccct ttgctcgccg tcccgggcgc cgtctaccgc 480
      cactgttggg gggaccgctg atgccaagct tggcaccct cggcggtcca cgcgcgcgc 540
      cacttgagcc tggagactct gaccgcttcc gtgcttccct tcgggctgat gctcggtgc 600
      tacagcgtga cgttggcagc gctgcggggc gcccgctggg gctccgggcg gcacggggcg 660
15     cgggtggggc ggctgggtgag cggcctcgtg ctgtgcttcc gcttgcctcg gggccctac 720
      cagcgagtca accttctgca ggcggtcgcg gccttggctc caccggaagg accttggcg 780
      aagctggggc gaggcgccca ggcggcgccg gcgggaacta cggccttggc ctctctcagt 840
      tctagcgtca accgggtgct ctacgctctc accgctggag atctgctgca cggcgaggt 900
      cccgcttccc tcacgctgct ctctgaaggc tctggggagg cccgaggggg cggcgctct 960
20     aggggaaggga ccatggagct ccgaactacc cctcagctga aagtggtggg gcagggccgc 1020
      ggcaatggag acccgggggg tgggagggag aaggacggtc cgggaatggga ccttctga 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

        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

        Ala Leu Ala Asp Gly Ala Val Leu Leu Leu Thr Pro Leu Phe Val Ala
        65             70             75             80

        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

        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 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

        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
5	His Ala Ala Ala His Leu Ser Leu Glu Thr Leu Thr Ala Phe Val Leu 180 185 190		
	Pro Phe Gly Leu Met Leu Gly Cys Tyr Ser Val Thr Leu Ala Arg Leu 195 200 205		
10	Arg Gly Ala Arg Trp Gly Ser Gly Arg His Gly Ala Arg Val Gly Arg 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 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 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		
35	Gly Phe Glu Trp Asp Leu 355		
	<210> 35		
	<211> 1005		
40	<212> DNA		
	<213> <i>Homo sapiens</i>		
	<400> 35		
45	atgctggggg tcatggcatg gaatgcaact tgcaaaaact ggctggcagc agaggctgcc 60		
	ctggaaaagt actacatttc cactttttat gggattgagt tcgttgtggg agtcccttga 120		
	aattaccattg ttgtttacgg ctacatcttc tctctgaaga actggaacag cagtaatat 180		
	tctctcttta acctctctgt ctctgactta gcttttctgt gcacctccc catgctgata 240		
	aggagttatg ccaatggaaa ctggatatac ggagacgtgc tctgcataag caaccgatat 300		
50	gtgcttcacg ccaacctcta taccagcatt ctctttctca cttttatcag cataqatcga 360		
	tacttgataa ttaagtatcc ttcccgagaa cactctctgc aaaagaaaga gtttgctatt 420		
	ttaatctctt tggccatttg ggttttagta accttagagt tactacctat acctccctt 480		
	ataaatctctg ttaraactga caatggcacc acctgtaatg atcttgcaag ttctggagac 540		
	cccaactacn acctcattta cagcatgtgt cttaacctgt tggggctcct tcttctctt 600		
55	tttgctgagt gtttctttta ttacaagatt gctctcttcc taaggcagag gaataggcag 660		
	gttgcctactg ctctgcccc ttgaaaagcct ctcaacttgg tcatcatggc agtggtatc 720		
	ttctctgtgc tttttacacc ctatcacgtc atgcggaatg tgaggatcgc ttcacgctg 780		
	gagagttgga agcagtatca gtgcactcag gtgcctatca actcctttta ctttgtgaca 840		
60	cggccttttg cctttctgna cagtgtcatt aacctgtct tctattttct ttggggagat 900		
	cacttcaggg acatgctgat gaatcaactg agacacaact tcaaatccct taccatcttt 960		
	agcagatggg ctcatgaact cctactttca ttacagagaaa 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	Trp	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	
	Glu	Phe	Val	Val	Gly	Val	Leu	Gly	Asn	Thr	Ile	Val	Val	Tyr	Gly	Tyr	35	40	45	
15	Ile	Phe	Ser	Leu	Lys	Asn	Trp	Asn	Ser	Ser	Asn	Ile	Tyr	Leu	Phe	Asn	50	55	60	
	Leu	Ser	Val	Ser	Asp	Leu	Ala	Phe	Leu	Cys	Thr	Leu	Pro	Met	Leu	Ile	65	70	75	80
20	Arg	Ser	Tyr	Ala	Asn	Gly	Asn	Trp	Ile	Tyr	Gly	Asp	Val	Leu	Cys	Ile	85	90	95	
25	Ser	Asn	Arg	Tyr	Val	Leu	His	Ala	Asn	Leu	Tyr	Thr	Ser	Ile	Leu	Phe	100	105	110	
	Leu	Thr	Phe	Ile	Ser	Ile	Asp	Arg	Tyr	Leu	Ile	Ile	Lys	Tyr	Pro	Phe	115	120	125	
30	Arg	Glu	His	Leu	Leu	Gln	Lys	Lys	Glu	Phe	Ala	Ile	Leu	Ile	Ser	Leu	130	135	140	
	Ala	Ile	Trp	Val	Leu	Val	Thr	Leu	Glu	Leu	Leu	Pro	Ile	Leu	Pro	Leu	145	150	155	160
35	Ile	Asn	Pro	Val	Ile	Thr	Asp	Asn	Gly	Thr	Thr	Cys	Asn	Asp	Phe	Ala	165	170	175	
40	Ser	Ser	Gly	Asp	Pro	Asn	Tyr	Asn	Leu	Ile	Tyr	Ser	Met	Cys	Leu	Thr	180	185	190	
	Leu	Leu	Gly	Phe	Leu	Ile	Pro	Leu	Phe	Val	Met	Cys	Phe	Phe	Tyr	Tyr	195	200	205	
45	Lys	Ile	Ala	Leu	Phe	Leu	Lys	Gln	Arg	Asn	Arg	Gln	Val	Ala	Thr	Ala	210	215	220	
	Leu	Pro	Leu	Glu	Lys	Pro	Leu	Asn	Leu	Val	Ile	Met	Ala	Val	Val	Ile	225	230	235	240
50	Phe	Ser	Val	Leu	Phe	Thr	Pro	Tyr	His	Val	Met	Arg	Asn	Val	Arg	Ile	245	250	255	
55	Ala	Ser	Arg	Leu	Gly	Ser	Trp	Lys	Gln	Tyr	Gln	Cys	Thr	Gln	Val	Val	260	265	270	
	Ile	Asn	Ser	Phe	Tyr	Ile	Val	Thr	Arg	Pro	Leu	Ala	Phe	Leu	Asn	Ser	275	280	285	
60	Val	Ile	Asn	Pro	Val	Phe	Tyr	Phe	Leu	Leu	Gly	Asp	His	Phe	Arg	Asp	290	295	300	
	Met	Leu	Met	Asn	Gln	Leu	Arg	His	Asn	Phe	Lys	Ser	Leu	Thr	Ser	Phe	305	310	315	320
65	Ser	Arg	Trp	Ala	His	Glu	Leu	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

```

10      atgcaggcgc ttaacattac cccggagcag ttctctcggc tgctgcggga ccacaacctg 60
      acgcgggagc agttcatcgc tctgtaccgg ctgcgaccgc tegtctacac cccagagctg 120
      ccgqgaecgc ccaagctggc cctcgtgctc accggcgctc tcatcttcgc cctggcgctc 180
      tttqgcaatg ctctgggtgtt ctacgtgggt acccgcagca aggccatgcg caccgtcacc 240
      aacatcttta tctgtctcctt ggcgctcagt gacctgctca tcacctctct ctgcattccc 300
15      gtcaccatgc tccagaacat ttccgacaae tggctggggg gtgcttccat ttgcaagatg 360
      gtgccatttg tccagtctac cgcctgttgt acagaaatgc tcactatgac ctgcattgct 420
      gtggaaaaggc accagggact tgtgcaccc tttaaaatga agtggcaata caccacccga 480
      agggctttca caatgctagg tgtggtcttg ctggtggcag tcatcgtagg atcaccatg 540
      tggcacgtgc aaacaactga gatcaaatat gacttcctat atgaanaagg acacatctgc 600
20      tgcttagaag agtggaccag cccgtgacac cagaagatct acaccacctt catccttgc 660
      atcctcttcc tctgtcctct tatggtgatg cttattctgt acagtaaaat tggttatgaa 720
      ccttggataa agaaaaaggt tggggatggt tcagtgcctc gaactattca tggaaaagaa 780
      atgtccaaaa tagccagga gaagaaacga gctgtcaltc tcatggggtc cgtgggtggc 840
      ctctttgctg tgtgctgggc accattccat gttgtccata tcatgattga atacagtatt 900
25      tttgaaaagg aatatgatga tgtcacnate aaagtgtatt ttgctatcgt gcaaatattt 960
      ggattttcca actccatctg taatcccatt gtctatgcat ttatgaatga aaacttcaaa 1020
      aaaaatgttt tgtctgcagt ttgttattgc atagtaaat aacacctctc tccagacaaa 1080
      aggcattgaa attcagggaat tacaatgatg cgggaagaaag caaagtcttc cctcagagag 1140
      aatccagtgg aggaaccaca aggagaagca ttcagtgatg gcaacattga agtcaaatg 1200
      tgtgaacaga cagaggagaa gaaaaagctc aaacgacatc ttgctctctt taggtctgaa 1260
30      ctggctgaga attctccttt agacagtggg cattaa 1296
  
```

<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
50
80
55
60
65
  
```

ES 2 289 610 T3

	35	40	45
5	Val Leu Thr Gly Val Leu Ile Phe Ala Leu Ala Leu Phe Gly Asn Ala 50 55 60		
	Leu Val Phe Tyr Val Val Thr Arg Ser Lys Ala Met Arg Thr Val Thr 65 70 75 80		
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 145 150 155 160		
	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		
	Leu Tyr Glu Lys Glu His Ile Cys Cys Leu Glu Glu Trp Thr Ser Pro 195 200 205		
30	Val His Gln Lys Ile Tyr Thr Thr Phe Ile Leu Val Ile Leu Phe Leu 210 215 220		
	Leu Phe Leu Met Val Met Leu Ile Leu Cys Ser Lys Ile Gly Val Glu 225 230 235 240		
35	Leu 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 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		
	Tyr Asp Asp Val Thr Ile Lys Met Ile Phe Ala Ile Val Gln Ile Ile 305 310 315 320		
50	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		

ES 2 289 610 T3

	Asn Lys Thr Phe Ser Pro Ala Gln Arg His Gly Asn Ser Gly Ile Thr 355 360 365	
5	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 Glu Val Lys Leu 385 390 395 400	
10	Cys Glu Gln Thr Glu Glu Lys 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 <212> DNA <213> <i>Homo sapiens</i> <400> 39	
25	ctgtgtacag cagttcgag agtg	24
	<210> 40 <211> 24 <212> DNA <213> <i>Homo sapiens</i> <400> 40	
35	gagtgccagg cagagcaggt agac	24
40	<210> 41 <211> 31 <212> DNA <213> <i>Homo sapiens</i> <400> 41	
50	cccgattcc tctgtctcc cagcttgccc c	31
	<210> 42 <211> 32 <212> DNA <213> <i>Homo sapiens</i> <400> 42	
60	tgtggatcct gctgtcaaag gtccattcc gg	32
	<210> 43 <211> 20 <212> DNA <213> <i>Homo sapiens</i>	

ES 2 289 610 T3

<400> 43

tcacaatgct aggtgtgggc

20

5

<210> 44

<211> 22

<212> DNA

10 <213> *Homo sapiens*

<400> 44

15 tgcatagaca atgggattac ag

22

<210> 45

<211> 511

20 <212> DNA

<213> *Homo sapiens*

<400> 45

25

```
tcacaatgct aggtgtgggc tggctgggagg cagtcacgtt aggatcacc atgtggcacg 60
tcacaatgct aggtgtgggc tggctgggagg cagtcacgtt aggatcacc atgtggcacg 120
tcacaatgct aggtgtgggc tggctgggagg cagtcacgtt aggatcacc atgtggcacg 180
tcacaatgct aggtgtgggc tggctgggagg cagtcacgtt aggatcacc atgtggcacg 240
tcacaatgct aggtgtgggc tggctgggagg cagtcacgtt aggatcacc atgtggcacg 300
tcacaatgct aggtgtgggc tggctgggagg cagtcacgtt aggatcacc atgtggcacg 360
tcacaatgct aggtgtgggc tggctgggagg cagtcacgtt aggatcacc atgtggcacg 420
tcacaatgct aggtgtgggc tggctgggagg cagtcacgtt aggatcacc atgtggcacg 480
tcacaatgct aggtgtgggc tggctgggagg cagtcacgtt aggatcacc atgtggcacg 511
```

35

<210> 46

<211> 21

<212> DNA

40 <213> *Homo sapiens*

<400> 46

45 ctgcttagaa gaggtagaca g

21

<210> 47

<211> 22

50 <212> DNA

<213> *Homo sapiens*

<400> 47

55

ctgtgcacca gaagatctac ac

22

<210> 48

60 <211> 21

<212> DNA

<213> *Homo sapiens*

65 <400> 48

caaggatgaa ggtggttag a

21

ES 2 289 610 T3

	<210> 49	
	<211> 23	
	<212> DNA	
5	<213> <i>Homo sapiens</i>	
	<400> 49	
10	gtgtagatct tccggtgcac agg	23
	<210> 50	
	<211> 21	
15	<212> DNA	
	<213> <i>Homo sapiens</i>	
	<400> 50	
20	gcaatgcagg tcatagtga c	21
	<210> 51	
25	<211> 27	
	<212> DNA	
	<213> <i>Homo sapiens</i>	
30	<400> 51	
	tggagcatgg tgacgggaat gcagaag	27
35	<210> 52	
	<211> 27	
	<212> DNA	
40	<213> <i>Homo sapiens</i>	
	<400> 52	
45	gtgatgagca ggctactgag cgccaag	27
	<210> 53	
	<211> 23	
	<212> DNA	
50	<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 acctgaaggg 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	tgcgtgttc tggaccctca cgtg	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	
	<212> DNA	
50	<213> <i>Homo sapiens</i>	
	<400> 59	
55	tgatgtgatg ccagatacta atagcac	27
	<210> 60	
60	<211> 27	
	<212> DNA	
	<213> <i>Homo sapiens</i>	
65	<400> 60	
	cctgattcat ttagtgaga ttgagac	27

ES 2 289 610 T3

	<210> 61	
	<211> 21	
	<212> DNA	
5	<213> <i>Homo sapiens</i>	
	<400> 61	
10	gacaggtacc ttgcatcaa g	21
	<210> 62	
	<211> 22	
15	<212> DNA	
	<213> <i>Homo sapiens</i>	
	<400> 66	
20	ctgcacaatg ccagtataa gg	22
	<210> 63	
25	<211> 27	
	<212> DNA	
	<213> <i>Homo sapiens</i>	
30	<400> 63	
	ctgacttctt gttcctggca gcagcgg	27
35	<210> 64	
	<211> 27	
	<212> DNA	
40	<213> <i>Homo sapiens</i>	
	<400> 64	
45	agaccagcca gggcacgctg aagagtg	27
	<210> 65	
	<211> 32	
50	<212> DNA	
	<213> <i>Homo sapiens</i>	
	<400> 65	
55	gatcaagctt ccatctact gaaacatgg tc	32
	<210> 66	
60	<211> 35	
	<212> DNA	
	<213> <i>Homo sapiens</i>	
65	<400> 66	
	gatcagatct cagttccaat attcacacca ccgtc	35

ES 2 289 610 T3

	<210> 67	
	<211> 22	
	<212> DNA	
5	<213> <i>Homo sapiens</i>	
	<400> 67	
10	ctggtgtgct ccatggcatc cc	22
	<210> 68	
	<211> 22	
15	<212> DNA	
	<213> <i>Homo sapiens</i>	
	<400> 68	
20	gtaagcctcc cagaacgaga gg	22
	<210> 69	
25	<211> 24	
	<212> DNA	
	<213> <i>Homo sapiens</i>	
30	<400> 69	
	cagcgcaggg tgaagcctga gage	24
35	<210> 70	
	<211> 24	
	<212> DNA	
40	<213> <i>Homo sapiens</i>	
	<400> 70	
45	ggcacctgct gtgacctgtg cagg	24
	<210> 71	
	<211> 22	
	<212> DNA	
50	<213> <i>Homo sapiens</i>	
	<400> 71	
55	gtcctgccac ttcgagacat gg	22
	<210> 72	
60	<211> 23	
	<212> DNA	
	<213> <i>Homo sapiens</i>	
65	<400> 72	
	gaaacttctc tgcccttacc gtc	23

ES 2 289 610 T3

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