

(19)



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS
ESPAÑA



(11) Número de publicación: **2 886 958**

(51) Int. Cl.:

A01K 67/027 (2006.01)
C12N 9/64 (2006.01)

(12)

TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Fecha de presentación y número de la solicitud internacional: **27.02.2017 PCT/US2017/019574**

(87) Fecha y número de publicación internacional: **08.09.2017 WO17151453**

(96) Fecha de presentación y número de la solicitud europea: **27.02.2017 E 17709888 (6)**

(97) Fecha y número de publicación de la concesión europea: **02.06.2021 EP 3422845**

(54) Título: **Roedores que tienen un gen TMPRSS humanizado**

(30) Prioridad:

29.02.2016 US 201662301023 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:

21.12.2021

(73) Titular/es:

**REGENERON PHARMACEUTICALS, INC.
(100.0%)
777 Old Saw Mill River Road
Tarrytown, NY 10591-6707, US**

(72) Inventor/es:

**PURCELL NGAMBO, LISA;
MUJICA, ALEXANDER O. y
TANG, YAJUN**

(74) Agente/Representante:

PONS ARIÑO, Ángel

ES 2 886 958 T3

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

DESCRIPCIÓN

Roedores que tienen un gen TMPRSS humanizado

5 Antecedentes

Las serina proteasas transmembrana de tipo II son una familia de proteasas caracterizadas por un dominio transmembrana aminoterminal (Bugge *et al.*, J. Biol. Chem. 284 (35): 23177-23181, 2009; Hooper *et al.*, J. Biol. Chem. 272(2): 857-860, 2001). Todos los miembros de esta familia se expresan como zimógenos monocatenarios y se activan proteolíticamente por escisión dentro de un motivo R/(IV)VGG altamente conservado. Un miembro de la familia, la proteasa transmembrana, serina de tipo 4 (TMPRSS4), se ha demostrado que activa el canal de sodio epitelial (ENaC, por sus siglas en inglés) que regula el flujo de sodio y agua a través de los epitelios (Guipponi *et al.* 2002 Hum. Mol. Genet. 11:2829; Vuagniaux *et al.* 2002 J. Gen. Physiol. 120:191). Se desconocen los activadores proteolíticos de TMPRSS4; sin embargo, los datos disponibles hasta la fecha sugieren que la proteína se activa a sí misma. Cuando se activa, el dominio catalítico de TMPRSS4 permanece unido al extremo N de la proteína a través de un enlace disulfuro. Se ha demostrado que TMPRSS4, TMPRSS2 y TMPRSS11D (o proteasa similar a tripsina de las vías respiratorias humana; "HAT, por sus siglas en inglés") escinden la hemaglutinina de la gripe A (HA) *in vitro*, que es la primera etapa esencial en el ciclo de vida vírico. Esta escisión es esencial para la actividad de HA, ya que la proteína se sintetiza como una proteína precursora (HA0) y necesita la escisión en HA1 y HA2 para su actividad. El ARNi con atenuación génica de TMPRSS4 en células Caco-2 dio como resultado una reducción de la propagación del virus. Así mismo, se demostró que TMPRSS4 está fuertemente regulada positivamente en los pulmones de ratones infectados con gripe (Böttcher *et al.* 2006 J. Virol. 80:9896; Böttcher *et al.* 2009 Vaccine 27: 6324; Böttcher-Friebershauser *et al.* 2010 J. Virol. 84: 5604; Bertam *et al.* 2010 J. Virol. 84:10016; Bertam *et al.* 2010 J. Virol. 84:10016; Böttcher-Friebershauser *et al.* 2011 J. Virol. 85: 1554; Bahgat *et al.* 2011 Virol. J. 8:27).

25 Se necesita la creación de un sistema *in vivo*, por ejemplo, un modelo de infección en roedores, para identificar y probar compuestos que incluyen anticuerpos que se dirigen específicamente a las serina proteasas transmembrana de tipo II humanas para el tratamiento y la prevención de infecciones víricas y otras enfermedades.

30 También se hace referencia a los siguientes documentos:

- Kuhn *et al* (2015) consultado de internet: URL: http://elib.tihhannover.de/dissertations/kuehnn_s15.pdf que se refiere a estudios sobre la respuesta del hospedador a las infecciones por el virus de la gripe A en mutantes nuligénicos de ratón;
- WO 2013/158516 A1 que se refiere a métodos para tratar o prevenir la infección por el virus de la gripe administrando un inhibidor de serina proteasas;
- Sol *et al* (2009) consultado de internet: URL: <http://www.dtic.mil/get-tr-doc/pdf?AD=ADA525092> que se refiere a la caracterización de la proteasa TMPRSS2 como moduladora de la metástasis de cáncer de próstata;
- Böttcher-Friebertshauser *et al* (2010) Journal of Virology 84 (11): 5605-5614, que se refiere a la escisión de la hemaglutinina del virus de la gripe por las proteasas de las vías respiratorias TMPRSS2 y HAT, que difiere en la localización subcelular y la susceptibilidad a los inhibidores de proteasas;
- Devoy *et al* (2012) Nature Reviews Genetics, 13: 14-20 que se refiere a ratones humanizados genómicamente: tecnologías y promesas.
- Murphy *et al* (2014) PNAS USA 111(14): 5153-5158, que se refiere a ratones con humanización de megabases de sus genes de inmunoglobulinas que generan anticuerpos tan eficazmente como ratones normales; y
- Maconald *et al* (2014) PNAS USA 111(14): 5147-5152 que se refiere a la humanización genética precisa e *in situ* de 6 Mb de genes de inmunoglobulina de ratón.

55 Sumario

La presente invención abarca el reconocimiento de que es deseable modificar por ingeniería animales roedores para proporcionar sistemas de identificación *in vivo* y creación de nuevas terapias. Por ejemplo, la presente invención abarca el reconocimiento de que los roedores que tienen un gen *Tmprss* humanizado son deseables para su uso en la identificación y creación de agentes terapéuticos para el tratamiento y la prevención de infecciones víricas.

En un aspecto, la invención proporciona un roedor cuyo genoma comprende un gen *Tmprss* humanizado, en donde el gen *Tmprss* humanizado:

- 65 - comprende una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS* humano afín,

- se encuentra en un locus *Tmprss* de roedor endógeno, y es el resultado de una sustitución de una secuencia genómica del gen *Tmprss* de roedor endógeno con dicha secuencia de nucleótidos del gen *TMPRSS* humano afín y está bajo el control del promotor del gen *Tmprss* de roedor endógeno, y
 - codifica una proteína *Tmprss* humanizada que comprende: (i) un ectodominio que tiene al menos un 85 % de identidad de secuencia con el ectodominio de la proteína *TMPRSS* humana codificada por el gen *TMPRSS* humano afín y (ii) una porción citoplasmática y transmembrana que tiene al menos un 85 % de identidad de secuencia con la porción citoplasmática y transmembrana de la proteína *Tmprss* de roedor endógena codificada por el gen *Tmprss* de roedor endógeno.
- 5 10 15 20 25 30 35 40 45 50 55 60
- El gen *Tmprss* humanizado en roedores divulgado en el presente documento codifica una proteína *Tmprss* humanizada que contiene un ectodominio al menos un 85 % (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico en secuencia al ectodominio de una proteína *TMPRSS* humana. La proteína *Tmprss* humanizada contiene una porción citoplasmática y transmembrana que es al menos un 85 % (p. ej., al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntica en secuencia a la porción citoplasmática y transmembrana de una proteína *Tmprss* de roedor endógena.
- Un roedor divulgado en el presente documento contiene un gen *Tmprss* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS* humano afín, en donde la secuencia de nucleótidos del gen *TMPRSS* humano afín codifica un polipéptido al menos un 85 % (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico en secuencia al ectodominio de la proteína *TMPRSS* humana codificada por el gen *TMPRSS* humano afín. Un roedor divulgado en el presente documento contiene un gen *Tmprss* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS* humano afín, en donde la secuencia de nucleótidos del gen *Tmprss* de roedor endógeno codifica un polipéptido al menos un 85 % (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico en secuencia a la porción citoplasmática y transmembrana de la proteína *Tmprss* de roedor endógena codificada por el gen *Tmprss* de roedor endógeno.
- En algunas realizaciones, un roedor divulgado en el presente documento contiene un gen *Tmprss* humanizado localizado en un locus *Tmprss* de roedor endógeno que es el resultado de una sustitución de una secuencia genómica contigua de un gen *Tmprss* de roedor endógeno con una secuencia genómica contigua de un gen *TMPRSS* humano afín. En realizaciones específicas, la secuencia genómica contigua de un gen *TMPRSS* humano afín que se está insertando incluye secuencias de exones que codifican un ectodominio sustancialmente idéntico al ectodominio de la proteína *TMPRSS* codificada por el gen *TMPRSS* humano. En algunas realizaciones, la secuencia genómica contigua de un gen *TMPRSS* humano afín también incluye la 3' UTR del gen *TMPRSS* humano afín.
- En algunas realizaciones, un roedor divulgado en el presente documento es heterocigoto para un gen *Tmprss* humanizado en un locus *Tmprss* de roedor endógeno. En otras realizaciones, un roedor es homocigoto para un gen *Tmprss* humanizado en un locus *Tmprss* de roedor endógeno.
- En realizaciones adicionales, un roedor contiene dos o más genes *Tmprss* humanizados en diferentes locus *Tmprss* de roedor endógenos estando cada locus *Tmprss* de roedor endógeno humanizado con un gen *TMPRSS* humano afín; por ejemplo, dos o más de los genes *Tmprss2* humanizado, *Tmprss4*, humanizado y *Tmprss11d* humanizado.
- En algunas realizaciones, un roedor divulgado en el presente documento contiene un gen *Tmprss2* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss2* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS2* humano, en donde el gen *Tmprss2* humanizado está bajo el control del promotor del gen *Tmprss2* de roedor endógeno.
- En algunas realizaciones, el gen *Tmprss2* humanizado codifica una proteína *Tmprss2* humanizada que contiene un ectodominio al menos un 85 % (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico en secuencia con el ectodominio de la proteína *TMPRSS2* humana codificada por el gen *TMPRSS2* humano utilizado en la humanización. La proteína *TMPRSS2* humana contiene, en algunas realizaciones, una secuencia de aminoácidos al menos un 85 % idéntica (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntica a la secuencia de aminoácidos expuesta en la SEQ ID NO: 4. En algunas realizaciones, una proteína *Tmprss2* humanizada contiene un ectodominio al menos un 85 % (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico a la secuencia de aminoácidos compuesta por los restos W106 a G492 o los 387 aminoácidos carboxiterminales de una proteína *TMPRSS2* humana como se expone en, por ejemplo, la SEQ ID NO: 4. En algunas realizaciones, el gen *Tmprss2* humanizado codifica una proteína *Tmprss2* humanizada que además contiene una porción citoplasmática y transmembrana que es al menos un 85 % (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntica a la porción citoplasmática y transmembrana de la proteína *Tmprss2* de roedor codificada por el gen *Tmprss2* de roedor endógeno que se está humanizando. Una proteína *Tmprss2* de roedor endógena ilustrativa se expone en la SEQ ID NO: 2.
- En algunas realizaciones, un roedor contiene un gen *Tmprss2* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss2* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS2* humano, en donde la secuencia de nucleótidos del gen *TMPRSS2* humano codifica un ectodominio al menos un 85 % (por ejemplo,

- menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico en secuencia con el ectodominio de la proteína TMPRSS2 humana codificada por el gen *TMPRSS2*. En realizaciones específicas, la secuencia de nucleótidos de un gen *TMPRSS2* humano es una secuencia genómica contigua de un gen *TMPRSS2* humano que contiene el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *TMPRSS2* humano. En realizaciones 5 particulares, la secuencia genómica contigua de un gen *TMPRSS2* humano contiene además la 3' UTR del gen *TMPRSS2* humano. En algunas realizaciones, la secuencia de nucleótidos de un gen *Tmprss2* de roedor endógeno incluido en un gen *Tmprss2* humanizado codifica una porción citoplasmática y transmembrana que es al menos un 85 % (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 % idéntica a la porción citoplasmática y transmembrana 10 de la proteína *Tmprss2* de roedor endógena codificada por el gen *Tmprss2* de roedor endógeno.
- En realizaciones particulares, un gen *Tmprss2* humanizado contiene los exones 1-2 codificantes de un gen *Tmprss2* de roedor endógeno y el exón 4 codificante hasta el exón 13 codificante de un gen *TMPRSS2* humano, en donde el 15 gen *Tmprss2* humanizado codifica una proteína *Tmprss2* humanizada que contiene una porción citoplasmática y transmembrana que es sustancialmente idéntica a la porción citoplasmática y transmembrana de la proteína *Tmprss2* de roedor codificada por el gen *Tmprss2* de roedor endógeno y un ectodominio que es sustancialmente idéntico al 20 ectodominio de la proteína *TMPRSS2* humana codificada por el gen *TMPRSS2* humano. El gen *Tmprss2* humanizado contiene un exón 3 que en algunas realizaciones codifica el exón 3 de un gen *TMPRSS2* humano y en otras realizaciones codifica el exón 3 de un gen *Tmprss2* de roedor endógeno. En algunas realizaciones, el gen *Tmprss2* humanizado contiene un exón 3 que incluye una porción 5' del exón 3 codificante de un gen *Tmprss2* de roedor endógeno y una porción 3' del exón 3 codificante de un gen *TMPRSS2* humano.
- En algunas realizaciones, un roedor divulgado en el presente documento contiene un gen *Tmprss4* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss4* de roedor endógeno y una secuencia de nucleótidos de un 25 gen *TMPRSS4* humano, en donde el gen *Tmprss4* humanizado está bajo el control del promotor del gen *Tmprss4* de roedor endógeno.
- En algunas realizaciones, el gen *Tmprss4* humanizado codifica una proteína *Tmprss4* humanizada que contiene un ectodominio al menos un 85 % (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico en secuencia 30 con el ectodominio de la proteína *TMPRSS4* humana codificada por el gen *TMPRSS4* humano utilizado en la humanización. La proteína *TMPRSS4* humana contiene, en algunas realizaciones, una secuencia de aminoácidos al menos un 85 % idéntica (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 % idéntica) a la secuencia de aminoácidos expuesta en la SEQ ID NO: 11. En algunas realizaciones, una proteína *Tmprss4* humanizada contiene 35 un ectodominio al menos un 85 % (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico a la secuencia de aminoácidos compuesta por los restos K54 a L437 o los 384 aminoácidos carboxiterminales de una proteína *TMPRSS4* humana como se expone en, por ejemplo, la SEQ ID NO: 11. En algunas realizaciones, el gen *Tmprss4* humanizado codifica una proteína *Tmprss4* humanizada que además contiene una porción citoplasmática y transmembrana que es al menos un 85 % (al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntica a la porción 40 citoplasmática y transmembrana de la proteína *Tmprss4* de roedor codificada por el gen *Tmprss4* de roedor endógeno que se está humanizando. Una proteína *Tmprss4* de roedor endógena ilustrativa se expone en la SEQ ID NO: 9.
- En algunas realizaciones, un roedor contiene un gen *Tmprss4* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss4* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS4* humano, en donde la 45 secuencia de nucleótidos de un gen *TMPRSS4* humano codifica un ectodominio al menos un 85 % idéntico al ectodominio de la proteína *TMPRSS4* humana codificada por el gen *TMPRSS4* humano. En realizaciones específicas, la secuencia de nucleótidos de un gen *TMPRSS4* humano es una secuencia genómica contigua que contiene el exón 4 codificante hasta el codón de parada en el exón 13 codificante de un gen *TMPRSS4* humano. En algunas realizaciones, la secuencia de nucleótidos de un gen *Tmprss4* de roedor endógeno incluido en un gen *Tmprss4* humanizado codifica una porción citoplasmática y transmembrana que es sustancialmente idéntica a la porción 50 citoplasmática y transmembrana de la proteína *Tmprss4* de roedor codificada por el gen *Tmprss4* de roedor endógeno.
- En realizaciones particulares, un gen *Tmprss4* humanizado contiene el exón 1 codificante hasta el exón 3 codificante de un gen *Tmprss4* de roedor endógeno y el exón 4 codificante hasta el codón de parada en el exón 13 codificante de un gen *TMPRSS4* humano.
- En algunas realizaciones, un roedor divulgado en el presente documento contiene un gen *Tmprss11d* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss11d* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS11D* humano, en donde el gen *Tmprss11d* humanizado está bajo el control del promotor del gen *Tmprss11d* de roedor endógeno.
- En algunas realizaciones, el gen *Tmprss11d* humanizado codifica una proteína *Tmprss11d* humanizada que contiene 55 un ectodominio al menos un 85 %, (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico en secuencia con el ectodominio de la proteína *TMPRSS11D* humana codificada por el gen *TMPRSS11D* humano utilizado en la humanización. La proteína *TMPRSS11D* humana contiene, en algunas realizaciones, una secuencia de aminoácidos al menos un 85 % idéntica (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 % idéntica) a la secuencia de aminoácidos expuesta en la SEQ ID NO: 18. En algunas realizaciones, una proteína *Tmprss11d* humanizada contiene 60 un ectodominio al menos un 85 % (al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntico a la secuencia de

- aminoácidos compuesta por los restos A42-I418 o los 377 aminoácidos carboxiterminales de una proteína TMPRSS11D humana como se expone en, por ejemplo, la SEQ ID NO: 18. En algunas realizaciones, el gen *Tmprss11d* humanizado codifica una proteína *Tmprss11d* humanizada que además contiene una porción citoplasmática y transmembrana que es al menos un 85 %, (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 %) idéntica a la porción citoplasmática y transmembrana de la proteína *Tmprss11d* de roedor endógena codificada por el gen *Tmprss11d* de roedor endógeno que se está humanizando. Una proteína *Tmprss11d* de roedor endógena ilustrativa se expone en la SEQ ID NO: 16.
- 5 En algunas realizaciones, un roedor contiene un gen *Tmprss11d* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss11d* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS11D* humano, en donde la secuencia de nucleótidos del gen *TMPRSS11D* humano codifica un ectodominio sustancialmente idéntico al ectodominio de la proteína *TMPRSS11D* humana codificada por el gen *TMPRSS11D* humano. En realizaciones específicas, la secuencia de nucleótidos de un gen *TMPRSS11d* humano es una secuencia genómica contigua que contiene el exón 3 codificante hasta el codón de parada en el exón 10 codificante de un gen *TMPRSS11D* humano. En realizaciones particulares, la secuencia genómica contigua de un gen *TMPRSS11D* humano contiene además la 3' UTR del gen *TMPRSS11D* humano. En algunas realizaciones, la secuencia de nucleótidos de un gen *Tmprss11d* de roedor endógeno incluido en un gen *Tmprss11d* humanizado codifica una porción citoplasmática y transmembrana que es al menos un 85 % idéntica a la porción citoplasmática y transmembrana de la proteína *Tmprss11d* de roedor codificada por el gen *Tmprss11d* de roedor endógeno.
- 10 En realizaciones particulares, un gen *Tmprss11d* humanizado contiene los exones 1-2 codificantes de un gen *Tmprss11d* de roedor endógeno y el exón 3 codificante hasta el exón 13 codificante de un gen *TMPRSS11D* humano.
- 15 En otro aspecto, la invención proporciona una célula o tejido de roedor aislado cuyo genoma contiene un gen *Tmprss* humanizado como se describe en el presente documento. En realizaciones específicas, el gen *Tmprss* humanizado se selecciona del grupo que consiste en un gen *Tmprss2* humanizado, un gen *Tmprss4* humanizado y un gen *Tmprss11d* humanizado.
- 20 En otro aspecto más, la invención proporciona una célula madre embrionaria de roedor cuyo genoma contiene un gen *Tmprss* humanizado como se describe en el presente documento. En realizaciones específicas, el gen *Tmprss* humanizado se selecciona del grupo que consiste en un gen *Tmprss2* humanizado, un gen *Tmprss4* humanizado y un gen *Tmprss11d* humanizado.
- 25 En otro aspecto, también se proporciona un embrión de roedor generado a partir de la célula madre embrionaria de roedor divulgada en el presente documento.
- 30 También se divulga un vector de ácido nucleico adecuado para su uso en la humanización de un gen *Tmprss* endógeno en un roedor. En algunos casos, el vector de ácido nucleico incluye una secuencia del ácido nucleico de *Tmprss* humano (por ejemplo, un ADN genómico humano que codifica el ectodominio de una proteína *TMPRSS* humana), flanqueado por un brazo de homología 5' y un brazo de homología 3'. Los brazos de homología 5' y 3' son secuencias de un ácido nucleico que se colocan en 5' y 3', respectivamente, de la secuencia del ácido nucleico de *Tmprss* humano y son homólogas a las secuencias de ADN genómico en un locus *Tmprss* endógeno en un roedor que flanquea un ADN genómico de roedor que codifica el ectodominio de una proteína *Tmprss* de roedor afín. Por tanto, los brazos de homología 5' y 3' son capaces de mediar la recombinación homóloga y la sustitución del ADN genómico de roedor que codifica el ectodominio de la proteína *Tmprss* de roedor afín con la secuencia del ácido nucleico de *Tmprss* humano para formar un gen *Tmprss* humanizado como se describe en el presente documento.
- 35 En un aspecto adicional, la invención se refiere a un método para proporcionar un roedor cuyo genoma contiene un gen *Tmprss* humanizado. El método incluye modificar el genoma de un roedor para reemplazar una secuencia genómica de un gen *Tmprss* de roedor endógeno con una secuencia genómica de un gen *TMPRSS* humano afín para formar un gen *Tmprss* humanizado.
- 40 En algunas realizaciones, la invención proporciona un método para hacer que un roedor (tal como un ratón o una rata) tenga un gen *Tmprss* humanizado, que comprende:
- 45 (a) insertar un fragmento genómico en un locus *Tmprss* de roedor endógeno en una célula madre embrionaria de roedor, comprendiendo dicho fragmento genómico una secuencia de nucleótidos de un gen *TMPRSS* humano afín, formando, de este modo, un gen *Tmprss* humanizado,
- 50 en donde el gen *Tmprss* humanizado está bajo el control del promotor del gen *Tmprss* de roedor en el locus *Tmprss* de roedor endógeno y codifica una proteína *Tmprss* humanizada, y
- 55 en donde la proteína *Tmprss* humanizada comprende (i) un ectodominio que tiene al menos un 85 % de identidad de secuencia con el ectodominio de la proteína *TMPRSS* humana codificada por el gen *TMPRSS* humano afín y (ii) una porción citoplasmática y transmembrana que tiene al menos un 85 % de identidad de secuencia con la porción citoplasmática y transmembrana de la proteína *Tmprss* de roedor codificada por el gen *Tmprss* de roedor en dicho locus *Tmprss* de roedor endógeno;

- (b) obtener una célula madre embrionaria de roedor que comprende el gen *Tmprss* humanizado de (a); y,
- (c) crear un roedor utilizando la célula madre embrionaria de roedor de (b).

- 5 En algunas realizaciones, el gen *Tmprss* humanizado se selecciona del grupo que consiste en un gen *Tmprss2* humanizado, un gen *Tmprss4*, humanizado y un gen *Tmprss11d* humanizado. En diversas realizaciones, el gen *Tmprss* humanizado codifica una proteína *Tmprss* humanizada que contiene un ectodominio al menos un 85 % idéntico (por ejemplo, al menos un 90 %, 95 %, 98 %, 99 % o 100 % idéntico en secuencia) al ectodominio de la proteína TMPRSS humana codificada por el gen *TMPRSS* humano utilizado para la humanización. En realizaciones 10 específicas, la proteína *Tmprss* humanizada contiene el ectodominio de una proteína TMPRSS humana seleccionada del grupo que consiste en una proteína TMPRSS2 humana, una proteína TMPRSS4 humana y una proteína TMPRSS 11D humana. En realizaciones específicas, la proteína *Tmprss* humanizada contiene además una porción citoplasmática y transmembrana que es sustancialmente idéntica a la porción citoplasmática y transmembrana de la proteína *Tmprss* de roedor codificada por el gen *Tmprss* de roedor endógeno que se está humanizando.
- 15 En otro aspecto, la invención proporciona un método para utilizar un roedor divulgado en el presente documento para evaluar la eficacia terapéutica de un compuesto (por ejemplo, inhibidores candidatos que se dirigen específicamente a una proteína TMPRSS humana) en el tratamiento de la infección por el virus de la gripe. En particular, la presente invención proporciona un método para evaluar la eficacia terapéutica de un compuesto en el tratamiento de una 20 infección por el virus de la gripe A, que comprende:

proporcionar un roedor de la presente invención;
 administrar un virus de la gripe A y un compuesto candidato al roedor, en donde opcionalmente el compuesto candidato es un anticuerpo o fragmento de unión a antígeno del mismo específico para una proteína TMPRSS humana; y
 25 monitorizar la presencia y la gravedad de la infección por el virus de la gripe A en el roedor para determinar la eficacia terapéutica del compuesto candidato.

- 30 En algunas realizaciones, el virus de la gripe se administra al roedor antes que el compuesto. En otras realizaciones, el virus de la gripe se administra al roedor después del compuesto.

- 35 En algunas realizaciones, el compuesto candidato es un anticuerpo o fragmento de unión a antígeno del mismo específico para una proteína TMPRSS humana. En realizaciones específicas, el compuesto candidato es un anticuerpo o fragmento de unión a antígeno del mismo específico para una proteína TMPRSS humana seleccionada del grupo que consiste en una proteína TMPRSS2 humana, una proteína TMPRSS4 humana y una proteína TMPRSS11D humana.

- Otras características, objetivos y ventajas de la presente invención son evidentes en la descripción detallada a continuación. Debe entenderse, sin embargo, que la descripción detallada, aunque indica realizaciones de la presente 40 invención, se proporciona solamente a modo de ilustración, no de limitación. Serán evidentes diversos cambios y modificaciones dentro del alcance de la invención para los expertos en la materia a partir de la presente descripción detallada.

Breve descripción de los dibujos

- 45 Los dibujos incluidos en el presente documento, que están compuestos por las siguientes figuras, son solamente para fines de ilustración y no de limitación.

Figuras 1A-1D. Estrategia ilustrativa para la humanización de *Tmprss2* de ratón.

- 50 La figura 1A muestra un diagrama, no a escala, de la organización genómica de los genes *Tmprss2* de ratón y *TMPRSS2* humano. Los exones están representados por barras delgadas colocadas a lo largo de las secuencias genómicas, indicando el primer exón codificante para ambos genes mediante el codón de inicio "ATG" por encima del exón e indicando el último exón codificante mediante el codón de "parada" por encima del exón. Se indican un fragmento genómico de ratón de aproximadamente 25.291 pb a eliminar y un fragmento genómico humano de aproximadamente 25.091 pb a insertar. Se indican las ubicaciones de las sondas utilizadas en un ensayo descrito en el ejemplo 1. DT: dominio transmembrana; SRCR: dominio de tipo rico en cisteína del receptor eliminador; LDLRa: receptor de lipoproteínas de baja densidad de clase A.

- 55 La figura 1B ilustra, no a escala, un vector BAC modificado ilustrativo para la humanización de un gen *Tmprss2* de ratón endógeno, junto con las secuencias de unión (SEQ ID NO: 22, 23 y 24).

- 60 La figura 1C ilustra, no a escala, un alelo *Tmprss2* humanizado después de que se haya eliminado el casete de neomicina, junto con las secuencias de unión (SEQ ID NO: 22 y 25).

- 65 La figura 1D establece una alineación de secuencias de una proteína TMPRSS2 humana (SEQ ID NO: 4), una

proteína Tmprss2 de ratón (SEQ ID NO: 2) y una proteína Tmprss2 humanizada ("pro 7010 mutante") (SEQ ID NO: 7).

Figuras 2A-2D. Estrategia ilustrativa para la humanización de *Tmprss4* de ratón.

- 5 La figura 2A muestra un diagrama, no a escala, de la organización genómica de los genes *Tmprss4* de ratón y *TMPRSS4* humano. Los exones están representados por barras delgadas colocadas a lo largo de las secuencias genómicas, indicando el primer exón (también el primer exón codificante) para ambos genes mediante el codón de inicio "ATG" por encima del exón e indicando el último exón codificante mediante el codón de "parada" por encima del exón. Se indican el fragmento genómico de ratón de aproximadamente 11.074 pb a eliminar y el fragmento genómico humano de aproximadamente 14.963 pb a insertar. Se indican las ubicaciones de las sondas utilizadas en un ensayo descrito en el ejemplo 2. DT: dominio transmembrana; SRCR: dominio de tipo rico en cisteína del receptor eliminador; LDLRa: receptor de lipoproteínas de baja densidad de clase A.
- 10 15 La figura 2B ilustra, no a escala, un vector BAC modificado ilustrativo para la humanización de un gen *Tmprss4* de ratón endógeno, junto con las secuencias de unión (SEQ ID NO: 38, 39 y 40).
- 20 La figura 2C ilustra, no a escala, un alelo *Tmprss4* humanizado después de que se haya eliminado el casete de neomicina, junto con las secuencias de unión (SEQ ID NO: 41 y 40).
- 25 La figura 2D establece una alineación de secuencias de una proteína TMPRSS4 humana (SEQ ID NO: 11), una proteína Tmprss4 de ratón (SEQ ID NO: 9) y una proteína Tmprss4 humanizada ("pro 7224 mutante") (SEQ ID NO: 14).

Figuras 3A-3D. Estrategia ilustrativa para la humanización de *Tmprss11d* de ratón.

- 30 La figura 3A muestra un diagrama, no a escala, de la organización genómica de los genes *Tmprss11d* de ratón y *TMPRSS11D* humano. Los exones están representados por barras delgadas colocadas a lo largo de las secuencias genómicas, indicando el primer exón (también el primer exón codificante) para ambos genes mediante el codón de inicio "ATG" por encima del exón e indicando el último exón codificante mediante el codón de "parada" por encima del exón. Se indican un fragmento genómico de ratón de aproximadamente 35.667 pb a eliminar y un fragmento genómico humano de aproximadamente 33.927 pb a insertar. Se indican las ubicaciones de las sondas utilizadas en un ensayo descrito en el ejemplo 3. DT: dominio transmembrana; SEA: dominio que se encuentra en la proteína, enterocinasa y agrina de esperma de erizo de mar.
- 35 40 La figura 3B ilustra, no a escala, un vector BAC modificado ilustrativo para la humanización de un gen *Tmprss11d* de ratón endógeno, junto con las secuencias de unión (SEQ ID NO: 57, 58 y 59).
- 45 La figura 3C ilustra, no a escala, un alelo *Tmprss11* humanizado después de que se haya eliminado el casete de neomicina, junto con las secuencias de unión (SEQ ID NO: 57 y 60).
- 50 La figura 3D establece una alineación de secuencias de una proteína TMPRSS11D humana (SEQ ID NO: 18), una proteína Tmprss11d de ratón (SEQ ID NO: 16) y una proteína Tmprss11d humanizada ("pro 7226 mutante") (SEQ ID NO: 21).

55 La **figura 4** muestra los resultados de un experimento que muestra que los ratones MAID7225 HumlnTMPRSS4 no difieren en su susceptibilidad a la exposición con dosis altas de gripe A H1N1 grave o H3N2 adaptada a ratón grave. Los ratones MAID7225 Humln TMRPSS4 expuestos con A/Puerto Rico/08/1934 (H1N1) (círculos de color gris claro, línea discontinua) mostraron tasas de supervivencia similares en comparación con los ratones de tipo silvestre (cuadrados de color gris claro, línea discontinua). De forma análoga, los ratones MAID7225 Humln TMRPSS4 expuestos con A/Aichi/02/1968-X31 (H3N2) (triángulos de color gris oscuro, línea discontinua) mostraron tasas de supervivencia similares en comparación con los ratones de tipo silvestre (triángulos invertidos de color gris claro, línea discontinua). Los ratones se infectaron IN el día 0 con 1150 UFP de A/Puerto Rico/08/1934 (H1N1) o 10.000 UFP de A/Aichi/02/1968-X31 (H3N2). El grupo de control incluía ratones MAID7225 Humln TMPRSS4 no infectado como control negativo y ratones de tipo silvestre (rombos negros, línea continua).

Descripción detallada de determinadas realizaciones

- 60 65 La presente invención se refiere a roedores genéticamente modificados (por ejemplo, ratones y ratas) que tienen un gen humanizado que codifica una serina proteasa transmembrana de tipo II (o "Tmprss", para proteasa/serina transmembrana). Los roedores genéticamente modificados son adecuados para su uso en la exploración de compuestos candidatos que se dirigen específicamente a una molécula de TMPRSS humana para el tratamiento y la prevención de enfermedades tales como la infección por el virus de la gripe. Por consiguiente, la presente invención proporciona roedores genéticamente modificados que tienen un gen *Tmprss* humanizado, células y tejidos aislados de roedores genéticamente modificados, métodos y composiciones para producir roedores genéticamente modificados

y el uso de roedores genéticamente modificados para explorar y probar compuestos terapéuticos. Las diversas realizaciones de la presente invención se describen con más detalle a continuación.

Serina proteasas transmembrana de tipo II ("Tmprss")

5 Las serina proteasas transmembrana de tipo II, también denominadas en el presente documento "Tmprss" para moléculas no humanas o "TMPRSS" para moléculas humanas ("proteasa/serina transmembrana"), son una familia de 10 proteínas caracterizadas por un dominio transmembrana aminoterminal y un dominio serina proteasa extracelular carboxiterminal. Se han identificado al menos 18 miembros en la familia, que se agrupan en cuatro subfamilias (Bugge 15 et al. (2009), *supra*). Todos los miembros comparten varias características estructurales comunes que definen a la familia, incluyendo (i) un dominio citoplasmático aminoterminal corto, (ii) un dominio transmembrana y (iii) un ectodominio que contiene un dominio proteasa y una región tallo que une el dominio transmembrana con el dominio proteasa. La región tallo contiene una combinación de dominios estructurales modulares de seis tipos diferentes: un dominio SEA (proteína/enteropeptidasa/agrina de esperma de erizo de mar), un dominio receptor eliminador del grupo 20 A, un dominio LDLA (receptor de lipoproteínas de *baja densidad de clase A*), un dominio CUB (factor de crecimiento embrionario de erizo Cls/Cir, proteína morfogenética ósea 1), un dominio MAM (meprina/antígeno A5/proteína receptora fosfatasa mu) y un dominio frizzled. Véase la revisión de Bugge et al. (2009), *supra*. Por ejemplo, TMPRSS2 y TMPRSS4, pertenecen ambas a la subfamilia hepsina/TMPRSS, tienen un dominio receptor eliminador del grupo A, precedido por un solo dominio LDLA en la región tallo. TMPRSS11D, también conocida como "HAT" para la proteasa similar a tripsina de las vías respiratorias humanas que pertenece a la subfamilia HAT/DESC, tiene un único dominio SEA. Véase la figura 1 de Bugge et al. (2009), *supra*.

25 Las serina proteasas transmembrana de tipo II se producen inicialmente como proenzimas inactivas que requieren activación por escisión después de un resto de aminoácido básico en un motivo de activación consenso que precede al dominio proteasa. Algunas de las proteasas activadas permanecen unidas a la membrana como resultado de un enlace disulfuro entre el prodominio y el dominio proteasa. Los dominios extracelulares se consideran cruciales para 30 la localización celular, activación, inhibición y/o especificidad de sustrato de estas proteasas (Bugge et al. (2009), *supra*; Szabo et al., Int. J. Biochem. Cell Biol. 40: 1297-1316 (2008)).

35 Se ha documentado diversa información bioquímica y fisiopatológica para miembros de las serina proteasas transmembrana de tipo II. Se han mostrado que TMPRSS2, TMPRSS4 y TMPRSS11D escinden la hemaglutinina de la gripe A (HA) *in vitro*, que es la primera etapa esencial en el ciclo de vida vírico. Animales roedores genéticamente modificados que tienen un gen *Tmprss* humanizado divulgado en el presente documento proporcionan sistemas *in vivo* que permiten una comprensión profunda de las funciones biológicas de las moléculas TMPRSS, así como la exploración de compuestos terapéuticos que se dirigen específicamente a moléculas de TMPRSS humanas.

40 En la presente solicitud se proporcionan secuencias Tmprss ilustrativas, incluyendo secuencias de los ácidos nucleicos y proteínas de Tmprss de ratón, humanas y humanizadas y se resumen en la siguiente tabla. También se incluyen en la tabla secuencias de cebadores y sondas utilizados en los ensayos descritos en la sección de ejemplos, y secuencias de unión de inserción de alelos Tmprss humanizados ilustrativos.

Descripción resumida de secuencias

| SEQ ID NO | Descripción | Características |
|-----------|---|---|
| 1 | <i>Tmprss2 de Mus musculus</i> , ARNm, NM_015775.2 | Longitud: 3175 pb CDS: 231-1703 Exones: 1-177; 178-245 (segundo exón y primer exón codificante); 246-465; 466-552; 553-672; 673-799; 800-910; 911-954; 955-1123; 1124-1299; 1300-1395; 1396-1538; 1539-1691; 1692-3161. |
| 2 | <i>Tmprss2 de Mus musculus</i> , proteína | Longitud: 490 aa |
| 3 | <i>TMPRSS2 de Homo sapiens</i> , variante 2 de transcripto, ARNm, NM_005656.3 | Longitud: 3212 pb CDS: 135-1613 Exones: 1-78; 79-149 (segundo exón y primer exón codificante); 150-372; 373-459; 460-579; 580-706; 707-817; 818-861; 862-1033; 1034-1209; 1210-1305; 1306-1448; 1449-1601; 1602-3204. |
| 4 | <i>TMPRSS2 de Homo sapiens</i> , variante 2 de transcripto, proteína | Longitud: 492 aa Ectodominio: comienza en W106. |

(continuación)

| SEQ ID NO | Descripción | Características |
|-----------|---|---|
| 5 | Fragmento genómico de humanización de <i>Tmprss2</i> | Longitud: 27947 pb 1-84: secuencia de ratón 85-25175: secuencia humana (25091 pb totales) 25176-27866: <i>Xhol-LoxP-Casete-loxP-/CeU1-Nhel</i> (269 1 pb totales) 27867-27947: secuencia de ratón |
| 6 | Fragmento genómico de humanización de <i>Tmprss2</i> con casete eliminado | Longitud: 25.333 pb 1-84: secuencia de ratón 85-25175: secuencia humana (25091 pb totales) 25176-25252: <i>Xhol-LoxP-/CeU1-Nhel</i> (77 pb) 25253-25333: secuencia de ratón |
| 7 | Proteína <i>Tmprss2</i> humanizada | Longitud: 491 aa |
| 8 | <i>Tmprss4</i> de <i>Mus musculus</i> , ARNm, NM_145403.2 | Longitud: 2267 pb CDS: 289-1596 Exones: 1-291 (primer exón y primer exón codificante); 292-325; 326-439; 440-592; 593-722; 723-824; 825-865; 866-1025; 1026-1192; 1193-1291; 1292-1434; 1435-1584; 1585-2267. |
| 9 | <i>Tmprss4</i> de <i>Mus musculus</i> , proteína | Longitud: 435 aa |
| 10 | <i>TMPRSS4</i> de <i>Homo sapiens</i> , variante 4 de transcripto, ARNm, NM_001173551.1 | Longitud: 3543 pb CDS: 292-1599 Exones: 1-294 (primer exón y primer exón codificante); 295-328; 329-442; 443-595; 596-725; 726-827; 828-868; 869-1028; 1029-1195; 1196-1294; 1295-1437; 1438-1587; 1588-3529. |
| 11 | <i>TMPRSS4</i> de <i>Homo sapiens</i> , variante 4 de transcripto, proteína | Longitud: 437 aa Ectodominio: comienza en K54. |
| 12 | Fragmento genómico de humanización de <i>Tmprss4</i> que contiene casete | Longitud: 20078 pb 1-18: secuencia de ratón 19-5014: <i>Sall/Xhol-LoxP-hUbi-EM7-Neo-Pm-Cre-loxP-/Ceul-Nhel</i> (4996 pb totales) 5015-19977: secuencia humana (14963 pb totales) 19978-20078: secuencia de ratón |
| 13 | Fragmento genómico de humanización de <i>Tmprss4</i> con casete eliminado | Longitud: 15159 pb 1-18: secuencia de ratón 19-95: <i>Sall/Xhol-LoxP-/Ceul-Nhel</i> (77 pb totales) 96-15058: secuencia humana (14963 pb totales) 15059-15159: secuencia de ratón |
| 14 | Proteína <i>Tmprss4</i> humanizada | Longitud: 435 aa |
| 15 | <i>Tmprss11d</i> de <i>Mus musculus</i> , ARNm, NM_145561.2 | Longitud: 2046 pb CDS: 36-1289 Exones: 1-43 (primer exón y primer exón codificante), 44-165, 166-284; 285-352; 353-507; 508-546; 547-724; 725-984; 985-1127; 1128-2046. |
| 16 | <i>Tmprss11d</i> de <i>Mus musculus</i> , proteína | Longitud: 417 aa |
| 17 | <i>TMPRSS11D</i> de <i>Homo sapiens</i> , ARNm, NM_004262.2 | Longitud: 2800 pb CDS: 66-1322 Exones: 1-73 (primer exón y primer exón codificante); 74-195; 196-314; 315-382; 383-540; 541-579; 580-757; 758-1017; 1018-1160; 1161-2783. |
| 18 | <i>TMPRSS11D</i> de <i>Homo sapiens</i> , proteína | Longitud: 418 aa El ectodominio comienza en A42. |
| 19 | Fragmento genómico de humanización de <i>Tmprss11d</i> que contiene casete | Longitud: 38.992 1-19: secuencia de ratón 20-33.946: secuencia humana (33927 pb totales) 33947-38942: <i>Xhol-LoxP-hUbi-EM7-Neo-Pm-Cre-loxP-/Ceul-Nhel</i> (4.996 pb totales) 38.943-38.992: secuencia de ratón |

(continuación)

| SEQ ID NO | Descripción | Características |
|-----------|---|---|
| 20 | Fragmento genómico de humanización de <i>Tmprss11d</i> con casete eliminado | Longitud: 34073 pb 1-19: secuencia de ratón 20-33.946: secuencia humana (33927 pb totales) 33947-34023: Xhol-LoxP-/CeuI-Nhel (77 pb) 34.024-34.073: secuencia de ratón |
| 21 | Proteína <i>Tmprss11d</i> humanizada | 418 aa |
| 22 | secuencia de unión 5' de ratón/5' humana para humanización de <i>Tmprss2</i> | 5' de ratón //5' humana |
| 23 | secuencia de unión 3' humana/casete para humanización de <i>Tmprss2</i> | Humana//Xhol//Casete loxP |
| 24 | secuencia de unión casete/3' de ratón para humanización de <i>Tmprss2</i> | Casete (loxP)/ICEUI//Nhel//de ratón |
| 25 | unión 3' humana/loxP/3' de ratón para humanización de <i>Tmprss2</i> | 3' humana//Xhol//(loxP)/ICEUI//Nhel//3' de ratón |
| 26-37 | Cebadores y sondas para ensayos de pérdida de alelos y ganancia de alelos para humanización de <i>Tmprss2</i> | Tabla 1 |
| 38 | secuencia de unión 5' de ratón/casete para humanización de <i>Tmprss4</i> | 5' de ratón//Sall-Xhol//casete(loxP) |
| 39 | Secuencia de unión casete/5' humana para humanización de <i>Tmprss4</i> | Casete (loxP)/ICEUI//Nhel//5' humana |
| 40 | secuencia de unión 3' humana/3' de ratón para humanización de <i>Tmprss4</i> | 3' humana/3' de ratón |
| 41 | unión 5' de ratón/loxP/5' humana para humanización de <i>Tmprss4</i> | 5' de ratón//Sall/Xhol//(loxP)/ICEUI//Nhel//5' humana |
| 42-56 | Cebadores y sondas para ensayos de pérdida de alelos y ganancia de alelos para humanización de <i>Tmprss4</i> | Tabla 2 |
| 57 | secuencia de unión 5' de ratón/5' humana para humanización de <i>Tmprss11d</i> | 5' de ratón //5' humana |
| 58 | secuencia de unión 3' humana/casete de para humanización de <i>Tmprss11d</i> | 3' humana//Xhol//Casete (loxP) |
| 59 | secuencia de unión casete/3' de ratón para humanización de <i>Tmprss11d</i> | casete (loxP)/ICEUI//Nhel//3' de ratón |
| 60 | unión 3' humana/loxP/3' de ratón para humanización de <i>Tmprss11d</i> | 3' humana//Xhol//(loxP)/ICEUI//Nhel//3' de ratón |
| 61-72 | Cebadores y sondas para ensayos de pérdida de alelos y ganancia de alelos para humanización de <i>Tmprss11d</i> | Tabla 3 |

Animales Roedores con *Tmprss* humanizado

En un aspecto, la presente invención proporciona animales roedores que contienen en la línea germinal un gen *Tmprss* humanizado que codifica una proteína *Tmprss* humanizada. En particular, la presente invención proporciona un roedor cuyo genoma comprende un gen *Tmprss* humanizado, en donde el gen *Tmprss* humanizado

- 5 - comprende una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS* humano afín,
- 10 - se encuentra en un locus *Tmprss* de roedor endógeno, y es el resultado de una sustitución de una secuencia

genómica del gen *Tmprss* de roedor endógeno con dicha secuencia de nucleótidos del gen *TMPRSS* humano afín y está bajo el control del promotor del gen *Tmprss* de roedor endógeno, y

- 5 - codifica una proteína *Tmprss* humanizada que comprende: (i) un ectodominio que tiene al menos un 85 % de identidad de secuencia con el ectodominio de la proteína *TMPRSS* humana codificada por el gen *TMPRSS* humano afín y (ii) una porción citoplasmática y transmembrana que tiene al menos un 85 % de identidad de secuencia con la porción citoplasmática y transmembrana de la proteína *Tmprss* de roedor endógena codificada por el gen *Tmprss* de roedor endógeno.
- 10 El término "humanizado/a", cuando se utiliza en el contexto de ácidos nucleicos o proteínas, se refiere a ácidos nucleicos o proteínas cuyas estructuras (es decir, secuencias de nucleótidos o aminoácidos) incluyen porciones que corresponden sustancial o idénticamente a las estructuras de un gen o proteína particular que se encuentra en la naturaleza en un animal roedor, y también incluyen porciones que difieren de las encontradas en el gen o proteína de roedor relevante y en cambio corresponden de manera más cercana o idéntica a estructuras encontradas en un gen o proteína humanos correspondiente. Un roedor que contiene un gen humanizado o que expresa una proteína humanizada es un roedor "humanizado".

- En algunas realizaciones, un roedor de la presente invención se selecciona de un ratón, una rata, y un hámster. En algunas realizaciones, un roedor de la presente invención se selecciona de la superfamilia Muroidea. En algunas realizaciones, un roedor genéticamente modificado de la presente invención es de una familia seleccionada de Calomyscidae (por ejemplo, hámsteres similares a ratones), Cricetidae (por ejemplo, hámster, ratas y ratones del nuevo mundo, campañoles), Muridae (ratones y ratas auténticos, jirbos, ratones espinosos, ratas con cresta), Nesomyidae (ratones trepadores, ratones de abazón de las rocas, ratas coliblancas, ratas y ratones de Madagascar), Platacanthomyidae (por ejemplo, lirón espinoso), y Spalacidae (por ejemplo, ratas topo, ratas del bambú, y zokores).
- 20 25 En algunas realizaciones determinadas, un roedor genéticamente modificado de la presente invención se selecciona de un auténtico ratón o rata (familia Muridae), un jirbo, un ratón espinoso y una rata con cresta. En algunas realizaciones determinadas, un ratón genéticamente modificado de la presente invención es de un miembro de la familia Muridae.
 - 30 35 En algunas realizaciones, el roedor divulgado en el presente documento contiene un gen *Tmprss* humanizado en el genoma que incluye una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS* humano, en donde la secuencia de nucleótidos del gen *Tmprss* de roedor endógeno y la secuencia de nucleótidos del gen *TMPRSS* humano están operativamente unidos entre sí de modo que el gen *Tmprss* humanizado codifica una proteína *Tmprss* y está bajo el control de un o unos elementos 5' reguladores, tal como el o los promotores y/o potenciadores, del gen *Tmprss* de roedor endógeno.

- La presente invención se refiere particularmente a la humanización similar; en otras palabras, una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno se une operativamente a una secuencia de nucleótidos de un gen *TMPRSS* humano afín para formar un gen humanizado. Por ejemplo, en algunas realizaciones, una secuencia de nucleótidos de un gen *Tmprss2* de roedor endógeno se une operativamente a una secuencia de nucleótidos de un gen *TMPRSS2* humano para formar un gen *Tmprss2* humanizado. En otras realizaciones, una secuencia de nucleótidos de un gen *Tmprss4* de roedor endógeno se une operativamente a una secuencia de nucleótidos de un gen *TMPRSS4* humano para formar un gen *Tmprss4* humanizado. En otras realizaciones más, una secuencia de nucleótidos de un gen *Tmprss11d* de roedor endógeno se une operativamente a una secuencia de nucleótidos de un gen *TMPRSS11D* humano para formar un gen *Tmprss11d* humanizado.

- En algunas realizaciones, un roedor genéticamente modificado de la presente invención contiene un gen *Tmprss* humanizado en su genoma, en donde el gen *Tmprss* humanizado codifica una proteína *Tmprss* humanizada que contiene un ectodominio que es sustancialmente idéntico al ectodominio de una proteína *TMPRSS* humana. El término "ectodominio" se refiere a la porción de una proteína transmembrana que se extiende fuera de la membrana celular, es decir, la porción extracelular de una proteína transmembrana. El ectodominio de una molécula de *TMPRSS* incluye un dominio proteasa y una región de tallo que une el dominio transmembrana con el dominio proteasa. Por un ectodominio o polipéptido que es "sustancialmente idéntico al ectodominio de una proteína *TMPRSS* humana", se entiende en algunas realizaciones, un polipéptido que es al menos un 85 %, 90 %, 95 %, 99 % o 100 % idéntico en secuencia al ectodominio de una proteína *TMPRSS* humana; en algunas realizaciones, un polipéptido que difiere del ectodominio de una proteína *TMPRSS* humana en no más de 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 aminoácidos; en algunas realizaciones, un polipéptido que difiere del ectodominio de una proteína *TMPRSS* humana solamente en el extremo N o C del ectodominio, por ejemplo, al carecer de aminoácidos o al tener aminoácidos adicionales en el extremo N o C del ectodominio; y en algunas realizaciones, un polipéptido que es sustancialmente el ectodominio de una proteína *TMPRSS* humana. Por "sustancialmente el ectodominio" de una proteína *TMPRSS* humana, se entiende un polipéptido que es idéntico al ectodominio, o difiere del ectodominio al carecer de 1-5 (es decir, 1, 2, 3, 4 o 5) aminoácidos o al tener 1-5 aminoácidos adicionales en el extremo N o C.

- En algunas realizaciones, el gen *Tmprss* humanizado codifica una proteína *Tmprss* humanizada que además contiene una porción citoplasmática y transmembrana que es sustancialmente idéntica a la porción citoplasmática y transmembrana de una proteína *Tmprss* de roedor endógena. Por una porción o polipéptido citoplasmáticos y

- transmembrana que es "sustancialmente idéntico a la porción citoplasmática y transmembrana de una proteína Tmprss de roedor endógena", se entiende en algunas realizaciones, un polipéptido que es al menos un 85 %, 90 %, 95 %, 95 %, 99 % o 100 % idéntico en secuencia a la porción citoplasmática y transmembrana de una proteína Tmprss de roedor endógena; en algunas realizaciones, un polipéptido que difiere de la porción citoplasmática y transmembrana de una proteína Tmprss de roedor endógena en no más de 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 aminoácidos; en algunas realizaciones, un polipéptido que difiere de la porción citoplasmática y transmembrana de una proteína Tmprss de roedor endógena solamente en el extremo C, por ejemplo, al carecer de aminoácidos o al tener aminoácidos adicionales en el extremo C del dominio transmembrana; y en algunas realizaciones, un polipéptido compuesto por el dominio citoplasmático y sustancialmente el dominio transmembrana de una proteína Tmprss de roedor endógena.
- 5 Por "sustancialmente el dominio transmembrana" de una proteína Tmprss de roedor endógena, se entiende un polipéptido que es idéntico al dominio transmembrana, o se diferencia del dominio transmembrana al carecer de 1-5 aminoácidos o al tener 1-5 aminoácidos adicionales en el extremo C.
- 10 El gen *Tmprss* humanizado en el genoma de un roedor genéticamente modificado incluye una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS* humano afín, en donde la secuencia de nucleótidos del gen *TMPRSS* humano afín codifica un polipéptido sustancialmente idéntico al ectodominio de la proteína *TMPRSS* humana codificada por el gen *TMPRSS* humano. En determinadas realizaciones, la secuencia de nucleótidos de un gen *TMPRSS* humano afín en un gen *Tmprss* humanizado codifica el ectodominio de la proteína *TMPRSS* humana codificada por el gen *TMPRSS* humano.
- 15 20 En algunas realizaciones, el gen *Tmprss* humanizado en el genoma de un roedor genéticamente modificado incluye una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS* humano afín, en donde la secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno codifica un polipéptido sustancialmente idéntico a la porción citoplasmática y transmembrana de la proteína Tmprss de roedor endógena codificada por el gen *Tmprss* de roedor endógeno. En realizaciones específicas, la secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno presente en un gen *Tmprss* humanizado codifica los dominios citoplasmático y transmembrana de la proteína Tmprss de roedor endógena codificada por el gen *Tmprss* de roedor endógeno.
- 25 30 Un gen *Tmprss* humanizado en la invención da como resultado una sustitución de una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno en un locus *Tmprss* de roedor endógeno con una secuencia de nucleótidos de un gen *TMPRSS* humano afín.
- 35 En algunas realizaciones, una secuencia genómica contigua de un gen *Tmprss* de roedor en un locus *Tmprss* de roedor endógeno se ha sustituido por una secuencia genómica contigua de un gen *TMPRSS* humano afín para formar un gen *Tmprss* humanizado.
- 40 En realizaciones específicas, una secuencia genómica contigua de un gen *TMPRSS* humano insertado en un gen *Tmprss* de roedor endógeno incluye exones, en su totalidad o en parte, de un gen *TMPRSS* humano, que codifican un ectodominio que es sustancialmente idéntico al ectodominio de la proteína *TMPRSS* humana codificada por el gen *TMPRSS* humano.
- 45 En determinadas realizaciones, la secuencia genómica de un gen *Tmprss* de roedor endógeno que permanece en un locus *Tmprss* de roedor endógeno después de la sustitución de humanización y está operativamente unido a la secuencia genómica de *TMPRSS* humano contigua codifica una porción citoplasmática y transmembrana que es sustancialmente idéntica a la porción citoplasmática y transmembrana de la proteína Tmprss de roedor endógena codificada por el gen *Tmprss* de roedor endógeno.
- 50 55 En circunstancias en las que una proteína Tmprss endógena y una proteína *TMPRSS* humana comparten aminoácidos comunes cerca de la unión entre el dominio transmembrana y el ectodominio, puede que no sea necesario insertar una secuencia genómica de *TMPRSS* humano que codifique precisamente el ectodominio de la proteína *TMPRSS* humana. Es posible insertar una secuencia genómica ligeramente más larga o más corta de un gen *TMPRSS* humano, que codifica sustancialmente el ectodominio de la proteína *TMPRSS* humana, en unión operativa a una secuencia genómica de un gen *Tmprss* de roedor endógeno que codifica el dominio citoplasmático y sustancialmente el dominio transmembrana de la proteína Tmprss de roedor endógena, de manera que la proteína Tmprss humanizada codificada por el gen *Tmprss* humanizado resultante incluye un ectodominio que es idéntico al ectodominio de la proteína *TMPRSS* humana y un dominio transmembrana que es idéntico al dominio transmembrana de la proteína Tmprss de roedor endógena.
- 60 65 En algunas realizaciones, la secuencia de nucleótidos de un gen *TMPRSS* humano incluido en un gen *Tmprss* humanizado también incluye la región 3' no traducida ("UTR") del gen *TMPRSS* humano. En determinadas realizaciones, además de la 3' UTR de un gen *TMPRSS* humano, un gen *Tmprss* humanizado también incluye una secuencia genómica humana adicional del locus del gen *TMPRSS* humano, después de la 3' UTR de *TMPRSS*. La secuencia genómica humana adicional puede consistir en al menos 10-200 pb, por ejemplo, 50 pb, 75 pb, 100 pb, 125 pb, 150 pb, 175 pb, 200 pb o más, encontradas en locus del gen *TMPRSS* humano inmediatamente cadena abajo de la 3' UTR del gen *TMPRSS* humano. En otras realizaciones, la secuencia de nucleótidos de un gen *TMPRSS* humano presente en un gen *Tmprss* humanizado no incluye una 3' UTR humana; en cambio, se incluye la 3' UTR de un gen

- 5 *Tmprss* de roedor endógeno y sigue inmediatamente al codón de parada del gen *Tmprss* humanizado. Por ejemplo, un gen *Tmprss* humanizado puede incluir una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno que contiene secuencias de exones que codifican los dominios citoplasmático y transmembrana de la proteína *Tmprss* de roedor endógena, seguida de una secuencia de nucleótidos de un gen *TMPRSS* humano que contiene secuencias de exones que codifican el ectodominio a través del codón de parada de la proteína *TMPRSS* humana, siguiendo la 3' UTR del gen *Tmprss* de roedor endógeno inmediatamente después del codón de parada.
- 10 En algunas realizaciones, un gen *Tmprss* humanizado da como resultado una expresión de la proteína *Tmprss* humanizada codificada en un roedor. En algunas realizaciones, una proteína *Tmprss* humanizada se expresa en un patrón comparable con, o sustancialmente el mismo que, una proteína *Tmprss* de roedor homóloga en un roedor de control (por ejemplo, un roedor sin el gen *Tmprss* humanizado). En algunas realizaciones, una proteína *Tmprss* humanizada se expresa a un nivel comparable con o sustancialmente igual que, una proteína *Tmprss* de roedor homóloga en un roedor de control (por ejemplo, un roedor sin el gen *Tmprss* humanizado). En determinadas realizaciones, se expresa y detecta una proteína *Tmprss* humanizada en la superficie celular. En determinadas 15 realizaciones, una proteína *Tmprss* humanizada o una forma soluble (por ejemplo, una forma secretada de ectodominio) se expresa y detecta en el suero de un roedor, por ejemplo, a un nivel comparable con o sustancialmente igual que, una proteína *Tmprss* de roedor homóloga o una forma soluble de la misma en un roedor de control. En el contexto de comparar un gen o proteína humanizados en un roedor humanizado con un gen o proteína de roedor endógeno en un roedor de control, el término "comparable" significa que las moléculas o niveles que se comparan 20 pueden no ser idénticos entre sí, pero son suficientemente similares para permitir la comparación entre ellos, de modo que se puedan extraer conclusiones razonablemente en función de las diferencias o similitudes observadas; y la expresión "sustancialmente el mismo" al referirse a los niveles de expresión significa que los niveles que se comparan no son diferentes entre sí en más de un 20 %, 19 %, 18 %, 17 %, 16 %, 15 %, 14 %, 13 %, 12 %, 11 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 % o 1 %. 25
- 30 En algunas realizaciones, la presente invención proporciona además una célula o tejido aislados de un animal roedor como se describe en el presente documento. En algunas realizaciones, una célula se selecciona de una célula dendrítica, linfocito (p. ej., un linfocito B o T), macrófago y monocito. En algunas realizaciones, un tejido se selecciona de adiposo, vejiga, cerebro, mama, médula ósea, ojo, corazón, intestino, riñón, hígado, pulmón, ganglio linfático, músculo, páncreas, plasma, suero, piel, bazo, estómago, timo, testículo, óvulo y una combinación de los mismos.
- 35 En algunas realizaciones, la presente invención proporciona una célula madre embrionaria de roedor cuyo genoma contiene un gen *Tmprss* humanizado como se describe en el presente documento. En algunas realizaciones, una célula madre embrionaria de roedor es una célula madre embrionaria de ratón. En otras realizaciones, una célula madre embrionaria de roedor que contiene un gen *Tmprss* humanizado en su genoma se puede utilizar para producir un animal roedor humanizado, como se describe adicionalmente en el presente documento a continuación.
- 40 En algunas realizaciones, un roedor proporcionado en el presente documento es heterocigoto para un gen *Tmprss* humanizado en su genoma. En otras realizaciones, un roedor proporcionado en el presente documento es homocigoto para un gen *Tmprss* humanizado en su genoma.
- 45 En determinadas realizaciones, un roedor incluye múltiples, es decir, dos o más, genes *Tmprss* humanizados en su genoma. En otras palabras, se han humanizado dos o más locus *Tmprss* endógenos diferentes en un roedor utilizando secuencias de nucleótidos de genes *TMPRSS* humanos afines. Por ejemplo, se ha humanizado un roedor en dos o más de los loci de genes seleccionados de: *Tmprss2*, *Tmprss4*, y *Tmprss11d*.
- 50 Se describen con más detalle a continuación roedores (tal como ratones) con *Tmprss2* humanizado, roedores (tal como ratones) con *Tmprss4* humanizado y roedores (tal como ratones) con *Tmprss11d* humanizado ilustrativos.
- 55 *Roedores con Tmprss2 humanizado*
- 60 En algunas realizaciones, la presente invención proporciona un roedor cuyo genoma contiene un gen *Tmprss2* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss2* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS2* humano, y que está bajo el control de un o unos elementos 5' reguladores, tal como el o los promotores y/o potenciadores, del gen *Tmprss2* de roedor endógeno. Ejemplos de roedores incluyen ratones y ratas.
- 65 En algunas realizaciones, un gen *Tmprss2* humanizado codifica una proteína *Tmprss2* humanizada que contiene un ectodominio que es sustancialmente idéntico al ectodominio de una proteína *TMPRSS2* humana.
- 70 En realizaciones específicas, la proteína *TMPRSS2* humana tiene una secuencia de aminoácidos que tiene al menos un 85 %, 90 %, 95 %, 98 %, 99 % o 100 % de identidad con la secuencia de aminoácidos expuesta en la SEQ ID NO: 4.
- 75 En algunas realizaciones, una proteína *Tmprss2* humanizada contiene los 387 aminoácidos carboxiterminales de una

- proteína TMPRSS2 humana, por ejemplo, los aminoácidos 106 a 492 de una proteína TMPRSS2 humana. En algunas realizaciones, una proteína *Tmprss2* humanizada contiene un ectodominio que es sustancialmente idéntico a la secuencia de aminoácidos compuesta de W106 a G492 de la SEQ ID NO: 4. En realizaciones específicas, una proteína *Tmprss2* humanizada contiene un ectodominio que tiene al menos un 85 %, 90 %, 95 %, 98 %, 99 % o 100 % de identidad con la secuencia de aminoácidos compuesta de W106 a G492 de la SEQ ID NO: 4; un ectodominio que difiere de la secuencia de aminoácidos compuesta de W106 a G492 de la SEQ ID NO: 4 en no más de 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 aminoácidos; o un ectodominio que difiere de la secuencia de aminoácidos compuesta de W106 a G492 de la SEQ ID NO: 4 solamente en el extremo N o C del ectodominio, por ejemplo, que carece de 1-5 aminoácidos o que tiene 1-5 aminoácidos adicionales en el extremo N o C.
- Una proteína *Tmprss2* humanizada contiene además una porción citoplasmática y transmembrana que es sustancialmente idéntica a la porción citoplasmática y transmembrana de una proteína *Tmprss2* de roedor endógena. En algunas realizaciones, una proteína *Tmprss2* humanizada incluye además el dominio transmembrana y el dominio citoplasmático de una proteína *Tmprss2* de roedor endógena.
- Una proteína *Tmprss2* humanizada contiene el dominio transmembrana y el dominio citoplasmático de una proteína *Tmprss2* de roedor endógena y el ectodominio de una proteína TMPRSS2 humana. En realizaciones particulares, un gen *Tmprss2* humanizado codifica una proteína *Tmprss2* humanizada que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 7.
- Un gen *Tmprss2* humanizado es el resultado de una sustitución de una secuencia de nucleótidos de un gen *Tmprss2* de roedor endógeno en un locus *Tmprss2* de roedor endógeno con una secuencia de nucleótidos de un gen *TMPRSS2* humano.
- En algunas realizaciones, una secuencia genómica contigua de un gen *Tmprss2* de roedor endógeno en un locus *Tmprss2* de roedor endógeno se ha sustituido por una secuencia genómica contigua de un gen *TMPRSS2* humano para formar un gen *Tmprss2* humanizado.
- En realizaciones específicas, la secuencia genómica contigua de un gen *TMPRSS2* humano insertado en un gen *Tmprss2* de roedor endógeno incluye secuencias de exones, es decir, exones en su totalidad o en parte, de un gen *TMPRSS2* humano, que codifican un ectodominio que es sustancialmente idéntico al ectodominio de la proteína TMPRSS2 humana codificada por el gen *TMPRSS2* humano. En circunstancias en las que una proteína *Tmprss2* endógena y una proteína *TMPRSS2* humana comparten aminoácidos comunes cerca de la unión del dominio transmembrana y el ectodominio, puede que no sea necesario insertar una secuencia genómica de *TMPRSS2* humano que codifique precisamente el ectodominio de la proteína TMPRSS2 humana, y es posible utilizar una secuencia genómica de *TMPRSS2* humano ligeramente más larga o más corta que codifique sustancialmente el ectodominio de una proteína *TMPRSS2* humana para producir una proteína *Tmprss2* humanizada que tiene un ectodominio que es idéntico al ectodominio de la proteína TMPRSS2 humana.
- En realizaciones específicas, una secuencia genómica contigua de un gen *TMPRSS2* humano que se está insertando en un gen *Tmprss2* de roedor endógeno contiene al menos el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *TMPRSS2* humano.
- En determinadas realizaciones, una secuencia genómica contigua de un gen *TMPRSS2* humano que se está insertando en un gen *Tmprss2* de roedor endógeno contiene el intrón 3 y el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *TMPRSS2* humano. En realizaciones particulares, una secuencia genómica contigua de un gen *TMPRSS2* humano que se está insertando en un gen *Tmprss2* de roedor endógeno contiene una porción 3' del exón 3 codificante, el intrón 3 y el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *TMPRSS2* humano. En realizaciones específicas, la porción 3' del exón 3 codificante de un gen *TMPRSS2* humano incluido en la humanización tiene una longitud de aproximadamente 5-10 pares de bases, es decir, aproximadamente 5, 6, 7, 8, 9 o 10 pares de bases del extremo 3' del exón 3 codificante.
- En algunas realizaciones, una secuencia genómica contigua de un gen *TMPRSS2* humano que se está insertando en un gen *Tmprss2* de roedor endógeno también contiene la 3' UTR del gen *TMPRSS2* humano. En realizaciones específicas, todo el exón 13 codificante de un gen *TMPRSS2* humano está incluido en la secuencia genómica de *TMPRSS2* contigua para la humanización, que incluye la 3' UTR del gen *TMPRSS2* humano. En realizaciones particulares, una secuencia genómica contigua de un gen *TMPRSS2* humano incluye una secuencia genómica humana adicional cadena abajo de la 3' UTR del gen *TMPRSS2* humano. La secuencia genómica humana adicional puede ser una secuencia de al menos 10-200 pb, o al menos 10, 20, 30, 40, 50, 75, 100, 125, 150, 175 o 200 pb, que se encuentra inmediatamente cadena abajo del 3' UTR del gen *TMPRSS2* humano en un locus *TMPRSS2* humano.
- En algunas realizaciones, la secuencia de nucleótidos de un gen *Tmprss2* de roedor endógeno que permanece en un locus *Tmprss2* humanizado codifica un polipéptido que es sustancialmente idéntico a la porción citoplasmática y transmembrana de la proteína *Tmprss2* de roedor endógena. En circunstancias en las que una proteína *Tmprss2* endógena y una proteína *TMPRSS2* humana comparten aminoácidos comunes cerca de la unión del dominio transmembrana y el ectodominio, puede que no sea necesario mantener la secuencia genómica de *Tmprss2* de roedor

endógeno que codifique precisamente el dominio transmembrana de la proteína *Tmprss2* de roedor endógena, y es posible mantener una secuencia genómica de *Tmprss2* de roedor ligeramente más larga o más corta que codifique sustancialmente el dominio transmembrana de la proteína *Tmprss2* de roedor endógena en la sustitución de humanización para codificar una proteína *Tmprss2* humanizada que tiene un dominio transmembrana que es idéntico

5 al dominio transmembrana de la proteína *Tmprss2* de roedor endógena. En algunas realizaciones, la secuencia de nucleótidos de un gen *Tmprss2* de roedor endógeno que permanece en un locus *Tmprss2* humanizado incluye los exones 1-2 y una porción 5' del exón 3 codificante de un gen *Tmprss2* de roedor endógeno, en donde la porción 5' del exón 3 codificante es una porción sustancial del exón 3 codificante, por ejemplo, todo el exón 3 codificante excepto 5-10 pares de bases en el extremo 3' del exón 3 codificante.

10 En realizaciones específicas, un gen *Tmprss2* humanizado contiene los exones 1-2 codificantes y una porción 5' del exón 3 codificante de un gen *Tmprss2* de roedor endógeno y una porción 3' del exón 3 codificante y del exón 4 codificante hasta el exón 13 codificante de un gen *TMPRSS2* humano, en donde el gen *Tmprss2* humanizado codifica una proteína *Tmprss2* humanizada que contiene una porción citoplasmática y transmembrana que es sustancialmente 15 idéntica a la porción citoplasmática y transmembrana de la proteína *Tmprss2* de roedor, y un ectodominio que es sustancialmente idéntico al ectodominio de la proteína *TMPRSS2* humana. En determinadas realizaciones, el gen *Tmprss2* humanizado codifica una proteína *Tmprss2* humanizada que contiene el dominio citoplasmático y el dominio transmembrana de la proteína *Tmprss2* de roedor codificada por un gen *Tmprss2* de roedor endógeno y el ectodominio 20 de la proteína *TMPRSS2* humana codificada por un gen *TMPRSS2* humano. En realizaciones particulares, un gen *Tmprss2* humanizado codifica una proteína *Tmprss2* humanizada que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 7.

En algunas realizaciones, los exones e intrones de un gen *TMPRSS2* humano y un gen *Tmprss2* de roedor utilizados en la humanización son los que se encuentran en las SEQ ID NO: 1, 3 y 5-6.

25 En algunas realizaciones, un gen *Tmprss2* humanizado da como resultado una expresión de la proteína *Tmprss2* humanizada codificada en un roedor. En algunas realizaciones, una proteína *Tmprss2* humanizada se expresa en un patrón comparable con, o sustancialmente igual que, una proteína *Tmprss2* de roedor homóloga en un roedor de control (por ejemplo, un roedor sin el gen *Tmprss2* humanizado). En algunas realizaciones, una proteína *Tmprss2* 30 humanizada se expresa a un nivel comparable con, o sustancialmente igual que, una proteína *Tmprss2* de roedor homóloga en un roedor de control (por ejemplo, un roedor sin el gen *Tmprss2* humanizado). En determinadas realizaciones, una proteína *Tmprss2* humanizada se expresa y detecta en la superficie celular. En determinadas realizaciones, una proteína *Tmprss2* humanizada o una forma soluble (por ejemplo, una forma secretada de ectodominio) se expresa y detecta en el suero de un roedor, por ejemplo, a un nivel comparable con o sustancialmente 35 igual que, una proteína *Tmprss2* de roedor homóloga o una forma soluble de la misma en un roedor de control.

Roedores con Tmprss4 humanizado

40 En algunas realizaciones, la presente invención proporciona un roedor cuyo genoma contiene un gen *Tmprss4* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss4* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS4* humano, y que está bajo el control de un o unos elementos 5' reguladores, tal como el o los promotores y/o un o unos potenciadores, del gen *Tmprss4* de roedor endógeno. Ejemplos de roedores incluyen ratones y ratas.

45 Un gen *Tmprss4* humanizado codifica una proteína *Tmprss4* humanizada que contiene un ectodominio que es sustancialmente idéntico al ectodominio de una proteína *TMPRSS4* humana. En realizaciones específicas, la proteína *TMPRSS4* humana tiene una secuencia de aminoácidos que tiene al menos un 85 %, 90 %, 95 %, 98 %, 99 % o 100 % de identidad con la secuencia de aminoácidos expuesta en la SEQ ID NO: 11.

50 En algunas realizaciones, una proteína *Tmprss4* humanizada contiene los 384 aminoácidos carboxiterminales de una proteína *TMPRSS4* humana, por ejemplo, los aminoácidos 54 a 437 de una proteína *TMPRSS4* humana. En algunas realizaciones, una proteína *Tmprss4* humanizada contiene un ectodominio que es sustancialmente idéntico a la secuencia de aminoácidos compuesta de K54 a L437 de la SEQ ID NO: 11. En realizaciones específicas, una proteína *Tmprss4* humanizada contiene un ectodominio que tiene al menos un 85 %, 90 %, 95 %, 98 %, 99 % o 100 % de 55 identidad con la secuencia de aminoácidos compuesta de K54 a L437 de la SEQ ID NO: 11; un ectodominio que difiere de la secuencia de aminoácidos compuesta de K54 a L437 de la SEQ ID NO: 11 en no más de 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 aminoácidos; o un ectodominio que difiere de la secuencia de aminoácidos compuesta de K54 a L437 de la SEQ ID NO: 11 solamente en el extremo N o C del ectodominio, por ejemplo, que carece de 1-5 aminoácidos o que tiene 1-5 aminoácidos adicionales en el extremo N o C.

60 Una proteína *Tmprss4* humanizada contiene además una porción citoplasmática y transmembrana que es sustancialmente idéntica a la porción citoplasmática y transmembrana de una proteína *Tmprss4* de roedor endógena. En algunas realizaciones, una proteína *Tmprss4* humanizada incluye además el dominio transmembrana y el dominio citoplasmático de una proteína *Tmprss4* de roedor endógena.

65 Una proteína *Tmprss4* humanizada contiene el dominio transmembrana y el dominio citoplasmático de una proteína

Tmprss4 de roedor endógena y el ectodominio de una proteína TMPRSS4 humana. En realizaciones particulares, un gen *Tmprss4* humanizado codifica una proteína Tmprss4 humanizada que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 14.

- 5 Un gen *Tmprss4* humanizado es el resultado de una sustitución de una secuencia de nucleótidos de un gen *Tmprss4* de roedor endógeno en un locus *Tmprss4* de roedor endógeno con una secuencia de nucleótidos de un gen *TMPRSS4* humano.
- 10 En algunas realizaciones, una secuencia genómica contigua de un gen *Tmprss4* de roedor endógeno en un locus *Tmprss4* de roedor endógeno se ha sustituido por una secuencia genómica contigua de un gen *TMPRSS4* humano para formar un gen *Tmprss4* humanizado.
- 15 En realizaciones específicas, la secuencia genómica contigua de un gen *TMPRSS4* humano insertado en un gen *Tmprss4* de roedor endógeno incluye secuencias de exones, es decir, exones en su totalidad o en parte, de un gen *TMPRSS4* humano que codifica un ectodominio que es sustancialmente idéntico al ectodominio de la proteína TMPRSS4 humana codificada por el gen *TMPRSS4* humano. En circunstancias en las que una proteína Tmprss4 endógena y una proteína TMPRSS4 humana comparten aminoácidos comunes cerca de la unión del dominio transmembrana y el ectodominio, puede que no sea necesario insertar una secuencia genómica de *TMPRSS4* humano que codifique precisamente el ectodominio de la proteína TMPRSS4 humana, y es posible utilizar una secuencia genómica de *TMPRSS4* humano ligeramente más larga o más corta que codifique sustancialmente el ectodominio de una proteína TMPRSS4 humana para producir una proteína Tmprss4 humanizada que tiene un ectodominio que es idéntico al ectodominio de la proteína TMPRSS4 humana.
- 20 En realizaciones específicas, una secuencia genómica contigua de un gen *TMPRSS4* humano que se está insertando en un gen *Tmprss4* de roedor endógeno contiene al menos el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *TMPRSS4* humano.
- 25 En determinadas realizaciones, una secuencia genómica contigua de un gen *TMPRSS4* humano que se está insertando en un gen *Tmprss4* de roedor endógeno incluye una porción 3' del intrón 3 y el exón 4 codificante hasta el codón de parada en el exón 13 codificante de un gen *TMPRSS4* humano. En realizaciones específicas, la porción 3' del intrón 3 de un gen *TMPRSS4* humano incluido en la humanización tiene una longitud de aproximadamente 140-160 pares de bases, es decir, aproximadamente 140, 145, 150, 155, 160 pares de bases del extremo 3' del intrón 3.
- 30 En algunas realizaciones, una secuencia genómica contigua de un gen *TMPRSS4* humano que se está insertando en un gen *Tmprss4* de roedor endógeno contiene la 3' UTR del gen *TMPRSS4* humano. En realizaciones específicas, una secuencia genómica contigua de un gen *TMPRSS4* humano que se está insertando en un gen *Tmprss4* de roedor endógeno no contiene la 3' UTR del gen *TMPRSS4* humano, y la 3' UTR del gen *Tmprss4* de roedor endógeno sigue inmediatamente después del codón de parada en el gen *Tmprss4* humanizado.
- 35 En algunas realizaciones, la secuencia de nucleótidos de un gen *Tmprss4* de roedor endógeno que permanece en un locus *Tmprss4* humanizado codifica un polipéptido que es sustancialmente idéntico a la porción citoplasmática y transmembrana de la proteína Tmprss4 de roedor endógena. En circunstancias en las que una proteína Tmprss4 endógena y una proteína TMPRSS4 humana comparten aminoácidos comunes cerca de la unión del dominio transmembrana y el ectodominio, puede que no sea necesario mantener la secuencia genómica de *Tmprss4* de roedor endógeno que codifique precisamente el dominio transmembrana de la proteína Tmprss4 de roedor endógena, y es posible mantener una secuencia genómica de *Tmprss4* de roedor ligeramente más larga o más corta que codifique sustancialmente el dominio transmembrana de la proteína Tmprss4 de roedor endógena en la sustitución de humanización para codificar una proteína Tmprss4 humanizada que tiene un dominio transmembrana que es idéntico al dominio transmembrana de la proteína Tmprss4 de roedor endógena.
- 40 En algunas realizaciones, un gen *Tmprss4* humanizado contiene los exones 1-3 codificantes de un gen *Tmprss4* de roedor endógeno y el exón 4 codificante hasta el codón de parada del exón 13 codificante de un gen *TMPRSS4* humano. En realizaciones particulares, un gen *Tmprss4* humanizado contiene los exones 1-3 codificantes y una porción 5' del intrón 3 de un gen *Tmprss4* de roedor endógeno y una porción 3' del intrón 3 y el exón 4 codificante hasta el codón de parada del exón 13 codificante de un gen *TMPRSS4* humano. En determinadas realizaciones, el gen *Tmprss4* humanizado codifica una proteína Tmprss4 humanizada que contiene el dominio citoplasmático y el dominio transmembrana de la proteína Tmprss4 de roedor codificada por un gen *Tmprss4* de roedor endógeno y el ectodominio de la proteína TMPRSS4 humana codificada por un gen *TMPRSS4* humano. En realizaciones particulares, un gen *Tmprss4* humanizado codifica una proteína Tmprss4 humanizada que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 14.
- 45 En algunas realizaciones, los exones e intrones de un gen TMPRSS4 humano y un gen *Tmprss4* de roedor utilizados en la humanización son los que se encuentran en las SEQ ID NO: 8, 10 y 12-13.
- 50 En algunas realizaciones, un gen *Tmprss4* humanizado da como resultado una expresión de la proteína Tmprss4 humanizada codificada en un roedor. En algunas realizaciones, una proteína Tmprss4 humanizada se expresa en un

patrón comparable con, o sustancialmente el mismo que, una proteína Tmprss4 de roedor homóloga en un roedor de control (p. ej., un roedor sin el gen *Tmprss4* humanizado que codifica la proteína Tmprss4 humanizada). En algunas realizaciones, una proteína Tmprss4 humanizada se expresa a un nivel comparable con, o sustancialmente igual que, una proteína Tmprss4 de roedor homóloga en un roedor de control (p. ej., un roedor sin el gen *Tmprss4* humanizado que codifica la proteína Tmprss4 humanizada). En determinadas realizaciones, una proteína Tmprss4 humanizada se expresa y detecta en la superficie celular. En determinadas realizaciones, una proteína Tmprss4 humanizada o una forma soluble (por ejemplo, una forma secretada de ectodominio) se expresa y detecta en el suero de un roedor, por ejemplo, a un nivel comparable con o sustancialmente igual que, una proteína Tmprss4 de roedor homóloga o una forma soluble de la misma en un roedor de control.

10 Roedores con *Tmprss11d* humanizado

En algunas realizaciones, la presente invención proporciona un roedor cuyo genoma contiene un gen *Tmprss11d* humanizado que incluye una secuencia de nucleótidos de un gen *Tmprss11d* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS11D* humano, y que está bajo el control de un o unos elementos 5' reguladores, tal como el promotor y/o el o los potenciadores del gen *Tmprss11d* de roedor endógeno. Ejemplos de roedores incluyen ratones y ratas.

Un gen *Tmprss11d* humanizado codifica una proteína Tmprss11d humanizada que contiene un ectodominio que es sustancialmente idéntico al ectodominio de una proteína TMPRSS11D humana.

En realizaciones específicas, la proteína TMPRSS11D humana tiene una secuencia de aminoácidos que tiene al menos un 85 %, 90 %, 95 %, 98 %, 99 % o 100 % de identidad con la secuencia de aminoácidos expuesta en la SEQ ID NO: 18.

25 En algunas realizaciones, una proteína Tmprss11d humanizada contiene los 377 aminoácidos carboxiterminales de una proteína TMPRSS11D humana, por ejemplo, los aminoácidos 42 a 418 de una proteína TMPRSS11D humana. En algunas realizaciones, una proteína Tmprss11d humanizada contiene un ectodominio que es sustancialmente idéntico a la secuencia de aminoácidos compuesta de A42 a I418 de la SEQ ID NO: 18. En realizaciones específicas, 30 una proteína Tmprss11d humanizada contiene un ectodominio que tiene al menos un 85 %, 90 %, 95 %, 98 %, 99 % o 100 % de identidad con la secuencia de aminoácidos compuesta de A42 a I418 de la SEQ ID NO: 18; un ectodominio que difiere de la secuencia de aminoácidos compuesta de A42 a I418 de la SEQ ID NO: 18 en no más de 10, 9, 8, 7, 6, 5, 4, 3, 2 o 1 aminoácidos; o un ectodominio que difiere de la secuencia de aminoácidos compuesta de A42 a I418 de la SEQ ID NO: 18 solamente en el extremo N o C, por ejemplo, al carecer de 1-5 aminoácidos o al tener 1-5 35 aminoácidos adicionales en el extremo N o C.

Una proteína Tmprss11d humanizada contiene además una porción citoplasmática y transmembrana que es sustancialmente idéntica a la porción citoplasmática y transmembrana de una proteína Tmprss11d de roedor endógena. En algunas realizaciones, una proteína Tmprss11d humanizada incluye el dominio transmembrana y el dominio citoplasmático de una proteína Tmprss11d de roedor endógena.

En realizaciones específicas, una proteína Tmprss11d humanizada contiene el dominio transmembrana y el dominio citoplasmático de una proteína Tmprss11d de roedor endógena y el ectodominio de una proteína TMPRSS11D humana. En realizaciones particulares, un gen *Tmprss11d* humanizado codifica una proteína Tmprss11d humanizada que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 21.

Un gen *Tmprss11d* humanizado es el resultado de una sustitución de una secuencia de nucleótidos de un gen *Tmprss11d* de roedor endógeno en un locus *Tmprss11d* de roedor endógeno con una secuencia de nucleótidos de un gen *TMPrSS11D* humano.

50 En algunas realizaciones, una secuencia genómica contigua de un gen *Tmprss11d* de roedor endógeno en un locus *Tmprss11d* de roedor endógeno se ha sustituido por una secuencia genómica contigua de un gen *TMPRSS11D* humano para formar un gen *Tmprss11d* humanizado. En realizaciones específicas, la secuencia genómica contigua de un gen *TMPRSS11D* humano insertado en un gen *Tmprss11d* de roedor endógeno incluye secuencias de exones, 55 es decir, exones en su totalidad o en parte, de un gen *TMPRSS11D* humano que codifica un ectodominio que es sustancialmente idéntico al ectodominio de la proteína *TMPRSS11D* humana codificada por el gen *TMPRSS11D* humano. En circunstancias en las que una proteína *Tmprss11d* endógena y una proteína *TMPRSS11D* humana comparten aminoácidos comunes cerca de la unión del dominio transmembrana y el ectodominio, puede que no sea necesario insertar una secuencia genómica de *TMPRSS11D* humano que codifique precisamente el ectodominio de 60 la proteína *TMPRSS11D* humana, y es posible utilizar una secuencia genómica de *TMPRSS11D* humano ligeramente más larga o más corta que codifique sustancialmente el ectodominio de una proteína *TMPRSS11D* humana para producir una proteína *Tmprss11d* humanizada que tiene un ectodominio que es idéntico al ectodominio de la proteína *TMPRSS11D* humana.

65 En realizaciones específicas, una secuencia genómica contigua de un gen *TMPrSS11D* humano que se está insertando en un gen *Tmprss11d* de roedor endógeno contiene al menos el exón 3 codificante hasta el codón de

parada en el exón 10 codificante de un gen *TMPRSS11D* humano.

En determinadas realizaciones, una secuencia genómica contigua de un gen *TMPRSS11D* humano que se está insertando en un gen *Tmprss11d* de roedor endógeno contiene, al menos, una porción 3' del intrón 2 y el exón 3 codificante hasta el codón de parada en el exón 10 codificante del gen *TMPRSS11D* humano. En realizaciones específicas, la porción 3' del intrón 2 de un gen *TMPRSS2* humano incluido en la humanización tiene una longitud de aproximadamente 444 pares de bases.

5 En algunas realizaciones, una secuencia genómica contigua de un gen *TMPRSS11D* humano que se está insertando en un gen *Tmprss11d* de roedor endógeno contiene la 3' UTR del gen *TMPRSS11D* humano. En realizaciones específicas, todo el exón 10 codificante de un gen *TMPRSS11D* humano está incluido en la secuencia genómica de *TMPRSS11D* contigua para la humanización, que incluye la 3' UTR de un gen *TMPRSS11D* humano. En realizaciones particulares, una secuencia genómica contigua de un gen *TMPRSS11D* humano incluye una secuencia genómica humana adicional cadena abajo de la 3' UTR del gen *TMPRSS11D* humano. La secuencia genómica humana adicional puede ser una secuencia de 10-200 pb, 50-200 pb, o aproximadamente 150, 160, 170, 180 pb, que se encuentra inmediatamente cadena abajo del 3' UTR del gen *TMPRSS11D* humano en un locus *TMPRSS11D* humano.

10 La secuencia de nucleótidos de un gen *Tmprss11d* de roedor endógeno que permanece en un locus *Tmprss11d* humanizado codifica un polipéptido que es sustancialmente idéntico a la porción citoplasmática y transmembrana de la proteína *Tmprss11d* de roedor endógena codificada por el gen *Tmprss11d* de roedor endógeno. En circunstancias en las que una proteína *Tmprss11d* endógena y una proteína *TMPRSS11D* humana comparten aminoácidos comunes cerca de la unión del dominio transmembrana y el ectodominio, puede que no sea necesario mantener la secuencia genómica de *Tmprss11d* de roedor endógeno que codifique precisamente el dominio transmembrana de la proteína *Tmprss11d* de roedor endógena, y es posible mantener una secuencia genómica de *Tmprss11d* de roedor ligeramente más larga o más corta que codifique sustancialmente el dominio transmembrana de la proteína *Tmprss11d* de roedor endógena en la sustitución de humanización para codificar una proteína *Tmprss11d* humanizada que tiene un dominio transmembrana que es idéntico al dominio transmembrana de la proteína *Tmprss11d* de roedor endógena.

15 20 25 30 35 En realizaciones específicas, un gen *Tmprss11d* humanizado contiene los exones 1-2 codificantes de un gen *Tmprss11d* de roedor endógeno y el exón 3 codificante hasta el exón 10 codificante de un gen *TMPRSS11D* humano. En determinadas realizaciones, el gen *Tmprss11d* humanizado codifica una proteína *Tmprss11d* humanizada que contiene el dominio citoplasmático y el dominio transmembrana de la proteína *Tmprss11d* de roedor codificada por un gen *Tmprss11d* de roedor endógeno y el ectodominio de la proteína *TMPRSS11D* humana codificada por un gen *TMPRSS11D* humano. En realizaciones particulares, un gen *Tmprss11d* humanizado codifica una proteína *Tmprss11d* humanizada que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 21.

35 En algunas realizaciones, los exones e intrones de un gen *TMPRSS11D* humano y un gen *Tmprss11d* de roedor utilizados en la humanización son los que se encuentran en las SEQ ID NO: 15, 17 y 19-20.

40 45 50 En algunas realizaciones, un gen *Tmprss11D* humanizado da como resultado una expresión de la proteína *Tmprss11d* humanizada codificada en un roedor. En algunas realizaciones, una proteína *Tmprss11d* humanizada se expresa en un patrón comparable con, o sustancialmente el mismo que, una proteína *Tmprss11d* de roedor homóloga en un roedor de control (p. ej., un roedor sin el gen *Tmprss11d* humanizado que codifica la proteína *Tmprss11d* humanizada). En algunas realizaciones, una proteína *Tmprss11d* humanizada se expresa a un nivel comparable con, o sustancialmente igual que, una proteína *Tmprss11d* de roedor homóloga en un roedor de control (p. ej., un roedor sin el gen *Tmprss11d* humanizado que codifica la proteína *Tmprss11d* humanizada). En determinadas realizaciones, una proteína *Tmprss11d* humanizada se expresa y detecta en la superficie celular. En determinadas realizaciones, una proteína *Tmprss11d* humanizada o una forma soluble (por ejemplo, una forma secretada de ectodominio) se expresa y detecta en el suero de un roedor, por ejemplo, a un nivel comparable con o sustancialmente igual que, una proteína *Tmprss11d* de roedor homóloga o una forma soluble de la misma en un roedor de control.

Métodos para producir animales roedores con *Tmprss* humanizados

55 Otros aspectos de esta divulgación se refieren a métodos para producir un roedor con *Tmprss* humanizado descrito anteriormente, así como vectores de ácido nucleico y células madre embrionarias no humanas adecuadas para su uso en la producción de un roedor con *Tmprss* humanizado. En particular, la presente invención también proporciona un método para producir un roedor que tenga un gen *Tmprss* humanizado, que comprende:

60 65 (a) insertar un fragmento genómico en un locus *Tmprss* de roedor endógeno en una célula madre embrionaria de roedor, comprendiendo dicho fragmento genómico una secuencia de nucleótidos de un gen *TMPRSS* humano afín, formando, de este modo, un gen *Tmprss* humanizado,
en donde el gen *Tmprss* humanizado está bajo el control del promotor del gen *Tmprss* de roedor en el locus *Tmprss* de roedor endógeno y codifica una proteína *Tmprss* humanizada, y
en donde la proteína *Tmprss* humanizada comprende (i) un ectodominio que tiene al menos un 85 % de identidad de secuencia con el ectodominio de la proteína *TMPRSS* humana codificada por el gen *TMPRSS* humano afín y (ii) una porción citoplasmática y transmembrana que tiene al menos un 85 % de identidad de secuencia con la

porción citoplasmática y transmembrana de la proteína Tmprss de roedor codificada por el gen *Tmprss* de roedor en dicho locus *Tmprss* de roedor endógeno;

(b) obtener una célula madre embrionaria de roedor que comprende el gen *Tmprss* humanizado de (a); y,

(c) crear un roedor utilizando la célula madre embrionaria de roedor de (b).

5 Los roedores proporcionados en el presente documento se pueden producir utilizando métodos conocidos en la técnica. En realizaciones ilustrativas, un clon de un cromosoma artificial bacteriano (BAC, por sus siglas en inglés) que lleva un gen *Tmprss* de roedor puede modificarse utilizando recombinación bacteriana homóloga y tecnología VELOCIGENE® (véase, por ejemplo, el documento US 6.586.251 y Valenzuela *et al.* (2003), High-throughput 10 engineering of the mouse genome coupled with high-resolution expression analysis, Nature Biotech. 21(6):652-659). Como resultado, se ha eliminado una secuencia de nucleótidos de *Tmprss* de roedor del clon BAC original y se ha insertado una secuencia de nucleótidos de *Tmprss* humano, dando como resultado un clon BAC modificado que lleva un gen *Tmprss* humanizado, flanqueado por brazos de homología 5' y 3' de roedor. El clon BAC modificado, una vez 15 linealizado, se puede introducir en células madre embrionarias de roedor (ES, por sus siglas en inglés) mediante, por ejemplo, electroporación. En la técnica se han descrito tanto células ES de ratón como células ES de rata. Véanse, por, ejemplo, los documentos US 7.576.259, US 7.659.442, US 7.294.754 y US 2008-0078000 A1 describen células ES de ratón y el método VELOCIMOUSE® para producir un ratón genéticamente modificado; los documentos US 2014/0235933 A1, US 2014/0310828 A1, Tong *et al.* (2010) Nature 467:211-215 y Tong *et al.* (2011) Nat Protoc. 6(6): 20 doi: 10.1038/nprot.2011.338 describen células ES de rata y métodos para producir una rata genéticamente modificada.

20 Se pueden seleccionar células ES que tienen un gen *Tmprss* humanizado integrado en el genoma. En algunas realizaciones, las células ES que tienen un *Tmprss* humanizado integrado en un locus *Tmprss* de roedor endógeno se pueden seleccionar basándose en ensayos de pérdida de alelos de roedor y/o ganancia de alelos humanos. Las células ES seleccionadas se utilizan después como células ES donantes para inyección en un embrión en etapa previa 25 a la mórula (p. ej., embrión en etapa de 8 células) mediante el método VELOCIMOUSE® (véase, por ejemplo, los documentos US 7.576.259, US 7.659.442, US 7.294.754 y US 2008-0078000 A1), o los métodos descritos en los documentos US 2014/0235933 A1 y US 2014/0310828 A1. El embrión que comprende las células ES donantes se incuba hasta la etapa de blastocisto y luego se implanta en una madre sustituta para producir un roedor F0 totalmente derivado de las células Es donantes. Las crías de roedores que llevan el gen *Tmprss* humanizado pueden identificarse 30 mediante genotipificación del ADN aislado de cortes de la cola utilizando ensayos de pérdida de alelos de roedor y/o ganancia de alelos humanos.

35 Se pueden cruzar roedores heterocigotos para un gen *Tmprss* humanizado con roedores homocigotos generados. Se pueden cruzar roedores que contienen un gen *Tmprss* humanizado con roedores que contienen otro gen *Tmprss* humanizado para hacer producir roedores que contienen múltiples genes *Tmprss* humanizados. Por ejemplo, se pueden cruzar roedores que contienen un gen *Tmprss2* humanizado con roedores que contienen un gen *Tmprss4* humanizado para producir roedores que contienen un gen *Tmprss2* humanizado y un gen *Tmprss4* humanizado.

Métodos que emplean roedores que tienen genes Tmprss humanizados

40 Los roedores divulgados en el presente documento proporcionan un sistema *in vivo* útil y una fuente de materiales biológicos (por ejemplo, células) que expresan proteínas *Tmprss* humanizadas para identificar y probar compuestos que se dirigen específicamente a proteínas TMPRSS humanas.

45 En un aspecto, un roedor divulgado en el presente documento se usa para determinar la capacidad de un compuesto candidato, tal como un inhibidor de una proteína TMPRSS humana, para tratar y/o prevenir la infección por el virus de la gripe. En particular, la presente invención proporciona un método para evaluar la eficacia terapéutica de un compuesto en el tratamiento de una infección por el virus de la gripe A, que comprende:

50 proporcionar un roedor de la presente invención;
administrar un virus de la gripe A y un compuesto candidato al roedor, en donde opcionalmente el compuesto candidato es un anticuerpo o fragmento de unión a antígeno del mismo específico para una proteína TMPRSS humana; y
55 monitorizar la presencia y la gravedad de la infección por el virus de la gripe A en el roedor para determinar la eficacia terapéutica del compuesto candidato.

60 En algunas realizaciones, un roedor que contiene un gen *Tmprss* humanizado y que expresa una proteína *Tmprss* humanizada divulgada en el presente documento se administra con un compuesto candidato antes de la infección experimental con el virus de la gripe. La eficacia profiláctica del compuesto se puede evaluar determinando si el roedor presenta menos síntomas y/o síntomas menos graves de infección por el virus de la gripe y/o una viabilidad mejorada, en comparación con el o los roedores de control.

65 En otras realizaciones, un roedor que contiene un gen *Tmprss* humanizado y que expresa una proteína *Tmprss* humanizada que comprende el ectodominio de una proteína TMPRSS humana se administra con un inhibidor candidato de esa proteína TMPRSS humana después de la infección experimental con el virus de la gripe. La eficacia profiláctica del inhibidor candidato se puede evaluar determinando si el roedor presenta menos síntomas y/o síntomas

menos graves de infección por el virus de la gripe y/o una viabilidad mejorada, en comparación con el o los roedores de control.

- 5 Los roedores de control adecuados incluyen, por ejemplo, roedores que contienen un gen *Tmprss* humanizado sin estar sometidos a la infección experimental; y roedores que contienen un gen *Tmprss* humanizado sometidos a la infección experimental sin ningún compuesto; y roedores que contienen un gen *Tmprss* humanizado sometidos a la infección experimental y un compuesto que se sabe que es terapéuticamente eficaz.
- 10 Los compuestos que pueden evaluarse en los métodos de la presente invención incluyen inhibidores de TMPRSS candidatos, por ejemplo, un inhibidor de proteasas de molécula pequeña, un inhibidor basado en ácidos nucleicos (p. ej., ARNip, ribozima, construcción antisentido, etc.), proteína de unión a antígeno (por ejemplo, anticuerpo o fragmento de unión a antígeno del mismo), o un péptido de bloqueo/inhibidor de péptido. Un inhibidor de TMPRSS puede funcionar inhibiendo o reduciendo la capacidad de una proteína TMPRSS de escindir proteolíticamente la proteína precursora de hemaglutinina (HA0) en las subunidades HA1 y HA2.
- 15 En algunas realizaciones, un inhibidor candidato es un anticuerpo o fragmento de unión a antígeno del mismo. Tanto los anticuerpos monoclonales como policlonales son adecuados para los fines de la presente invención. En realizaciones específicas, el anticuerpo se une específicamente a una proteína TMPRSS e inhibe la actividad proteasa de esa proteína TMPRSS y no inhibe sustancialmente la actividad proteasa de otra proteína TMPRSS. Por ejemplo, 20 un inhibidor de anticuerpos anti-TMPRSS2 se une específicamente a una proteína TMPRSS2 e inhibe la actividad proteasa de la proteína TMPRSS2, y no tiene ningún efecto sobre la actividad proteolítica de TMPRSS4 o TMPRSS1 ID, o reduce la actividad proteolítica de TMPRSS4 o TMPRSS1 ID en no más de un 25 % (p. ej., 20 %, 15 %, 10 %, 5 %, o menos) con respecto a una molécula de control no inhibidora probada en condiciones experimentales idénticas o sustancialmente idénticas.
- 25 En algunas realizaciones, el inhibidor es un anticuerpo anti-TMPRSS2 o fragmento de unión a antígeno del mismo. En algunas realizaciones, el inhibidor es un anticuerpo anti-TMPRSS4 o fragmento de unión a antígeno del mismo. En otras realizaciones, el inhibidor es un anticuerpo anti-TMPRSS11D o fragmento de unión a antígeno del mismo.
- 30 La infección con el virus de la gripe experimental se puede inducir y controlar siguiendo protocolos conocidos. Véase, por ejemplo, el documento US 2013/0273070 A1. Por ejemplo, los animales roedores se pueden administrar por vía intranasal con el virus de la gripe. Los animales infectados pueden evaluarse para determinar los síntomas y la gravedad de la infección. Por ejemplo, los animales pueden analizarse para determinar(1) cambio de peso y supervivencia, (2) cambios celulares a través de citometría de flujo, (3) inmunohistoquímica, tinción de PAS y H&E de pulmones completos y (4) niveles de citocinas en suero. Los animales de control que se sabe que son susceptibles al virus presentan un aumento significativo en la frecuencia de células dendríticas, los niveles de macrófagos alveolares positivos para gripe, neutrófilos o células epiteliales en los pulmones y los niveles de IFNy, en comparación con los animales no infectados.
- 35

40 Ejemplos

Los siguientes ejemplos se proporcionan con el fin de describir para los expertos en la materia como preparar y usar los métodos y composiciones de la invención, y no pretenden limitar el alcance de lo que los inventores de la presente invención consideran como su invención. A menos que se indique lo contrario, la temperatura se indica en grados Celsius y la presión es la atmosférica o casi atmosférica.

Ejemplo 1. Humanización de un gen *Tmprss2* endógeno.

Este ejemplo ilustra métodos ilustrativos de humanización de un gen endógeno que codifica *Tmprss2* en un roedor (por ejemplo, un ratón). Los métodos descritos en este ejemplo se pueden emplear para humanizar un gen *Tmprss2* endógeno de un roedor utilizando cualquier secuencia humana, o combinación de secuencias (o fragmentos de secuencias) humanas según se deseé.

Un vector de direccionamiento para la humanización de un gen *Tmprss2* endógeno se construyó utilizando clones de cromosomas artificiales bacterianos (BAC) y tecnología VELOCIGENE® (véase, por ejemplo, la patente de los Estados Unidos N.º 6.586.251 y Valenzuela *et al.* (2003) High-throughput engineering of the mouse genome coupled with high-resolution expression analysis, Nature Biotech. 21(6):652-659).

Brevemente, se utilizó el clon bMQ-264A15 de cromosoma artificial bacteriano (BAC) de ratón que contiene un gen *Tmprss2* de ratón y se modificó de la siguiente manera. Se generó un fragmento de ADN para incluir una secuencia de nucleótidos de homología de ratón 5', un ADN genómico de *TMPRSS2* humano de aproximadamente 25.091 pb (que contiene los últimos 7 pb del exón 3 codificante, el intrón 3 y el exón 4 codificante hasta el exón 13 (incluida la 3' UTR que forma parte del exón 13 codificante), de un gen *TMPRSS2* humano), un casete de neomicina autoeliminable de aproximadamente 2.691 pb y una secuencia de homología 3' de ratón. Este fragmento de ADN se utilizó para modificar el clon BAC bMQ-264A15 mediante recombinación homóloga en células bacterianas. Como resultado, un ratón con un fragmento genómico de *Tmprss2* de ratón que codifica un ectodominio (de aproximadamente 25.291 pb)

en el clon BAC se sustituyó con el fragmento genómico de *TMPRSS2* humano de aproximadamente 25.091 pb, seguido de un casete de neomicina autoeliminable de aproximadamente 2691 pb. Específicamente, el fragmento genómico de *Tmprss2* de ratón que se sustituyó incluía los últimos 7 pb del exón 3 codificante, el intrón 3 y el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *Tmprss2* de ratón (figuras 1A-1B). El fragmento genómico de *TMPRSS2* humano que se insertó incluía los últimos 7 pb del exón 3 codificante, el intrón 3 y el exón 4 codificante hasta el exón 13 codificante de un gen *TMPRSS2* humano (incluida la 3' UTR de *TMPRSS2* humano) y una secuencia genómica 3' humana de 131 pb cadena abajo de la 3' UTR de *TMPRSS2* humano (figuras 1A-1B). El clon BAC modificado resultante incluía, de 5' a 3', (i) un brazo de homología 5' de ratón que contiene aproximadamente 12 kb de ADN genómico de ratón que incluye la 5' UTR de *Tmprss2* de ratón, el exón 1 (no codificante) de *Tmprss2* de ratón, los exones 1-3 codificantes (excepto los últimos 7 pb del exón 3 codificante); (ii) un fragmento genómico de *TMPRSS2* humano de aproximadamente 25.091 pb que incluye los últimos 7 pb del exón 3 codificante humano, el intrón 3, los exones 4 a 13 codificantes humanos (incluida la 3' UTR de *TMPRSS2* humano), y una secuencia genómica 3' humana; (iii) un casete de neomicina autoeliminable de aproximadamente 2691 pb, seguido de (iv) un brazo de homología 3' de ratón de 45 kb que contiene la 3' UTR de *Tmprss2* de ratón y el ADN genómico de ratón restante en el clon BAC original. Véase las figuras 1A-1B. Las secuencias de unión también se exponen en la parte inferior de la figura 1B. La parte del clon BAC modificado que contiene el fragmento genómico de *TMPRSS2* humano y el casete de neomicina, así como las uniones de inserción cadena arriba y cadena abajo, se expone en la SEQ ID NO:5. La secuencia de aminoácidos de la proteína codificada por el gen *Tmprss2* humanizado se expone en la SEQ ID NO: 7. Se proporciona una alineación de esta proteína *Tmprss2* humanizada ("proteína 7010 mutante"), una proteína *Tmprss2* de ratón (SEQ ID NO: 2) y una proteína *TMPRSS2* humana (SEQ ID NO: 4) en la figura 1D.

El clon BAC modificado que contiene el gen *Tmprss2* humanizado, como se describe anteriormente, se utilizó para someter a electroporación a células madre embrionarias (ES) de ratón para crear células ES modificadas que comprenden un gen *Tmprss2* humanizado. Las células ES dirigidas positivamente que contienen un gen *Tmprss2* humanizado se identificaron mediante un ensayo (Valenzuela *et al.*, *supra*) que detectó la presencia de secuencias de *TMPRSS2* humano (p. ej., los exones 4-13 codificantes de *TMPRSS2* humano) y confirmó la pérdida y/o conservación de secuencias de *Tmprss2* de ratón (p. ej., pérdida de los exones 4-13 codificantes de *Tmprss2* de ratón). La tabla 1 muestra los cebadores y las sondas que se utilizaron para confirmar la humanización de un gen *Tmprss2* endógeno como se describe anteriormente (figuras 1A-1B). Una vez que se ha seleccionado un clon de células ES correctamente dirigido, el casete de selección de neomicina se puede escindir introduciendo una recombinasa Cre, por ejemplo, a través de electroporación. Como alternativa, el casete de selección de neomicina puede eliminarse cruzando la descendencia generada a partir del clon ES con una cepa de roedor eliminador que expresa una recombinasa Cre. El locus *Tmprss2* humanizado después de la eliminación del casete se muestra en la figura 1C, con las secuencias de unión que se muestran en la parte inferior de la figura 1C.

Se utilizaron clones de células ES seleccionados (con o sin el casete) para implantar en ratones hembra utilizando el método VELOCIMOUSE® (véase, por ejemplo, la patente de los EE.UU. N.º 7.294.754 y Poueymirou *et al.*, F0 generation mice that are essentially fully derived from the donor gene-targeted ES cells allowing immediate phenotypic analyses, 2007, Nature Biotech. 25(1):91-99) para generar una camada de crías que contenía un alelo *Tmprss2* humanizado en el genoma. Se pueden confirmar e identificar nuevamente ratones que llevan un alelo *Tmprss2* humanizado mediante la genotipificación del ADN aislado de los cortes de la cola utilizando una modificación del ensayo de alelos (Valenzuela *et al.*, *supra*) que detecta la presencia de las secuencias del gen *TMPRSS2* humano. Las crías se genotipifican y se seleccionan las cohortes de animales heterocigotos para el locus *Tmprss2* humanizado para la caracterización. Se producen animales homocigotos para el locus *Tmprss2* humanizado cruzando animales heterocigotos.

TABLA 1

| Nombre | Cebador | Secuencia (5'-3') | SEQ ID NO |
|---------|-------------|----------------------------|-----------------|
| 7010U | Directo | GCCGTGACTGTGACCTTCTC | (SEQ ID NO: 26) |
| | Sonda (BHQ) | TGGAGGAGGCCACCTGTATGCCTC | (SEQ ID NO: 27) |
| | Inverso | GCCTTGCCCTCAATGGAAAC | (SEQ ID NO: 28) |
| 7010D | Directo | GGTTGCACAGCAAGGAAGAAG | (SEQ ID NO: 29) |
| | Sonda (BHQ) | CCAGGAGTTCCCTGTGAGCCTACCC | (SEQ ID NO: 30) |
| 7010hU | Inverso | TGGAATGGAAGGAGCTGGAG | (SEQ ID NO: 31) |
| | Directo | GTCCCCACCTCCTGCAACTG | (SEQ ID NO: 32) |
| 7010hTD | Sonda (BHQ) | TGAGCCTTCCCACATCAGCCTGGG | (SEQ ID NO: 33) |
| | Inverso | CCACAATGGCACATGGCTCG | (SEQ ID NO: 34) |
| 7010hTD | Directo | GGTGCTTGCCTCCCAAGA | (SEQ ID NO: 35) |
| | Sonda (BHQ) | <u>CCTAAAAGGTGTTGAATGG</u> | (SEQ ID NO: 36) |
| | Inverso | GGCAATAAAGAAGGAAGACGTTT | (SEQ ID NO: 37) |

Ejemplo 2. Humanización de un gen *Tmprss4* endógeno.

Este ejemplo ilustra métodos ilustrativos de humanización de un gen endógeno que codifica *Tmprss4* en un roedor (por ejemplo, un ratón). Los métodos descritos en este ejemplo se pueden emplear para humanizar un gen *Tmprss4* endógeno de un roedor utilizando cualquier secuencia humana, o combinación de secuencias (o fragmentos de

secuencias) humanas según se deseé.

Un vector de direccionamiento para la humanización de un gen *Tmprss4* endógeno se construyó utilizando clones de cromosomas artificiales bacterianos (BAC) y tecnología VELOCIGENE® (véase, por ejemplo, la Patente de los Estados Unidos N.º 6.586.251 y Valenzuela *et al.* (2003), *supra*).

Brevemente, se utilizó el clon RP23-71M15 de cromosoma artificial bacteriano (BAC) de ratón que contiene un gen *Tmprss4* de ratón y se modificó de la siguiente manera. Se generó un fragmento de ADN para incluir una secuencia de nucleótidos de homología de ratón 5', un casete de neomicina autoeliminable de aproximadamente 4996 pb, un ADN genómico humano de aproximadamente 14.963 pb (que contiene el exón 4 codificante hasta el codón de parada en el exón 13 codificante de un gen *TMPRSS4* humano), y una secuencia de homología 3' de ratón. Este fragmento de ADN se utilizó para modificar el clon BAC RP23-71M15 mediante recombinación homóloga en células bacterianas. Como resultado, un fragmento genómico de ratón que codifica un ectodominio (de aproximadamente 11.074 pb) en el clon BAC se sustituyó con un casete de neomicina autoeliminable de aproximadamente 4.996 pb, seguido del ADN genómico humano de aproximadamente 14.963 pb. Específicamente, el fragmento genómico de ratón que se eliminó y sustituyó incluía las 130 pb 3' del intrón 3 de ratón, el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *Tmprss4* de ratón (figuras 2A-2B). El fragmento genómico humano que se insertó incluía una porción 3' del intrón 3 de *TMPRSS4* humano de aproximadamente 150 pb, y el exón 4 codificante de *TMPRSS4* humano hasta el codón de parada en el exón 13 codificante (figuras 2A-2B). El clon BAC modificado resultante incluía, de 5' a 3', un brazo de homología 5' de ratón que contiene aproximadamente 44,8 kb de ADN genómico de ratón (incluida una 5' UTR de *Tmprss4* de ratón, los exones 1 a 3 codificantes de *Tmprss4* de ratón, el intrón 3 de *Tmprss4* de ratón en parte (sin los 130 pb 3'), un casete de neomicina autoeliminable de aproximadamente 4996 pb, una porción 3' del intrón 3 de *TMPRSS4* de aproximadamente 150 pb, el exón 4 codificante de *TMPRSS4* hasta el codón de parada en el exón 13 codificante, seguido directamente por la 3' UTR de *Tmprss4* de ratón y el ADN genómico de ratón restante en el clon BAC original (un brazo de homología 3' de ratón de aproximadamente 118 kb en total). Véase las figuras 2A-2B. Las secuencias de unión también se exponen en la parte inferior de la figura 2B. La parte del clon BAC modificado que contiene el casete de neomicina y el fragmento genómico de *TMPRSS4* humano, así como las uniones de inserción cadena arriba y cadena abajo, se expone en la SEQ ID NO:12. La secuencia de aminoácidos de la proteína codificada por el gen *Tmprss4* humanizado se expone en la SEQ ID NO: 14. Se proporciona una alineación de esta proteína *Tmprss4* humanizada ("pro 7224 mutante"), una proteína *Tmprss4* de ratón (SEQ ID NO: 9) y una proteína *TMPRSS4* humana (SEQ ID NO: 11) en la figura 2D.

El clon BAC modificado que contiene el gen *Tmprss4* humanizado, como se describe anteriormente, se utilizó para someter a electroporación a células madre embrionarias (ES) de ratón para crear células ES modificadas que comprenden un gen *Tmprss4* humanizado. Las células ES dirigidas positivamente que contienen un gen *Tmprss4* humanizado se identificaron mediante un ensayo (Valenzuela *et al.*, *supra*) que detectó la presencia de secuencias de *TMPRSS4* humano (p. ej., los exones 4-13 codificantes de *TMPRSS4* humano) y confirmó la pérdida y/o conservación de secuencias de *Tmprss4* de ratón (p. ej., pérdida de los exones 4-13 codificantes de *Tmprss4* de ratón). La tabla 2 muestra los cebadores y las sondas que se utilizaron para confirmar la humanización de un gen *Tmprss4* endógeno como se describe anteriormente (figuras 2A-2B). Una vez que se ha seleccionado un clon de células ES correctamente dirigido, el casete de selección de neomicina se puede escindir introduciendo una recombinasa Cre, por ejemplo, a través de electroporación. Como alternativa, el casete de selección de neomicina puede eliminarse cruzando la descendencia generada a partir del clon ES con una cepa de roedor eliminador que expresa una recombinasa Cre. El locus *Tmprss4* humanizado después de la eliminación del casete se muestra en la figura 2C, con las secuencias de unión que se muestran en la parte inferior de la figura 2C.

Se utilizaron clones de células ES seleccionados (con o sin el casete) para implantar en ratones hembra utilizando el método VELOCIMOUSE® (véase, por ejemplo, la Patente de los Estados Unidos N.º 7.294.754 y Poueymirou *et al.* (2007), *supra*) para generar una camada de cachorros que contenía un alelo *Tmprss4* humanizado en el genoma. Se pudieron confirmar e identificar nuevamente ratones que llevaban un alelo *Tmprss4* humanizado mediante la genotipificación del ADN aislado de los cortes de la cola utilizando una modificación del ensayo de alelos (Valenzuela *et al.*, *supra*) que detectaba la presencia de las secuencias del gen *TMPRSS4* humano. Las crías se genotipificaron y se seleccionaron las cohortes de animales heterocigotos para el locus *Tmprss4* humanizado para la caracterización. Se produjeron animales homocigotos para el locus *Tmprss4* humanizado cruzando animales heterocigotos.

55

TABLA 2

| Nombre | Cebador | Secuencia (5'-3') | SEQ ID NO |
|----------|-------------|-------------------------|-----------------|
| 7224mTU | Directo | GAGCAGGGCCATGACACAT | (SEQ ID NO: 42) |
| | Sonda (BHQ) | ACCATTAGATCCCGACGTGGACA | (SEQ ID NO: 43) |
| | Inverso | AAACCCCTCCCGAGAGAGAA | (SEQ ID NO: 44) |
| | Directo | GAGGAACACTGTGTCAAGGACTT | (SEQ ID NO: 45) |
| 7224mTU2 | Sonda (BHQ) | CCTGAAAAGCCGGAGTGGCAG | (SEQ ID NO: 46) |
| | Inverso | GGGCAGAGACCACATCTGA | (SEQ ID NO: 47) |

(continuación)

| Nombre | Cebador | Secuencia (5'-3') | SEQ ID NO |
|---------------|----------------|---|------------------|
| 7224mTD | Directo | GGAAGCCCTCTCGATACTTG | (SEQ ID NO: 48) |
| | Sonda (BHQ) | TTCTACCCTGAGGGCATGCAGC | (SEQ ID NO: 49) |
| | Inverso | TGGGATGTAGAAGGTTGTCAGA | (SEQ ID NO: 50) |
| 7224hTU | Directo | CTGAGCCTGGAACTCACACATG | (SEQ ID NO: 51) |
| | Sonda (BHQ) | TCTGAGAGCCCAGCACTATGCC | (SEQ ID NO: 52) |
| | Inverso | GCTGAGGGTCAGGCTTGAG | (SEQ ID NO: 53) |
| 7224hTD | Directo | TCTGCAGGGTAGGGAGAGAAG | (SEQ ID NO: 54) |
| | Sonda (BHQ) | TGTTTCAGAAAAGGAAGACTCACGTTA CA | (SEQ ID NO: 55) |
| | Inverso | GAGACCGATGAAGAGAAAGTCAGA | (SEQ ID NO: 56) |

Ejemplo 3. Humanización de un gen *Tmprss11d* endógeno.

- 5 Este ejemplo ilustra métodos ilustrativos de humanización de un gen endógeno que codifica *Tmprss11d* en un roedor (por ejemplo, un ratón). Los métodos descritos en este ejemplo se pueden emplear para humanizar un gen *Tmprss11d* endógeno de un roedor utilizando cualquier secuencia humana, o combinación de secuencias (o fragmentos de secuencias) humanas según se deseé.
- 10 Un vector de direccionamiento para la humanización de un gen *Tmprss11d* endógeno se construyó utilizando clones de cromosomas artificiales bacterianos (BAC) y tecnología VELOCIGENE® (véase, por ejemplo, la Patente de los Estados Unidos N.º 6.586.251 y Valenzuela *et al.* (2003), *supra*).
- 15 Brevemente, se utilizó el clon RP23-95N22 de cromosoma artificial bacteriano (BAC) de ratón que contiene un gen *Tmprss11d* de ratón y se modificó de la siguiente manera. Se generó un fragmento de ADN para incluir una secuencia de nucleótidos de homología de ratón 5', un ADN genómico de *TMPRSS11D* humano de aproximadamente 33.927 pb (que contiene 444 pb en el extremo 3' del intrón 2, y el exón 3 codificante hasta el exón 10 codificante (incluida la 3' UTR que es parte del exón 10 codificante), de un gen *TMPRSS11D* humano), un casete de neomicina autoeliminable de aproximadamente 4.996 pb y una secuencia de homología de ratón 3'. Este fragmento de ADN se utilizó para
- 20 modificar el clon BAC RP23-95N22 mediante recombinación homóloga en células bacterianas. Como resultado, un ratón con un fragmento genómico de *Tmprss11d* de ratón que codifica un ectodominio (de aproximadamente 35.667 pb) en el clon BAC se sustituyó con el fragmento genómico de *TMPRSS11D* humano de aproximadamente 33.927 pb, seguido de un casete de neomicina autoeliminable de aproximadamente 4996 pb. Específicamente, el fragmento genómico de *Tmprss11d* de ratón que se sustituyó incluía una porción 3' del intrón 2 y el exón 3 codificante hasta el codón de parada en el exón 10 codificante del gen *Tmprss11d* de ratón (figuras 3A-3B). El fragmento genómico de *TMPRSS11D* humano que se insertó incluía 444 pb en el extremo 3' del intrón 2 y el exón 3 codificante hasta el exón 10 codificante de un gen *TMPRSS11D* humano (incluida la 3' UTR de *TMPRSS11D* humano) y una secuencia genómica 3' humana de aproximadamente 172 pb cadena abajo de la 3' UTR de *TMPRSS11D* humano (figuras 3A-3B). El clon BAC modificado resultante incluía, de 5' a 3', (i) un brazo de homología 5' de ratón que contiene
- 25 aproximadamente 143 kb de ADN genómico de ratón que incluye la 5' UTR de *Tmprss11d* de ratón, los exones 1-2 codificantes de *Tmprss11d* de ratón y una porción 5' del intrón 2; (ii) un fragmento genómico de *TMPRSS11D* humano que incluía una porción 3' del intrón 2 y los exones 3 a 10 codificantes (incluida la 3' UTR) de *TMPRSS11D* humano y una secuencia genómica 3' humana; (iii) un casete de neomicina autoeliminable de aproximadamente 4996 pb, seguido de (iv) un brazo de homología 3' de ratón de 10 kb que contiene la 3' UTR de *Tmprss11d* de ratón y el ADN
- 30 genómico de ratón restante en el clon BAC original. Véase las figuras 3A-3B. Las secuencias de unión también se exponen en la parte inferior de la figura 3B. La parte del clon BAC modificado que contiene el fragmento genómico de *TMPRSS11D* humano y el casete de neomicina, así como las uniones de inserción cadena arriba y cadena abajo, se expone en la SEQ ID NO:19. La secuencia de aminoácidos de la proteína codificada por el gen *Tmprss11d* humanizado se expone en la SEQ ID NO: 21. Se proporciona una alineación de esta proteína *Tmprss11d* humanizada ("pro 7226 mutante"), una proteína *Tmprss11d* de ratón (SEQ ID NO: 16) y una proteína *TMPRSS11D* humana (SEQ ID NO: 18) en la figura 3D.
- 35 El clon BAC modificado que contiene el gen *Tmprss11d* humanizado, como se describe anteriormente, se utiliza para someter a electroporación a células madre embrionarias (ES) de ratón para crear células ES modificadas que comprenden un gen *Tmprss11d* humanizado. Las células ES dirigidas positivamente que contienen un gen *Tmprss11d* humanizado se identifican mediante un ensayo (Valenzuela *et al.*, *supra*) que detecta la presencia de las secuencias de *TMPRSS11D* humano (p. ej., los exones 3-10 codificantes de *TMPRSS11D* humano) y confirma la pérdida y/o conservación de secuencias de *Tmprss11d* de ratón (p. ej., pérdida de los exones 3-10 codificantes de *Tmprss11d* de ratón). La tabla 3 muestra los cebadores y las sondas que se utilizaron para confirmar la humanización de un gen *Tmprss11d* endógeno como se describe anteriormente (figuras 3A-3B). Una vez que se ha seleccionado un clon de
- 40
- 45
- 50

células ES correctamente dirigido, el casete de selección de neomicina se puede escindir introduciendo una recombinasa Cre, por ejemplo, a través de electroporación. Como alternativa, el casete de selección de neomicina puede eliminarse cruzando la descendencia generada a partir del clon ES con una cepa de roedor eliminador que expresa una recombinasa Cre. El locus *Tmprss11d* humanizado después de la eliminación del casete se muestra en la figura 3C, con las secuencias de unión que se muestran en la parte inferior de la figura 3C.

5 Se utilizan clones de células ES seleccionados (con o sin el casete) para implantar en ratones hembra utilizando el método VELOCIMOUSE® (véase, por ejemplo, la Patente de los Estados Unidos N.º 7.294.754 y Poueymirou *et al.* (2007), *supra*) para generar una camada de cachorros que contenía un alelo *Tmprss11d* humanizado en el genoma.

10 10 Se pueden confirmar e identificar nuevamente ratones que llevaba un alelo *Tmprss11d* humanizado mediante la genotipificación del ADN aislado de los cortes de la cola utilizando una modificación del ensayo de alelos (Valenzuela *et al.*, *supra*) que detecta la presencia de las secuencias del gen *TMPRSS11D* humano. Las crías se genotipifican y se seleccionan las cohortes de animales heterocigotos para el locus *Tmprss11d* humanizado para la caracterización. Se producen animales homocigotos para el locus *Tmprss11d* humanizado cruzando animales heterocigotos.

15

TABLA 3

| Nombre | Cebador | Secuencia (5'-3') | SEQ ID NO |
|---------|-------------|----------------------------------|-----------------|
| 7226mTU | Directo | TCCTCTCCAGACAAGAAAGCT | (SEQ ID NO: 61) |
| | Sonda (BHQ) | TCATAGCAGCTTCAAATCTAAACGT TGA | (SEQ ID NO: 62) |
| | Inverso | TCGTGTGAGCTGGTGAGTT | (SEQ ID NO: 63) |
| 7226mTD | Directo | CATGCGATCACAGGAGGAGATC | (SEQ ID NO: 64) |
| | Sonda (BHQ) | AATTGGGCCCGAAGCCAGATGC | (SEQ ID NO: 65) |
| | Inverso | CGGAAGGCTCTGTGACTTC | (SEQ ID NO: 66) |
| 7226hTU | Directo | GTCTCCCACCTCTGACATAATGAAC | (SEQ ID NO: 67) |
| | Sonda (BHQ) | CCCAGTGTAAACCCTACATCTGGTTCC | (SEQ ID NO: 68) |
| | Inverso | TGGGAAGAGACTCTTGGACA | (SEQ ID NO: 69) |
| 7226hTD | Directo | ATGAGCTCCTAGTACAGCTAAAGTT | (SEQ ID NO: 70) |
| | Sonda (MGB) | ATGCATGATCATCTATGCGTCAGAGC | (SEQ ID NO: 71) |
| | Inverso | TGCCAGATGCAGGGAGTTAG | (SEQ ID NO: 72) |

Ejemplo 4. Evaluación de los virus de la gripe A del grupo 1 y del grupo 2 en ratones MAID7225 Humln frente a ratones *Tmprss4* de tipo silvestre

20 20 Para validar el uso de roedores con *Tmprss* humanizados como modelo animal de infección, se realizaron experimentos para evaluar la supervivencia de ratones MAID7225 Humln TMPRSS4 frente a compañeros de camada de tipo silvestre (TS) en un modelo de infección grave por gripe A del grupo 1 y del grupo 2.

25 25 Los ratones MAID7225 Humln TMPRSS4 son homocigotos para un gen *Tmprss4* humanizado en su genoma y se generaron como se describe en el ejemplo 2. Las cepas víricas utilizadas en estos estudios incluyeron el aislado histórico del virus de la gripe A del grupo 1 A/Puerto Rico/08/1934 (H1N1) y un aislado del virus de la gripe A del grupo 2 propio adaptado a ratón A/Aichi/02/1968 (HA, NA) X-31 (H3N2). Todos los experimentos se realizaron en ratones machos y hembras MAID7225 Humln TMPRSS4 de 6-8 semanas de edad o compañeros de camada de TS. Los ratones se expusieron con 1150 unidades formadoras de placa (UFP) de A/Puerto Rico/08/1934 (H1N1) o 10.000 UFP de A/Aichi/02/1968-X31 (H3N2). En estos modelos de supervivencia, los ratones se expusieron por vía intranasal (IN) el día 0 después de la infección (p.i.). Los ratones se pesaron y observaron diariamente hasta el día 14 p.i. y se sacrificaron cuando perdieron el 20 % de su peso inicial. Los resultados se indican como porcentaje de supervivencia (tabla 4).

35

Tabla 4

| ID del grupo | Número de ratones por grupo | Porcentaje de supervivencia (n.º de ratones supervivientes/n.º de ratones total en el grupo) |
|---|-----------------------------|--|
| Control no infectado (2 Humln, 2 ratones de TS) | 4 | 100 (4/4) |
| TMPRSS4 TS; infectado con H1 PR34 | 10 | 20 (2/10) |
| Humln TMPRSS4; infectado con H1 PR34 | 8 | 25 (2/8) |
| TMPRSS4 TS; infectado con H3 X31 | 9 | 11,1 (1/9) |

(continuación)

| ID del grupo | Número de ratones por grupo | Porcentaje de supervivencia (n.º de ratones supervivientes/n.º de ratones total en el grupo) |
|-------------------------------------|-----------------------------|--|
| Humln TMPRSS4; infectado con H3 X31 | 8 | 25 (2/8) |

La supervivencia de los ratones MAID7225 Humln TMPRSS4 se comparó con los compañeros de camada de TS después de la exposición tanto con el virus de la gripe A del grupo 1 grave [A/Puerto Rico/08/1934 (H1N1)] como con el virus de la gripe A del grupo 2 adaptado a ratón grave [A/Aichi/02/1968-X31 (H3N2)] (figura 4). La supervivencia de los ratones MAID7225 Humln TMPRSS4 no fue diferente de la de los ratones de tipo silvestre con la exposición a H1N1 (25 %; n = 8 y 20 %; n = 10, respectivamente) o con la exposición a H3N2 (25 %; n = 8 y 11,1 %; n = 9, respectivamente).

10 LISTADO DE SECUENCIAS

<110> REGGENERON PHARMACEUTICALS, INC.

<120> ROEDORES QUE TIENEN UN GEN TMPRSS HUMANIZADO

<130> 33093PCT (10234WO01)

<150> 62/301.023

<151> 29/02/2016

20 <160> 72

<170> PatentIn versión 3.5

25 <210> 1

<211> 3175

<212> ADN

<213> *Mus musculus*

30 <400> 1

ES 2 886 958 T3

| | |
|--|------|
| gccttcctg gccgttccct cttctggcc gaggtgcctg cgtttagggg tgtcaccctg | 60 |
| gctcccgaaa cgccgcctcc ggagattaa gcgagaactg gagtaggtcg tgtacttggaa | 120 |
| gcggacgagg aagccaagag ctcggacaga ggcggagagg ggcgggaagc gcaacaggc | 180 |
| acctggagga agccccatac tgacctcctc atgctgctga cacaggcagg atggcattga | 240 |
| actcagggtc acctccagga atcggacctt gctatgagaa ccacgggtat cagtctgagc | 300 |
| acatctgtcc tccagagacca ccagtggctc ccaatggcta caacttgtat ccagcccagt | 360 |
| actacccatc tccagtgcct cagtatgctc cgaggattac aacgcaagcc tcaacatctg | 420 |
| tcatccacac acatcccaag tcctcaggag cactgtgcac ctcaaagtct aagaaatcgc | 480 |
| tgtgttagc cctcgccctg ggcactgtcc tcacgggagc tgctgtggct gctgtcttgc | 540 |
| tttggaggtt ctgggacagc aactgttcta cgtctgagat ggagtgtggg tcttcaggca | 600 |
| catgcatcag ctcttctctc tggtgtgacg gggtagcaca ttgtcccaac ggagaagatg | 660 |
| agaaccgtt tgttcgtctc tacggacaaa gcttcatcct ccaggtttac tcatactcaga | 720 |
| ggaaagcctg gtatcccgta tgccaggatg attggagtga gagctacggg agagcagcat | 780 |
| gtaaagacat gggataacaag aacaattttt attctagcca agggatacca gaccagagcg | 840 |
| ggcaacgag ctttatgaag ctgaatgtga gctcaggcaa cggtgacctc tataaaaaac | 900 |
| tctaccacag tgactcatgt tcatacccgca tggtggtttc tttgcgtgt atagaatgcg | 960 |
| gggttcgctc agtgaaacgc cagagcagga ttgtgggtgg attgaatgcc tcaccaggag | 1020 |
| actggccctg gcaggtcagc ctgcacgtcc aaggagtcca cgtctgcggg ggctccatca | 1080 |
| tcaccccccga gtggattgtg acggccgccc actgtgtgga agaaccctc agcagcccgaa | 1140 |
| ggtactggac ggcatttgcg ggaattctga gacagtctct catgttctat ggaagtagac | 1200 |
| accaggtaga aaaagtaatt tcccatccaa attacgactc taagaccaag aataacgaca | 1260 |

ES 2 886 958 T3

| | |
|--|------|
| ttgctctcat gaagctgcag acaccttgg ctttaatga tctagtgaag ccagtgtgtc | 1320 |
| tgccgaaccc aggcatgatg ctagacccatg accaggaatg ctggatttcg ggggggggg | 1380 |
| ccacctatga gaaagggaag acctcgacg tggtaatgc tgccatggta cccttgatcg | 1440 |
| agccctccaa atgtataatg aaatacatat acaacaacct aatcacacca gccatgatct | 1500 |
| gtgccggctt cctccagggg tctgtcgact cttgccaggg agacagtgga gggccgctgg | 1560 |
| ttactttgaa gaatgggatc tggggctga ttggggacac gagctgggc tcgggctgtg | 1620 |
| ccaaggcact cagacctgga gtatacggg acgtgacggt atttacagat tggatctacc | 1680 |
| agcaaatgag ggcgaacagc taatccacgt ggcttgcac cagacttcct ttgtctcaa | 1740 |
| caaccttctg caagaaaacc aagggcctga attttaactt cctgtgcaca atgtacctt | 1800 |
| ttagatgatt cgaagggcct ttcactttt taaaacagtg acttgttga ctgtgctccc | 1860 |
| tggtcctgtg agggcttcag tgccccaccc ctgggccact tctgcagctc ccaccagaat | 1920 |
| ggatgaccag attctgttg gtttggcac ataggccaa aggagagga gggtggcact | 1980 |
| ctcatgttgg aacttctttt gggctcatgc tcaggcctt tttggatcac taaggactat | 2040 |
| gacctctgag taacctgtatg acctgagaaa gagtaaggag gccaggcagg gccttggcc | 2100 |
| caggaacagg taccttgaga gtgagagcta cccattgcct gtggctaaa tctgctgtgc | 2160 |
| aggttggct ggtcatactg tcatgatttc attaacagcc tgggtgaaca tggctggag | 2220 |
| taaaggcctt gcttcctgc atgttgacat gacggccctt tccaagggtg atggaggctt | 2280 |
| tcccaagcta agggcctagg cagatctctc agagcaagaa gctaatgccg gcatgtccct | 2340 |
| tgggtgagct ctacatggtg ttattcagtc tggttcttg ctccccacta ctgtttctct | 2400 |
| cagcctctca gagcctgaaa cttacctctt agcttggct acaggcatgg cctagtcac | 2460 |
| gatggagcct gtatagctca gctaataaaa tggaggctca ggtccatcag aatcaggac | 2520 |
| ttgtgatttc agtcacctt cttctgggtt gtgtttcttc tcttactacc tcactgcacc | 2580 |
| tggacactag agtggatgaa tgtctggagt tcacctgcat ttggactgtg tgattgtgcc | 2640 |
| tcagacacta gacctcttcc agatggtag gttgttctgt agactggcaa ttagattaga | 2700 |
| agttccttagc ttcaaaaaaa gatgaaagag aggagatcat tgtcttctgt cttcttctgg | 2760 |
| ccctgggttt ataccagaa agccatgcca gaattacca atatgaagta tgaatgtctt | 2820 |
| acccacggtg aggctctgcc tccttctctc tgcctggttc ttcagaaggc agtgaatggg | 2880 |
| tcataactgg gactccatct ttgctggga aagtctccca cctaggaat ggttaccact | 2940 |
| ccatgtaaag aaaactccct catgcgtcct ctgggacctt cttagatgct gtaaggtacc | 3000 |
| tacatacaga ctaaatgtgc aagcacctt aagtgtgaga acctgtcccc tccttagctc | 3060 |
| tccttgcctt tgctgttggt tggttatttc ctgctttgtg tctgttctga gctgtgagat | 3120 |
| tccactgtga aatatatgaa taaagtataat aattttttt aaaaaaaaaaaaaaaa | 3175 |

ES 2 886 958 T3

<210> 2
<211> 490
<212> PRT
<213> *Mus musculus*

5

<400> 2

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Ala | Leu | Asn | Ser | Gly | Ser | Pro | Pro | Gly | Ile | Gly | Pro | Cys | Tyr | Glu |
| 1 | | | | | 5 | | | | 10 | | | | | 15 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | His | Gly | Tyr | Gln | Ser | Glu | His | Ile | Cys | Pro | Pro | Arg | Pro | Pro | Val |
| | | | | | 20 | | | 25 | | | | 30 | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Pro | Asn | Gly | Tyr | Asn | Leu | Tyr | Pro | Ala | Gln | Tyr | Tyr | Pro | Ser | Pro |
| | | | | | 35 | | 40 | | | | 45 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Val | Pro | Gln | Tyr | Ala | Pro | Arg | Ile | Thr | Thr | Gln | Ala | Ser | Thr | Ser | Val |
| | | | | | 50 | | 55 | | | 60 | | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ile | His | Thr | His | Pro | Lys | Ser | Ser | Gly | Ala | Leu | Cys | Thr | Ser | Lys | Ser |
| | | | | | 65 | | 70 | | | 75 | | | 80 | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Lys | Lys | Ser | Leu | Cys | Leu | Ala | Leu | Ala | Leu | Gly | Thr | Val | Leu | Thr | Gly |
| | | | | | 85 | | | 90 | | | 95 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ala | Ala | Val | Ala | Ala | Val | Leu | Leu | Trp | Arg | Phe | Trp | Asp | Ser | Asn | Cys |
| | | | | | 100 | | | 105 | | | 110 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Thr | Ser | Glu | Met | Glu | Cys | Gly | Ser | Ser | Gly | Thr | Cys | Ile | Ser | Ser |
| | | | | | 115 | | | 120 | | | 125 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Leu | Trp | Cys | Asp | Gly | Val | Ala | His | Cys | Pro | Asn | Gly | Glu | Asp | Glu |
| | | | | | 130 | | | 135 | | | 140 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Arg | Cys | Val | Arg | Leu | Tyr | Gly | Gln | Ser | Phe | Ile | Leu | Gln | Val | Tyr |
| | | | | | 145 | | | 150 | | | 155 | | | 160 | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Ser | Ser | Gln | Arg | Lys | Ala | Trp | Tyr | Pro | Val | Cys | Gln | Asp | Asp | Trp | Ser |
| | | | | | 165 | | | 170 | | | 175 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Glu | Ser | Tyr | Gly | Arg | Ala | Ala | Cys | Lys | Asp | Met | Gly | Tyr | Lys | Asn | Asn |
| | | | | | 180 | | | 185 | | | 190 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Phe | Tyr | Ser | Ser | Gln | Gly | Ile | Pro | Asp | Gln | Ser | Gly | Ala | Thr | Ser | Phe |
| | | | | | 195 | | | 200 | | | 205 | | | | |

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Lys | Leu | Asn | Val | Ser | Ser | Gly | Asn | Val | Asp | Leu | Tyr | Lys | Lys | Leu |
| | | | | | 210 | | | 215 | | | 220 | | | | |

ES 2 886 958 T3

Tyr His Ser Asp Ser Cys Ser Ser Arg Met Val Val Ser Leu Arg Cys
225 230 235 240

Ile Glu Cys Gly Val Arg Ser Val Lys Arg Gln Ser Arg Ile Val Gly
245 250 255

Gly Leu Asn Ala Ser Pro Gly Asp Trp Pro Trp Gln Val Ser Leu His
260 265 270

Val Gln Gly Val His Val Cys Gly Gly Ser Ile Ile Thr Pro Glu Trp
275 280 285

Ile Val Thr Ala Ala His Cys Val Glu Glu Pro Leu Ser Ser Pro Arg
290 295 300

Tyr Trp Thr Ala Phe Ala Gly Ile Leu Arg Gln Ser Leu Met Phe Tyr
305 310 315 320

Gly Ser Arg His Gln Val Glu Lys Val Ile Ser His Pro Asn Tyr Asp
325 330 335

Ser Lys Thr Lys Asn Asn Asp Ile Ala Leu Met Lys Leu Gln Thr Pro
340 345 350

Leu Ala Phe Asn Asp Leu Val Lys Pro Val Cys Leu Pro Asn Pro Gly
355 360 365

Met Met Leu Asp Leu Asp Gln Glu Cys Trp Ile Ser Gly Trp Gly Ala
370 375 380

Thr Tyr Glu Lys Gly Lys Thr Ser Asp Val Leu Asn Ala Ala Met Val
385 390 395 400

Pro Leu Ile Glu Pro Ser Lys Cys Asn Ser Lys Tyr Ile Tyr Asn Asn
405 410 415

Leu Ile Thr Pro Ala Met Ile Cys Ala Gly Phe Leu Gln Gly Ser Val
420 425 430

Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Leu Lys Asn
435 440 445

Gly Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly Cys Ala
450 455 460

Lys Ala Leu Arg Pro Gly Val Tyr Gly Asn Val Thr Val Phe Thr Asp

ES 2 886 958 T3

465

470

475

480

Trp Ile Tyr Gln Gln Met Arg Ala Asn Ser
 485 490

<210> 3

<211> 3212

<212> ADN

<213> *Homo sapiens*

<400> 3

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|------------|------|
| gagtaggcgc | gagctaagca | ggaggcggag | gcggaggcgg | agggcgaggg | gcggggagcg | 60 |
| ccgcctggag | cgcggcaggt | catattgaac | attccagata | cctatcatta | ctcgatgctg | 120 |
| ttgataaacag | caagatggct | ttgaactcaag | ggtcaccacc | agctattgga | ccttactatg | 180 |
| aaaaccatgg | ataccaacccg | gaaaacccct | atcccgcaca | gcccaactgtg | gtccccactg | 240 |
| tctacgaggt | gcatccggct | cagtactacc | cgtcccccgt | gccccagtagc | gccccgaggg | 300 |
| tcctgacgca | ggcttccaac | cccgtcgtct | gcacgcagcc | caaatcccc | tccgggacag | 360 |
| tgtgcacctc | aaagactaag | aaagcactgt | gcatcacctt | gaccctgggg | accttcctcg | 420 |
| tgggagctgc | gctggccgct | ggcctactct | ggaagttcat | gggcagcaag | tgctccaact | 480 |
| ctgggataga | gtgcgactcc | tcaggtacct | gcatcaaccc | ctctaactgg | tgtatggcg | 540 |
| tgtcacactg | ccccggcggg | gaggacgaga | atcggtgtgt | tcgcctctac | ggaccaaact | 600 |
| tcatccttca | ggtgtactca | tctcagagga | agtcctggca | ccctgtgtgc | caagacgact | 660 |
| ggaacgagaa | ctacgggcgg | gcggcctgca | gggacatggg | ctataagaat | aattttact | 720 |
| ctagccaagg | aatagtggat | gacagcggat | ccaccagctt | tatgaaactg | aacacaagtg | 780 |
| ccggcaatgt | cgatatctat | aaaaaactgt | accacagtga | tgcctgttct | tcaaaagcag | 840 |
| tggtttcttt | acgctgtata | gcctgcgggg | tcaacttgaa | ctcaagccgc | cagagcagga | 900 |
| ttgtggcgg | cgagagcgcg | ctccccgggg | cctggccctg | gcaggtcagc | ctgcacgtcc | 960 |
| agaacgtcca | cgtgtgcgga | ggctccatca | tcaccccccga | gtggatcgtg | acagccgccc | 1020 |
| actgcgtgga | aaaacctctt | aacaatccat | ggcattggac | ggcatttgcg | gggattttga | 1080 |
| gacaatctt | catgttctat | ggagccggat | accaagtaga | aaaagtgatt | tctcatccaa | 1140 |
| attatgactc | caagaccaag | aacaatgaca | ttgcgtgtat | gaagctgcag | aagcctctga | 1200 |
| ctttcaacga | cctagtgaaa | ccagtgtgtc | tgcccaaccc | aggcatgatg | ctgcagccag | 1260 |
| aacagctctg | ctggatttcc | gggtgggggg | ccaccgagga | gaaagggaaag | acctcagaag | 1320 |
| tgctgaacgc | tgccaagggt | cttctcattg | agacacagag | atgcaacagc | agatatgtct | 1380 |
| atgacaacct | gatcacacca | gccatgatct | gtgccggctt | cctgcagggg | aacgtcgatt | 1440 |
| cttgccaggg | tgacagtgg | gggcctctgg | tcacttcgaa | gaacaatatc | tggtggctga | 1500 |

5

10

ES 2 886 958 T3

| | |
|--|------|
| tagggatac aagctgggt tctggctgtg ccaaagctta cagaccagga gtgtacggga | 1560 |
| atgtgatggt attcacggac tggatttac gacaaatgag ggcagacggc taatccacat | 1620 |
| ggtcttcgtc cttgacgtcg ttttacaaga aaacaatggg gctggtttg cttccccgtg | 1680 |
| catgatttac tcttagagat gattcagagg tcacttcatt tttattaaac agtgaacttg | 1740 |
| tctggctttg gcactctctg ccattctgtg caggctgcag tggctccct gcccagcctg | 1800 |
| ctctccctaa ccccttgcgc gcaagggtg atggccggct ggttgtggc actggcggc | 1860 |
| aagtgtggag gagaggggtg gaggctgccc cattgagatc ttccctgctga gtccttcca | 1920 |
| ggggccaatt ttggatgagc atggagctgt cacctctcag ctgctggatg acttgagatg | 1980 |
| aaaaaggaga gacatggaaa gggagacagc caggtggcac ctgcagcggc tgccctctgg | 2040 |
| ggccacttgg tagtgtcccc agcctacctc tccacaaggg gatttgctg atgggttctt | 2100 |
| agagccttag cagccctgga tggtggccag aaataaaggg accagccctt catgggtgg | 2160 |
| gacgtggtag tcacttgtaa gggaaacaga aacatttttgc ttcttatggg gtgagaatat | 2220 |
| agacagtgcc cttggtgca gggaaagcaat tgaaaaggaa cttgcctga gcactcctgg | 2280 |
| tgcaggtctc cacctgcaca ttgggtgggg ctccctggag ggagactcag cttccctcct | 2340 |
| catcctccct gaccctgctc ctgcaccctt ggagagtgc catgcccctt ggtcctggca | 2400 |
| gggcgccaag tctggcacca tggtggcctc ttcaaggcctg ctgcactg gaaattgagg | 2460 |
| tccatggggg aaatcaagga tgctcagttt aaggtacact gtttccatgt tatgtttcta | 2520 |
| cacattgcta cctcagtgcct cctggaaact tagctttga tgtctccaag tagtccacct | 2580 |
| tcatttaact cttgaaaact gtatcatctt tgccaagtaa gagtggtggc ctatttcagc | 2640 |
| tgctttgaca aaatgactgg ctcctgactt aacgttctat aaatgaatgt gctgaagcaa | 2700 |
| agtgcctcatg gtggcggcga agaagagaaa gatgtgtttt gtttggact ctctgtggc | 2760 |
| ccttccaatg ctgtgggtt ccaaccaggga aagggtccc tttgcatttgc ccaagtgc | 2820 |
| taaccatgag cactactcta ccatggttct gcctcctggc caagcaggct ggtttgc | 2880 |
| aatgaaatga atgattctac agctaggact taaccttgaa atggaaagtc atgcaatccc | 2940 |
| atttgcagga tctgtctgtg cacatgcctc tgcgtggc agcattccca gggaccttgg | 3000 |
| aaacagttgg cactgttaagg tgcttgctcc ccaagacaca tcctaaaagg tggtgtatg | 3060 |
| gtgaaaacgt cttccttctt tattgcctt tcttattttat gtgaacaact gtttgc | 3120 |
| ttttgtatct tttttaact gtaaagttca attgtgaaaa tgaatatcat gcaaataaat | 3180 |
| tatgcaattt tttttcaaa gtaaaaaaaaaaa aa | 3212 |

<210> 4
<211> 492
<212> PRT
<213> *Homo sapiens*

ES 2 886 958 T3

<400> 4

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

ES 2 886 958 T3

Cys Ile Ala Cys Gly Val Asn Leu Asn Ser Ser Arg Gln Ser Arg Ile
245 250 255

Val Gly Gly Glu Ser Ala Leu Pro Gly Ala Trp Pro Trp Gln Val Ser
260 265 270

Leu His Val Gln Asn Val His Val Cys Gly Gly Ser Ile Ile Thr Pro
275 280 285

Glu Trp Ile Val Thr Ala Ala His Cys Val Glu Lys Pro Leu Asn Asn
290 295 300

Pro Trp His Trp Thr Ala Phe Ala Gly Ile Leu Arg Gln Ser Phe Met
305 310 315 320

Phe Tyr Gly Ala Gly Tyr Gln Val Glu Lys Val Ile Ser His Pro Asn
325 330 335

Tyr Asp Ser Lys Thr Lys Asn Asn Asp Ile Ala Leu Met Lys Leu Gln
340 345 350

Lys Pro Leu Thr Phe Asn Asp Leu Val Lys Pro Val Cys Leu Pro Asn
355 360 365

Pro Gly Met Met Leu Gln Pro Glu Gln Leu Cys Trp Ile Ser Gly Trp
370 375 380

Gly Ala Thr Glu Glu Lys Gly Lys Thr Ser Glu Val Leu Asn Ala Ala
385 390 395 400

Lys Val Leu Leu Ile Glu Thr Gln Arg Cys Asn Ser Arg Tyr Val Tyr
405 410 415

Asp Asn Leu Ile Thr Pro Ala Met Ile Cys Ala Gly Phe Leu Gln Gly
420 425 430

Asn Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Ser
435 440 445

Lys Asn Asn Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly
450 455 460

Cys Ala Lys Ala Tyr Arg Pro Gly Val Tyr Gly Asn Val Met Val Phe
465 470 475 480

Thr Asp Trp Ile Tyr Arg Gln Met Arg Ala Asp Gly
485 490

ES 2 886 958 T3

<210> 5
 <211> 27947
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Polinucleótido recombinante
 10 <400> 5

| | | | | | | |
|-------------|-------------|------------|-------------|-------------|-------------|------|
| gcagagtcta | agaaatcgct | gtgttagcc | ctcgccctgg | gcactgtct | cacgggagct | 60 |
| gctgtggctg | ctgtcttgct | ttggaagttc | agtaagtgc | gggagcctcg | atcccaccat | 120 |
| gtgctcctgc | agtccccagt | gctctgagcc | agaccctgtct | ctctgggcta | ttgagacctc | 180 |
| tggaggccct | ccgtgagggtt | cctcttttac | ataacgaggc | tgtctcttctt | cccttcttctt | 240 |
| gttagctat | gagattgaca | catcatgggg | aaagcattta | aatgtaccc | agtgcgtttgg | 300 |
| ggtgcttgggt | gccacccagc | actgtgagca | caggttcttc | taccttgggg | ccacacccag | 360 |
| ttacctgtat | ctcaactgcac | agcagtggt | gttggggacc | aggcccaccc | ctccatgtcc | 420 |
| cacccctgc | aactgcagcc | tgagccttcc | catcagcctg | gggtggtgca | gaccatgtg | 480 |
| ccatttgtga | tccttcaagt | tacctgtgt | gcagagagga | cgtgtgagtg | ccgtccaaac | 540 |
| ccaaacactg | agagggtcct | tcccattgcc | cccacggaag | taaggtgccc | cagtgctaatt | 600 |
| tccacttata | cttgctggtg | gcaaggacac | tttctcttct | tattaaagtg | ggggattggc | 660 |
| tgggtgaggt | ggctcacgccc | tgttatccca | gcactttaag | aggccaaggc | aggtggacca | 720 |
| cctgaggtca | ggagtttgag | accacaagcc | tggccaacat | gttgaactc | catctctact | 780 |
| aaaaatacaa | aaattagtca | ggcgtggtgg | cgtgcacctg | taatcccagc | tacttaggag | 840 |
| gctggggcag | gaggatca | tgaaccagg | agttggaggt | tgcagtgc | caagattgt | 900 |
| cccctgcact | ccagcctggg | tgacagaatg | agacttcatc | tcaaaaacaa | aacaaaacaa | 960 |
| aacacagtgg | ggccaggagt | tggaggctgc | agcgagctac | agtaatgcca | cggtgttccct | 1020 |
| cactccatga | ggctcattgc | gtttctcagc | ctgaaggca | cctctttctt | gttttctctg | 1080 |
| caagtggca | gcaagtgc | caactctggg | atagagtgc | actcctcagg | tacctgcac | 1140 |
| aaccctcta | actggtgtga | tggcgtgtca | cactgcccc | gcggggagga | cgagaatcg | 1200 |
| tgtggtgagt | cagccttgac | cttgggaagg | gactcctctg | ctcaccttgg | agacagcagc | 1260 |
| cgggtccagg | ggccttggg | tgactggcc | tggcgtgcgt | ccagtacgt | gacacatgtat | 1320 |
| gtcattgaat | ccctgctcca | ggctgagccc | tggggctcag | agaggttgc | tttccggccc | 1380 |
| aacctcaccc | gcaggtggg | agatgacagg | gccaccgagg | actgtgtcat | tggaaccaca | 1440 |
| cgtgctctga | actgccacag | gaagtcagtt | aagatgagca | aactgtttat | aaagttggag | 1500 |
| atgcaggcta | ggaacgggtgg | ctcatgcctg | taatcccagc | actttggag | gccgaggcag | 1560 |
| atggatcacc | tgaggtcagg | agtttgagac | cagcctgacc | aatatggtga | aaccttatct | 1620 |

ES 2 886 958 T3

| | |
|---|------|
| ccactaaaaa tacaaaaatt agccaagcgc ggtggcggtt gcctgttaatt ccagctattc | 1680 |
| aggaggctga ggcaggagaa tcacttgaac ctgggaggcg gaggttgcag tgagctgaga | 1740 |
| tcacgccact gcattccagc ctggagaca gagctggctc aaaaataaa ttaattaatt | 1800 |
| aaaaacaaaa ttggagatgc actatgttat tttcaaaaca agctgcctt aaagatctat | 1860 |
| ctgttgtcac agggtggct catctgttcc attttatTTT ctgtggttt tctatTTT | 1920 |
| catttaatg aactaggaag cattgctcct atttatggca taccacatga tgTTTggata | 1980 |
| cgtgtatgcc tgtggcatgg ctaagtcaag ctagaacatg ggccttacct catatacgtg | 2040 |
| tcttattaag aacacataaa acctactttt gttagtgattt tcaaatatgc aacatatagt | 2100 |
| ttattaactg cagtcactat gatgtacaat agattgctcg aacttattcc tcctgtctaa | 2160 |
| ctaagatttt gtgacctctg accaacatct ccccagtgtt gtcaccccccc gccccagcc | 2220 |
| tctgatagct gccttctac tctctgcttc tgtgagttt atgttatac attccacatg | 2280 |
| taagtggcct catgcagtgt ttctgtctct gtgtctggct tgTTTcaTTT gCGTAATGTC | 2340 |
| ctccagcttc atctatgttgg ttggaaatga caggatttcc ttcttcttgg tggctgaata | 2400 |
| gtattgcctt gtgcataaac accacatTTT cttatccct tcattcactg atggactctt | 2460 |
| aggttgcgtt catgtcttgg ctgttgaa aatgcccga gtgaggcgtgg gcgtgcaggt | 2520 |
| ccctcttcaa cacacggatt tccttcctt tgatataaa cccagcagtg agattgctgg | 2580 |
| atcacatggc agttctgttt ctcacccccc gaggaaactc catactgttt tccataatgg | 2640 |
| ctgttagcaac ttccactccc acccccacgg tgcaaaagtct ccatttctct tctacaacct | 2700 |
| caccaactcc tggatTTTcat tgcattttt atttgcattt ttcatgaatt tttgagggtg | 2760 |
| tacctcattt tggTTTcat ttgcattttt atttgcattt ttcatgaatt tttgagggtg | 2820 |
| atttcaagg tagtttagtga ctcgaacagg gaaacgatcc tgagtatgag ggTTGTGCTA | 2880 |
| atcatcccccc tcctgccagc tgctacggat atggggctct gcagatggca gggagctggc | 2940 |
| tcgtttctct ttaagagctg cttttactt ttcttcctct tcctttaaaa cttatTTTCT | 3000 |
| ggccggacgc agtggctcat gcctgttaatc ccagcacttt gggaggccga ggtggggcgg | 3060 |
| tcacgaggc aggaattcca gaccagcctg gccaacatgg tgaaaccccg tctctactaa | 3120 |
| aaataaaaaa attagccaga cgtgggttg cgggcctata gtcccagcta ctcgggaggc | 3180 |
| tgaggcagga gaatcacttgg aacctggag gagggggttg cagtgagccg agattgcgc | 3240 |
| actgcactcc agcctggcg acagagccag actccatctc aaaaacaaaa aaaaagttat | 3300 |
| ttcccaagca cagccatgtt ttccaggctt gtggatcagc gttgggttg gtgtgtgctc | 3360 |
| tcatacttta gttccagctt agcacactct gacatgttta cactagaacc atttgcgg | 3420 |
| tctagaaaaa gaaatttcag aattgttagag tcagaggact taccagaaat ctcttaggtt | 3480 |

| | |
|--|------|
| gttctcctcc cctccctcaa gtgcagtcct aacctcctgg agttttctgt agaaaccaca | 3540 |
| agcctcagag ctggccgaga attctagcca aagattttc catgccaaag taatcccccc | 3600 |
| tctcctaagg gccatccttg gtggggactg gttcctgtt aagccctcgc tgtcagtccct | 3660 |
| ggctgtggaa tttcctggtg aggagcactg gcccggtggag ctccggccctc gtgcggcct | 3720 |
| tgagcaggcc caagtgttcc gtgttcttga taccttcct ccagcacagt ctgtttccc | 3780 |
| agaaaaaggt ttgcacttga aaatgatgca tttgctgatt aaacatagtt ctttgcttt | 3840 |
| atttggtttc taaaataaaag tgggagttt tgagatttag taacgtgagg ttaagatagc | 3900 |
| acgtggaatg gctttttctt ttcttttat ttttttttt ttttcctgg agacagggtt | 3960 |
| tcactctgtt gcccaggctg gagtgccagag gcatgaccat ggctcactgc aacttcgatg | 4020 |
| tcctgggtt aagcgatccc ccagcctcag cccccaagt ggctggact acaggtgctc | 4080 |
| gccaccacac ctggctaatt tttgtatTTT ttgtagaaaa tgggttcat caatgttgc | 4140 |
| cagactggtc tcgaactcct gacctaaggc aattctcctg cctcaggccctc ccagactgct | 4200 |
| gggattacag gcgtgaacta ccacgcctgg cctggaatgg ctttgatgt tctcctatgt | 4260 |
| gcacatgtgg gtgaataaac accaacaag tccttatgtt acctgaagag ttgctctt | 4320 |
| cttaatattt aagtctgttatt tattttaaaat ctttaatagt tgtacactat taaagtatta | 4380 |
| ttaggtcaaa atcaaggaag tacaaaaggg tatgctgtga aaaatctt cttccttgct | 4440 |
| ctgcttactt acctaccccg catccccca tacacccag acacacacac acacacacac | 4500 |
| acacacacac acacacgcac cactccata catgcccacc tgTTTaccag ccaatcacat | 4560 |
| ttcttgggc aactcatctg agttgcttct cttccagag agttttgca taaaagaagca | 4620 |
| caggtatTTT tgcgttacca tgaccattt tcccaagtggt tccttagccag ttgactctcc | 4680 |
| tgcactggat accatcctgg acagcattcc ttagggaaat gagccccctg tttttccca | 4740 |
| ccatggcaca gttggcctt tgcattggcaccattt gcccctgtct ctcttggtg | 4800 |
| gaccttaagg ttttctccat cttttgctg taacacacac tgctccaagt gtgtgagcat | 4860 |
| atcagtagga aacgcttcca ggagttagaac tgcttagtca gagggcgtgt ggatctgtaa | 4920 |
| cctgacagac ctagaccggc ttcaagtgg ttttatccag tttccatatt gattattcat | 4980 |
| ataaaaaggaa acagacaaac ataacgctgt gcatgtattc tctcttagac cagaacaggc | 5040 |
| atagggtgca cttttatTTT gtccatTTG tagagtagaa attgttttg ctgaaatgaa | 5100 |
| caccttagga tgctgaagaa tatgaccgt cccatggaaa acattaaaa atgtgtgttag | 5160 |
| cgctttcttc ccaagggtgt gtgtgcgcac attttAACAC taattcactt tctacttccg | 5220 |
| ttgctatcct ttctgtgagt ctttctcaga atctcagaaa agaaactaaa ttgttcaactc | 5280 |
| tagttatcaa tgctgtactc tataccctgga atttgctaaa agggcagatt ttaagtattc | 5340 |
| tcaccacaga aaagagaaaa gaaaatggta attatgtgac gtggggaca tggtaacttag | 5400 |

ES 2 886 958 T3

| | |
|---|------|
| ctttattatg gtgagcattt cacagcggat atccagtcat cacgctgtac acattaaaca | 5460 |
| tgtacaattg ggttttttg agacaaggtc tccttctgtc acccagtctg gagtgcagtg | 5520 |
| gctcagtcat ggctcattgc agcctcgacc tcctgggctc aatccatcct tccccctca | 5580 |
| cctcctgaaa agctggggcc acaggcatgt accatcatgc caggctaattg catatatatt | 5640 |
| tatattttt ggtggagatg gggttggtct cgaactctgg gctcaagtga tcctcccggc | 5700 |
| ttgcccttcc aaagtgtga gattacaggc atgaaccaca gcaccaggcc tacatgtaaa | 5760 |
| atttttattt gtcaactata ctgtacaaa gctgagaaaaaaaatcctaa tattnnnnnnnn | 5820 |
| aaaaaaaaaaa aggacttagct tgagaccttt tccagctctc tggcttatca gctgccgtct | 5880 |
| cttccgggtg cagatagctg gaagggaaag aaaatcccta aaattaccca caagccaaga | 5940 |
| atgaagtgtc tccctttgag ccacagtggc agttttgttt ttaatcatag aagtgtattt | 6000 |
| tgagccgggt gtgctggctc acgcctgtaa tccccgcact ttgggaggcc gaggtggggg | 6060 |
| gcggagggggg tggggatcgc ctgaggtcag gagttcgaga ccagcctgac caacatggag | 6120 |
| aaaccccgtc tctactaaaa atacaaaatt agccggcgtg gtggcgtatg cctgtaatcc | 6180 |
| cagctactca tgaggctgag tcaggagaat ctcttgaacc caggaggtgg aggttgcgt | 6240 |
| gagctgagat catgccattt cactccagcc tggaaacaag aaaaaaaaaag aagaagaaga | 6300 |
| agaagtgtat tcatttcagt tactttaaa aaagtgaaca gactttatata tttagagcgg | 6360 |
| tttaggttt acagaaaatg aaacagacag ggcagcgcgc tccttgcact cctccccagc | 6420 |
| acacagttgc cctgttatga acatcccaca tcagtgtgt gcgttcatta acaccgtatga | 6480 |
| acctgatgca tacattatga tgaactgaag tcctggactt caccctttctt cttgtacagt | 6540 |
| tctgtggat ttgacaaaatg cataatgctg tacagccaca atgatagtat cgtccagagt | 6600 |
| agttctcctg cctaaaaacc tttttgcgtg cacctgtttc tctctccccca ctcacccca | 6660 |
| ctatctgatc ttcttagtgc ctccgaagtt ttgggtttt caggatgttg tagcgttgaa | 6720 |
| atcatggagt atgtgcctt caccacatac accttccttc actttgttg cttcccttac | 6780 |
| ttagtaatat gcattcaagt ttctccatg cctttcatg gcttgatagc tcatttctt | 6840 |
| ttagcaccaa ataatattcc gttgtccaga tgttagcacaa tgtttatcca ttcatgtaac | 6900 |
| ctgtgaccga ctcacagata ggttgtggaa tcactcacca cagaggcatt agacaataat | 6960 |
| cagacccaaag tcatttcatg gggaaacaag cccacaggta ccagactgtc cagtgagtca | 7020 |
| gggccactcg taggaagtaa gaagagaggc tagagcatag ccaggtcctc actttataact | 7080 |
| ttaagcccat gtgtatttct cccaaaccac acagcattgt ttccatgttt tcagctttgc | 7140 |
| atgaataacg tgataacttga acgcatcatt tatcacttgc tctcttccc acagcgctgt | 7200 |
| tttcaagctt ctccctgttc atgatgtct gcttaaccct taagctgtcat gggattctgt | 7260 |

| | |
|--|------|
| tctgtgaata cgcccacccc atgtattatc ctgcccagca aaaagtcccc | 7320 |
| atggtgttta cctctaggga gggagagaag agattggaa tagggagcga cttcaacgg | 7380 |
| gtttgtaatg tttgtttct ttaaataaaa gagctgagat catttcagca gaatgttgat | 7440 |
| ttagagtctc ctggacaatt tggtgctcaa agtgctctct taaagagcac ttaaaaaaaaa | 7500 |
| aaaacccttt atcttattat ttatttattt atttatttag acggagttt gctctgtcac | 7560 |
| ccaggctgga gtggagtgggt gtgatctcag ctcactgcaa ccttacaccc ctgggttcaa | 7620 |
| gcaattcccc tgcctcagcc tcccaagtag gtgggattac agatgcgtgc caccacactt | 7680 |
| ggctaatttt tgcatttttag tagagatcggtttctccatgttggccaggc tgatctcaa | 7740 |
| cgcctgacct caggtgatct gcccgccttg gcctccaaa gtgctggtat tacaggcgtg | 7800 |
| agctaccatg cctggcttat cttatatatt ttaaaaaaaaa gcttatttagatcttataattt | 7860 |
| tgtaccataa aattcaagta tataattcag tgctttataa tataaaacat atatatgaaa | 7920 |
| tagcttattt agatataatt ttttatataa aacagcttataatgtaa tgatgtacc | 7980 |
| ataaaaattta aatatataat tcactggctt ttatataattc acgaatatgt gcaactatca | 8040 |
| ccacagtcaa ttttagcata ttttcatcag ctcataaaga aaccccaagc ctttgaacta | 8100 |
| tcacccata tccctcctcc cagccgtcc ctcctactca taagcaacca ctaatctact | 8160 |
| tagtgcata agatttccta ctctaggcat tccatgttagtggcatatgttgcata | 8220 |
| gctcacacaa tataagtggc attccatgtg agtcggctca tgcagtatgt ccggctcctt | 8280 |
| tcactgagca taaggcttc agcactcatc caggttgcag cctgtgtctg aatttcattt | 8340 |
| cctcttctgg ctgaatcgta ttccattgtg tatcttggac atatcctatt ctgctcaccc | 8400 |
| agccgttgggt gggcggttgg agtgtttcg ccttcagct gttttaaagag ggttgcagt | 8460 |
| aacatttgcata caagtttgg acccaatgcc tgtttcaat tctcttgcgtt agagagcact | 8520 |
| tttttagcaga aaaagaatag atttgcgttcc tccctttgtg tgcggctagt gccttgagaa | 8580 |
| gagtgaactg tgctgccacc tcggagccg tggagagcgc ggggcttggg tagcagctag | 8640 |
| gacgatacaa gttggacaa ggccaggtgc aatggctcac gcctgttaatt ccaacacttt | 8700 |
| gggagaccga ggcaggggga tcacctgagg tcaggagttc aagaccagcc tggccaacat | 8760 |
| ggtgaaaccc catctctaataa aaaaacagaaaa aattaactgg acgggggtggg ggacgcctgt | 8820 |
| aatcccagct actcgggagg ctgaggcagg agaatcactt gaacctggga ggcggaggct | 8880 |
| gcagtgcgtg gagatcagac cactgcactt cagcctaggt gacagagcga gactccgtct | 8940 |
| caaaaaaaaaaag aaagaaactc atggataatc ctccctctcg tgcagttcgc | 9000 |
| ctctacggac caaacttcat cttcaggtg tactcatctc agaggaagtc ctggcaccc | 9060 |
| gtgtgccaag acgactggaa cgagaactac gggcgccgg cctgcaggaa catggctat | 9120 |
| aagtgagtat gggcagcac cggccagtg acagtaacag acagcagaaaa cacgagaaga | 9180 |

ES 2 886 958 T3

| | | | | | |
|-----------------------|------------|------------|------------|-------------------|-------|
| ccctctctct gcctccctgt | gaaagcaccg | gcacatgagt | gctggggaca | attgtcacct | 9240 |
| tccaaaagct gagccctata | accagcaggt | ggaatttgc | ctgctagggc | tgtgccagc | 9300 |
| acacagacct tggctca | ctgac | ccac | cttggc | cctct atagactc | 9360 |
| ggttgctcg | gagt | ggcc | tgctgtgg | tc atctgg | 9420 |
| agg | tttcca | agg | tgcc | tcag ccagggtc | 9480 |
| ttggaccaga | acag | ctgg | aggagg | ttg ggctggg | 9540 |
| tgccagttc | actt | catt | ggag | ccatgg aagc | 9600 |
| aatcagactg | atact | aaaa | aatgt | cattc ctgctcg | 9660 |
| gcttaacagg | gtgt | gtgtcg | cctg | acagtg attcctaa | 9720 |
| atttaccaggc | actg | cctggg | gagat | gcggc agcc | 9780 |
| ggatgctact | gcc | actttgt | tttccat | gggg aggtcccc | 9840 |
| agggtattca | atat | gccc | agt | tttcagaatg aattac | 9900 |
| tagctagtca | ctgt | gacatc | cctaa | acagc aggggt | 9960 |
| tcacagtcca | gga | acctgg | ggaatt | tgata actggggc | 10020 |
| attagattaa | atgt | gtgttat | gattata | caa tcctatgt | 10080 |
| taacctggat | aaga | aaacacg | acca | atgaaag gaatttt | 10140 |
| atcgaaaaat | cgtt | acaata | ttct | tagcact tggtt | 10200 |
| tgctaagg | tttcc | cctt | attc | taatg tcgtatg | 10260 |
| gatttgtgtt | tgctt | atgccc | ttaacc | attatgatc ttcc | 10320 |
| attttactc | tagc | caagg | atgtgg | atg acagcgg | 10380 |
| acacaagtgc | cgg | caatg | tcgat | ccaccat | 10440 |
| cttctgaaa | aat | tttg | aaaatca | taggac | 10500 |
| agtctgattt | ctct | gagct | caata | catggg | 10560 |
| agagtctgac | cagg | cctgg | tact | tttgcatt | 10620 |
| aggtttctct | gctat | tttca | acc | tttgc | 10680 |
| acttgcaaca | tcg | acatt | actctat | tttgc | 10740 |
| gtacctattt | tgtc | agg | tttgc | tttgc | 10800 |
| agggtgaccg | gctc | attc | tttgcc | tttgc | 10860 |
| gctccattcc | aggc | agctgg | tttgcc | tttgc | 10920 |
| ccagagggtg | ctct | gtgagg | tttgcc | tttgc | 10980 |
| tttccagggg | gtaga | agtgg | tttgc | tttgc | 11040 |

ES 2 886 958 T3

| | |
|---|-------|
| gggagggcag aagaggggat ggacttaggc tcttcctttt gcttgacaca tattnaggat | 11100 |
| gtttgcagac ttgctatgtat tgttgtgtt atgtgtttt tggatgtgaaa gatacacagt | 11160 |
| gtcccttgcc catgagctct ccttgacccca caggtccccca gggcttatgc ctgggtgtcta | 11220 |
| ggcatcacct ccctgcctgc caggtgccag gtgctgcatt tcgggggagg atgaactaat | 11280 |
| caccccgccg cacctttctt ctgagtgaaa gcctggggca ggtttgcatt cctggaggcc | 11340 |
| gctgggtggag gggcttgaaa gcctgacttc cactgcagcc tgctgtccctg gggaaatgtgg | 11400 |
| cagggcaagc ccagtggggaa gggctgtgca cggccaggtg cacccatcaa aacagcaggg | 11460 |
| ctgcggtttg tccctgtgga gaagctaaac acagctgcct gggcactttt taaatgctga | 11520 |
| gtggttcttt gtctttctgg gttacacacg gaatcaggga gccaagtcca gccgggcagg | 11580 |
| gacggggggaa ggggaggagg tgctgcccgc cttggcaag agccttgaaa actcacaagg | 11640 |
| aggctggagg gcttggaaaga aagaagagaa ggcattgtc tggtaggctc tattctatct | 11700 |
| cggtggttgtt ggtggggggaa ggcgcacttc ttttcccttt tctgtgcagc agttggccctt | 11760 |
| tgtgcctga gttcttggtc tgttttctgt cgggcttctg tgaataacca catgtgcctt | 11820 |
| ggcgctgtga ccacacaggg cttccctac cgaccttagg attcttagga aatgtcttct | 11880 |
| cttaaagggg acatgtcttc acttggccgt gtcagtgcctt cagagccaga gtccacctgg | 11940 |
| aatgcacctg tagtcactga gaaccgggg ggtgtgcatt agtaagaagg tgtcaggaag | 12000 |
| gacctattat tgtagggcct gggctcctgc aaggtggttt ggggggtgggtt ggaggaagca | 12060 |
| gagatttgct ctggatttggaa tgctgtcagg aagcagggggtt aattctgtga ggctgctttt | 12120 |
| ttatTTTTT tctaggagga ggttggaaatg aggctaggct aaagctgtga ttggtaaaga | 12180 |
| aacgtccgtc gctcaagtta gccaggacag gaggagacat cagatgtga ttttgggtt | 12240 |
| gtgagcacaa gttccctgtt ctgtctgttc agacatcatt tcggaggagg ctccctgtgt | 12300 |
| cttgcggcat ctcaggcatg gagggcccta gtccgatatt gacgctcagt gaaataattc | 12360 |
| aggttccgca gacacacgg cccagctatc agggcgggccc agctctgcat gccaggggcc | 12420 |
| gcgtcttccc ttctcagcat agcctggaa attcactgca ggacaaaatg catcagttac | 12480 |
| ttcctcttca tccataacct gggatgtttt actccaaat gagtaactct tacgtttctt | 12540 |
| ctaattcttag gaaaaactatt gtttatattt cttcaacac tacaattta aagcagttat | 12600 |
| aggagccag agtttccaa atggcttcct taaaaattttt aagatgattt taaattccaa | 12660 |
| gaggaaaaac aaaactagca ttattgtata ctaccctca caaccgtcctt aggagctgg | 12720 |
| acaattttaa gagaggttaa gtaacttgcc caaggtcaca ctgtggggat gtgagccgcg | 12780 |
| taccttggct cagtgtctgg tctttgccac tgccctata tggattttact taccttattt | 12840 |
| gagttgtaac tagcagaccc ttctatgtct cagaagacag gagagggaaac atcggaaagaa | 12900 |
| atgactgatt tctaaggcatg tgagaggcag gtgactccgc actatcgtga ccagaatttc | 12960 |

ES 2 886 958 T3

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| ccctgttctt | tttgcagtga | tgcctgttct | tcaaaagcag | tggtttctt | acgctgtata | 13020 |
| ggtaagttca | tctggagtcc | ccctttgat | acttctaact | aggaaaagct | ctctactttc | 13080 |
| agaacagtac | tccctgtgtc | tctggggcg | tgggagggaa | gaaggtgggg | tcacgggttg | 13140 |
| gaatgtgcc | agcggcgtct | cgctcttcc | aaggagctcc | tggtttagat | ttccatggcc | 13200 |
| tgttagacacc | ttcagcctt | ggtccaaggg | acaccccctg | agatcaggca | cgctcaagaa | 13260 |
| gctgacaag | ccctacactt | tatgccaccc | atgagctgga | ggcccgccag | gtctcttct | 13320 |
| ccagaaagca | aaggggggtg | gcgttagtga | gccctggcag | ccacctaacf | tggacttgga | 13380 |
| gcatctgcgg | ggctgtggc | cagcaccacc | gtgtggccac | caggtgctca | tcagccagtg | 13440 |
| ggacccggga | ggagggacaa | gaccagagaa | caacagtgt | cttgcctt | ctctctgaa | 13500 |
| ttttggacgg | tggcttagac | ttgggtgtcc | ccatctctgt | gttagagtg | cttacagttt | 13560 |
| ccaaactgtt | tgcaaatgtg | gaagccaccc | tccctctcct | ctggatggc | ccagtgtgt | 13620 |
| cgtggggccg | tggcctgag | ctcagctttt | catttgaaga | ggtggaagga | gctgacacccg | 13680 |
| tcccattccg | gcagggctgg | ctcaggtt | ctttaggtcc | tgagtgggg | tccagcacag | 13740 |
| ccccaaagggt | gcgtggcacc | cgccctgccc | tctgcccatt | cactcatctc | ctgggtggaga | 13800 |
| agacactcac | acacaggaag | cagggaaaggc | agcagaccc | actcaccctt | cacccctca | 13860 |
| ctcaccctt | actcaccctt | tcaacctctc | attcaccacc | cacccctctg | ccccctcact | 13920 |
| cacccctca | ctccctcaac | cctcactcac | ctcctcactc | cctcaaccct | cactcacctc | 13980 |
| ctcaccttct | cactctcccc | ctcatccctc | cctcaccctt | ccccgtcacc | tcctcactca | 14040 |
| cctcctcacc | ccctcactca | cccttcaccc | cctcactcac | cacctcacct | cctcactcac | 14100 |
| ccccctactca | acccttcatt | cacccctcac | cccctcactc | accctgcac | cccctcactc | 14160 |
| acccttcat | ccactcaccc | acctgctcac | ctcctcactc | aaccctcac | cccctcacta | 14220 |
| atccctcact | ccctcaccctt | ctcacgccc | cactcacacc | ttcaccttct | cactcacccc | 14280 |
| ctcaccctt | caacccctta | cttacccctt | cactcatccc | ttcacccctc | actcaccctt | 14340 |
| tctctcaccc | attcaccctt | tcactcatgc | ttcacccctt | tcactcacct | cctcactcac | 14400 |
| accttcaccc | ctcagtcacc | ccctcactca | ccccttcacc | ccctcaatca | tgccttact | 14460 |
| ccctcactca | ccccttcacc | ctctgaatta | ctccctcactc | ccctcactca | ccccctcact | 14520 |
| caccccttca | ccccctcacc | caccacctca | cccacccctc | acccaccccc | tcacctcatt | 14580 |
| acccctcacc | cccctcactc | acccctcacc | ccctcactca | ccacccctacc | cacccctcac | 14640 |
| ccacccctc | actcactccc | tcacccctc | actcaccctt | tcacccctc | actcaccctt | 14700 |
| tcacccaccc | ctcacccacc | ccctcacc | ctcactcacc | ccttcaccc | ctcactcacc | 14760 |
| ccctcactca | ccccttcacc | ccctcactca | ccacccctacc | cacccctcac | ccaccccttc | 14820 |

ES 2 886 958 T3

| | |
|---|-------|
| actcactccc tcacccctc actcaccccc tcacccctc actcaccccc tcatctcctc | 14880 |
| actcaccccc tcacccctc actcacccgc tcacccctc actcaccccc tggccccc | 14940 |
| actcacccct caccctca cccctca caccctcac cccctcgccc cctca | 15000 |
| ccctcgccc cctca caccctcac ccctcaccccc ctacccct cactcatccc | 15060 |
| cactcaccccc ctacccctcactcaccccc ctacccctcactcaccccc | 15120 |
| ccaccaccccc ctca caccctcac ccctcaccccc ctacccctcactcaccccc | 15180 |
| cactcaccccc ctacccctcactcaccccc ctca caccctcac ccctcaccccc | 15240 |
| cactcaccccc ctacccctcactcaccccc ctacccctcactcaccccc | 15300 |
| cactcatgcc ctacccctcactcaccccc ttacccctcacttgctcatccc | 15360 |
| ccctca caccctcacccgtcaatcacc ccccccgtcaatcacc ccctcacctt | 15420 |
| ccctca caccctcacccgtcaatcacc ccccaactcacccctcaccc | 15480 |
| cccaactcaccc accctcaccc ccccaactcacccctcaccc | 15540 |
| tcacccctcactcaccccc tcacccctcacttatcccc tcacccctcaattaccccc | 15600 |
| tcacccctcaattactccc tcatccccc aattacccac tcacccctcactcac | 15660 |
| tcctca caccctcaactcacccctca cccctcaactcacccctcac | 15720 |
| ccctca cttccagccc tgccccccc atttccctt tctttgtgtg agaatctggg | 15780 |
| gtccctgagt ggtgtcagtc cctccaagac tcaaggagtc cccagggct | 15840 |
| tgttatccag aacacccca cctgggtccc gggagaccccc atgggatcac | 15900 |
| aggagtgttc aggaaagtgg tgcttcctgg gtctgggtgg gctggagggg | 15960 |
| catactcccttccatcc ccagcagtgg tgccccctgac ctgtccctgc | 16020 |
| ggagccccc aagtccatcc ccagcagtgg tgccccctgac ctgtccctgc | 16080 |
| acccttggga gggcgcccg ctgggtggct gggcgccctc tgctggctc | 16140 |
| acccttggga gggcgcccg ctgggtggct gggcgccctc tgctggctc | 16200 |
| actgcgttca cgtgtgcgga ggctccatca tcacccca gttggatcgtg | 16260 |
| acagccccc acatgttca cgtgtgcgga ggctccatca tcacccca gttggatcgtg | 16320 |
| actgcgttca cgtgtgcgga ggctccatca tcacccca gttggatcgtg | 16380 |
| acatgttca cgtgtgcgga ggctccatca tcacccca gttggatcgtg | 16440 |
| ggcagtgcacgc ggctgttatgc actgagtgac ccctgcacgc | 16500 |
| aggcacagg gtgcagggtgg aacattctcc acatgttca cgtgacc | 16560 |
| gttcctccatca cgtgaccagg gttccatca gccactgtcc tctgagctct | 16620 |
| gttcctccatca cgtgaccagg gttccatca gccactgtcc tctgagctct | 16680 |
| gttcctccatca cgtgaccagg gttccatca gccactgtcc tctgagctct | 16740 |

ES 2 886 958 T3

| | |
|--|-------|
| cagtattagc atttatggac gctaccaccc cctccccc ttccctaaaca catagtgctt | 16800 |
| ttggtcacat gctgctttgg aggaggcctc acttggcgga tgtatTTTC tgccttagag | 16860 |
| agaggctgaa ctgggTTTGA ctgttgccc agccctctct tgctgcgtgc ccttagacga | 16920 |
| ttcactcaac gtctctgatc catggcatgt acaactataa gatgggcatt cccttctcct | 16980 |
| ctcgggctgt tatgaaggta aaggaagcaa gggctgttac ccaagggtgc tcccttctct | 17040 |
| ccccctttc acaccccccag gtgctctggg ccctcttagga actgggTTTc tctcaaggc | 17100 |
| tgttacccaa gggtgctccc ttctctcccc ctctcacac cactgggtgc tctggccca | 17160 |
| ctaggagctg ggattctctt aagagggaaa ctcttgata aaggaaatgg tttgattgat | 17220 |
| atcgacaaag tctgttcatt agtattcatt tattaagcac ctaccatgtg ccaggaaatg | 17280 |
| ctttggcgta caaaggaaaa taagggccag tcctgctaga aatggccttg aaacccagg | 17340 |
| gagggatgtc ggccatttgt gggtgctgca gattccttga aggtgatgca agagccagaa | 17400 |
| agaaggatga tggggggggc tgaggcaggg agtcgggggtt gggggagtgt gggggagaag | 17460 |
| gggagaccga gcaccccttc cactatctcc ctgtgtgggt tttgtgaac catcctgcct | 17520 |
| ctgggtgtct tgcctccagc ttctgacgtt ggaagttcat ccactgagag ctctgtgttt | 17580 |
| atggctctga gatactgagt ctttcttctc tcccagaccc cttacaatac catggcattg | 17640 |
| gacggcattt gcggggattt tgagacaatc tttcatgttc tatggagccg gataccaagt | 17700 |
| agaaaaagtg atttctcatc caaattatga ctccaagacc aagaacaatg acattgcgt | 17760 |
| gatgaagctg cagaagcctc tgactttcaa cggtaacgtgt ggctcaggct tggcaagcag | 17820 |
| gttggcagaa tcttaaagag atgttGattt gaaatgacac ttgtgctatg ccaaattggaa | 17880 |
| gggaggcatt tgcgttgagc gaggtagcg tgcagcgggt ggccaatggg agaggctcac | 17940 |
| agaggctaag agcacctgcc gcattttggg ggaggcagca gccaccacat ctgttctgt | 18000 |
| ctgtactgag tgggtgtat tcaagccagg catggaaaag gctagaacag ggctttccca | 18060 |
| ctgcagcacc cttgacatct gggtggtct ctgttgtagg gctctttgtt gccttgtagg | 18120 |
| atgtttaaca gcgtccccag cctctaccca ctggaggcca gtagctacca agctgtgaca | 18180 |
| accagtgttgc cctgctgaca ttgccaaaca tccgctttga ggcaaagtca cttccagttt | 18240 |
| agaactactg gcctaaaatg tgtaaagatc cttgatTTTT aaagatacat tctaaaacca | 18300 |
| agttgcttaa ttcaaggacaa acatgcttcc tcttagcctc ttattcggtc ccactctgg | 18360 |
| ccatccaaagg gtctggaaatg ttctagcccc atgtggatac agaagaagca aaacccatc | 18420 |
| cctccctaca gcatgtctgtt attcacattt gggaaatgggtt cacatataga agagcgaatg | 18480 |
| cctgagcaat ggcgtggcgc ctctggggcg aaagctgact ccattgactc catcggttt | 18540 |
| ttggctgttg ctcctgtgt gtctttcccg tcttgatcac ctggagatat gtaattttgg | 18600 |

ES 2 886 958 T3

| | |
|---|-------|
| aagcagagct agcaaataat tcctcttata agcagagcta gcaaataatt ctacttataa | 18660 |
| gtagcataac gtcttcctg ccagaaggag aggtctggca gggggagaaa gtgagaatgt | 18720 |
| gggacttgtt gggatgcagg gtccctctggg cagggtgcc agggtgccag gcccagcagc | 18780 |
| ctgcattgtgg gaaggccagg tggagacata ggtgataccc gcctggctca ctgtgtttc | 18840 |
| tcttcttgcac acagacccatg tgaaaccagt gtgtctgccc aaccaggca tgatgctgca | 18900 |
| gccagaacag ctctgcttga tttccgggtg gggggccacc gaggagaaag gtgaggctgc | 18960 |
| tcctggcac acaggactgc agggccaca gatggagcat tgggttcggg agtgggaggt | 19020 |
| ccaggtttta atccccatgc tactactcaa tgactggatg actttgggtt attcccccaag | 19080 |
| tccttgcac tcagtttctc catctgctaa gtggagaaa tcctgcccag cctacctaatt | 19140 |
| acactgtgtt cttatcgta tcacacagag cagcatgtgg aatggctttt gaagtatctg | 19200 |
| ggccatacga gtttagaggt gcaggatctc ctgtgttgca ctcattgtga gtttagagct | 19260 |
| ggccctggaga tcccaccaag gcctgcgtgg ctgagtgaca gggggcttgg tgaggacggg | 19320 |
| catcctggac ccatggtggc cacatctaag cctgtctctc gccctgataa ccacagagag | 19380 |
| aggctctctc caaccacttc ctttgcataatc tgcatcttc tctgacagtc tttcaaatga | 19440 |
| agggagcctg gctgcttcat ttttatggag gtttggaaatgt gcttagtggc aggcacaaag | 19500 |
| gttcattttt catattgttt atatccttct caaaagcgta taggcatac agacaacaaa | 19560 |
| tcctttcaaa caagggaaa agtacaaagg ttgggtgatt tctggggagc gtcagggaaag | 19620 |
| gtagtgggg gcatcctggc tcctcatcg cagaaactta ctacagttaga gccacaggct | 19680 |
| ggccaaaaga cctcatggaa tccaagatga aggaatatc gacaaatatt tgtgcgcacc | 19740 |
| tgcacctagt acaggctggg tgctactcgat gtgctggaa tgcagaagtg aacagagtaa | 19800 |
| gacaaatgtc tctgctgtca ggagcttac ctctcttctg gatgtcgggtg gtggggacgg | 19860 |
| ggcaggtgtg gtcagacaga tggagacaa acaactgago gaggtacttc caaacatctg | 19920 |
| agggtggga tcacaaggta cccgctattt tgaagggtg gtcagggaaag gcttctcgga | 19980 |
| agaggtggca tttgagctga gactcaaattt gcaaaaatgt gtacacatca aaaaggctag | 20040 |
| tgcattgtatc ttcatgtgtg gtcaaggggc caaggaggtg ggctggggcc agattgcata | 20100 |
| ggtcctgtg gattatggtg aagacaccag ctctcatct gcttgggtg gggagatcgt | 20160 |
| gagccgggaa gtccatgtatc ctggcagctg cgtggggagt gggatgaat ggatggagac | 20220 |
| gaggatgtatc gtgacaagtc cattgtgtg gttcatttgc acaggaagcc agctcatagc | 20280 |
| agagtgcggg cgtggatgtg aagagatgag ggtacactag ggctagagcc accagactta | 20340 |
| ctgatgggtt gcatgtctgt gggagagaga gtgagaagtc agggacatgt gctttccact | 20400 |
| ctgtggctga agccccaggg tggcgggtgg tgccattttt caagccagga aatattggtt | 20460 |
| ggtgagaatt tgggtggga gaaggtgtga cgagggttc tggtttgca cactaagccc | 20520 |

ES 2 886 958 T3

| | | | | | | |
|-------------|-------------|------------|------------|------------|------------|-------|
| acggtgccca | gaagatgccc | gaggggaggc | agcaaagcga | gagtggaaa | tgcagaggtg | 20580 |
| gcaagtgcag | gccgtgtctt | gagaagctct | aatgtgcagg | ggagccgaga | agcaggcgcc | 20640 |
| ctagggaggg | tcacgtgtgc | tccagaagag | tgtgtcatg | ccagagggga | aacaggcgcc | 20700 |
| tgtgtgtcct | gggtggggtt | cagtgaggag | tggaaattt | gttcagcaga | accaagccgt | 20760 |
| tgggtgaata | agagggggat | tccatggcac | tgatagagcc | ctatagtttc | agagctggga | 20820 |
| atttcttcc | ctgaagctga | actccagagc | tgcattcagc | acaggcaccg | ccagttgtaa | 20880 |
| ggagaatcca | ggttcccag | gagaggggtt | ggtgctggga | tgagctgacc | ggggcagggc | 20940 |
| tggaaaatag | ggctgtgacc | atctgttag | tgcgtgtgga | ggtctcaggg | aggaaagtgt | 21000 |
| gctctccctg | cgagagctgc | aggcaacact | gggagctcaa | caagtctccc | tgtccttagg | 21060 |
| gaagacctca | gaagtgctga | acgctccaa | ggtgcttctc | attgagacac | agagatgcaa | 21120 |
| cagcagatat | gtctatgaca | acctgatcac | accagccatg | atctgtgccg | gcttcctgca | 21180 |
| gggaaacgtc | gattttgcc | aggtaattca | acattttat | tctaccttt | gtccttacca | 21240 |
| gatcctactg | aaccccccatt | gagagagagg | gcattttgg | ggtcagcaga | gcctcctcag | 21300 |
| tgacacggag | ccagctcggt | gcagtcatgg | gaagtgacgg | ccacaaacag | tgcgaacgct | 21360 |
| tctggggca | gaaggaagta | cagtcaacaa | atcacacaca | ccctctgaaa | aaccggattt | 21420 |
| tggtaaaatg | gccagtggaa | cagaaacaag | tatttagact | atttaaatt | atgaacggca | 21480 |
| atttattnag | taacttttag | cttgaacaga | ttaaaattca | ggatgggggc | tatctcttt | 21540 |
| ggggttacat | ctctgttacc | atcaccctt | gatggggag | attcgaagcc | cacacagtca | 21600 |
| ctcgtaactc | acactgcgac | cccccccccc | caactcctct | aggcctggc | agtgggtgc | 21660 |
| ggcagattgt | gacttgattt | tctgctctct | gtaccttgct | gtgtcccaca | gggtgacagt | 21720 |
| ggagggcctc | tggtcacttc | gaagaacaat | atctggggc | tgatagggga | tacaagctgg | 21780 |
| ggttctggct | gtgccaaagc | ttacagacca | ggagtgtacg | ggaatgtgat | ggtattcacc | 21840 |
| gactggattt | atcgacaaat | gaggtaact | atcctgtct | ccttctgact | gtgttctccg | 21900 |
| atcctcgag | ccaaagccag | acatctgtta | ggcgtggttc | tgctgtggaa | agctgactgg | 21960 |
| tgaccactgg | tcagcatgaa | gcaaactctg | cttcctccag | ccacagcccc | atccccccat | 22020 |
| tgtccaccca | ttgcccattt | cctctactg | gttcacttg | catattccc | ctggtgttt | 22080 |
| gatgaaaagc | gctggggctc | agcttgtgt | aaattcctt | gtgctctgcc | aaccacactt | 22140 |
| cgttctggct | cagctgactc | agctgttcca | cccaggccac | ctcacatcaa | acttttttt | 22200 |
| tttttttttg | agatggagtc | tcactgtgtc | gccaggctg | gagtgcagtg | gcacaatctc | 22260 |
| gactcaactgc | aaccttgcc | tcctgggtt | aagtgattct | cctgcctcag | cctcccaagt | 22320 |
| agctgggact | acaggcatgc | gccaccacgc | ccagctactt | tttgtat | tagtagagat | 22380 |

ES 2 886 958 T3

| | | | | | | |
|------------|-------------|--------------|-------------|-------------|-------------|-------|
| ggggtttctc | catgttggcc | aggctggtct | cgaagccctg | acctcaggtg | attcacccac | 22440 |
| ctcagcctcc | cacagtgctg | ggattacaag | tgtgaaccac | ggtgcccgac | ctcacatgaa | 22500 |
| acttttgatt | tatagagagc | agagggaaaga | gccggctgtg | cccatcctt | tctggggcca | 22560 |
| tcgagtggtc | cctgggcagc | ccccaaagggtt | aggaagggca | ggagcagcca | gggttctctg | 22620 |
| atgcccaga | ctcaaggcacr | agggaaaggtc | tcaggggttc | catgtgagcc | tcatggatgt | 22680 |
| ctctgcttag | cagagccctg | gctttggca | ttgtccagat | aggggggtgag | aaccagatct | 22740 |
| tctcatctcc | aggacctcag | acgtatagtt | ttctcagatt | tctgtgctt | ctggggctgg | 22800 |
| gctactagt | gaagaaaagca | gtctattctg | tcttctccca | aatctccca | atgcccagtc | 22860 |
| tgttgaagga | ggagcagaac | cagggggcct | ttcccgctga | ggcccgacct | gtgtctccctt | 22920 |
| caaatacgac | gcgggactca | gggccttccc | atgaccatgg | ggcccagggg | gcgtcacctg | 22980 |
| gcccaggggc | cagtgttaga | aacagatgac | cccaggagga | ggagggcaggg | caggagggaa | 23040 |
| gctggcaggg | ctgggatgg | cagccaggct | gagggggcgg | ctcgcaccag | gatggagcta | 23100 |
| ggaaatgatc | caggtgtgtt | tggcggtgc | aggtgggtcc | gcatggctgt | gcagggaggg | 23160 |
| aagggctgcg | tggcaggaga | gcagccgggg | gaggcccaga | ctctgtgaa | gagatgcctg | 23220 |
| ttgtgccggc | ctccacatcc | gctgcccgt | ccttccggag | ctcctgcccc | gccatgctca | 23280 |
| gcctgactct | gaccaacacg | ttggagagaa | aatgtatccc | tttgtgttat | taagttgtct | 23340 |
| tatTTggttt | ctaagtgttt | catgcgaacc | tagaggaaaa | aattatTTc | cacctttgtt | 23400 |
| tgtcttaaga | aaataaacaca | ctttttttt | tcctatttga | acaggcagac | ggctaattcca | 23460 |
| catggtcttc | gtccttgacg | tcgtttaca | agaaaacaat | ggggctggtt | ttgcttcccc | 23520 |
| gtgcattgt | tactttttaga | gatgattttag | aggtcacttc | atTTTatta | aacagtgaac | 23580 |
| ttgtctggct | ttggcactct | ctgcattct | gtgcaggctg | cagtggctcc | cctgcccagc | 23640 |
| ctgctctccc | taaccccttg | tccgcaaggg | gtgatggccg | gctgggtgtg | ggcactggcg | 23700 |
| gtcaagtgt | gaggagaggg | gtggaggctg | ccccattttag | atcttctgc | tgagtccctt | 23760 |
| ccaggggcca | atTTggatg | agcatggagc | tgtcacctct | cagctgctgg | atgacttgag | 23820 |
| atgaaaaagg | agagacatgg | aaagggagac | agccaggtgg | cacctgcagc | ggctgccctc | 23880 |
| tggggccact | tgttagtgc | cccagctac | ctctccacaa | ggggattttg | ctgatgggtt | 23940 |
| cttagagcct | tagcagccct | ggatgggtgc | cagaaataaa | gggaccagcc | cttcatgggt | 24000 |
| ggtgacgtgg | tagtcaactt | taagggaaac | agaaacattt | ttgttcttat | gggggtgagaa | 24060 |
| tatagacagt | gcccttggtg | cgagggaaagc | aattgaaaag | gaacttgcctc | tgagcactcc | 24120 |
| tggtgcaggt | ctccacactgc | acattgggtg | gggctcctgg | gagggagact | cagccttcct | 24180 |
| cctcatcttc | cctgaccctg | ctcctagcac | cctggagagt | gcacatgccc | cttggtcttg | 24240 |
| gcagggcgcc | aagtctggca | ccatgttggc | ctttcaggc | ctgctagtca | ctggaaattt | 24300 |

ES 2 886 958 T3

| | |
|--|-------|
| aggccatgg gggaaatcaa ggatgctcag ttaaggtag actgtttcca tggttatgtt | 24360 |
| ctacacattg ctacactcagt gtcctggaa acttagctt tgatgtctcc aagtagtcca | 24420 |
| ccttcattta actctttgaa actgtatcat ctttgc当地 taagagtggt ggcctatttc | 24480 |
| agctgcttg acaaaaatgac tggctcctga cttaacgttc tataaatgaa tgtgctgaag | 24540 |
| caaagtgc当地 atggtggcgg cgaagaagag aaagatgtgt ttgttttgg actctctgtg | 24600 |
| gtcccttcca atgctgtggg tttccaaacca gggaaagggt ccctttgca ttgccaagtg | 24660 |
| ccataaccat gagcactact ctaccatggc tctgc当地 ccccaaggcag gctggttgc | 24720 |
| aagaatgaaa tgaatgattc tacagctagg acttaaccctt gaaatggaaa gtcatgcaat | 24780 |
| cccatttgca ggatctgtct gtgcacatgc ctctgttag agcagcattc ccaggacact | 24840 |
| tggaaacagt tggcactgta aggtgcttgc tccccaaagac acatcctaaa aggtgttgta | 24900 |
| atggtaaaaa cgtcttcctt ctttatttgc ccttcttatt tatgtgaaca actgtttgtc | 24960 |
| ttttttgtt tcttttttaa actgtaaagt tcaattgtga aatgaatat catgcaaata | 25020 |
| aattatgcaa ttttttttcc aaagtaacta ctgc当地 cttt gaagttctgc ctggtagta | 25080 |
| ggaccagcct ccatttcctt ataaggggt gatgttgagg ctgctggc当地 gaggacaaa | 25140 |
| ggtgaggcaa ggccagactt ggtgctcctg tggctctcga gataacttcg tataatgtat | 25200 |
| gctatacgaa gtttatgca tggcctccgc gccggggttt ggccctccc gccccgccc | 25260 |
| ccctccctcac ggc当地 cgct gccacgtcag acgaaggcgc cagc当地 cgt cctt | 25320 |
| ccgccccggac gctcaggaca gc当地 cccgct gctcataaga ctc当地 cttt gaaccccagt | 25380 |
| atcagcagaa ggacatttttta ggacggact tgggtgactc tagggactg gtttttttcc | 25440 |
| cagagagcgg aacaggcgg gaaaagttagt cc当地 tc当地 cgattctgcg gagggatotc | 25500 |
| cgtggggcgg tgaacgccga tgattatata aggacgc当地 gggtagggca cagctagttc | 25560 |
| cgtc当地 cagcc gggatttggg tc当地 cggc当地 tt当地 tgat cgtacttgg | 25620 |
| tgagtagcgg gctgctggc tggccgggac tt当地 cggc当地 gccggccgc tc当地 gggac | 25680 |
| ggaaggcgtgt ggagagaccg ccaaggcctg tagtctgggt cc当地 gagcaa ggttgc当地 | 25740 |
| aactgggggt tggggggagc gc当地 aaaat ggccgctgtt cccgacttca gaatggaaaga | 25800 |
| cgcttctgag gc当地 cgtgt aggtcgtga aacaaggcgg gggcatggt gggccggcaag | 25860 |
| aacccaaaggc cttgaggcct tc当地 aatgc gggaaaggctc tt当地 cgggt gagatgggct | 25920 |
| ggggcaccat ctggggaccc tgacgtgaag tt当地 cactg actggagaac tc当地 gggc当地 | 25980 |
| gtctgttgcg gggccggcag tt当地 cggc当地 gccgttggc agtgc当地 cccg taccttgg | 26040 |
| agc当地 cggcc ctc当地 cgtgt cgtacgtca cccgctctgt tggcttataa tgc当地 ggggttgg | 26100 |
| ggccacactgc cggtaggtgt gggtaggtct tt当地 cggc当地 gcaggacgc当地 gggttc当地 | 26160 |

ES 2 886 958 T3

| | |
|--|-------|
| ctaggtagg ctccctgaa tcgacaggcg ccggacctct ggtgagggga gggataagtg | 26220 |
| aggcgtcagt ttcttggtc gttttatgt acctatctc ttaagtagct gaagctccgg | 26280 |
| ttttaacta tgcgtcggg gttggcgagt gtgtttgtg aagttttta ggcaccttt | 26340 |
| gaaatgtaat catttggtc aatatgtaat tttcagtgtt agactagtaa attgtccgct | 26400 |
| aaattctggc cgttttggc tttttgtta gacgtgttga caattaatca tcggcatagt | 26460 |
| atatcggcat agtataatac gacaagggtga ggaactaaac catgggatcg gccattgaac | 26520 |
| aagatggatt gcacgcaggt tctccggccg cttgggtgga gaggctattc ggctatgact | 26580 |
| gggcacaaca gacaatcgcc tgctctgatg ccggcgtgtt ccggctgtca ggcgcaggggc | 26640 |
| gccccgttct ttttgtcaag accgacctgt ccgggtccct gaatgaactg caggacgagg | 26700 |
| cagcgcggct atcgtggctg gccacgacgg gcgttccttg cgcaagctgtg ctgcacgttg | 26760 |
| tcactgaagc gggaaaggac tggctgctat tgggcgaagt gcccggcag gatctcctgt | 26820 |
| catctcacct tgctctgccc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc | 26880 |
| atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcag | 26940 |
| cacgtactcg gatggaagcc ggtcttgcg atcaggatga tctggacgaa gagcatcagg | 27000 |
| ggctcgcgcc agccgaactg ttgcgcaggc tcaaggcgcg catgcccgc ggcgatgatc | 27060 |
| tctgtgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgcttt | 27120 |
| ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg | 27180 |
| ctaccgtga tattgctgaa gagcttggcg gcaaatggc tgaccgcttc ctgcgtctt | 27240 |
| acggtatcgc cgctcccgat tcgcagcgc tcgccttcta tcgccttctt gacgagttct | 27300 |
| tctgagggga tccgctgtaa gtctgcagaa attgatgatc tattaaacaa taaagatgtc | 27360 |
| cactaaaatg gaagttttc ctgtcatact ttgttaagaa gggtgagaac agagtaccta | 27420 |
| cattttgaat ggaaggattg gagctacggg ggtgggggtg gggtgggatt agataaatgc | 27480 |
| ctgctcttta ctgaaggctc tttactattt ctttatgata atgttcata gttggatatc | 27540 |
| ataatttaaa caagcaaaac caaattaagg gccagctcat tcctccact catgatctat | 27600 |
| agatctatacg atctctcg ggtcattgt ttttctctt attccactt tgtggttcta | 27660 |
| agtactgtgg tttccaaatg tgtcagttc atagcctgaa gaacgagatc agcagcctct | 27720 |
| gttccacata cacttcattc tcagtattgt tttgccaagt tctaattcca tcagacctcg | 27780 |
| acctgcagcc cctagataac ttctgtataat gtatgctata cgaagttatg ctgtacta | 27840 |
| taacggtcct aaggttagcga gctagctcca cgtggcttg tcccagactt ctttgcgtt | 27900 |
| caacaacctt ctgcaagaaa accaaggggcc tgaattttaa ctccctg | 27947 |

ES 2 886 958 T3

<211> 25333
<212> ADN
<213> Secuencia artificial

5 <220>
<223> Polinucleótido recombinante
<400> 6

ES 2 886 958 T3

| | |
|---|------|
| gcagagtcta agaaatcgct gtgttagcc ctgcacctgg gcactgtctt cacgggagct | 60 |
| gctgtggctg ctgttttgct ttggaagttc agtaagtgc gggagcctcg atccccat | 120 |
| gtgctcctgc agtccccagt gctctgagcc agaccctgct ctctgggcta ttgagacctc | 180 |
| tggaggccct ccgtgagggtt cctcttttac ataacgaggc tgcgtcttcc ttgcattttt | 240 |
| gttagctat gagattgaca catcatgggg aaagcatttta gaatgtaccc agtgcattgg | 300 |
| ggtgcttggt gccaccccagc actgtgagca caggttcttc taccttgggg ccacacccag | 360 |
| ttacctgtat ctcaactgcac agcagtggct gttggggacc aggcccaccc ctccatgtcc | 420 |
| cacccctgc aactgcagcc tgagccttcc catcagcctg gggtggtgca gacccatgtg | 480 |
| ccattgtgga tccttcaagt tacctgtgtg gcagagagga cgtgtgagtg cgtccaaac | 540 |
| ccaaacactg agagggtcct tcccatgcc cccacggaag taaggtgccc cagtgcataat | 600 |
| tccacttata cttgctggtg gcaaggacac ttctcccttattaaagtg ggggattggc | 660 |
| tgggtgaggt ggctcacgcc tggttatccc gcactttaag aggccaaggc aggtggacca | 720 |
| cctgaggtca ggagttttag accacaagcc tggccaacat gttgaaactc catctctact | 780 |
| aaaaatacaa aaattagtca ggcgtggtgg cgtgcacctg taatcccagc tacttaggag | 840 |
| gctggggcag gaggatcact tgaaccagg agttggaggt tgcagtgagc caagatttg | 900 |
| ccctgcact ccagcctggg tgacagaatg agacttcatc tcaaaaacaa aacaaaacaa | 960 |
| aacacagtgg ggcaggagt tggaggctgc agcgagctac agtaatgcca cggtgttct | 1020 |
| cactccatga ggctcattgc gtttctcagc ctgaaggcga cctctttct gtttctctg | 1080 |
| caagtggcga gcaagtgctc caactctggg atagagtgcg actcctcagg tacctgcac | 1140 |
| aacccctcta actgggtgtga tggcgtgtca cactgccccg gcggggagga cgagaatcgg | 1200 |
| tgtggtgagt cagccttgac cttgggaagg gactcctctg ctcaccttgg agacagcagc | 1260 |
| cgggtccagg ggccttggg tgactggcc tggcgtgcgt ccagtagct gacacatgat | 1320 |
| gtcattgaat ccctgctcca ggctgagccc tggggctcaag agaggttgg tttccggccc | 1380 |
| aacctcaccc agcaggtggg agatgacagg gccaccgagg actgtgtcat tggaaaccaca | 1440 |
| cgtgctctga actgcccacag gaagtcaattt aagatgagca aactgtttat aaagttggag | 1500 |
| atgcaggcta ggaacggtgg ctcattgcctg taatcccagc actttggag gccgaggcag | 1560 |
| atggatcacc tgaggtcagg agtttgagac cagcctgacc aatatggtga aaccttatct | 1620 |
| ccactaaaaa tacaaaaatt agccaagcgc ggtggcgggt gcctgtaatt ccagctattc | 1680 |

ES 2 886 958 T3

| | |
|---|------|
| aggaggctga ggcaggagaa tcacttgaac ctgggaggcg gaggttgcag tgagctgaga | 1740 |
| tcacgccact gcattccagc ctgggagaca gagctggctc aaaaaataaa ttaattaatt | 1800 |
| aaaaacaaaa ttggagatgc actatgttat tttcaaaaca agctgcctt aaagatctat | 1860 |
| ctgttgcac agggtggct catctgttc attttattt ctgtggttt tctatttatt | 1920 |
| catttaatg aactaggaag cattgctcct atttatggca taccacatga tgtttgata | 1980 |
| cgtgtatgcc tgtggcatgg ctaagtcaag ctagaacatg ggccttacct catatactg | 2040 |
| tcttattaag aacacataaa acctactctt gttagtattt tcaaatatgc aacatatagt | 2100 |
| ttattaactg cagtcactat gatgtacaat agattgctcg aacttattcc tcctgtctaa | 2160 |
| ctaagattt gtgacctctg accaacatct ccccagtgtt gtcacccccc gccccagcc | 2220 |
| tctgatact gccttctac tctctgcttc tgtgagttt atgtttatac attccacatg | 2280 |
| taagtggcct catgcagtgt ttctgtctct gtgtctggct tgttcactta gcgtaatgtc | 2340 |
| ctccagcttc atctatgtt ttggaaatga caggattcc ttcttctt tggctgaata | 2400 |
| gtattgcctt gtgcatacac accacattt ctttatccct tcattcactg atggactctt | 2460 |
| aggtttagtgt catgtcttgg ctgttgtgaa aaatgccgca gtgagcgtgg gcgtgcaggt | 2520 |
| ccctcttcaa cacacggatt tccttcctt tggatataaa cccagcagtg agattgctgg | 2580 |
| atcacatggc agttctgtt ctcacccttt gaggaaactc catactgtt tccataatgg | 2640 |
| ctgtagcaac ttccactccc acccccacgg tgcaaagtct ccattctct tctacaacct | 2700 |
| caccaactcc ttttatttcc catcttctg atagtagcca tttgaagagg tatgagatga | 2760 |
| tacctcattt tggtttcat ttgcattttt atttgtattt ttcatgaatt tttgagggtg | 2820 |
| atttcaaggg tagtttagtga ctcgaacagg gaaacgatcc tgagtagttag ggttgtgcta | 2880 |
| atcatcccc tcctgccagc tgcgtacgga atggggctct gcagatggca gggagctggc | 2940 |
| tcgtttctct ttaagagctg cctttactt ttcttcctct tcctttaaaa cttatttct | 3000 |
| ggccggacgc agtggctcat gcctgtaatc ccagcacttt gggaggccga ggtgggcccga | 3060 |
| tcacgaggc aggaattcca gaccagcctg gccaacatgg tgaaaccccg tctctactaa | 3120 |
| aaataaaaaa attagccaga cgtgggtgtg cgggcctata gtcccagcta ctcgggaggc | 3180 |
| tgaggcagga gaatcacttg aacctggag gagggggttg cagtgagccg agattgcgc | 3240 |
| actgcactcc agcctggcg acagagccag actccatctc aaaaaacaaa aaaaagttat | 3300 |
| ttcccaagca cagccatgtt ttccaggctt gtggatcagc gttgggtgtg gtgtgtgctc | 3360 |
| tcatatctt gttccagcta agcacactct gacatgttta cactagaacc atttgtttt | 3420 |
| tctagaaata gaaatttcag aattgttagag tcagaggact taccagaaat ctcttagtta | 3480 |
| gttctcctcc cctccctcaa gtgcagtctt aacctcctgg agtttctgt agaaaccaca | 3540 |
| agcctcagag ctggccgaga attctagcca aagattttc catgccaaag taatcccccc | 3600 |

ES 2 886 958 T3

tctcctaagg gccatcctt gttgggactg gtttcctgtt aagccctcg tgcgttcct 3660
ggctgtggaa ttccctggtg aggagcaactg gcccgtggag ctccggccct gtgccggcc 3720
tgagcaggcc caagtgttcc gtgttcttga tacctttcct ccagcacagt cttgcttccc 3780
agaaaaaggt ttgcacttga aaatgatgca ttgtctgatt aaacatagtt ctttgcttt 3840
atttggtttc taaaataaaag tgggagttt tgagatttag taacgtgagg ttaagatagc 3900
acgtggaatg gcttttctt ttctttctat tttttttt ttttcctgg agacagggtt 3960
tcactctgtt gcccaggctg gagtgcaagag gcatgaccat ggctcaactgc aacttcgatg 4020
tcctggggtt aagcgatccc ccagcctcag ccccccaagt ggctggact acaggtgctc 4080
gccaccacac ctggctaatt tttgtatTT ttgttagaaaa tgggtttcat caatgttgc 4140
cagactggtc tcgaactcct gacctcaagc aattctcctg cctcagcctc ccagactgct 4200
gggattacag gcgtgaacta ccacgcctgg cctggaatgg cttttgatgt tctcctatgt 4260
gcacatgtgg gtgaataaac accaacaag tccttatgtt acctgaagag ttgctcttctt 4320
cttaatattt aagtcgtatt tatTTaaata cttaatagt tgtacactat taaagtatta 4380
tttaggtcaaa atcaaggaag tacaaaagg tatgctgtga aaaatctt cttccttgc 4440
ctgcttactt acctaccccg catccccca tacaccccaag acacacacac acacacacac 4500
acacacacac acacacgcat cactccata catgcccacc tgTTTaccag ccaatcacat 4560
ttcttgggc aactcatctg agttgcttct cttccagag agttttgca taaaagaagca 4620
caggtatTT tgctttacca tgaccattt tcccaagtgg tccctagccag ttgactctcc 4680
tgcactggat accatcctgg acagcattcc tttagggaaat gagccccctg tttttccca 4740
ccatggcaca gttggtcctt tgcatggacg caccattatt gcccctgtct cttcttggtg 4800
gaccttaagg ttttctccat cttttgctg taacacacac tgctccaagt gtgtgagcat 4860
atcagtagga aacgcttcca ggagttagaac tgcttaggtca gagggcgtgt ggatctgtaa 4920
cctgacagac ctagaccggc ttcaagtgg ttttatccag tttccatatt gattattcat 4980
ataaaaggaa acagacaaac ataacgtgt gcatgtattc tctcttagac cagaacaggc 5040
atagggtgca ctttaattt gtccatttcg tagagtagaa attgttttg ctgaaatgaa 5100
caccttagga tgctgaagaa tatgaccgt cccatggaaa acattaaaa atgtgtgtag 5160
cgctttcttc ccaagggtgt gtgtgcgcattttaacac taattcactt tctacttccg 5220
ttgctatcct ttctgtgagt ctttctcaga atctcagaaa agaaactaaa ttgttcactc 5280
tagttatcaa tgctgtactc tatacctggat atttgctaaa agggcagatt ttaagtattc 5340
tcaccacaga aaagagaaaa gaaaatggta attatgtgac gtggggaca tggtaacttag 5400
ctttattatg gtgagcattt cacagccggat atccagtcatt cacgctgtac acattaaaca 5460

ES 2 886 958 T3

| | |
|---|------|
| tgtacaattg ggtttttttgc agacaaggc tccttcgtc acccagtctg gagtgca | 5520 |
| gctcagtcat ggctcattgc agcctcgacc tcctgggctc aatccatcct tccccctc | 5580 |
| cctcctgaaa agctggggcc acaggcatgt accatcatgc caggcta atg catatatatt | 5640 |
| tatattttt ggtggagatg ggggtggctc cgaactctgg gctcaagtga tcctccc | 5700 |
| ttgcccttcc aaagtgcgtg gattacaggc atgaaccaca gcaccaggcc tacatgtaaa | 5760 |
| atttttatgt gtcaactata ctttgacaaa gctgagaaaaaaa aaaatccaa tatttaaaaaa | 5820 |
| aaaaaaaaaaa aggacttagct tgagacctt tccagctctc tggcttatca gctgccgtct | 5880 |
| cttccgggtc cagatagctg gaaggaaag aaaatccctaa aaattaccca caagccaaga | 5940 |
| atgaagtgtc tcccttgag ccacagtggc agtttgc ttatcatag aagtgtattt | 6000 |
| tgagccgggt gtgtggctc acgcctgtaa tccccgcact ttgggaggcc gaggtgggg | 6060 |
| gcggagggggg tgggatcgc ctgaggtcag gagttcgaga ccagcctgac caacatggag | 6120 |
| aaaccccgctc tctactaaaa atacaaaatt agccggcgtg gtggtgcatt cctgtatcc | 6180 |
| cagctactca tgaggctgag tcaggagaat ctcttgaacc caggaggtgg aggttgcgg | 6240 |
| gagctgagat catgccatttgc cactccagcc tggacaacaag aaaaaaaaaag aagaagaaga | 6300 |
| agaagtgtat tcatttcagt tactttaaa aaagtgaaca gactttat ttttagagcgg | 6360 |
| tttttaggtt acagaaaatg aaacagacag ggcagcgcgc tccttgact cctccccc | 6420 |
| acacagttgc cctgttatga acatcccaca tcagtgcgt gcgttcatta acaccgatga | 6480 |
| acctgatgca tacattatga tgaactgaag tcctggactt caccctttctt cttgtacagt | 6540 |
| tctgtggat ttgacaaatg cataatgctg tacagccaca atgatagtat cgtccagagt | 6600 |
| agttctctg cctaaaacc tctttgctg cacctgtttc tctctccccca ctcaccc | 6660 |
| ctatctgatc ttcttagtgc ctccgaagtt ttggctttt caggatgttg tagcgttgg | 6720 |
| atcatggagt atgttagcctt caccacatac accttccttc actttgttg cttcctttac | 6780 |
| ttagtaatat gcattcaagt ttccctccatg cctttcatg gcttgatagc tcatttctt | 6840 |
| ttagcaccaa ataatattcc gttgtccaga tgttagcacaa tgtttatcca ttcatgtaac | 6900 |
| ctgtgaccga ctcacagata ggatgtggaa tcactcacca cagaggcatt agacaataat | 6960 |
| cagacccaaatg tcatttcattg gggaaacaag cccacaggta ccagactgtc cagtgagtca | 7020 |
| ggcccaactcg taggaagtaa gaagagaggc tagagcatag ccaggtcctc actttatact | 7080 |
| ttaagccat gtgtatttctt cccaaaccac acagcattgt ttccatgtt tcagctttgc | 7140 |
| atgaataacg tgatacttgc acgcatcatt tatcacttgc tctctttccc acagcgctgt | 7200 |
| tttcaagctt ctccctgttc atgatgcctt gcttaaccct taagctgcattt gggattctgt | 7260 |
| tctgtgaata cgcccacccc atgtattatc ctgcccagca aaaagtcccc aaaactctgg | 7320 |
| atggtggta cctctaggaa gggagagaag agattggaa tagggagcga cttaacgg | 7380 |

ES 2 886 958 T3

| | |
|--|------|
| gtttgtaatg ttttgttct ttaaataaaa gagctgagat catttcagca gaatgttgat | 7440 |
| ttagagtctc ctggacaatt tggtgctcaa agtgctctct taaagagcac ttaaaaaaaaa | 7500 |
| aaaacccccc atcttattat ttatTTTATTt atttatttag acggagtttt gctctgtcac | 7560 |
| ccaggctgga gtggagtgggt gtgatctcg ctcactgcaa cctttacctc ctgggttcaa | 7620 |
| gcaattcccc tgcctcagcc tcccaagtag gtgggattac agatgcgtgc caccacactt | 7680 |
| ggctaatttt tgcatTTTATG tagagatcgg tttctccatg ttggccaggc tgatctcaa | 7740 |
| cgcctgacct caggtgatct gcccgccctg gcctccaaa gtgctggtat tacaggcgtg | 7800 |
| agctaccatg cctggcttat cttatatatt ttAAAAAACA gcttatttag atctaattt | 7860 |
| tgtaccataa aattcaagta tataattcag tgctttata tataAAACAT atatatgaaa | 7920 |
| tagcttattt agatataatt ttTTTATAA aacagcttatt tgatatgtaa tgtatgtacc | 7980 |
| atAAAATTt aatatataat tcactggctt ttatataattc acgaatatgt gcaactatca | 8040 |
| ccacagtcaa ttttagcata ttttcatcag ctcataaaga aaccccaagc ctttgaacta | 8100 |
| tcacccata tccctcctcc cagcccggtcc ctcctactca taagcaacca ctaatctact | 8160 |
| tagtgtctat agatttccta ctctaggcat tccatgtgag cgggatcatg caatacgtgg | 8220 |
| gctcacacaa tataagtggc attccatgtg agtcggctca tgcaGTatgt ccggctcctt | 8280 |
| tcactgagca taaggcttc agcactcatc caggttgcag cctgtgtctg aatttcattt | 8340 |
| cctttctgg ctgaatcgta ttccatgtg tatcttggac atatcctatt ctgctcaccc | 8400 |
| agccgttggt gggcgTTGG agtgtttcg cctttcagct gtttaAGAG gttgcagtg | 8460 |
| aacatttGTA caagTTTGG acccaatGCC tGTTTCAAT tctcttGtGt agagAGCAct | 8520 |
| ttttAGCAGA AAAAGAATAG atttGtGGCC tcccttGtG tGCGGTcAGt GccttGAGAA | 8580 |
| gagtGAActG tgctGCCAcc tccggagccg tggagAGCgc ggggcttggg tagcagctag | 8640 |
| gacgatacaa gttgggacaa ggccagggtgc aatggctcac GcctGtaatt ccaacacttt | 8700 |
| gggagaccga ggcaggGGGA tcacctgagg tcaggAGTtC aagaccAGCC tggccaacat | 8760 |
| ggtgaaACCC catctctaAt AAAACAGAAA aattaACTGg acggggTGGt ggacGcCTGt | 8820 |
| aatcccAGtCt actcGGGAGG ctgaggcagg agaAtcActt GAACCTGGt ggcggaggct | 8880 |
| gcagtGAGtG gagatcAGAC cactGcActt cAGCCTAGtG gacAGAGCgA GACTCCGtCt | 8940 |
| caaaaaAAAG aaaaaaaaaaAG aaAGAAACtC atggataAtC ctccctctG tGcAGtTCGc | 9000 |
| ctctacggac caaaACTtCAT cttcAGGTG tactcatctc agAGGAAGtC ctggCACtCt | 9060 |
| gtgtGCCAg acgactGGAA cgagaACTAC gggcgggCgg cctGcAGGGA catGGGctat | 9120 |
| aagtGAGtAt gggcAGCAc ccGCCGAGtG acAGTAACAG acAGCAGAAA cacGAGAAGA | 9180 |
| ccctctctct GcctCCtGt gaaAGCACCg GcACAtGAGt GctGGGGAcA attGtCACtCt | 9240 |

ES 2 886 958 T3

| | |
|--|-------|
| tccaaaagct gagccctata accagcaggt ggaatttgc tcgtctaggc tgtgcccagc | 9300 |
| acacagacct tggctcaactg ccacacctgcc ctgccttcctc cttggcctct atagactcct | 9360 |
| ggttgctcg gagtgcggcag tgctgtggc atctggtcag agggtaggc tgagggcggt | 9420 |
| aggtgcctct ttttccaagg tgcctctcag ccagggtcca ttcacccccc tgggttagagg | 9480 |
| ttggaccaga acagctggcg aggagggttg ggctggggag agcagcagag acaaatcctg | 9540 |
| tgccagtttc acttcattcg ggagccatgg aagcctttt agctggggag agaatcaatc | 9600 |
| aatcagactg atacttaaaa aatgtcattc ctgctcgttag ctctgagggg aggtggaaag | 9660 |
| gcttaacagg gtgtgtgtcg cctgacagtg attcctaacc ggggtggggc ggtggttacc | 9720 |
| atttaccaggc actgcctggg gagatgcggc agccctcagg catcgaaaaa gagggtggta | 9780 |
| ggatgtact gccactttgt tttccatggg agggccccca ggtgatttct atgcaacttt | 9840 |
| agggtattca atatgccagt tttcagaatg aattaccact cggtgagaaa gttggcatct | 9900 |
| tagctagtca ctgtgacatc cctaaacagc aggggtgaat tacacagcaa agccccccca | 9960 |
| tcacagtcca ggaacctgg ggaattgata actggggcca tgttaacatc tgtacctttt | 10020 |
| attagattaa atgtgtgtat gattatacaa tcctatgtcc ttctcatagt ttcttgatcc | 10080 |
| taacctggat aagaaacacg accaatgaag gaattttgtc tgacacttta gggttattga | 10140 |
| atcgaaaaat cgttacaata ttctagcaact tggttagaac gtgtgatttt tttcctaa | 10200 |
| tgctaagggtt tttccctctt attctgaatg tcgtatgagc ggtattatga catagtata | 10260 |
| gatttgtgtt tgcttatgcc ttaaccatta tcacaaataa ggtttcttt tttaggaata | 10320 |
| attttactc tagccaagga atagtggatg acagcggatc caccagctt atgaaactga | 10380 |
| acacaagtgc cgccaatgtc gatatctata aaaaactgtt ccacaggtat gcagcaattt | 10440 |
| cttcttgaaa aattttggaa tgaaatcaac taggagacac catggggat cggtgtcctg | 10500 |
| agtctgattt ctctgagctg caatactcgg tctggatggg tttgcattt ggaggagatt | 10560 |
| agagtctgac caggcctgg tactctaagg agccctggg ttattcatag gaagtggctg | 10620 |
| aggtttctct gctatttcat tttcagccctc taccgtctgc cttgttggt agcggctcac | 10680 |
| acttgcaaca tcgacattca actctatttta gtttctttc ctcttcagac atttagaggt | 10740 |
| gtacctattt tgtcagggcg tggttctagg aatccaagat aatgtctcag tgtcccagcc | 10800 |
| agggtgaccg gtcatttcca gtttgcagg gacttcactg gtttgcggaa gggaaagtcc | 10860 |
| gctccattcc aggcagctgg gctggctgg cccgttagcc ccaaccccg gacagcagtg | 10920 |
| ccagagggtg ctctgtgagg gatgggcagc attctggcgg cctggaaatg agttgtgggt | 10980 |
| tttccagggg gtagaaagtgg gtacaagcca caggtcacat gatgagtggc tgacctggct | 11040 |
| gggagggcag aagaggggat ggacttaggc tcttcctttt gctttgcaca tatttaggat | 11100 |
| gtttgcagac ttgctatgtat tggtgtgtt atgtgttttgc tgatgtgaaa gatacacag | 11160 |

ES 2 886 958 T3

| | | | | | | |
|------------|-------------|-------------|-------------|-------------|-------------|-------|
| gtcctttgcc | catgagctct | cattgcctcc | caggtccccca | gggcttatgc | ctggtgtcta | 11220 |
| ggcatcacct | ccctgcctgc | caggtgccag | tgctgcatt | tcgggggagg | atgaactaat | 11280 |
| caccccgcbc | cacccccc | ctgagtgaaa | gcctggggca | ggtttgcatt | cctggaggcc | 11340 |
| gctggtgag | gggtctgggg | gcctgacttc | cactgcagcc | tgctgtcctg | ggaaatgtgg | 11400 |
| cagggcaagc | ccagtgggg | gggctgtgca | cggccagggt | caccatcaa | aacagcaggg | 11460 |
| ctgcgggtt | tccctgtgga | gaagctaaac | acagctgcct | gggcacttt | taaatgtga | 11520 |
| gtggttcttt | gtctttctgg | gttacacacg | gaatcaggga | gccaagtcca | gccgggcagg | 11580 |
| gacgggggg | ggggaggagg | tgctggcgtc | ccttggcaag | agccttgg | actcacaagg | 11640 |
| aggctggagg | gcttggaaaga | aagaagagaa | ggccattgtc | tggtaggctc | tattctatct | 11700 |
| cggtgggt | ggtgggggg | ggcgcaactc | ttttcctt | tctgtgcagc | agttgcctt | 11760 |
| tgatgcctga | gttcttgct | tgtttctgt | cgggcttctg | tgaataacca | catgtgcct | 11820 |
| ggcgctgtga | ccacacaggg | ctatccctac | cgaccttagg | attcttagga | aatgtttct | 11880 |
| cttaaagggg | acatgtcttc | acttggccgt | gtcagtgc | cagagccaga | gtccacctgg | 11940 |
| aatgcacctg | tagtactga | gaaccgggg | ggtgtgc | agtaagaagg | tgtcaggaag | 12000 |
| gacctattat | tgttaggcct | gggctcctgc | aaggtgg | gggggtgg | ggaggaagca | 12060 |
| gagatttgct | ctggattgga | tgctgtcagg | aagcagggg | aattctgtga | ggctgctt | 12120 |
| ttatTTTTT | tctaggagga | ggttggaaatg | aggctaggct | aaagctgtga | ttggtaaaga | 12180 |
| aacgtccgtc | gctcaagtta | gccaggacag | gaggagacat | cagatcgtga | ttttgtgg | 12240 |
| gtgagcacaa | ggttcctgtt | ctgtctgttc | agacatcatt | tcggaggagg | ctcctgtgt | 12300 |
| cttgcoccat | ctcaggcatg | gagggccta | gtccgatatt | gacgctcagt | gaaataattc | 12360 |
| aggttccgca | gagcacacgg | cccagctatc | agggcggcc | agctctgcat | gccagggg | 12420 |
| gcgtcttccc | ttctcagcat | agcctggaa | attcaactgca | ggacaaaatg | catcagttac | 12480 |
| ttcctttca | tccataacct | ggatgttt | actccaaat | gagtaactct | tacgttt | 12540 |
| ctaattctag | ggaaactatt | ggttatatt | cttcaacac | tacaaattt | aagcagttat | 12600 |
| aggagcccag | aggttccaa | atggctcct | taaaaattag | aagatgattt | taaattccaa | 12660 |
| gaggaaaaac | aaaactagca | ttattgtata | cttaccctca | caaccgtcct | aggagctgg | 12720 |
| acaattttaa | gagaggtaa | gtaaactgccc | caaggtcaca | ctgtgggat | gtgagccgcg | 12780 |
| taccttggct | cagtgtctgg | tcttgcac | tgtccctata | tggatttact | taccttatt | 12840 |
| gagttgtaac | tagcagaccc | ttctatgtct | cagaagacag | gagagggaaac | atcggaaagaa | 12900 |
| atgactgatt | tctaagcatg | tgagaggcag | gtgactccgc | actatcgtga | ccagaattc | 12960 |
| ccctgttctt | tttgcagtga | tgcctgttct | tcaaaagcag | tggtttctt | acgctgtata | 13020 |

ES 2 886 958 T3

| | |
|---|-------|
| ggtaagttca tctggagtcc ccctttgat acttctaact aggaaaagct ctctactttc | 13080 |
| agaacagtac tccctgtgtc tctggggcg tgggagggaa gaaggtgggg tcacgggttg | 13140 |
| gaatgtgccc agcggcgtct cgctttcc aaggagctcc tggtttagat ttccatggcc | 13200 |
| tgttagacacc ttcagcctt ggtccaaggg acaccccctg agatcaggca cgctcaagaa | 13260 |
| gctgacaaag ccctacactt tatgccaccc atgagctgga ggccggcag gtctttct | 13320 |
| ccagaaagca aaggggggtg gcgttagtga gccctggcag ccacctaacf tggacttgga | 13380 |
| gcacatgcgg ggctgtggtc cagcaccacc gtgtggccac caggtgctca tcagccagtg | 13440 |
| ggaccggga ggagggacaa gaccagagaa caacagtgtc cttgcctt ctctcctgaa | 13500 |
| ttttggacgg tggcttagac ttgggtgtcc ccatctctgt gtttagagtg cttacagttt | 13560 |
| ccaaactgtt tgcaaatgtg gaagccaccg tcccttcctt ctggatggc ccagtgtgt | 13620 |
| cgtggggccg tggcctgag ctcagttt catttgaaga ggtgaaagga gctgacaccg | 13680 |
| tcccattcccg gcagggctgg ctcaggtttt ctttaggtcc ttagtgggg tccagcacag | 13740 |
| ccccaaagggt gcgtggcacc cgcctgccc tctgcccattt cactcatctc ctggtgagaa | 13800 |
| agacactcac acacaggaag cagggaaaggc agcagacccctt actcaccctt cacccttca | 13860 |
| ctcaccctt actcacccttca tcaacctctc attcaccacc cacccttca ccccttact | 13920 |
| cacccttca ctccctcaac ctcactcac ctcctactc cctcaaccctt cactcacctc | 13980 |
| ctcaccctt cactctcccc ctcatccctc ctcacccttca ccccttccacc tccactca | 14040 |
| cctcactca cccttccacc ccttcaccc ctcactcac caccccttccacc cctcactcac | 14100 |
| cccttactca acccccttattt cacccttca ccccttactc acccccttccacc ccccttactc | 14160 |
| accccttcat ccacttccacc acctgtcac ctcactcac acccccttccacc ccccttactca | 14220 |
| atcccttactca cccttccacc ctcacccctt cactcacacc ttccacttccacttccacc | 14280 |
| ctcaccctt caacccttta cttaccctt cactcatccc ttccacttccacttccacc | 14340 |
| tcttacttccacc attcacccttccacttccacc ttccacttccacttccacc | 14400 |
| accccttccacc ctcacttccacc cccttccacc ctcacttccacc acccccttccacc | 14460 |
| cccttactca cccttccacc ctcacttccacc ctcacttccacc tccacttccacc | 14520 |
| cacccttca ccccttccacc caccactca cccacccttccacc acccccttccacc | 14580 |
| accccttccacc ccccttccacc ctcacttccacc ctcacttccacc acccccttccacc | 14640 |
| ccacccttccacc actcacttccacc tccacttccacc actcacccttccacc | 14700 |
| tcacccttccacc ctcacttccacc ctcacttccacc ctcacttccacc tccacttccacc | 14760 |
| cccttactca ccccttccacc ctcacttccacc ctcacttccacc cacccttccacc | 14820 |
| actcacttccacc tcacccttccacc actcacccttccacc tcacccttccacc | 14880 |
| actcacccttccacc tcacccttccacc actcacccttccacc tcacccttccacc | 14940 |

ES 2 886 958 T3

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| actcaccctt | cacccctca | ccccctact | cacccctcac | cccctcgccc | cctcaactcac | 15000 |
| ccccctcgccc | cctcaactcac | ccctcacccc | ctcaccctt | cactcatccc | ctcacctct | 15060 |
| cactcacccc | ctcacctct | cactcacccc | ctcacctct | cactcacccc | ctcacctct | 15120 |
| cacccacccc | ctcaactact | ccctcacccc | ctcaccctt | cactcacccc | ctcacctct | 15180 |
| cactcacccc | ctcacctct | cacccacccc | ctcaactact | ccctcacccc | ctcaccctt | 15240 |
| cactcacccc | ctcacctct | cactcacccc | ctcacctct | cactcacccc | ctcacctct | 15300 |
| cactcatgcc | ctcaccctt | cactcacctt | ttcacctct | tgctcatccc | ctcaacttacc | 15360 |
| ccctcaacttc | gtcaatcacc | cccccacctc | gtcaatcacc | ccctcacctt | ttcacctacc | 15420 |
| ccctcaactca | cccccttact | tcctcaactta | cctcctcacc | ccccactcac | ccctcacccc | 15480 |
| ccccactcac | ccctcacccc | acactcaccc | cctcacccccc | cactcacccc | ctcacccctc | 15540 |
| tcaccccttc | actcacccccc | tcaccccttc | acttatcccc | tcaccccttc | aattacccccc | 15600 |
| tcaccccttc | aattactccc | tcatccttcc | aattacccac | tcaccccttc | acctcctcac | 15660 |
| tcctcaactca | ctccctcaact | caccccttca | ccttctcaact | cacccctcg | tctcctcacc | 15720 |
| ccctcaactca | cttccagccc | tgcccccctcc | atttccctt | tctttgtgtg | agaatctggg | 15780 |
| gtccctgagt | ggtgtcagtc | cctccaagac | tcaaggagtc | cccagggct | tgttatccag | 15840 |
| aacaccccca | cctgggtccc | gggagacccc | atgggatcac | aggagtgtt | aggaaagtgg | 15900 |
| tgttcctgg | gtctgggtgg | gctggagggg | catcctccct | tcccaagag | gagaccccca | 15960 |
| ggagccccc | aagtccatcc | ccagcagtgg | tgcccttgcc | ctgtccttgc | agcctggag | 16020 |
| acccttggga | ggggcggggcg | ctgggtggct | gggcggcttc | tgctggtctc | accccactgg | 16080 |
| cctcctgttt | gtcatcctca | gcctgcgggg | tcaacttgaa | ctcaagccgc | cagagcagga | 16140 |
| ttgtggcgg | cgagagcgcg | ctccgggggg | cctggccctg | gcaggtcagc | ctgcacgtcc | 16200 |
| agaacgtcca | cgtgtgcgga | ggctccatca | tcaccccccga | gtggatcgtg | acagccgccc | 16260 |
| actgcgtgga | aaagtatgcc | aggggcggcg | cgggcgggggt | gggggctcag | ggctggccta | 16320 |
| cagccaccct | gtgaccttga | gcaggtctca | acccttgcag | ccccggcatc | cttgtgttta | 16380 |
| aatggggaga | gtattgcacc | tgcttcctag | ggctgtgaga | catcaagtgc | gctcatgcca | 16440 |
| ggcagtgcac | ggctgtatgc | actgagtgac | ccctgcacgc | agggcacagg | gtgcaggtgg | 16500 |
| aacattctcc | acgatgtcgc | cgtgaccagc | gttccttcca | gccactgtcc | tctgagctct | 16560 |
| gtcctgcctt | tgagcaaagc | ccctgcgggg | tgaggtatcc | tgtctccggg | acgctagtc | 16620 |
| caggagaggg | cacactcaga | caggcttcag | gtgccttgc | tggaaaggccc | ctggggttaa | 16680 |
| gcgttcttgg | ccacagcatt | gctcatgcag | agggtaggt | aggggtgagg | ctagccgtga | 16740 |
| cagtattagc | atttatggac | gctaccaccc | cctccccc | tccttaaaca | catagtgctt | 16800 |

ES 2 886 958 T3

| | |
|---|-------|
| ttggtcacat gctgcttgg aggaggcctc acttggcgga tgtattttc tgccttagag | 16860 |
| agaggctgaa ctgggttga ctgttgccc agccctctc tgctgcgtgc ccttagacga | 16920 |
| ttcactcaac gtctctgatc catggcatgt acaactataa gatgggcattg cccttctcct | 16980 |
| ctcgggtgt tatgaaggta aaggaagcaa gggctgttac ccaagggtgc tcccttctct | 17040 |
| ccccctttc acaccccccag gtgctctggg ccctcttagga actgggtttc tctcaagggc | 17100 |
| tgttacccaa gggtgctccc ttctctcccc ctctcacac cactgggtgc tctgggccc | 17160 |
| ctaggagctg ggattcttta aagaggaaaa ctcttgata aaggaaatgg tttgattgat | 17220 |
| atcgacaag tctgttcatt agtatccatt tattaagcac ctaccatgtg ccaggaaatg | 17280 |
| ctttggcgta caaaggaaaa taagggccag tcctgctaga aatggccttg aaaccccgagg | 17340 |
| gagggatgtc ggcccattgt gggtgctgca gattccttga aggtgatgca agagccagaa | 17400 |
| agaaggatga tggggggggc tgaggcaggg agtcgggggtt gggggagtgt gggggagaag | 17460 |
| gggagaccga gcacctcttc cactatctcc ctgtgtgggtt tttggtaac catcctgcct | 17520 |
| ctgggtgtct tgctccagc ttctgacgtt ggaagttcat ccactgagag ctctgtgttt | 17580 |
| atggctctga gatactgagt ctttcttctc tccagacac cttacaatac catggcattt | 17640 |
| gacggcattt gcggggattt tgagacaatac tttcatgttc tatggagccg gataccaagt | 17700 |
| agaaaaatg atttctcatc caaattatga ctccaagacc aagaacaatg acattgcct | 17760 |
| gatgaagctg cagaagcctc tgactttcaa cggtacgtgt ggctcaggct tggcaagcag | 17820 |
| gttggcagaa tcttaaagag atgttggattt gaaatgacac ttgtgctatg ccaaattggaa | 17880 |
| gggaggcatt tgcgttgagc gagggtagcg tgcagcgggt gccaatggg agaggctcac | 17940 |
| agaggctaag agcacctgcc gcattttggg ggaggcagca gccaccacat ctgttctgt | 18000 |
| ctgtactgag tgggtgtat tcaagccagg catggaaaag gctagaacag ggctttccca | 18060 |
| ctgcagcacc cttgacatct gggtgggtct ctgtttagg gctctcttgc gcctttagg | 18120 |
| atgtttaaca gcgtccccag cctctaccca ctggaggcca gtagctacca agctgtgaca | 18180 |
| accagtgttgc cctgctgaca ttgccaaaca tccgcttga ggcaaagtca cttccagttt | 18240 |
| agaactactg gcctaaaatg tgtaaagatc cttgattttt aaagatacat tctaaaacca | 18300 |
| agttgcttaa ttcaaggacaa acatgcttcc tttagccctc ttattggc ccactctgg | 18360 |
| ccatccaaagg gtctggaatg ttctagcccc atgtggatac agaagaagca aaacctcagc | 18420 |
| cctccctaca gcatgtctgt attcacattt gggaaatgggtt cacatataga agagcgaatg | 18480 |
| cctgagcaat ggcgtggcgc ctctggggcg aaagctgact ccattgactc catggcttt | 18540 |
| ttggctgttgc cctcctgtgt gtctttcccg tcttgcac ctggagatat gtaattttgg | 18600 |
| aagcagagct agcaaataat tcctcttata agcagagcta gcaaataatt ctacttataa | 18660 |
| gtagcataac gtcttgactg ccagaaggag aggtctggca gggggagaaaa gtgagaatgt | 18720 |

ES 2 886 958 T3

| | | | | | | |
|-------------|------------|------------|-------------|------------|-------------|-------|
| gggacttgtt | gggatgcagg | gtcctctggg | cagggtgccc | agggtgccag | gcccagcagc | 18780 |
| ctgcattgtgg | gaaggccagg | tggagacata | ggtgatacc | gcctggctca | ctgtgtttc | 18840 |
| tcttcctgaa | acagacctag | tgaaccagt | gtgtctgcc | aaccaggca | tgtgctgca | 18900 |
| gccagaacag | ctctgctgga | tttccgggtg | ggggccacc | gaggagaaag | gtgaggctgc | 18960 |
| tcctggcac | acaggactgc | agggcccaca | gatggagcat | tgggtcgga | agtgggaggt | 19020 |
| ccaggttta | atcccagttc | tactactcaa | tgactggatg | actttggttg | attccccag | 19080 |
| tccttgc | tcagttctc | catctgctaa | gtgggagaaa | tcctgcccag | cctaccta | 19140 |
| acactgtgtt | cttatcgtga | tcacacagag | cagcatgtgg | aatggcttt | gaagtatctg | 19200 |
| ggccatacga | gttagaggt | gcaggatctc | ctgtgttgca | ctcattgtga | gtttagagct | 19260 |
| gcccctggaga | tcccaccaag | gcctgcgtgg | ctgagtgaca | gggggcttgg | tgaggacggg | 19320 |
| catcctggac | ccatggtggc | cacatctaag | cctgtcctct | gcccgtataa | ccacagagag | 19380 |
| aggctctctc | cacccacttc | cttgcaatc | tgcatattctc | tctgacagtc | tttcaaata | 19440 |
| agggagcctg | gctgcttcat | tttatggag | ggttggaaat | gcttagtggc | aggcacaaag | 19500 |
| gttcatttta | catattgttt | atatccttct | caaaagcgtc | taggcatac | agacaacaaa | 19560 |
| tcctttcaaa | caagggaaaa | agtacaaagg | ttgggtgatt | tctggggagc | gtcagggaaag | 19620 |
| gtagtgggg | gcatcctggc | tcctcatcag | cagaaactta | ctacagtaga | gccacaggct | 19680 |
| ggcaaaaga | cctcatggaa | tccaagatga | agggatatc | gacaaatatt | tgtgcgcacc | 19740 |
| tgcacctagt | acaggctggg | tgctactcag | gtgctggaa | tgcagaagt | aacagagtaa | 19800 |
| gacaaatgtc | tctgctgtca | ggagctttac | ctctttctg | gatgtcggtg | gtggggacgg | 19860 |
| ggcaggtgtg | gtcagacaga | tgggagacaa | acaactgagc | gaggtacttc | caaacatctg | 19920 |
| agggtgggaa | tcacaaggc | ccggctattt | tgaagggtg | gtcaggaaag | gcttcgtgga | 19980 |
| agaggtggca | tttagctga | gactcaaata | gcaaaaatgt | gtacacatca | aaaaggctag | 20040 |
| tgcattgtatc | ttcaggtgtg | gtcaaggggc | caaggaggtg | ggctggggcc | agattgcata | 20100 |
| ggtcctgtg | gattatggtg | aagacaccag | cttctcatct | gcttgaggtg | gggagatcgt | 20160 |
| gagccgggaa | gtgccatgt | ctggcagctg | cgtggggagt | ggggatgaat | ggtggagac | 20220 |
| gaggatgtg | gtgacaagtc | cattgctgtg | gttcctttag | acaggaagcc | agctcatagc | 20280 |
| agagtgcggg | cgtggatgtg | aagagatgag | ggtacactag | ggctagagcc | accagactta | 20340 |
| ctgatgggtt | gcatgtctgt | gggagagaga | gtgagaagtc | agggacgatg | gctttccact | 20400 |
| ctgtggctga | agccccaggg | tggcgggtgg | tgccattttt | caagccagga | aatattggtt | 20460 |
| ggtgagaatt | tgggtggga | gaaggtgtga | cggagggttc | tggtttgca | cactaagccc | 20520 |
| acggtgccca | gaagatgccc | gaggggaggc | agcaaagcga | gagtggaaaa | tgcagaggtg | 20580 |

ES 2 886 958 T3

| | | | | | | |
|-------------|-------------|-------------|------------|------------|-------------|-------|
| gcaagtgcag | gccgtgtctt | gagaagctct | aatgtgcagg | ggagccgaga | agcaggcgac | 20640 |
| ctaggagggg | tcacgtgtgc | tccagaagag | tgtgtgcatt | ccagagggga | aacaggcgcc | 20700 |
| tgtgtgtcct | gggtggggtt | cagtgaggag | tggaaattt | gttcagcaga | accaagccgt | 20760 |
| tgggtgaata | agagggggat | tccatggcac | tgatagagcc | ctatagtttc | agagctggga | 20820 |
| atttcttcc | ctgaagctga | actccagagc | tgcattcagc | acaggcaccg | ccagttgtaa | 20880 |
| ggagaatcca | ggttcccag | gagaggggtt | ggtgctggga | tgagctgacc | ggggcagggc | 20940 |
| tggaaaatag | ggctgtgacc | atctgtgtag | tgcgtgtgg | ggtctcaggg | agggaaagtgt | 21000 |
| gctctccctg | cgagagctgc | aggcaacact | gggagctcaa | caagtctccc | tgtccttagg | 21060 |
| gaagacctca | gaagtgctga | acgctgccaa | ggtgcttctc | attgagacac | agagatgcaa | 21120 |
| cagcagatat | gtctatgaca | acctgatcac | accagccatg | atctgtgccg | gcttcctgca | 21180 |
| ggggaacgtc | gattcttgcc | agtaattca | acattttat | tctacctttg | gtccttacca | 21240 |
| gatcctactg | aaccccccatt | gagagagagg | gcattcttg | ggtcagcaga | gcctcctcag | 21300 |
| tgacacggag | ccagctcggt | gcagtcatgg | gaagtgacgg | ccacaaacag | tgcgaacgct | 21360 |
| tctggggca | gaaggaagta | cagtcaacaa | atcacacaca | ccctctgaaa | aaccggatt | 21420 |
| tggtaaaagt | gccagtggaa | cagaaacaag | tattagact | atttaaatt | atgaacggca | 21480 |
| atttatttag | taacttttag | tttgaacaga | ttaaaattca | ggatgggggc | tatctcttt | 21540 |
| ggggttacat | ctctgttacc | atcaccctt | gatggtgag | attcgaagcc | cacacagtca | 21600 |
| ctcgtaactc | acactgcgac | cccccccccc | caactcctct | aggcctggc | agtggtgtgc | 21660 |
| ggcagattgt | gacttgattt | tctgctctct | gtaccttgct | gtgtcccaca | gggtgacagt | 21720 |
| ggagggcctc | tggtcacttc | gaagaacaat | atctggggc | tgatagggga | tacaagctgg | 21780 |
| ggttctggct | gtgccaaagc | ttacagacca | ggagtgtacg | ggaatgtgat | ggtattcacf | 21840 |
| gactggattt | atcgacaaat | gaggtaact | atcctgtct | ccttctgact | gtgttctccg | 21900 |
| attcctcgag | ccaaagccag | acatctgtta | ggcgtggttc | tgctgctgg | agctgactgg | 21960 |
| tgaccactgg | ttagcatgaa | gcaaactctg | cttcctccag | ccacagcccc | atccccccat | 22020 |
| tgtccaccca | ttgcccattt | cctctcaactg | gcttcacttg | catattcccc | ctgggtttt | 22080 |
| gatgaaaagc | gctggggctc | agcttggtg | aaattcctt | gtgctctgcc | aaccacactt | 22140 |
| cgttctggct | cagctgactc | agctgttcca | cccaggccac | ctcacatcaa | acttttttt | 22200 |
| ttttttttt | agatggagtc | tcactgtgtc | gcccaggctg | gagtgcagtg | gcacaatctc | 22260 |
| gactcaactgc | aaccttgc | tcctgggttc | aagtgattct | cctgcctcag | cctcccaagt | 22320 |
| agctggact | acaggcatgc | gccaccacgc | ccagctactt | tttgttattt | tagtagagat | 22380 |
| ggggttctc | catgttggcc | aggctggct | cgaagccctg | acctcaggtg | attcacccac | 22440 |
| ctcagcctcc | cacagtgctg | ggattacaag | tgtgaaccac | ggtgcccggc | ctcacatgaa | 22500 |

ES 2 886 958 T3

| | |
|--|-------|
| acttttgatt tatagagagc agagggaaga gccggctgtg cccatcctt tctggggcca | 22560 |
| tcgagtggct cctgggcagc ccccaagggtt aggaagggcgaa ggagcagccg gggttctctg | 22620 |
| atgccccaga ctcaagcacg aggaaaggc tcaggggttc catgtgagcc tcatggatgt | 22680 |
| ctctgcttag cagagccctg gcttggca ttgtccagat agggggtag aaccagatct | 22740 |
| tctcatctcc aggacctcg acgtatagtt ttctcagatt tctgtgcttt ctggggctgg | 22800 |
| gctactatgt gaagaaagca gtctattctg tcttctccca aatctcccag atgcccagtc | 22860 |
| tgttgaagga ggagcagaac cagggggct ttcccgtga ggcccacct gtgtctccctt | 22920 |
| caaatgacac gcgggactca gggccttccc atgaccatgg ggcccagggg gcgtcacctg | 22980 |
| gcccaaggcc cagtgttaga aacagatgac cccaggagga ggaggcaggg caggaggaa | 23040 |
| gctggcaggg ctggatggt cagccaggct gagggggcggg ctcgcaccag gatggagcta | 23100 |
| ggaaatgatc caggtgtgtt tggcggctgc aggtgggtcc gcatggctgt gcagggaggg | 23160 |
| aagggctgctg tggcaggaga gcagccgggg gaggcccaga ctctgctgaa gagatgcctg | 23220 |
| ttgtgccggc ctccacatcc gctgcccgt ccttccggag ctccctgcccc gccatgctca | 23280 |
| gcctgactct gaccaacacg ttggagagaa gaatgatccc tttgtgctat taagcttgc | 23340 |
| tatttggttt ctaagtgcctt catgcgaacc tagaggaaaa aattatttc cacctttgtt | 23400 |
| tgtcttaaga aaataacaca cttttttttt tcctatttga acaggcagac ggctaattcca | 23460 |
| catggtcttc gtccttgacg tcgtttaca agaaaacaat gggctgggtt ttgcttcccc | 23520 |
| gtgcatgatt tactctttaga gatgattcag aggtcacttc atttttatta aacagtgaac | 23580 |
| ttgtctggct ttggcactct ctgccattct gtgcaggctg cagtggctcc cctgcccagc | 23640 |
| ctgctctccc taacccttg tccgcaaggg gtgatggccg gctgggtgtg ggcactggcg | 23700 |
| gtcaagtgtg gaggagaggg gtggaggctg ccccatttagat ctttactgc tgagtccctt | 23760 |
| ccaggggcca attttggatg agcatggagc tgtcacctct cagctgctgg atgactttag | 23820 |
| atgaaaaagg agagacatgg aaagggagac agccaggtgg cacctgcagc ggctgccctc | 23880 |
| tggggccact tggtagtgac cccagctac ctctccacaa ggggatttttgc tggatgggtt | 23940 |
| cttagagcct tagcagccct ggttgtggc cagaaataaa gggaccagcc cttcatgggt | 24000 |
| ggtgacgtgg tagtcacttg taagggaaac agaaacattt ttgttcttat ggggtgagaa | 24060 |
| tatagacagt gcccttgggt cgagggaaagc aattgaaaag gaacttgccc tgagcactcc | 24120 |
| tggtgcaaggctt ccacccctgc acattgggtg gggctcctgg gagggagact cagccttcct | 24180 |
| cctcatccctc cctgaccctg ctcctagcac cctggagagt gcacatgccc cttgggtcctg | 24240 |
| gcagggcgcc aagtctggca ccatgttggc ctcttcaggc ctgctagtca ctggaaatttgc | 24300 |
| aggccatgg gggaaatcaa ggatgctcag tttaaggtac actgtttcca tgttatgttt | 24360 |

ES 2 886 958 T3

| | |
|--|-------|
| ctacacattg ctacacctagt gctcctggaa acttagcttt tgatgtctcc aagtagtcca | 24420 |
| ccttcattta actctttgaa actgttatcat cttgccaaag taagagtggg ggcctattc | 24480 |
| agctgcttg acaaaaatgac tggctcctga cttAACGTTc tataaatgaa tgtgctgaag | 24540 |
| caaagtgcCcC atggTggcgg cgaagaagag aaagatgtgt tttgtttgg actctctgtg | 24600 |
| gtcccttcca atgctgtggg ttccaacca gggaaagggt ccctttgca ttgccaagtg | 24660 |
| ccataaccat gagcactact ctaccatggt tctgcctcct ggccaagcag gctggttgc | 24720 |
| aagaatgaaa tgaatgattc tacagctagg acttaacctt gaaatggaaa gtcatgcaat | 24780 |
| cccatttgca ggatctgtct gtgcacatgc ctctgttagag agcagcattc ccagggacct | 24840 |
| tggaaacagt tggcactgta aggtgcttgc tccccaaagac acatcctaaa aggtgttgc | 24900 |
| atggtgaaaa cgcttcctt ctttattgcc ccttcttatt tatgtgaaca actgtttgc | 24960 |
| ttttttgtta tcttttttaa actgtaaagt tcaattgtga aaatgaatat catgcaaata | 25020 |
| aattatgcaa ttttttttc aaagtaacta ctgcattttt gaagttctgc ctggtgagta | 25080 |
| ggaccagcct ccatttcctt ataagggggat gatgttgagg ctgctggtca gaggacaaa | 25140 |
| ggtgaggcaa ggccagactt ggtgctcctg tggttctcgataacttcg tataatgtat | 25200 |
| gctatacgaa gttatgctag taactataac ggtcctaagg tagcgagcta gctccacgtg | 25260 |
| gctttgtccc agacttcctt tgtcttcaac aaccttctgc aagaaaacca agggcctgaa | 25320 |
| ttttaacttc ctg | 25333 |

<210> 7
<211> 491
5 <212> PRT
<213> Secuencia artificial

<220>
<223> Proteína recombinante
10 <400> 7

ES 2 886 958 T3

Met Ala Leu Asn Ser Gly Ser Pro Pro Gly Ile Gly Pro Cys Tyr Glu
1 5 10 15

Asn His Gly Tyr Gln Ser Glu His Ile Cys Pro Pro Arg Pro Pro Val
20 25 30

Ala Pro Asn Gly Tyr Asn Leu Tyr Pro Ala Gln Tyr Tyr Pro Ser Pro
35 40 45

Val Pro Gln Tyr Ala Pro Arg Ile Thr Thr Gln Ala Ser Thr Ser Val
50 55 60

Ile His Thr His Pro Lys Ser Ser Gly Ala Leu Cys Thr Ser Lys Ser
65 70 75 80

ES 2 886 958 T3

Lys Lys Ser Leu Cys Leu Ala Leu Ala Leu Gly Thr Val Leu Thr Gly
85 90 95

Ala Ala Val Ala Ala Val Leu Leu Trp Lys Phe Met Gly Ser Lys Cys
100 105 110

Ser Asn Ser Gly Ile Glu Cys Asp Ser Ser Gly Thr Cys Ile Asn Pro
115 120 125

Ser Asn Trp Cys Asp Gly Val Ser His Cys Pro Gly Gly Glu Asp Glu
130 135 140

Asn Arg Cys Val Arg Leu Tyr Gly Pro Asn Phe Ile Leu Gln Val Tyr
145 150 155 160

Ser Ser Gln Arg Lys Ser Trp His Pro Val Cys Gln Asp Asp Trp Asn
165 170 175

Glu Asn Tyr Gly Arg Ala Ala Cys Arg Asp Met Gly Tyr Lys Asn Asn
180 185 190

Phe Tyr Ser Ser Gln Gly Ile Val Asp Asp Ser Gly Ser Thr Ser Phe
195 200 205

Met Lys Leu Asn Thr Ser Ala Gly Asn Val Asp Ile Tyr Lys Lys Leu
210 215 220

Tyr His Ser Asp Ala Cys Ser Ser Lys Ala Val Val Ser Leu Arg Cys
225 230 235 240

Ile Ala Cys Gly Val Asn Leu Asn Ser Ser Arg Gln Ser Arg Ile Val
245 250 255

Gly Gly Glu Ser Ala Leu Pro Gly Ala Trp Pro Trp Gln Val Ser Leu
260 265 270

His Val Gln Asn Val His Val Cys Gly Gly Ser Ile Ile Thr Pro Glu
275 280 285

Trp Ile Val Thr Ala Ala His Cys Val Glu Lys Pro Leu Asn Asn Pro
290 295 300

Trp His Trp Thr Ala Phe Ala Gly Ile Leu Arg Gln Ser Phe Met Phe
305 310 315 320

Tyr Gly Ala Gly Tyr Gln Val Glu Lys Val Ile Ser His Pro Asn Tyr

ES 2 886 958 T3

325

330

335

Asp Ser Lys Thr Lys Asn Asn Asp Ile Ala Leu Met Lys Leu Gln Lys
 340 345 350

Pro Leu Thr Phe Asn Asp Leu Val Lys Pro Val Cys Leu Pro Asn Pro
 355 360 365

Gly Met Met Leu Gln Pro Glu Gln Leu Cys Trp Ile Ser Gly Trp Gly
 370 375 380

Ala Thr Glu Glu Lys Gly Lys Thr Ser Glu Val Leu Asn Ala Ala Lys
 385 390 395 400

Val Leu Leu Ile Glu Thr Gln Arg Cys Asn Ser Arg Tyr Val Tyr Asp
 405 410 415

Asn Leu Ile Thr Pro Ala Met Ile Cys Ala Gly Phe Leu Gln Gly Asn
 420 425 430

Val Asp Ser Cys Gln Gly Asp Ser Gly Gly Pro Leu Val Thr Ser Lys
 435 440 445

Asn Asn Ile Trp Trp Leu Ile Gly Asp Thr Ser Trp Gly Ser Gly Cys
 450 455 460

Ala Lys Ala Tyr Arg Pro Gly Val Tyr Gly Asn Val Met Val Phe Thr
 465 470 475 480

Asp Trp Ile Tyr Arg Gln Met Arg Ala Asp Gly
 485 490

<210> 8

<211> 2267

<212> ADN

<213> *Mus musculus*

<400> 8

| | | | | | | |
|-------------|------------|------------|-------------|-------------|------------|-----|
| ccgggttgtt | tataggactt | gaccagcccc | aatagtccctc | aagtcaactcc | tagatacagt | 60 |
| ggcaggtgg | tggt | agctggcttg | cggaaggaag | aggaagaaga | aatgtgggc | 120 |
| caaggccagc | cttgcacttg | ggccccctct | gctcagtgt | gaccagggt | ttctgagccg | 180 |
| cttcctaattg | aggctcattt | gaagaccccc | ccccaccccc | ctcctgctgt | cttgggtggc | 240 |
| agagctagct | ccaggctgta | agaaaattag | gaggattacc | aaagcagtat | ggagtcagac | 300 |
| agtggccaac | ccctcaacaa | ccgtgatatt | gttcccttcc | gcaaaccccg | aaggccccag | 360 |
| gagacccatc | aaaaggtggg | gatccccatc | attgcagtgc | tgctgagcct | gatagccctc | 420 |

ES 2 886 958 T3

| | |
|---|------|
| gtgattgtgg cccttctcat caaggtgatt ctggataaat actacttcat ctgcggcagt | 480 |
| ccccctgacct tcattcagag gggccagttg tgtgacggcc accttgactg cgcctcaggg | 540 |
| gaggatgagg aacactgtgt caaggacttc cctgaaaagc ccggagtggc agtccggctc | 600 |
| tccaaggaca gatccaccct gcaggtgctg gatgcagcca cagggacctg ggcctcagtc | 660 |
| tgtttcgaca acttcacaga agcactggcc aagacagcct gcagacagat gggctatgac | 720 |
| agccagcccg ct当地cagagc agtggagatc cgtccagatc agaacctccc tggctcaaa | 780 |
| gtcacaggaa acagccagga acttcaggtg cagaatggaa gcagatcctg cctctcaggc | 840 |
| tccctgggtt ccttgcgctg ctttgcactgt ggaaagagcc tgaagactcc tcgtgtggtg | 900 |
| gggtgggtgg aggcccctgt ggattcttgg ccgtggcagg tcagcatcca gtacaacaag | 960 |
| cagcatgtct gtgggtggag catcctggat ccccactgga tcctcacagc agcccaactgc | 1020 |
| ttcaggaagt atcttgcgtgt gtcaagctgg aaggtcaggg caggctaaa catactgggt | 1080 |
| aactctccat ctttgcctgt ggccaagatc ttcatcgctg aaccctatcc tctgtacccc | 1140 |
| aaagagaagg acattgccct tgttaagctg cagatgccac tcacattctc aggctcagtc | 1200 |
| aggcccatct gcctgccctt ctctgatgag gtgcttgtcc cagccacacc agtctgggtc | 1260 |
| attggatggg gctttacaga agaaaacgga ggaaagatgt ctgacatgct actgcaggca | 1320 |
| tcagtcagg tcattgacag cacacggtgc aatgcagagg atgcctacga agggaaagtg | 1380 |
| accgctgaga tgctgtgtgc aggtacccca cagggtggca aggacacctg ccagggtgac | 1440 |
| agtgggtggc ctttgcgtgtta ccattctgac aagtggcagg tagtaggcat cgtgagctgg | 1500 |
| ggccatggat gccccggccc aagtactcct ggagtgtata ccaaggtcac tgcctatctc | 1560 |
| aactggatct acaatgttcg gaagtctgag atgtaacgct gccgtcccc acatccagaa | 1620 |
| gctgcttccc ttcaagaccta cctacggcat gaccctcaa agtcagatata gggacaagag | 1680 |
| cctccttcaa caaactctgg tatccctgca gcaagcaagg atacattgca gaggtgcccc | 1740 |
| gagtggagtc agatgggcta gctcagccac ccctgcacatct cccaaaccct gggagacatg | 1800 |
| tggcccatgg gagtaaatcc aggacattga ctcaactctc agaagtgtta ttcaagtcaag | 1860 |
| gaggctctcc ttccactga aggaaggaaa gtcagctctc tcctgaaagg ccagatcaact | 1920 |
| ggctgagtag atgagacaag ggtatgaaag gcctttgcca tcttcttgc ccagtcctga | 1980 |
| aagcaactgac gtaagagacc agtcagttct aatgtaaggt gtatattta gtgtcagggt | 2040 |
| attgcaatttgc acctctgt ggtcaatatac attaaacagg tatgagaatt cgctggcata | 2100 |
| gacttcctgg tctgcttaat aagaatccaa ctaaggatgt cacatgacag tttccagaa | 2160 |
| aatgtgaaca agtgcctatc tgacacacgg caccaatgac aaaccaaaga agttattctg | 2220 |
| cctgagtctc agttgctgaa ctaataaatt agctgcgggtt tcttgca | 2267 |

<210> 9
<211> 435
<212> PRT
<213> *Mus musculus*

5

<400> 9

ES 2 886 958 T3

Met Glu Ser Asp Ser Gly Gln Pro Leu Asn Asn Arg Asp Ile Val Pro
1 5 10 15

Phe Arg Lys Pro Arg Arg Pro Gln Glu Thr Phe Lys Lys Val Gly Ile
20 25 30

Pro Ile Ile Ala Val Leu Leu Ser Leu Ile Ala Leu Val Ile Val Ala
35 40 45

Leu Leu Ile Lys Val Ile Leu Asp Lys Tyr Tyr Phe Ile Cys Gly Ser
50 55 60

Pro Leu Thr Phe Ile Gln Arg Gly Gln Leu Cys Asp Gly His Leu Asp
65 70 75 80

Cys Ala Ser Gly Glu Asp Glu Glu His Cys Val Lys Asp Phe Pro Glu
85 90 95

Lys Pro Gly Val Ala Val Arg Leu Ser Lys Asp Arg Ser Thr Leu Gln
100 105 110

Val Leu Asp Ala Ala Thr Gly Thr Trp Ala Ser Val Cys Phe Asp Asn
115 120 125

Phe Thr Glu Ala Leu Ala Lys Thr Ala Cys Arg Gln Met Gly Tyr Asp
130 135 140

Ser Gln Pro Ala Phe Arg Ala Val Glu Ile Arg Pro Asp Gln Asn Leu
145 150 155 160

Pro Val Ala Gln Val Thr Gly Asn Ser Gln Glu Leu Gln Val Gln Asn
165 170 175

Gly Ser Arg Ser Cys Leu Ser Gly Ser Leu Val Ser Leu Arg Cys Leu
180 185 190

Asp Cys Gly Lys Ser Leu Lys Thr Pro Arg Val Val Gly Gly Val Glu
195 200 205

Ala Pro Val Asp Ser Trp Pro Trp Gln Val Ser Ile Gln Tyr Asn Lys
210 215 220

ES 2 886 958 T3

Gln His Val Cys Gly Gly Ser Ile Leu Asp Pro His Trp Ile Leu Thr
 225 230 235 240

 Ala Ala His Cys Phe Arg Lys Tyr Leu Asp Val Ser Ser Trp Lys Val
 245 250 255

 Arg Ala Gly Ser Asn Ile Leu Gly Asn Ser Pro Ser Leu Pro Val Ala
 260 265 270

 Lys Ile Phe Ile Ala Glu Pro Asn Pro Leu Tyr Pro Lys Glu Lys Asp
 275 280 285

 Ile Ala Leu Val Lys Leu Gln Met Pro Leu Thr Phe Ser Gly Ser Val
 290 295 300

 Arg Pro Ile Cys Leu Pro Phe Ser Asp Glu Val Leu Val Pro Ala Thr
 305 310 315 320

 Pro Val Trp Val Ile Gly Trp Gly Phe Thr Glu Glu Asn Gly Gly Lys
 325 330 335

 Met Ser Asp Met Leu Leu Gln Ala Ser Val Gln Val Ile Asp Ser Thr
 340 345 350

 Arg Cys Asn Ala Glu Asp Ala Tyr Glu Gly Glu Val Thr Ala Glu Met
 355 360 365

 Leu Cys Ala Gly Thr Pro Gln Gly Gly Lys Asp Thr Cys Gln Gly Asp
 370 375 380

 Ser Gly Gly Pro Leu Met Tyr His Ser Asp Lys Trp Gln Val Val Gly
 385 390 395 400

 Ile Val Ser Trp Gly His Gly Cys Gly Gly Pro Ser Thr Pro Gly Val
 405 410 415

 Tyr Thr Lys Val Thr Ala Tyr Leu Asn Trp Ile Tyr Asn Val Arg Lys
 420 425 430

Ser Glu Met
 435

<210> 10
 <211> 3543
 <212> ADN
 <213> Homo sapiens

<400> 10

ES 2 886 958 T3

atcattccag ttggcaact tcactttag ggctgtttta atcaagctgc ccaaagtccc 60

ES 2 886 958 T3

| | |
|--|------|
| ccaatcactc ctggaataca cagagagagg cagcagcttgc | 120 |
| ggcgtgaggg accaaggcct gccctgcact cgggcctcct ccagccagtgc | 180 |
| acttctgacc tgctggccag ccaggacctg tgtgggagg ccctcctgct gccttgggt | 240 |
| gacaatctca gctccaggct acagggagac cgggaggatc acagagccag catggatcct | 300 |
| gacagtgtac aacctctgaa cagcctcgat gtcaaaccccc tgcgcaaacc ccgtatcccc | 360 |
| atggagacct tcagaaaggt gggatcccc atcatcatag cactactgag cctggcgagt | 420 |
| atcatcattg tgggtgtcct catcaaggtg attctggata aatactactt cctctgcggg | 480 |
| cagcctctcc acttcatccc gaggaagcag ctgtgtgacg gagagctgga ctgtcccttg | 540 |
| ggggaggacg aggagcactg tgtcaagagc ttccccaag ggcctgcagt ggcagtcgc | 600 |
| ctctccaagg accgatccac actgcagggtg ctggactcgg ccacaggaa ctggttctct | 660 |
| gcctgtttcg acaacttcac agaagctc gctgagacag cctgttaggca gatggctac | 720 |
| agcagcaaac ccacttcag agctgtggag attggccag accaggatct ggatgttgg | 780 |
| gaaatcacag aaaacagcca ggagcttcgc atgcggaaact caagtggcc ctgtctctca | 840 |
| ggctccctgg tctccctgca ctgtcttgcc tgtggaaaga gcctgaagac ccccccgttg | 900 |
| gtgggtgggg aggaggcctc tgtggattct tggccttggc aggtcagcat ccagtacgac | 960 |
| aaacagcacg tctgtggagg gagcatcctg gaccccaact gggtcctcac ggcagcccac | 1020 |
| tgcttcagga aacataccga tgtgttcaac tggaaagggtgc gggcaggctc agacaaaactg | 1080 |
| ggcagcttcc catccctggc tgtggccaag atcatcatca ttgaattcaa ccccatgtac | 1140 |
| cccaaagaca atgacatcgc cctcatgaag ctgcagttcc cactcacttt ctcaggcaca | 1200 |
| gtcaggccca tctgtctgcc cttcttgat gaggagctca ctccagccac cccactctgg | 1260 |
| atcattggat gggctttac gaagcagaat ggagggaaaga tgtctgacat actgctgcag | 1320 |
| gcgtcagtcc aggtcattga cagcacacgg tgcaatgcag acgatgcgta ccagggggaa | 1380 |
| gtcaccgaga agatgatgtg tgcaaggatc ccggaaagggg gtgtggacac ctgcagggt | 1440 |
| gacagtggtg ggcccctgat gtaccaatct gaccagtggc atgtggtggg catcgtagt | 1500 |
| tggggctatg gctgcggggg cccgagcacc ccaggagtat acaccaaggt ctcagcctat | 1560 |
| ctcaacttgg a tctacaatgt ctggaaaggct gagctgtaat gctgctgccc cttgcagtg | 1620 |
| ctgggagccg cttcccttcct gccctgcaca cctggggatc ccccaaagtc agacacagag | 1680 |
| caagagtccc cttgggtaca cccctctgcc cacagcctca gcatttcttg gagcagcaaa | 1740 |
| ggcctcaat tcctataaga gaccctcgca gcccagaggc gcccagagga agtcagcagc | 1800 |
| cctagctcgg ccacacttgg tgctcccagc atcccaggaa gagacacagc ccactgaaca | 1860 |
| aqgtctcaqq qqtattqcta aqccaqaqaq qaactttccc acactactqa atqqaqaqcq | 1920 |

| | | | | | | |
|-------------|--------------|-------------|-------------|-------------|-------------|------|
| gctgtcttgt | aaaagcccag | atcactgtgg | gctggagagg | agaaggaaag | ggtctgcgcc | 1980 |
| agccctgtcc | gtcttcaccc | atccccaaagc | ctactagagc | aagaaaccag | ttgtaatata | 2040 |
| aatgcactg | ccctactgtt | ggtatgacta | ccgttaccta | ctgttgtcat | tgttattaca | 2100 |
| gctatggcca | ctattattaa | agagctgtgt | aacatctctg | gcataggcta | gctggaatgc | 2160 |
| ttgataagaa | ctgagctggg | atgattgaac | tttcattctt | tggcttgggg | agaaaagaag | 2220 |
| tcctggggaa | gcaattgagt | ctcaaagttag | aggcagggga | aaaaagagtt | agggagacca | 2280 |
| gatctgctga | gtggcagcaa | gagtgagctg | cagattacag | aaaccagggt | gagcaagttt | 2340 |
| gagtcccaca | cagggccttc | tcccttgcc | tcttccctc | cctccctgcc | tgtgataatc | 2400 |
| agccaggagc | cagggataac | ctatgacttg | gaaaaagagat | gagttaggca | gtcaagggtg | 2460 |
| acattcaatc | aggatccac | aagtggctgg | aaagaaatgc | tggtcctgtg | tcctaacttt | 2520 |
| ttccgcctgg | agagccctca | gtgtggcttc | ttacatttaa | aaaacaaaaaa | ggatcagctg | 2580 |
| ccaggtgtga | ggcagtc(ccc) | aagctgagtt | gtgaggatgt | aagcatgaat | aagtccctgc | 2640 |
| actcaaaatg | gtcaaagaat | taaacccat | ggactttttt | ggcatctgta | tgaaagcttg | 2700 |
| ggtttctga | ggactgtctt | gctatagtt | agtcatatcc | tagatgaaat | atacttgttc | 2760 |
| atactgtact | aggttcttag | gaaacaacag | aattcctcaa | atgccaaaaaa | caaagaaaaat | 2820 |
| agaaacccag | aaaacaaaaac | aaaataaaac | aaaaccatca | gaactgtgag | tggaaactaa | 2880 |
| ggtgatgatc | tgggagcaat | acactaaaat | cttgggtcga | gacctatatg | aaggctggca | 2940 |
| gtggagctaa | acctggacac | actgaagaca | agggagctga | accagggctc | ctacatgaag | 3000 |
| cagggataac | tcatggcagt | aatgtggtc | tcaaattgca | gatggctgg | aggaaaattt | 3060 |
| cccaaattta | gagcctcagg | attccaaag | atcctccaaa | tatgagctca | caatcaaaga | 3120 |
| tcagagacgt | tgaaaaataa | aaaacacctt | aagtggcag | cataaaaaac | agctaattta | 3180 |
| gaaccccaaa | ggcttcagat | gtcagaatat | tagagactta | tgataataag | caatatttgc | 3240 |
| agagtatttgc | tatgtgccag | acactattgt | aagtgcctca | tcatgtactg | attcatttaa | 3300 |
| tactcacaga | aatctgtgag | atgggtatta | ttcttatcct | cactctatgg | attaaaaaaaa | 3360 |
| ctaaggcaca | aagtggtaa | gctccttgcc | tgagattata | gactgtaagt | tgaacgtgag | 3420 |
| cacttggaaat | acagagttca | tgctgtaaac | taccacacta | tagggcctcc | aatatgataaa | 3480 |
| tttataaaaat | atttgaataa | aaaatgaata | ctagttccac | atttaaaaaa | aaaaaaaaaa | 3540 |
| aaa | | | | | | 3543 |

<210> 11
<211> 437
<212> PRT
<213> *Homo sapiens*

<400> 11

ES 2 886 958 T3

Met Leu Gln Asp Pro Asp Ser Asp Gln Pro Leu Asn Ser Leu Asp Val
1 5 10 15

Lys Pro Leu Arg Lys Pro Arg Ile Pro Met Glu Thr Phe Arg Lys Val
20 25 30

Gly Ile Pro Ile Ile Ala Leu Leu Ser Leu Ala Ser Ile Ile Ile
35 40 45

Val Val Val Leu Ile Lys Val Ile Leu Asp Lys Tyr Tyr Phe Leu Cys
50 55 60

Gly Gln Pro Leu His Phe Ile Pro Arg Lys Gln Leu Cys Asp Gly Glu
65 70 75 80

Leu Asp Cys Pro Leu Gly Glu Asp Glu Glu His Cys Val Lys Ser Phe
85 90 95

Pro Glu Gly Pro Ala Val Ala Val Arg Leu Ser Lys Asp Arg Ser Thr
100 105 110

Leu Gln Val Leu Asp Ser Ala Thr Gly Asn Trp Phe Ser Ala Cys Phe
115 120 125

Asp Asn Phe Thr Glu Ala Leu Ala Glu Thr Ala Cys Arg Gln Met Gly
130 135 140

Tyr Ser Ser Lys Pro Thr Phe Arg Ala Val Glu Ile Gly Pro Asp Gln
145 150 155 160

Asp Leu Asp Val Val Glu Ile Thr Glu Asn Ser Gln Glu Leu Arg Met
165 170 175

Arg Asn Ser Ser Gly Pro Cys Leu Ser Gly Ser Leu Val Ser Leu His
180 185 190

Cys Leu Ala Cys Gly Lys Ser Leu Lys Thr Pro Arg Val Val Gly Gly
195 200 205

Glu Glu Ala Ser Val Asp Ser Trp Pro Trp Gln Val Ser Ile Gln Tyr
210 215 220

Asp Lys Gln His Val Cys Gly Ser Ile Leu Asp Pro His Trp Val
225 230 235 240

Leu Thr Ala Ala His Cys Phe Arg Lys His Thr Asp Val Phe Asn Trp
245 250 255

ES 2 886 958 T3

Lys Val Arg Ala Gly Ser Asp Lys Leu Gly Ser Phe Pro Ser Leu Ala
 260 265 270

Val Ala Lys Ile Ile Ile Glu Phe Asn Pro Met Tyr Pro Lys Asp
 275 280 285

Asn Asp Ile Ala Leu Met Lys Leu Gln Phe Pro Leu Thr Phe Ser Gly
 290 295 300

Thr Val Arg Pro Ile Cys Leu Pro Phe Phe Asp Glu Glu Leu Thr Pro
 305 310 315 320

Ala Thr Pro Leu Trp Ile Ile Gly Trp Gly Phe Thr Lys Gln Asn Gly
 325 330 335

Gly Lys Met Ser Asp Ile Leu Leu Gln Ala Ser Val Gln Val Ile Asp
 340 345 350

Ser Thr Arg Cys Asn Ala Asp Asp Ala Tyr Gln Gly Glu Val Thr Glu
 355 360 365

Lys Met Met Cys Ala Gly Ile Pro Glu Gly Val Asp Thr Cys Gln
 370 375 380

Gly Asp Ser Gly Gly Pro Leu Met Tyr Gln Ser Asp Gln Trp His Val
 385 390 395 400

Val Gly Ile Val Ser Trp Gly Tyr Gly Cys Gly Pro Ser Thr Pro
 405 410 415

Gly Val Tyr Thr Lys Val Ser Ala Tyr Leu Asn Trp Ile Tyr Asn Val
 420 425 430

Trp Lys Ala Glu Leu
 435

<210> 12

<211> 20078

<212> ADN

<213> Secuencia artificial

<220>

<223> Polinucleótido recombinante

10

<400> 12

ccaccggcac acactacagt cgagataact tcgtataatg tatgtataac gaagttatat 60
 gcatggcctc cgcgcgggt tttggcgctt cccgcggcgccccctcct cacggcgagc 120

ES 2 886 958 T3

| | |
|---|------|
| gctgccacgt cagacgaagg ggcgcagcgag cgtcctgatc cttccgcccc gacgctcagg | 180 |
| acagcggccc gctgctata agactcggcc ttagaacccc agtatacagca gaaggacatt | 240 |
| ttaggacggg acttgggtga ctctagggca ctgggtttct ttccagagag cgaaacaggc | 300 |
| gaggaaaagt agtcccttct cggcgttct gcggaggat ctccgtgggg cggtgaacgc | 360 |
| cgtatgattat ataaggacgc gccgggtgtg gcacagctag ttccgtcgca gccgggattt | 420 |
| gggtcgcggt tcttgggtgt ggatcgctgt gatcgtaact tggtgagtag cgggctgctg | 480 |
| ggctggccgg ggcttcgtg gccgcgggc cgctcggtgg gacggaagcg tgtggagaga | 540 |
| ccgccaaggc ctgtatgtc ggtccgcgag caaggttgcc ctgaactggg gttgggggg | 600 |
| agcgcagcaa aatggcggct gttcccgagt cttgaatgga agacgcttgt gaggcggcgt | 660 |
| gtgaggcgt tgaaacaagg tggggggcat ggtggcggc aagaacccaa ggtcttggg | 720 |
| ccttcgctaa tgcgggaaag ctcttattcg ggtgagatgg gctggggcac catctggga | 780 |
| ccctgacgtg aagtttgtca ctgactggag aactcggtt gtcgtctgtt gccccggcgg | 840 |
| cagttatggc ggtgccgtt ggcagtgcac ccgtacctt gggagcgcgc gccctcgctg | 900 |
| tgtcggtacg tcaccgttc tgggtgtta taatgcaggg tggggccacc tgccggtagg | 960 |
| tgtcggttag gctttctcc gtcgcaggac gcagggttcg ggcctagggt aggctctcct | 1020 |
| gaatcgacag gcccggacc tctggtgagg ggaggataa gtgaggcgac agtttcttg | 1080 |
| gtcggttta tgtacctatc ttcttaagta gctgaagctc cggtttgaa ctatgcgc | 1140 |
| gggggtggcg agtgtgttt gtgaagttt ttaggcacct tttgaatgt aatcatttg | 1200 |
| gtcaatatgt aatttcagt gttagactag taaattgtcc gctaaattct gggcgtttt | 1260 |
| ggctttttt ttagacgtgt tgacaattaa tcatcgcat agtatacgg catagtataa | 1320 |
| tacgacaagg tgaggaacta aaccatggga tcggccattt aacaagatgg attgcacgc | 1380 |
| ggttctccgg ccgcttgggt ggagaggcta ttcggctatg actgggcaca acagacaatc | 1440 |
| ggctgctctg atgcgcgcgt gttccggctg tcaagcgcagg ggcccccgt tcttttgc | 1500 |
| aagaccgacc tgcgggtgc cctgaatgaa ctgcaggacg aggcaagcg gctatcggt | 1560 |
| ctggccacga cgggcgttcc ttgcgcagct gtgcgtcgacg ttgtcaactga agcgggaagg | 1620 |
| gactggctgc tattggcga agtgcgggg caggatctcc tgcatacgat cttgtctcct | 1680 |
| gccgagaaag tatccatcat ggctgtatgca atgcggcggc tgcatacgat tgatccggct | 1740 |
| acctgccccat tcgaccacca agcgaacat cgcacgcgc gacgcgtac tcggatggaa | 1800 |
| gccggcttt tcgatcgatc tgatctggac gaagagcatc aggggctcgc gccagccgaa | 1860 |
| ctgttcgcca ggctcaaggc ggcgcgtcccc gacggcgtatc atctcgctgt gacccatggc | 1920 |
| gatgcctgct tgccgaatat catggtgaa aatggccgct tttctggatt catcgactgt | 1980 |
| ggccggctgg gtgtggcggc ccgctatcag gacatagcgt tggctacccg tgatattgct | 2040 |

ES 2 886 958 T3

| | |
|--|------|
| gaagagcttgcggcgaatggctgaccgc ttccctcgatcttacggtat cgccgcgtccc | 2100 |
| gattcgacgc gcatcgccctt ctatcgcctt ctgtacgagt ttttctgagg ggatccgctgtca | 2160 |
| taagtctgca gaaattgatg atctattaaa caataaaagat gtccactaaa atgaaagttt | 2220 |
| ttccctgtcat actttgttaa gaagggttag aacagagtagtac ctacatggat aatgaaagga | 2280 |
| ttggagctac ggggggtgggg gtgggggtggg attagataaa tgcctgctctt tactgaagg | 2340 |
| ctctttacta ttgctttatg ataatgttc atagttggat atcataattt aaacaagcaa | 2400 |
| aaccaaatta agggccagct cattcctccc actcatgatc tatagatcta tagatctctc | 2460 |
| gtgggatcat tgaaaaatcc ttgattccca ctttgtggtt ctaagtactg tggtttccaa | 2520 |
| atgtgtcagt ttcatagcct gaagaacgag atcagcagcc tctgttccac atacacttca | 2580 |
| ttctcagttat tgaaaaatccca agttctaatt ccatcagacc tcgacctgca gcccttagcc | 2640 |
| cgggcgccag tagcagcacc cacgtccacc ttctgtcttag taatgtccaa cacctccctc | 2700 |
| agtccaaaca ctgctctgca tccatgtggc tcccatggat acctgaagca cttgtatgggg | 2760 |
| cctcaatgtt ttactagagc ccacccccc gcaactctga gaccctctgg atttgcgtgt | 2820 |
| cagtgccctca ctggggcggtt ggataatttc ttaaaaggc aagttccctc agcagcattc | 2880 |
| tctgagcagt ctgaagatgt gtgcgtttca cagttcaaat ccatgtggct gttcacccca | 2940 |
| cctgcctggc ctggggttat ctatcaggac ctgcctaga agcaggtgtg tggcacttaa | 3000 |
| cacctaagct gagtgactaa ctgaacactc aagtggatgc catcttgatc acttcttgac | 3060 |
| tgtgacacaa gcaactcctg atgccaaggc cctgcccacc cctctcatgc ccatatttgg | 3120 |
| acatggtaca ggtcctact ggcattggc tggaggtcc tggcctctt tgacttcata | 3180 |
| attccttaggg gccacttagta tctataagag gaagagggtg ctggctccca ggcacagcc | 3240 |
| cacaaaatttcc cacctgctca caggttggct ggctcgacc aggtgggtgc ccctgctctg | 3300 |
| agccagctcc cggccaaagcc agcaccatgg gtaccccaaa gaagaagagg aaggtgcgtta | 3360 |
| ccgatttaaa ttccaaatttca ctgaccgtac accaaaattt gcctgcattt ccggcgtatg | 3420 |
| caacgagtga tgaggttcgc aagaacctga tggacatgtt cagggatcgc caggcggttt | 3480 |
| ctgagcatac ctggaaaatg cttctgtccg tttgcggcgtc gtggcggca tggcgtcaatgt | 3540 |
| tgaataaccg gaaatggttt cccgcagaac ctgaagatgt tcgcgttattt cttctatatac | 3600 |
| ttcaggcgcg cggctctggca gtaaaaacta tccagcaaca tttggccag ctaaacatgc | 3660 |
| ttcattcgatcg gtccgggctg ccacgaccaaa gtgacagacaa tgctgtttca ctgggttatgc | 3720 |
| ggcggatccg aaaagaaaac gttgatgccc gtgcacgtgc aaaacaggct ctgcgttcc | 3780 |
| aacgcactga ttgcgaccag gttcggttccac tcatggaaaa tagtgcgtgc tgccaggata | 3840 |
| tacgtaatct ggcatttctg gggattgctt ataacaccct gttacgtata gccgaaatttgc | 3900 |

ES 2 886 958 T3

| | |
|--|------|
| ccaggatcg ggttaaagat atctcacgta ctgacggtgg gagaatgtta atccatattg | 3960 |
| gcagaacgaa aacgctggtt agcaccgcag gttagagaa ggcacttagc ctggggtaa | 4020 |
| ctaaactggt cgagcgatgg atttccgtct ctggtgtagc tgatgatccg aataactacc | 4080 |
| tgtttgccg ggtcagaaaa aatggtgtt ccgcgcac tcgcaccagc cagctatcaa | 4140 |
| ctcgccccct ggaagggatt tttgaagcaa ctcatcgatt gattacggc gctaaggtaa | 4200 |
| atataaaatt tttaagtgtta taatgtgtta aactactgat tctaattgtt tgtgtatTTT | 4260 |
| aggatgactc tggtcagaga tacctggcct ggtctggaca cagtgcggcgt gtcggagccg | 4320 |
| cgcgagatat ggcccgcgt ggagttcaa taccggagat catgcaagct ggtggctgg | 4380 |
| ccaatgtaaa tattgtcatg aactatatcc gtaacctgga tagtgaarda gggcaatgg | 4440 |
| tgcgcctgct ggaagatggc gattgatcta gataagtaat gatcataatc agccatatca | 4500 |
| catctgtaga ggtttactt gctttaaaaa acctcccaca cctccccctg aacctgaaac | 4560 |
| ataaaatgaa tgcaattgtt gttgttaaac ctgccttagt tgccgccaat tccagctgag | 4620 |
| cgtgcctccg caccattacc agttggctg gtgtaaaaaa taataataac cgggcagggg | 4680 |
| ggatctaagc tctagataag taatgtatcat aatcagccat atcacatctg tagaggttt | 4740 |
| acttgcttta aaaaacctcc cacacctccc cctgaacctg aaacataaaa tgaatgcaat | 4800 |
| tgttgtgtt aacttgttta ttgcagctta taatggttac aaataaagca atagcatcac | 4860 |
| aaatttcaca aataaagcat tttttcact gcattctagt tgtggttgt ccaaactcat | 4920 |
| caatgtatct tatcatgtct ggaataactt cgataatgt atgctatacg aagttatgct | 4980 |
| agtaactata acggcctaa ggtgcgagc tagccaagtc tgtgtgctac caagtagcaa | 5040 |
| aactgagcct ggaactcaca catgcgtgtc tgagagccca gcaactatcgc caggaaaacc | 5100 |
| cagcgtctcc ctgctcaagc ctgaccctca gccctctctg cctcccccctg cacttgcctt | 5160 |
| ccagtcaagg tgattctgga taaatactac ttctctgctg ggcagctct ccacttcatc | 5220 |
| ccgaggaagc agctgtgtga cggagagctg gactgtccct tggggagga cgaggagcac | 5280 |
| tgtgtcaaga gcttccccga agggcctgca gtggcaggtg agtgcagggc ctgaggcaca | 5340 |
| agagaagtgg gcccagcagg aggtctgctc aggccccac gcccactgc atagtatctg | 5400 |
| ccccctactt gtcacttttca tccttggttataagggtt tttgtttgtt tggtttgtt | 5460 |
| tgttttgagg cagagtgtctc tgtggccaa gatggagtgc agtgtcttgg tctcgctca | 5520 |
| ctgcaacctc tgcctcccaag tttcaagtga ttcttctgcc tcagcctcat gagtagctgg | 5580 |
| gattacaggt gccagccacc acgcctggct aattttata ttttagtag agacggggtt | 5640 |
| ttgccacatt ggtcaggctg atcttgaact cctgacctca ggtgatctgc ccgcctcagc | 5700 |
| ctcccaaagt gctgggatta caggcgtgag ccaccgtgcc cagctgtgt aagttcttga | 5760 |
| gagcaggacc ctgtcttgta tacctttaaa tccttagtact taacacacag caaacagtaa | 5820 |

ES 2 886 958 T3

| | |
|--|------|
| ctatttgatg accaaatgtg agccagaaaag gacaggaaat tgtaactgag gctgccccat | 5880 |
| gcgtgctcg cctgggtggat ttcaggcaga gggctagact gggtagac ttggcattcc | 5940 |
| tcctttctat gaaatttggat atttcaagga gactagaaaa gagacttctc agccacttcg | 6000 |
| ccagctattg gtccttctat tcattagtgt ttgctgagac atgctatgtg acaggactga | 6060 |
| gccaggtcct ttcaatggat aggagatgtt ttgagcataa aatccacgtt ctctttggg | 6120 |
| ctgggctctt ctaccttctt ccccctggtg ctgggctct gaagaaaaaa agataggtag | 6180 |
| gagatgagtg atggggcttc tgagggcagg gctgagtgac tttctgtgtt ttgctctt | 6240 |
| ctttatcaga agtcaaatgc ccacaggcac ctgtcatcct actgccagta ggacttctca | 6300 |
| ctcaaccccttc ccctctgacc ttacttggag aaggacttag gtcctctct cagacatttc | 6360 |
| cccaggctgg gcaagttgtg tggaccatgg atgggtatgt ggtccataca attaaacaa | 6420 |
| gctgtatatg gtcgctgggt agagtgacca cataattgat catcaaaact gatacctgta | 6480 |
| agagcaaaag gggcactat taaccattgg gtcagggcaa caggtcaaaa tggagaccta | 6540 |
| ccctgggact tctggtcaca ctagctactg tcaaaatggg gcccaaatacg acaaagccaa | 6600 |
| atggaagaaa ttcccttgac attgaaatgt ttggggctct gtggcacccc cagttctagg | 6660 |
| ttgggggagc ttgggctggt ctcatgatga gttctgaggg ggatggccca gttgggcccc | 6720 |
| ccgttccatc taactcaggt tccttcctc ccagtccgccc tctccaagga ccgatccaca | 6780 |
| ctgcaggtgc tggactcggc cacagggAAC tggttctctg cctgttcga caacttcaca | 6840 |
| gaagctctcg ctgagacagc ctgttaggcag atgggtacca gcaggttaacc aacctggcc | 6900 |
| tcttccttt ttccctcctt ctccttcctc ctttccttc tcccttcctc | 6960 |
| tctttccct aaaaattacg ggcattggag ccaggcagaa tggctttga atcccagcat | 7020 |
| ttcacttata agcaacatga agttaaattt cctaagcctc agttcctca ggagttaaatt | 7080 |
| gggggaacta atgccaacct cataggatag ttttgcattt ccagttagag aatgtgtgct | 7140 |
| gccctccaac acacacacac acacttcctag cgtctatgca gtcctctct tccctttact | 7200 |
| cctcaacccct cactcctttg tgctggcttt gcaagaaact gttcctgccc agtaatacaa | 7260 |
| aagctaagtt aacttattca aagttcggtt agttaagatt tagcttaagt gaggctagtt | 7320 |
| tcagtggggc cccatcttca gcaatcccag ctctctctgc aaatttcaaa agcagttcca | 7380 |
| aatctggagt ggatgaaaag gtgttaagatg atagtaagag taatttgcatt tctatataatt | 7440 |
| tatattcact tgattttggc agaaaaccaa aaagatagtt attatatctt atatataat | 7500 |
| atatattata tctatttcat aaataggctc aaacaaaagta agtaacttgc tagggtaacta | 7560 |
| gctgggaggt agagggctag aatttgagcc caagacccct aattcttgcg cattaggagt | 7620 |
| tcccacatttgc ttctgtttc tagactgagt aattctttat tctcatgttag gacatcatct | 7680 |

ES 2 886 958 T3

| | |
|--|------|
| ctaagggaag gggctaatga gatgggtgat cactcagaga gtttagctgg agaggatgga | 7740 |
| aaagaaccca tacattcagt tgcagattga gatagcctat ctctggcagg cctcagattt | 7800 |
| cttcaggatt ctaacagact ggaccaggag actaggccaa acaaacaac aaacaaaaac | 7860 |
| tctactaggc agacatcacc aaccaatcac agaactctct cccatggatc cctaatacag | 7920 |
| cctcaaagtc ctttcagta aatgctccag gcagccatta caaatcaatc agaattattt | 7980 |
| gcctttctct tctctgctca acgggcttct gctgctctct actttccata gggggcaact | 8040 |
| tccattaccc tctagaaagc acaccccacc actttcattt caaggagagt gaggaactca | 8100 |
| tgcccagcac ctgctattct cccctttcc tgcagccacg gagccagcc tcgctgcagc | 8160 |
| cagccctgcc tccccactgt agtccagtca actgctgcat cagccgttcc tggcacagca | 8220 |
| ggctgagcct tgattatgaa acctgggtgt ctccaggggt tcttaagatg ataggctcct | 8280 |
| ggaatttctg tccttttggg gctcagtaag gcaccaaacc acctgagtct ttgtcattcac | 8340 |
| aaaatcaaag ttcatcagaa tcattcattt ggatgaaatt ggtgaacaga agttaacttt | 8400 |
| cctgggaatg tccatttcca ccatattccg tccttctagg tctcagactt ctctactttc | 8460 |
| tttcctctct cttagatcgga ggccttctt gtcctagaac cataggcatt tcaagatgtg | 8520 |
| ggagacccta gggatcatct agtccacgca tctttttttt ttttttttga cagagtctca | 8580 |
| ctctgtcacc caggctggag tgcaatggca ccatctctgc ttactgcaac ctccacctcc | 8640 |
| caggttcaag tgatttttc gcctcagcct cccaaatggc tgggattaca ggcacgcacc | 8700 |
| atcatgcccc gctaattttt atattttgt agagaccgag tttcaccatg ttggccaggc | 8760 |
| tggcttgaa ctccctgacct caggtgatcc acccacctcg gcctccaaa gtgctggat | 8820 |
| tacaggcgtg agccactgca cccagccccg tgcattttt tatagagggg gaaactgagg | 8880 |
| cttggagaga cccagaaaaaa gaatatgacc tgcccaaggc cacacatcaa actagtgcac | 8940 |
| gagccagggc cagaacctag atcatgagga ctctaaaat gcactctagt cctccaggt | 9000 |
| ctgagacttg ggtcatttcca ggaagtgcac gcattcctgc ctgagaatgt gccaatccac | 9060 |
| cagtattgcc aatgactcag ccctccatgg agagcttcta ctaacattac tagcatagtt | 9120 |
| agggatggaa ggaaaagatt tagaagaggc agattcagta aaggaacaat cagagagatg | 9180 |
| gaattaaatca aggaaggctt cctggaggag gaaaaacttc aacccaaggt ttgaaagtag | 9240 |
| caagcatgga ttagcagggc gaaagaggga gagtggtcca gttgagagaa acgtttgtct | 9300 |
| ggattcataat gaagacagat ctatcctgt tctattaaat atctctaagg gggccaaaaaa | 9360 |
| catacccccg ctatcaaagt cagaccagat gctttgtttt gagaacgaaa tatccacatt | 9420 |
| ccaaactccct cccaggttag aaggagacta acctgagccc ctatgcctct ttgtttccct | 9480 |
| gctgtgaacc agaagacatt gctggatat ttgaaatagg gacagagctg ggaatatgga | 9540 |
| aaggagaccc ctaacatttc tccagggctc tgggttctgg atttggattc cccacccaag | 9600 |

ES 2 886 958 T3

| | |
|---|-------|
| aaagcaagtt acatcagcaa tgcactgagg gttgagtccct gggatgccaa gggtcggttc | 9660 |
| tttattgtat agcaaagcag gccccatctt cactgactaa gaccatctcc actccctggc | 9720 |
| cactccccac caagcattct ctgccactct ttctcctgaa agtgggggcc aactctacca | 9780 |
| tcttgttcta acccccgtcc ccagctcaca actctctctc cctcttgatg tgagcagcaa | 9840 |
| acccacttcc agagctgtgg agattggccc agaccaggat ctggatgttgg ttgaaatcac | 9900 |
| agaaaacagc caggagcttc gcatgcggaa ctcaagtggg taagtgaggg gacacccttct | 9960 |
| ggcctacaga aggccccac atggacgctg ctcttcaggt tgcaaccagc tcacctggaa | 10020 |
| ccccaaagcag ccaggggaat gtaagcagac atcaggaaga actcctagcc agatggatca | 10080 |
| ttcaatgcca agagctatacg actcacattt tggagaggtt ttctgtgttgg acttgggggg | 10140 |
| aatacaatgg acagctggac aaagtgtgtt gtctactca gagccagagg gatggataat | 10200 |
| gtgaccccttc catcaatctg gatagtaaat agttttgct actgctgttag gttttctaat | 10260 |
| aaattgcccataaggcaaga ttccaaagtc actttgtcct tccctaccac ttacccagcc | 10320 |
| agagctcccc accttcttga tgctccaggg aagaggctcc atggcccttg tgggtggcct | 10380 |
| gttcctgagc ctcgccaccc tgtgttagag cagagcatcc agatgaaatc tgacacttg | 10440 |
| tggcaaagtgc gctcagagag gaggctggct tcctagcatt cagggacggtt gctgagggcc | 10500 |
| gcttattcac cgaaaataaa tcttggaaag gacagggctg gtagcagaat gatcctttac | 10560 |
| ctaaaattct atcaaaatcc cattctcca tttggaaagc ccacagtgtc acagactctg | 10620 |
| ttccgggctc tgtcctcttc cctcttggtt cccaggagcc caggctggc tttgaagcag | 10680 |
| gcagggccca gcacacagta ggtactcagc agtgggggtt ttgaatccaa tcaaacggaa | 10740 |
| gtgtcaatgc agaaaaatgca atggatgtca atgcagtctc caaatgttcc ccactgtgca | 10800 |
| gcttccacat tcccgaggtt ttgggggggg acttgaatta acagcttcgg gaggcctgag | 10860 |
| tccctgcctc ccagctgagg aagaagctt aatcacaggg cgctgtgtct gtcttccagg | 10920 |
| ccctgtctct caggctccct ggtctccctg cactgtcttgg gtgagttaccc ccaatctctg | 10980 |
| agggtttggg gcctggccca gcaatgagca gggaggaaga cttcatctt cactcctaaa | 11040 |
| tttctgggac tccaagtttcc attctgcctt ggtctacagc cttgggctt gtcggtcaat | 11100 |
| gccccctcga gttgttgggtt gccttggca ggtcacattc ttttctggg tctttccaag | 11160 |
| ccccagtttc ccccttctac catctgtgca tggctccatg acctaagtgg agacctggga | 11220 |
| gagagtgtta ggaagaccga aaagggcagg acggggcctc cactgcctcc catccctgg | 11280 |
| ccgggcccac atagccttct ttgtcacaat cagctcaggt atccaagatc agattaccca | 11340 |
| cattcattat ttgagcaact attcattgaa cagttagaat atgtctcaact ctgtcagttg | 11400 |
| ctggctagaa gtagaaagta ccagatgagt gaaataattt gccactatcc ttggtagctg | 11460 |

ES 2 886 958 T3

| | |
|---|-------|
| atgactaagt aagagagaga tgcaagacaa catgtggaaa atgccaaact gagtagcagt | 11520 |
| cacagttgac atgctgcaga gagagctggc cgggggtcag aagacctggg caccagtcct | 11580 |
| gttcatttcc agtgtggcct cgagtcattc acctgacctc cctgaagttc attttcccaa | 11640 |
| gaagttgttt agtccaactg cccatcaagg atcttttaggg acccttctag ctctaacaga | 11700 |
| ggagatcaga aaagaaaaca agcaatgtgg ctcagctcat cctacaagct tcataagagaa | 11760 |
| ctgagactgg cctggaagca tagccagaaa ttagaacgcc taagggaga aggtcacaac | 11820 |
| gctgcctctg caatttagga gtgtatatgc tttcctgcag gatgttgaga gtttcattca | 11880 |
| ttatcgatg ccccctaccc cggccccaca atacctagtg cgtggatct gacacgtgg | 11940 |
| ggctggtcaa tgaatgaatg aatgaatggt cacaccatct gaggttctgc actgagtagc | 12000 |
| cctgaaggct tgaagcagca taagtgacag gtcccccctt gaggggcctc tgtttacca | 12060 |
| ataagccaag acctaagctc aacaacactg aaagggtggc caataccag gacagcctgt | 12120 |
| gggaattcca gagaaaaggga gattcccagg gactggggc ccaggctaaa cactaaaaaa | 12180 |
| tgcacatctgta ggctcaagga ggaaaagccc atgtctgtct gtcttgccca ccactctctc | 12240 |
| ccagcaccca gcactgcccc aggacagaga gcacttgaca caagttgggt agattaatga | 12300 |
| atgattttaga gttcagtggt ccccaacctt tttggcacaa gagactgggt gcatggaga | 12360 |
| caattttcc gcaaaccaag agggggatag agagcattag attctctctt tttttttttt | 12420 |
| ttgagaccaa gtctggctct tgtcaacttag cctggagtaa agtgttgca tctcggtca | 12480 |
| ctgcaacctc cgccctcctgg attcaagcga ttctcctgcc tcagccccct aaatagctgg | 12540 |
| gattacaggc acccggtcacc agcccagctg ggactatagg catgtgccac catgcccggc | 12600 |
| taattttgt atttttagta gagacggcgt ttcaccatgt tggccaggct agtctcgaac | 12660 |
| tcctgacctc aggtgatctg cccgcctgag cctcccaaag tgctggatt acaggcatga | 12720 |
| gctgcctcac ccagcctaaa gtctcataag gaacgtacag catagatccc tcacatgtgc | 12780 |
| agttcacaat aagggttgtgc tcctacaaga atctaacgcc acctctgtac tgacaggagg | 12840 |
| tgaagctcag gtggtcatgc tcgcttgtcc ctgccactca cttcctaattg tacagccagg | 12900 |
| ttcctaacag gccacgaacc agtgggaagg gcatctttt ggtcaaaaaa cagaattact | 12960 |
| tttagagaa ctacaagcag atcaatttgg ctagacagag actttatatg aaacagcagg | 13020 |
| aggctgctag gaggagtgg aactctactt tgccctcaag ggagatcccg aagggtttg | 13080 |
| caggagcggg caaggtggca tgaagaaagc agtgtttgaa atcaggtgg atttgaaaag | 13140 |
| cccgccctt ccccttagaa tggcccttct accatctgtg catggctcca caaccgtgg | 13200 |
| ggtggctgcc agaagaattg gaaaggcaga gcatgggtgg agagggggga cctgagggct | 13260 |
| ttacaggagt tccgggggtg gtgagggtgt gaaagccagg tcagtcagta ggaagacagg | 13320 |
| atgtcagatt gagagactcc cctggccggg gaaacagact tggagaaggg ggagttttgg | 13380 |

ES 2 886 958 T3

| | |
|--|-------|
| atgagacagt ccacttccga gtcacaaaat agcttggtgg tgtctgttta ctgttactca | 13440 |
| gtgggagtgg ctggggcacac gccacctggg cagggctttc gtaattctgc atcacttgc | 13500 |
| aaggtcacag attcccagca caacggacac acccatgttc atagtctgaa ctcctaaaca | 13560 |
| catcttaaac caaaaataaaa aaaaaagaaaa gaaagaaaaga aaaaggagag ggagggttga | 13620 |
| ggaaagccta tggctggga cactcaatac ctcccatgaa tatctcatat tgggctggc | 13680 |
| ctctctccac tctggccccca gccataaggg ccctgcttag agcagattt gggtgctgag | 13740 |
| tggaggcagc ctcatccccca acagcctgac ttccctgcctc ctccctgcct ctgcctgtgt | 13800 |
| ccagcctgtg ggaagagcct gaagaccccc cgtgtggtgg gtgtggagga ggcctctgtg | 13860 |
| gattcttggc ctggcaggt cagcatccag tacgacaaac agcacgtctg tggagggagc | 13920 |
| atcctggacc cccactgggt cctcacggca gcccactgct tcaggttaaga cccagctgt | 13980 |
| aaggaggct ctggggacca aggccagtca gggaccagag agcttgggt cctgtctcct | 14040 |
| ggcacccgtcc ttctcttcac tctccacta gagacgtttt ccaggtgtg gtggccccaa | 14100 |
| tgagacaatg gccatgatgc cctttgttag gctttgggt gtctgagcag agggtgctgg | 14160 |
| tcaccaagca tggccttccctc ctgggtggac accagcagat acccagagtc ctcacccac | 14220 |
| ccccatatacg ttcaagctac aaaagctttt cccacctgcc tcaacttcca agaactcact | 14280 |
| ctcttttgc ttgtttccag gaagttgttc cagggtctag agtcatagcc acgtcctcat | 14340 |
| tatgtctgga aactttaaaa aaattaaaga gcataggttc ctttcagtcc acagagaagc | 14400 |
| ctggccttac ctcaggaaag ggctactccc agacccctt cactttttt tttttttttt | 14460 |
| ttttttttt ttttgagaca gagtcttgct ctgttgctta ggctggagcg cagcagcatg | 14520 |
| atcttggctc actgcaacct ccgcctcctg agttcaagca attctcctgc ctcagcttcc | 14580 |
| caagtagctg ggactatagg catggccac catgcccggc taattttgt atttttggta | 14640 |
| gagacaggggt ttccatgt tggccaggct gatctctaac tcctgaccc aagtgtatcg | 14700 |
| cccacctcag cctccaaac tgctgggatt acaggcatga gccaggcat ccggcttttta | 14760 |
| tttattcatt cattcaatat ctaatgagca cctaccaggta accaaacacc agatgatgcg | 14820 |
| cccaagttca ttagacccca ccgctgtctt caaggcactc atgatctagg ccagcgtttt | 14880 |
| ttaaccactt tttttttttt ttttttgag attctggtga gagctataaa ttcttcctg | 14940 |
| aaaaaacatc tctgcacact aagcttgcc tggcattggg aaaaagaaaag cacgtaatgt | 15000 |
| aactgacagc atgagtaaca cagtgagaaa ggttggagga gagagcgcca ggacctcaga | 15060 |
| actcaggcat tagaggagcc cttccccag ccctccttga gtttcgttg ggcaggtttc | 15120 |
| actgagggaaa aagggtcaaa tcccttttc gaatttgact tcttgtaagt gccagaagac | 15180 |
| tgccccttct ccaccatccc tgcctcacca tcacatccctc tcccaaggca gtgacatcca | 15240 |

ES 2 886 958 T3

| | |
|---|-------|
| gcaccccgat ccctagggcc ctggggaccc agccttggc aaagtctcct caggcttgg | 15300 |
| tcagggctga acccagctgt ctctacccccc aggaaacata ccgatgtgtt caactggaag | 15360 |
| gtgcgggcag gctcagacaa actggcagc ttcccatccc tggctgtggc caagatcatc | 15420 |
| atcattgaat tcaaccccat gtacccaaa gacaatgaca tcgcccctat gaagctgcag | 15480 |
| ttcccactca ctttctcagg tgagaagcag ggcccaaggc cactcaagcc tcttacatca | 15540 |
| gtttcacgc ccactctgct attagctcac tgaccgcct tggcacataa tgtctcctct | 15600 |
| caagtccctca gcttgcccat ttgtctctaa tacgtcagcc taacatcaact gatgccatga | 15660 |
| ggcctccctca agctgtcagc taacacctcc actccattcc ctgcccagaga ttcttccaag | 15720 |
| gcctgtcttc cctatgtgga gcccctcgag tgagaactgg agttcatcc aatcttggag | 15780 |
| ttttaggaga ccttttaaaa agattatcga gctaattccc caccactgac caacacgcaa | 15840 |
| gagcctgctc agtatccctg ccaaggagtc attgtgcctt tggtgcctt cctccagggg | 15900 |
| cagggAACCC attacctgtg aggccccc cagagtctt gaacagctt gttggatgcc | 15960 |
| tttgtccttat actgaaatgt atttagatca ggattcccaa ctgtgggtc cacaagacac | 16020 |
| tggcccttg gagaagagag gattccattt tcaaataagt ttgggaaca tttcataact | 16080 |
| acagctccct tcttggaca cattagttt ttaaaggtag gagaagttt taaaataatc | 16140 |
| tgttttattt cgtttaacct acattttta aatttatttg accacagaat cctttttca | 16200 |
| tgctacttct attagcatcc catagaacaa gtgttctaga gaccctggtg tgacccttt | 16260 |
| cagagagctt aactgccagg ctctcctgag ccctgggttg tgttcaaga tttgtgcctg | 16320 |
| ggaattttt taatcaggta tggcaaggtag acagatacag acacagctat ctttggaaaga | 16380 |
| agagtttattt atttataattt cctgagagaa agggacatac cccacccccc aacacaggga | 16440 |
| cacccggggaa agcagctggg tccaccagga ggcaggagtg agggaaaggc atggcccaga | 16500 |
| gccacccgtg gcttccatgg gcaggtctgg ccaaggtagg gtaggcaaga ttgagcatgc | 16560 |
| tcaggattgg atagtgtgga caattctcta ggctatagat gtcagcctct ggttgtctag | 16620 |
| tatctgtccc tgggggtatt tagggcaggg aaaatattgg cttgggtgtt gagagtca | 16680 |
| taaaggaaat ggttggggat atgggcattt ggttggctgg tttgcctatt aaaggcgtgc | 16740 |
| ccaaagccaa gttgtttact atctgcagga attagctaac ccagtcctc ccagaccagc | 16800 |
| aagatccccca taatcataaa gcatcataat ttacagaaaa ttaacactta tgatgaataa | 16860 |
| aagatctcct tcttcctctg tgctcctggc aggcacagtc aggccatct gtctgcctt | 16920 |
| ctttgtgatgag gagtcactc cagccacccc actctggatc attggatggg gctttacgaa | 16980 |
| gcagaatgga ggttaagtccct ggggtcagga ccacaggcga ggagatgcc ttgttatgagg | 17040 |
| gagcagcttc cagaagtaat gggaaaggagg accacccttc agagaaaccc atcctggagg | 17100 |
| accaagcacc aaggcgccag gcagaaagca aagtggttt gcaatccagg gctggggat | 17160 |

ES 2 886 958 T3

ES 2 886 958 T3

| | | |
|--|--|-------|
| | cgcctcagcc tcccaagggtg ctgggattac aggtgtgagc caccgctcct ggccaacaac | 19080 |
| | caggctttt taagacatca ctcagagcct ttaatttgct aatgtgagtt gtgaatctct | 19140 |
| | gagagaaggc taacggcatg cttgcaactt acttgtccac agacaagcct ttctgcccc | 19200 |
| | gaagagaaga ccattctagg gtgctaata gcaaagaggg tgagggtgga atatcgaga | 19260 |
| | gcagcagggg gtgcagggg acagataggc cagttcaggg agcagagaag gagaagcccc | 19320 |
| | cccacctcac ctgccctccc cagcagtctc tttctggc tctcacaggg tgacagtgg | 19380 |
| | gggcccctga tgtaccaatc tgaccagtgg catgtggtgg gcacgttag ttggggctat | 19440 |
| | ggctgcgggg gcccggcac cccaggagta tacaccaagg tctcagccta tctcaactgg | 19500 |
| | atctacaatg tctgaaagg aaggtacctt tgccctaccc actgtgcctt ccctccagtc | 19560 |
| | ctctacctgg ggggtgccaa tccatcctca gtttggatt aaatggttt gacaactctt | 19620 |
| | tacatcccaa ataactttcc ctccaagcaa gggacagcct gagattgcac tattaaggct | 19680 |
| | gaaattcctt aggtcagaga ttctgataa atgcaaatac cttaggaaat agaacacacacc | 19740 |
| | aaggctttct ttctcttttc tgacagaatg agactatcag atccttcta gagagaagat | 19800 |
| | tctgataagg aagagagtgg aaaggctcat gagacctcct ggccctctgc aggtaggg | 19860 |
| | gagaagcaaa gtgttcaga aaaggaagac tcacgttaca catgtcacca cttgtccag | 19920 |
| | tttcagataa tctgactttc tttcatcggt tctcttttat tctaggctga gctgtAACGC | 19980 |
| | tgccgtcccc cacatccaga agctgctcc cttcagacacct acctacggca tgaccctca | 20040 |
| | aagtcaagata tgggacaaga gcctccttga acaaactc | 20078 |

<210> 13

<211> 15159

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Polinucleótido recombinante

10 <400> 13

| | |
|---|-----|
| ccacccgcac acactacagt cgagataact tcgtataatg tatgtataac gaagttatgc | 60 |
| tagtaactat aacggtccta aggttagcgag ctggccaaat ctgtgtgcta ccaagtagca | 120 |
| aaactgagcc tggaaactcac acatgcgtgt ctgagagccc agcaactatcg ccaggaaaac | 180 |
| ccagcgtctc cctgctcaag cctgaccctc agccctctct gcctctccct gcacttgcct | 240 |
| tccagtcaag gtgattctgg ataaatacta cttectctgc gggcagccta tccacttcat | 300 |
| cccgaggaag cagctgtgtg acggagagct ggactgtccc ttgggggagg acgaggagca | 360 |
| ctgtgtcaag agttcccccg aaggccctgc agtggcaggt gagtgcaggg tctgaggcac | 420 |
| aagagaagtg ggcccagcag gaggtctgct caggccccca cggcccaactg catagtatct | 480 |
| gccccctact tgcactttt catccttgcgtt gtataagggtt ctttgggtt ttgtttgttgc | 540 |

ES 2 886 958 T3

| | |
|---|------|
| ttgttttagag gcagagtgc ctgtggccca agatggagtg cagtgtcttgc | 600 |
| actgcaacct ctgcctccca gtttcaagtg attcttctgc ctcagcctca ttagtagctg | 660 |
| ggattacagg tgccagccac cacgcctggc taatttttat attttagta gagacggggt | 720 |
| tttgccacat tggtcaggct gatcttgaac tcctgaccc aggtgatctg cccgcctcag | 780 |
| cctcccaaag tgctgggatt acaggcgtga gccaccgtgc ccagctgtgt aagtttcttgc | 840 |
| agagcaggac cctgtcttgc ctaccttaa atccttagtac ttaacacaca gcaaacagta | 900 |
| actatttgcat gacaaatgt gagccagaaa ggacaggaaa ttgttaactga ggctgccccca | 960 |
| tgcgtgtgc gcctgggatg tttcaggcag agggctagac tgggtgaccc tggggcattc | 1020 |
| ctcctttcta tgaatattgt tatttcaagg agactagaaa agagacttct cagccacttc | 1080 |
| gccagctatt ggtccttcta ttcatttagtgc ttgtgatacatgatgt gacaggactg | 1140 |
| agccaggtcc tttcaatggat taggagatgt ttgtgacata aaatccacgt tctctttgg | 1200 |
| gctgggctct tctaccttct tccccctggt gcttgggctc tgaagaaaaa aagataggta | 1260 |
| ggagatgagt gatggggctt ctgagggcag ggctgagtga ctttctgtgt atttgcctt | 1320 |
| tctttatcgt aagtcaaatttgc cccacaggca cctgtcatcc tactgccagt aggacttctc | 1380 |
| actcaacattt cccctctgac cttaattggat gaaggactta ggtccctctc tcagacattt | 1440 |
| ccccaggctg ggcacaggatgt gtggaccatg gatgggtatg tggccatatac aattttaaaca | 1500 |
| agctgtatata ggtcgctggg tagagtgacc acataattgtatcatcaaaac tgataccgt | 1560 |
| aagagcaaaa gggggcacta ttaaccatttgc ggtcagggca acaggtcaaa atggagaccc | 1620 |
| accctgggac ttctggtcac actagctact gtcaaaatgg ggcaccaata gacaaagcca | 1680 |
| aatggaaagaa attccatttgc cattgaaatgt gttggggctc tgtggcaccc ccagttctag | 1740 |
| gttggggag ctggggctgg tctcatgtatg agttctgagg gggatggccc agttggccc | 1800 |
| cccggttccat ctaactcagg ttcccttcttcc cccagtcgc ctctccaaagg accgatccac | 1860 |
| actgcaggtg ctggactcgg ccacaggaa ctggttcttgc gctgtttcg acaacttcac | 1920 |
| agaagcttc gctgagacag cctgtaggca gatgggtac agcaggtaac caacctggc | 1980 |
| ctctcttctt tttcccttcttcc tcctcttcttcc cctttcttcc ctcccttctt | 2040 |
| ctctcttcc taaaaatttgc gggcatttgc gccaggcaga atggcttttgc aatcccagca | 2100 |
| tttcacttat aagcaacatg aagttaaatttgc tcctaaggcct caggttccctc aggagttat | 2160 |
| tgggggaaact aatgccaacc tcataaggata gtttgcaat gccagtgaga gaatgtgtgc | 2220 |
| tgccctccaa cacacacaca cacacttcttgc gctgtatgc agtctctcc ttcccttac | 2280 |
| tcctcaacct tcactccttgc tgctggcatt tgcaagaaac tgttcctgccc cagtaataca | 2340 |
| aaagcttaatgt taacttatttgc aaagtttcgt tagttaatgt ttagcttaag tgagcctagt | 2400 |

ES 2 886 958 T3

| | |
|--|------|
| ttcagtgggg ccccatcttc agcaatccca gctctctctg caaattcaa aagcagttcc | 2460 |
| aaatctggag tggatgaaaa ggtgtaagat gatagttaga gtaatttgca ttcttatatat | 2520 |
| ttatattcac ttgattttgg cagaaaacca aaaagatagt tattatatct tatatataga | 2580 |
| tatatattat atctatttca taaataggct caaacaaagt aagtaacttg cttagggtact | 2640 |
| agctgggagg tagagggcta gaatttgagc ccaagacccc taattcttgc gcattaggag | 2700 |
| ttcccacatt gtttctgttt cttagactgag taattcttta ttctcatgta ggacatcatc | 2760 |
| tctaaggaa ggggctaattg agatggttga tcactcagag agtttagctg gagaggatgg | 2820 |
| aaaagaaccc atacattcag ttgcagattt agatagccta tctctggcag gcctcagatt | 2880 |
| tcttcaggat tctaacagac tggaccaga gactaggcca aacaacaaa caaacaaaaa | 2940 |
| ctctactagg cagacatcac caaccaatca cagaactctc tcccatggat ccctaataca | 3000 |
| gcctcaaagt cctttcagt aaatgctcca ggcagccatt acaaatcaat cagaattatt | 3060 |
| tgccttctc ttctctgctc aacgggcttc tgctgctctc tactttccat agggggcaac | 3120 |
| ttccattacc ctctagaaaag cacacccac cacttcatt tcaaggagag tgaggaactc | 3180 |
| atgcccagca cctgctattc tcccccttcc ctgcagccac ggagccagc ctcgctgcag | 3240 |
| ccagccctgc ctccccactg tagtccagtc aactgctgca tcagccgttc ctggcacagc | 3300 |
| aggctgagcc ttgattatga aacctgggtg tctccagggg ttcttaagat gataggctcc | 3360 |
| tggaatttct gtcctttgg agctcagtaa ggcaccaaac cacctgagtc ttgtgcttca | 3420 |
| caaaatcaaa gttcatcaga atcattcatt gggatggaat tggtaacag aagttactt | 3480 |
| tcctggaaat gtccatttcc accatattcc gtccttctag gtctcagact tctctacttt | 3540 |
| cttcctctc tctagatcg aggcccttct tgccttagaa ccataggcat ttcaagatgt | 3600 |
| gggagacctt agggatcatc tagtccacgc atctttttt tttttttt acagagtctc | 3660 |
| actctgtcac ccaggctgga gtgcaatggc accatctctg cttactgcaa cctccacctc | 3720 |
| ccaggttcaa gtgattctt cgcctcagcc tcccaagtag ctgggattac aggcacgcac | 3780 |
| catcatgccc agctaatttt tatattttt tagagaccga gtttccat gttggccagg | 3840 |
| ctggtcttga actcctgacc tcaggtgatc cacccacctc ggcctccaa agtgcgtggg | 3900 |
| ttacaggcgt gagccactgc acccagcccc gtgcattttt ttatagaggg ggaaactgag | 3960 |
| gcttggagag acccagaaaa agaatatgac ctgccccagg ccacacatca aactagtgcc | 4020 |
| agagccaggg acagaaccta gatcatgagg actctaaaa tgcactctag tcctcccagg | 4080 |
| tctgagactt gggccttcc aggaagtgcc agcattctg cctgagaatg tgccaatcca | 4140 |
| ccagtattgc caatgactca gccctccatg gagagttct actaacatta ctgcatagt | 4200 |
| tagggatgga aggaaaaagat ttagaagagg cagattcagt aaaggaacaa tcagagagat | 4260 |
| ggaattaatc aaggaaggct tcctggagga ggaaaaactt caacccaagg tttgaaagta | 4320 |

ES 2 886 958 T3

| | |
|---|------|
| gcaagcatgg attagcaggg agaaagaggg agagtggtcc agttgagaga aacgtttgtc | 4380 |
| tggattcata tgaagacaga tctagtcctg ttctattaaa tatctctaag ggggc当地 | 4440 |
| acatacccccc gctatcaaag tcagaccaga tgctttgtt ggagaacgaa atatccacat | 4500 |
| tccaaactccc tcccagggtga gaagggagct aacctgagcc cctatgcctc tttgtttccc | 4560 |
| tgctgtgaac cagaagacat tgctggata tttgaaatag ggacagagct gggaatatgg | 4620 |
| aaaggagacc cctaacadttt ctccagggtct ctgggttctg gatttggatt ccccacccaa | 4680 |
| gaaagcaagt tacatcagca atgcactgag gttttagtcc tggtatgcc aaggctcggtt | 4740 |
| ctttattgtta tagcaaagca ggccccatct tcactgacta agaccatctc cactccctgg | 4800 |
| ccactccccca ccaagcattc tctgcccactc tttctcctga aagtggggc caactctacc | 4860 |
| atcttggttt aacccctgc cccagctcac aactctctct ccctcttgat gtgagcagca | 4920 |
| aacccacttt cagagctgtg gagattggcc cagaccagga tctggatgtt gttgaaatca | 4980 |
| cagaaaacag ccaggagctt cgcatgcgga actcaagtgg gtaagtgagg ggacacccccc | 5040 |
| tggcctacag aaggccccca catggacgct gctcttcagg ttgcaaccag ctcacctgga | 5100 |
| accccaagca gccaggggaa tgtaagcaga catcaggaag aactccttagc cagatggatc | 5160 |
| attcaatgcc aagagctata gactcacatt ttggagaggt tttctgtgtt gacttgtttt | 5220 |
| taatacaatg gacagctgga caaagtgtgt tgcctactc agagccagag ggatggataa | 5280 |
| tgtgaccttt ccatcaatct ggatagtaaa tagttttgc tactgctgta gttttctaa | 5340 |
| taaattgccc aataggcaag attccaaagt cactttgtcc ttccctacca cttacccagc | 5400 |
| cagagctccc caccttcttg atgctccagg gaagaggctc catggccctt gtgggtggcc | 5460 |
| tgttctgag cctcgccacc ctgtgttaga gcagagcatc cagatgaaat ctgtcacact | 5520 |
| gtggcaaaagt ggctcagaga ggaggctggc ttccctagcat tcagggacgt tgctgaggc | 5580 |
| cgcttattca ccgaaaataa atctgaaaa ggacagggt ggttagcagaa tgatcctta | 5640 |
| cctaaaattc tatcaaatac ccattctcc atttggaaag cccacagtgt cacagactct | 5700 |
| gttccgggct ctgtcctt ccctttggg tcccaggagc ccaggctggg ctttgaagca | 5760 |
| ggcagggccc agcacacagt aggtactcag cagtgggggt gttgaatcca atcaaacgga | 5820 |
| agtgtcaatg cagggaaatgc aatggatgtc aatgcagtct ccaaattttc cccactgtgc | 5880 |
| agcttccaca ttcccggaggt attggggagg gacttgaatt aacagctcg ggaggcctga | 5940 |
| gtccctgcct cccagctgag gaagaagctt aaatcacagg gcgctgtgtc tgtcttccag | 6000 |
| gccctgtctc tcaggctccc tggtctccct gcactgtctt ggtgagtaacc cccaatctct | 6060 |
| gagggtttgg ggcctgggccc agcaatgagc agggaggaag accttcatct tcactcctaa | 6120 |
| atttctggga ctccaagttt cattctgcct tggtctacag cccttggct tgtcggtaa | 6180 |

ES 2 886 958 T3

| | |
|---|------|
| tgccccctcg agttgttgtt ggccttggc aggtcacatt cttttctgg gtcttccaa | 6240 |
| gccccagttt cccccttcta ccatctgtgc atggctccat gacctaagtg gagacctggg | 6300 |
| agagagtgtt aggaagaccg aaaagggcag gacggggcct ccactgcctc ccatccctgg | 6360 |
| tccgggccca catagccttc tttgtcacaa tcagctcagg tatccaagat cagattaccc | 6420 |
| acattcatta ttttagcaac tattcattga acagttagaa tatgtctcac tctgtcagtt | 6480 |
| gctggctaga agtagaaaagt accagatgag tgaataatt ggccactatc cttggtagct | 6540 |
| gatgactaag taagagagag atgcaagaca acatgtggaa aatgccaaac tgagtagcag | 6600 |
| tcacagttga catgctgcag agagagctgg ccgggggtca gaagacctgg gcaccagtcc | 6660 |
| tgttcatttc cagtgtggcc tcgagtcatt cacctgacct ccctgaagtt cattttccca | 6720 |
| agaagttgtt tagtccaact gcccatacg gatctttagg gacccttcta gctctaacag | 6780 |
| aggagatcag aaaagaaaaac aagcaatgtg gctcagctca tcctacaagc ttcatagaga | 6840 |
| actgagactg gcctggaagc atagccagaa attagaacgc ctaagggaa aaggtcaca | 6900 |
| cgctgcctct gcaattttagg agtgtatatg ctttcctgca ggatgtttagg agtttcattc | 6960 |
| attatcgtat gccccctacc ccggccccac aatacctagt gcgtggatc tgacacgtgg | 7020 |
| tggctggta atgaatgaat gaatgaatgg tcacaccatc tgaggtctg cactgagtag | 7080 |
| ccctgaaggc ttgaagcagc ataagtgaca ggtcctccct tgaggggcct ctgtttacc | 7140 |
| aataagccaa gacctaagct caacaacact gaaagggtgg ccaataccca ggacagcctg | 7200 |
| tgggaattcc agagaaaaggg agattccag ggactggggg cccaggctaa acactgaaaa | 7260 |
| atgcatctgt aggctcaagg aggaaaagcc catgtctgtc tgtctgccc accactctct | 7320 |
| cccagcaccc agcaactgccc caggacagag agcaacttgac acaagttgg tagattaatg | 7380 |
| aatgatttag agttcagtgg tccccaacct ttttggcaca agagactggg tgcatggaag | 7440 |
| acaattttc cgcaaaccas gagggggata gagagcatta gattctctt tttttttttt | 7500 |
| tttgagacca agtctggctc ttgtcactca gcctggagta aagtgttgcg atctcggctc | 7560 |
| actgcaacct ccgcctcctg gattcaagcg attctcctgc ctcagcccc taaatagctg | 7620 |
| ggattacagg caccgtcac cagccagct gggactatacg gcatgtgcca ccatgcccgg | 7680 |
| ctaatttttgc tatttttagt agagacggcg tttcaccatg ttggccaggc tagtctcgaa | 7740 |
| ctcctgacct caggtgatct gccccctga gcctccaaa gtgctggat tacaggcatg | 7800 |
| agctgcctca cccagcctaa agtctcataa ggaacgtaca gcatagatcc ctcacatgtg | 7860 |
| cagttcacaa taagggtgtg ctcttacaag aatctaacgc cacctctgat ctgacaggag | 7920 |
| gtgaagctca ggtggctatg ctgccttgc cctgccactc acttcataat gtacagccag | 7980 |
| gttccttaaca ggccacgaac cagtggaaag ggcattttt tggatcaaaa acagaattac | 8040 |
| tttttagaga actacaagca gatcaatttgc gctagacaga gactttatat gaaacagcag | 8100 |

ES 2 886 958 T3

| | |
|--|------|
| gaggctgcta ggaggagtgg aaactctact ttgccctcaa gggagatccc gaaggggctt | 8160 |
| gcaggagcgg gcaagggtggc atgaagaaag cagtgttg aatcaggtgg tatttgaaaa | 8220 |
| gcccgccct tccccctaga atggcccttc taccatctgt gcatggctcc acaaccgtgg | 8280 |
| tggtggctgc cagaagaatt ggaaaggcag agcatgggtg gagagggggg acctgaggc | 8340 |
| tttacaggag ttccgggggt ggtgaggggtg taaaagccag gtcagtcagt aggaagacag | 8400 |
| gtatgcagat tgagagactc ccctggccgg gaaaacagac ttggagaagg gggagtttg | 8460 |
| gtatgagacag tccacttccg agtcacaaaa tagttgtgg gtgtctgtt actgttactc | 8520 |
| agtggagtg gctggggaca cgccacactgg gcagggcttt cgtaattctg catcaactgt | 8580 |
| gaaggtcaca gattcccagc acaacggaca cacccatgtt catagtctga actcctaacc | 8640 |
| acatcttaaa cccaaataaa aaaaaaaagaa agaaagaaag aaaaaggaga gggaggttg | 8700 |
| aggaaaggcct atggtctggg acactcaata cctcccatga atatctcata ttgggcttgt | 8760 |
| cctctctcca ctctggcccc agccataagg gccctgctta gagcagattt tgggtgctga | 8820 |
| gtggaggcag ctcatcccc aacagcctga ctccctgcct cctccctgcc tctgcctgtg | 8880 |
| tccagcctgt gggaaagagcc tgaagacccc ccgtgtggtg ggtgtggagg aggcccttgt | 8940 |
| ggattcttgg ctttggcagg tcagcatcca gtacgacaaa cagcacgtct gtggagggag | 9000 |
| catcctggac ccccaactggg tcctcacggc agcccaactgc ttcaggttaag accccagctg | 9060 |
| taaggaggtc tctggggacc aaggccagtc agggaccaga gagcttgggg tctgtctcc | 9120 |
| tggcaccgtc cttctttca ctctccact agagacgtt tccaggttgt ggtggccccca | 9180 |
| atgagacaat ggccatgatg ccctttgtta ggctttggg tgtctgagca gagggtgctg | 9240 |
| gtcaccaagc atggcctctt cttggggaa caccaggaga tacccagagt ctcacccca | 9300 |
| ccccatatac gttcaagcta caaaagctct tcccacactgc ctcaacttcc aagaactcac | 9360 |
| tcttttttgc ttgtttcca ggaagttgtt ccagggtcta gagtcatacg cacgtcctca | 9420 |
| ttatgtctgg aaactttaaa aaaattaaag agcataggtt ctttcagtc cacagagaag | 9480 |
| cctggcccta ctcagggaa gggctactcc cagacccct tcacttttt ttttttttt | 9540 |
| tttttttttt ttttgagac agagtcttgc tctgttgctt aggctggagc gcagcagcat | 9600 |
| gatcttggct cactgcaacc tccgcctcct gagttcaagc aattctcctg ctcagcttc | 9660 |
| ccaagtagct gggactatag gcatgggca ccatgcccgg ctaattttt tatttttgt | 9720 |
| agagacaggg tttcaccatg ttggccaggc tgatctctaa ctccgtaccc caagtgtatct | 9780 |
| gcccacctca gcctccaaa ctgctggat tacaggcatg agccagggca tccggcttt | 9840 |
| atttattcat tcattcaata tctaattgagc acctaccagg taccaaacac cagatgtgc | 9900 |
| gcccaagttc attagacccc accgctgtct tcaaggcact catgatctag gccagcgtt | 9960 |

ES 2 886 958 T3

| | | | | | | |
|-------------|-------------|------------|-------------|-------------|-------------|-------|
| ttaaccact | ttttttttt | tttttttga | gattctggtg | agagctataa | atttttcct | 10020 |
| ggaaaaacat | ctctgcacac | taagctgtgc | ctggcattgg | aaaaaagaaa | gcacgtaatg | 10080 |
| taactgacag | catgagtaac | acagtgagaa | aggttggagg | agagagcgcc | aggacctcag | 10140 |
| aactcaggca | ttagaggagc | ccctcccca | gccctcctt | aggttcgtt | ggcaggttt | 10200 |
| cactgaggaa | aaagggtcaa | atccctttt | cgaatttgc | ttcttgtaag | tgccagaaga | 10260 |
| ctgccccttc | tccaccatcc | ctgcctcacc | atcatcttc | ctcccaaggc | agtgacatcc | 10320 |
| agcacccccga | tccctagggc | cctggggacc | cagccttgg | caaagtctcc | tcaggcttgg | 10380 |
| atcaggcctg | aacccagctg | tctctacccc | cagggaaacat | accgatgtgt | tcaactggaa | 10440 |
| ggtgtcgggca | ggctcagaca | aactgggcag | cttcccatcc | ctggctgtgg | ccaagatcat | 10500 |
| catcattgaa | ttcaacccca | tgtacccaa | agacaatgac | atcgccctca | tgaagctgca | 10560 |
| gttcccactc | actttctcag | gtgagaagca | gggcccagg | ccactcaagc | ctcttacatc | 10620 |
| agttttcacg | cccaactctgc | tattagctca | ctgaccgccc | ttggcacata | atgtctcctc | 10680 |
| tcaagtcctc | agtttgccca | tttgtctcta | atacgtcagc | ctaacatcac | tgatgccatg | 10740 |
| aggcctccctc | aagctgtcag | ctaacacctc | cactccattc | cctgcccagag | attcttccaa | 10800 |
| ggcctgtctt | ccctatgtgg | agcccctcga | gtgagaactg | gagtttcatc | caatcttggaa | 10860 |
| gttttaggag | acctttaaa | aagattatcg | agctaattcc | ccaccactga | ccaacacgca | 10920 |
| agagoctgct | cagtatccct | gccaggagt | cattgtgccc | ctgtttgctc | tcctccaggg | 10980 |
| gcagggAACCC | cattacctgt | gaggcagccc | acagagtctt | tgaacagctc | tgttggatgc | 11040 |
| cttgtgctta | tactgaaatg | tattnagatc | aggattccca | actgtgggt | ccacaagaca | 11100 |
| ctggccctt | ggagaagaga | ggattccatt | gtcaaataag | tttggggAAC | attttcatac | 11160 |
| tacagctccc | ttcttggAAC | acattagttt | attaaaggta | ggagaagttt | ttaaaataat | 11220 |
| ctgttttatt | gcgtttaacc | tacattttt | aaatttattt | gaccacagaa | tcctttttc | 11280 |
| atgctacttc | tattagcatc | ccatagaaca | agtgttctag | agaccctggt | gtgaccctt | 11340 |
| tcagagagct | taactgccag | gctctcctga | gccctgggt | gtgtttcaag | atttgcct | 11400 |
| gggaattgtt | ttaatcaggt | atggcaaggt | gacagataca | gacacagcta | tctttgaaag | 11460 |
| aagagtttat | tattnataat | tcctgagaga | aagggacata | ccccacccccc | caacacaggg | 11520 |
| acacccgggg | aagcagctgg | gtccaccagg | aggcaggagt | gaggggaagg | catggcccag | 11580 |
| agccacctgt | ggcttccatg | ggcaggtctg | gccaaggtag | ggtaggcaag | attgagcatg | 11640 |
| ctcaggattg | gatagtgtgg | acaattctct | aggctataga | tgtcagcctc | tggttgtcta | 11700 |
| gtatctgtcc | ctggggtgat | ttagggcagg | aaaaatattt | gcttgggtgc | tgagagtca | 11760 |
| ataaaggaag | tgggggggat | tatggcattt | gggttggctg | gtttgcctat | taaaggcgtg | 11820 |
| cccaaagcca | agttgtttac | tatctgcagg | aattagctaa | cccagtctct | cccagaccag | 11880 |

ES 2 886 958 T3

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|-------------|-------|
| caagatcccc | ataatcataa | agcatcataa | tttacagaaa | attaacactt | atgatgaata | 11940 |
| aaagatctcc | ttttcctct | gtgctcctgg | caggcacagt | caggcccatt | tgtctgccct | 12000 |
| tcttgatga | ggagctca | ccagccaccc | cactctggat | cattggatgg | ggctttacga | 12060 |
| agcagaatgg | aggttaagtcc | tgggtgcagg | accacaggc | aggagatgcc | cttgtatgag | 12120 |
| ggagcagctt | ccagaagtaa | tgggaaggag | gaccaccctt | cagagaacc | catcctggag | 12180 |
| gaccaagcac | caaggcgcca | ggcagaaagc | aaagtggttt | ggcaatccag | ggctggggga | 12240 |
| tagaaggcaa | ggatggaaat | gtgagtgttt | ttaccctccc | aggaaagatg | tctgacatac | 12300 |
| tgctgcaggc | gtcagtccag | gtcattgaca | gcacacggtg | caatcagac | gatgcgtacc | 12360 |
| agggggaa | caccgagaag | atgatgtgt | caggcatccc | ggaagggggt | gtggacacct | 12420 |
| gccaggtggg | gcctccaaga | atcatggga | gttctaagaa | tagggttag | gtcctagaga | 12480 |
| gatgagaaaa | cccagaggct | gcatgcccta | caggaagcct | tgcataatcat | ggcactcaa | 12540 |
| tgtgtatga | tgggaggaag | agagggaggg | aaggaaagga | tagtcagata | aaagtgtacc | 12600 |
| aatagatgag | tgggtggatg | gatggatgca | gacaagcaga | gagattcaa | atgtctctt | 12660 |
| cacattcgaa | gatgatgtta | ctggcctggc | atggggctc | acgcttgtaa | tcccagcact | 12720 |
| ttgggaggct | gaggcgggca | ggtgattga | ggtcaggaat | tcaagaccag | cctggccaac | 12780 |
| atggtaatc | ccagctactt | gggaggctga | ggcaggagaa | ttgcttgaac | ccaggaggca | 12840 |
| gaggttgcag | taagctgaga | ttgcggccact | gcactccagc | ctgggtgacc | cagcaagact | 12960 |
| ccatctgaaa | acaacaacaa | caacaaagat | gacattactc | atccacccca | cccaccccttc | 13020 |
| tcactagcta | cagaatgatt | agcccttga | ggtcaggaat | cccaggtcta | ttttctctgt | 13080 |
| gactctcccc | aagctgctga | actacactag | gaaagaatta | ccgcctgcag | aatgctggaa | 13140 |
| gcacatctgt | gtgtccctc | acccggcct | cattggccat | caggactgct | tagcaatccc | 13200 |
| tgttagacctt | tttcctcccc | catacttcca | gaggatcttc | tgaactat | tcttttttta | 13260 |
| ttttttcttt | tatgtttttt | aacagagaca | gggtcactat | gttgcccagt | ctggctcaa | 13320 |
| actcctgggt | tcaagggatt | ctcccacctc | agcttccaa | aatgctggga | ttacaggcat | 13380 |
| gagccatcgt | gtttggcctg | aaccatttc | ataaaaaccc | ctaccctact | ctcacctcca | 13440 |
| tttccagtca | ttaaattcct | tcatttaaga | ggcatctt | agtcatcgca | tgtgtgccat | 13500 |
| gaacatggta | gtctttggag | accctcagg | gagctcacag | tgggggggg | aaaggggggc | 13560 |
| attaaacaga | catttaagct | atagtttgg | gttcagaggg | aggaagcccc | aggggctaaa | 13620 |
| acagctgata | aggactccca | gataagtgc | cttttca | tctggcattt | tcttgttttg | 13680 |
| ttatggctt | gttcactgtc | tctcacccca | tttgatccta | agcttctga | ggcaggat | 13740 |

ES 2 886 958 T3

| | | | | | |
|-----------------------|-----------------------|--------------------------|---------------------|----------------|-------|
| ctttgtttt ttcatca | g tggatccaa ttgc | ttagaa cactac | ctgg caca | aaaatag | 13800 |
| gcactctata agtgattaca | caaatttgg aacgactagg | ttaaacaatg ataacc | aggc | | 13860 |
| ttttttttt tttttgaga | ctgagtctca ctctgttgc | caggctagag tgaagtgg | tt | | 13920 |
| tgatctcggc tcactgcagc | ctccgcctct gggttca | aat gattctccac | ctcagcctcc | | 13980 |
| tgagtagctg ggattacagg | tgccgttgcac tatgccc | agc taattttgt attt | gttagta | | 14040 |
| gagacgggtt tcaccatgtt | ggccaggctg gtcttga | act cctgac | ctca agtgattcac | | 14100 |
| ccgcctcagc ctcccaaggt | gctgggatta caggtgtg | ag ccaccg | ctcc tggccaacaa | | 14160 |
| ccaggctttt ttaagacatc | actcagagcc tttaatttgc | taatgtg | agt ttagt | atctc | 14220 |
| tgagagaagg ctaacggcat | gcttgcact tacttgc | ca | gacaagcc tttctgcccc | | 14280 |
| agaagagaag accattctag | ggtgcta atg agcaa | agagg gtgagggtgg | aatatcg | gag | 14340 |
| agcagcaggg agtgcagggg | aacagatagg ccagttc | agg gacagagaa ggagaagccc | | | 14400 |
| ccccacactca cctgccc | ctcc ccagc | agtct ctgttctgg | ctctcacagg gtgac | agtgg | 14460 |
| tggcccttg atgtaccaat | ctgaccagtg gcatgtgg | tg ggc | atcg ttttta gtt | ggggct | 14520 |
| tggctgcggg ggcccgagca | ccccaggagt atacacca | ag gtctcagc | ct atctca | actg | 14580 |
| gatctacaat gtctggaagg | taaggtacct ttgc | ccctt acc cactgtgc | ctcc | tccatcc | 14640 |
| cctctacctg ggggtgcca | atccatc | ctc aggtt | gatt taaatgg | ttc tgacaactct | 14700 |
| ttacatccca aataacttcc | cctccaa | gca agggac | agcc tgagatt | gca ctattaaggc | 14760 |
| tgaaattcct taggtcagag | atttctgata aatgca | aaata cctt | agg gaa taga | acacac | 14820 |
| caagccccc tttctttt | ctgacagaat gagactat | ca gatc | c ttcttctt | ctt agagagaaga | 14880 |
| ttctgataag gaagagagt | gaaaggctca tgagac | ctcc tggcc | ctctc tctg cagg | gttaggg | 14940 |
| agagaagcaa agtgttccag | aaaaggaaga ctcac | gtt tac acatgt | cacc actttgt | ccca | 15000 |
| gtttcagata atctgactt | ctttcatcg gtctcttta | ttctt | aggctg agctgt | taacg | 15060 |
| ctgccgtccc ccacatccag | aagctg | tttc | cctt cagacc tac | ctacggc atgacc | 15120 |
| aaagtca | gat atgggaca | aaag | ctcc ttg aaca | aaactc | 15159 |

<210> 14

<211> 435

<212> PRT

<213> Secuencia artificial

<220>

<223> Proteína recombinante

10

<400> 14

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Met | Glu | Ser | Asp | Ser | Gly | Gln | Pro | Leu | Asn | Asn | Arg | Asp | Ile | Val | Pro |
| 1 | | | | 5 | | | | | 10 | | | | 15 | | |

ES 2 886 958 T3

Phe Arg Lys Pro Arg Arg Pro Gln Glu Thr Phe Lys Lys Val Gly Ile
20 25 30

Pro Ile Ile Ala Val Leu Leu Ser Leu Ile Ala Leu Val Ile Val Ala
35 40 45

Leu Leu Ile Lys Val Ile Leu Asp Lys Tyr Tyr Phe Leu Cys Gly Gln
50 55 60

Pro Leu His Phe Ile Pro Arg Lys Gln Leu Cys Asp Gly Glu Leu Asp
65 70 75 80

Cys Pro Leu Gly Glu Asp Glu Glu His Cys Val Lys Ser Phe Pro Glu
85 90 95

Gly Pro Ala Val Ala Val Arg Leu Ser Lys Asp Arg Ser Thr Leu Gln
100 105 110

Val Leu Asp Ser Ala Thr Gly Asn Trp Phe Ser Ala Cys Phe Asp Asn
115 120 125

Phe Thr Glu Ala Leu Ala Glu Thr Ala Cys Arg Gln Met Gly Tyr Ser
130 135 140

Ser Lys Pro Thr Phe Arg Ala Val Glu Ile Gly Pro Asp Gln Asp Leu
145 150 155 160

Asp Val Val Glu Ile Thr Glu Asn Ser Gln Glu Leu Arg Met Arg Asn
165 170 175

Ser Ser Gly Pro Cys Leu Ser Gly Ser Leu Val Ser Leu His Cys Leu
180 185 190

Ala Cys Gly Lys Ser Leu Lys Thr Pro Arg Val Val Gly Val Glu Glu
195 200 205

Ala Ser Val Asp Ser Trp Pro Trp Gln Val Ser Ile Gln Tyr Asp Lys
210 215 220

Gln His Val Cys Gly Gly Ser Ile Leu Asp Pro His Trp Val Leu Thr
225 230 235 240

Ala Ala His Cys Phe Arg Lys His Thr Asp Val Phe Asn Trp Lys Val
245 250 255

Arg Ala Gly Ser Asp Lys Leu Gly Ser Phe Pro Ser Leu Ala Val Ala
260 265 270

ES 2 886 958 T3

Lys Ile Ile Ile Glu Phe Asn Pro Met Tyr Pro Lys Asp Asn Asp
275 280 285

Ile Ala Leu Met Lys Leu Gln Phe Pro Leu Thr Phe Ser Gly Thr Val
290 295 300

Arg Pro Ile Cys Leu Pro Phe Phe Asp Glu Glu Leu Thr Pro Ala Thr
305 310 315 320

Pro Leu Trp Ile Ile Gly Trp Gly Phe Thr Lys Gln Asn Gly Gly Lys
 325 330 335

Met Ser Asp Ile Leu Leu Gln Ala Ser Val Gln Val Ile Asp Ser Thr
340 345 350

Arg Cys Asn Ala Asp Asp Ala Tyr Gln Gly Glu Val Thr Glu Lys Met
 355 360 365

Met Cys Ala Gly Ile Pro Glu Gly Gly Val Asp Thr Cys Gln Gly Asp
370 375 380

Ser Gly Gly Pro Leu Met Tyr Gln Ser Asp Gln Trp His Val Val Gly
385 390 395 400

Ile Val Ser Trp Gly Tyr Gly Cys Gly Gly Pro Ser Thr Pro Gly Val
405 410 415

Tyr Thr Lys Val Ser Ala Tyr Leu Asn Trp Ile Tyr Asn Val Trp Lys
 420 425 430

Ala Glu Leu
435

<210> 15

<211> 2046

<212> ADN

<213> *Mus musculus*

5

<400> 15

ES 2 886 958 T3

| | |
|---|-----|
| cagaaacaag gacctttca ttattcaaga gtaaaatgta taggccaaga ccaatgctat | 60 |
| caccgtcaag attcttcaact cccttgcag tagcttcgt tgtcataata acggtagggc | 120 |
| tcctggccat gatggcaggt ctacttattc actttttagc ttttgacaag aaagcttact | 180 |
| tttatcatag cagcttcaa atcctaaacg ttgaatacac tgaggctta aactcaccag | 240 |
| ctacacacga atacagaacc ttgagtgaaa gaattgaggc tatgattact gatgaattc | 300 |
| gaggatcaag tctaaaaagt gagtttatca ggacacatgt tgtcaaacta agaaaagaag | 360 |

ES 2 886 958 T3

| | | | | | | |
|------------|------------|-------------|------------|------------|-------------|------|
| ggactggtgt | ggttgcggat | gttgtcatga | aatttcgatc | tagtaaacgt | aacaacagaa | 420 |
| aggtaatgaa | aaccagaatt | caatctgtgc | tacgaagact | cagcagctct | ggaaacttgg | 480 |
| aaatagcccc | ttcgaatgag | ataacatcac | tcactgacca | ggatacagaa | aatgtttga | 540 |
| ctcaagaatg | tggagcacgt | ccagaccta | taacactgtc | agaagagaga | atcattggag | 600 |
| gcatgcaagc | tgagcccggt | gactggccct | ggcaagtca | tctacagctc | aataatgtcc | 660 |
| accactgtgg | aggtgccctg | atcagtaaca | tgtgggtcct | gacagcagct | cattgcttca | 720 |
| aaagctatcc | taatcctcaa | tattggacag | ccaccttgg | ggtttctaca | atgagcccta | 780 |
| ggctgagagt | gagagtaagg | gctattttag | cccacgacgg | gtacagctcc | gtaactcg | 840 |
| acaatgacat | cgcagttgta | caacttgaca | gatctgtcgc | ctttccaga | aatatccata | 900 |
| gggtatgtct | cccagcagca | acccaaaata | tcatccctgg | ttctgtcgca | tatgttacag | 960 |
| gatggggatc | tctcacatat | ggaggcaacg | cagtcacaaa | tctacggcaa | ggagaggtca | 1020 |
| gaataataag | ttcagaggaa | tgcaataacgc | cagctggta | cagtggaa | gtcttgccag | 1080 |
| gaatgctgtg | tgctggaatg | cgttcagggg | ccgtggatgc | atgccagggt | gattcaggtg | 1140 |
| gcccgttagt | acaagaagac | tcaaggcggc | tttggtttgt | tgtggcatt | gtgagctgg | 1200 |
| gatatcagtg | tggcctccca | aataagccag | gctgtatac | tcgagtgaca | gcctaccgca | 1260 |
| actggatcag | acagcagacg | ggaatctagt | gcaaccgagg | aaaaaacgtg | ccatgagg | 1320 |
| tctgtatcca | agtgtgactg | actcgatgc | catggcttca | cattcaact | gcaaaggaga | 1380 |
| ctggaaatgc | cccttctgaa | cgtcccatta | cataaatatg | gtttaactgt | ttagtattc | 1440 |
| tttgcgta | cagattttta | ctttcttgag | aaaaaaaaaa | acatgaacat | ggctaagtaa | 1500 |
| gaattatgtt | aggctagtaa | caggaagaca | tttattacat | gggtggtcag | gtgttagtagt | 1560 |
| gagaagtcag | gtaagttaag | tcaataattt | acagaaaata | atgtcaggt | gtcctaacgt | 1620 |
| taaatatgtg | aggccacaga | acaaatagt | ttagaactga | agccatccca | agtatttaac | 1680 |
| atttgtttc | aagtgaaact | aagaaacaga | cttacatata | gtttaatgg | tgaattttca | 1740 |
| ttttaaatat | tttatctaca | tagaaaagac | atatctcctt | catgaagaag | ctgaggtgat | 1800 |
| gaatcaacac | agccttca | gctatgttg | caaccacaag | atttgggaa | aagaaatccc | 1860 |
| tactaccaac | ttcctactgt | tggcattatt | ttttagagta | acacgacgca | caatagcaa | 1920 |
| attnaagtaa | caaattaaaa | gttaatgatg | aagaagaagt | aaagagttt | tttgcaaaga | 1980 |
| caaaaattaa | acagattaat | atcaataat | ctggagacag | aagggtctca | gattcatatt | 2040 |
| ctctct | | | | | | 2046 |

<210> 16

<211> 417

<212> PRT

ES 2 886 958 T3

<213> *Mus musculus*

<400> 16

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

ES 2 886 958 T3

Phe Gly Val Ser Thr Met Ser Pro Arg Leu Arg Val Arg Val Arg Ala
245 250 255

Ile Leu Ala His Asp Gly Tyr Ser Ser Val Thr Arg Asp Asn Asp Ile
260 265 270

Ala Val Val Gln Leu Asp Arg Ser Val Ala Phe Ser Arg Asn Ile His
275 280 285

Arg Val Cys Leu Pro Ala Ala Thr Gln Asn Ile Ile Pro Gly Ser Val
290 295 300

Ala Tyr Val Thr Gly Trp Gly Ser Leu Thr Tyr Gly Gly Asn Ala Val
305 310 315 320

Thr Asn Leu Arg Gln Gly Glu Val Arg Ile Ile Ser Ser Glu Glu Cys
325 330 335

Asn Thr Pro Ala Gly Tyr Ser Gly Ser Val Leu Pro Gly Met Leu Cys
340 345 350

Ala Gly Met Arg Ser Gly Ala Val Asp Ala Cys Gln Gly Asp Ser Gly
355 360 365

Gly Pro Leu Val Gln Glu Asp Ser Arg Arg Leu Trp Phe Val Val Gly
370 375 380

Ile Val Ser Trp Gly Tyr Gln Cys Gly Leu Pro Asn Lys Pro Gly Val
385 390 395 400

Tyr Thr Arg Val Thr Ala Tyr Arg Asn Trp Ile Arg Gln Gln Thr Gly
405 410 415

Ile

<210> 17
<211> 2800
<212> ADN
<213> *Homo sapiens*

<400> 17

5

10

ES 2 886 958 T3

| | |
|---|-----|
| atttgagtgg gaatctcaa gcagttgagt aggcagaaaa aagaacctct tcattaagga | 60 |
| ttaaaaatgta taggccagca cgtgtaactt cgacttcaag atttctgaat ccatatgttag | 120 |
| tatgtttcat tgcgtcgca gggtagtga tcctggcagt caccatagct ctacttgtt | 180 |
| acttttttagc ttttgatcaa aaatcttact ttataggag cagtttcaa ctcctaaatg | 240 |
| ttgaatataa tagtcagttt aattcaccag ctacacagga atacaggact ttgagtggaa | 300 |

ES 2 886 958 T3

| | |
|--|------|
| gaattgaatc tctgattact aaaacattca aagaatcaaa tttaagaat cagttcatca | 360 |
| gagctcatgt tgccaaactg aggcaagatg gtagtgggt gagagcggat gttgtcatga | 420 |
| aatttcaatt cactagaaat aacaatggag catcaatgaa aagcagaatt gagtctgtt | 480 |
| tacgacaaat gctgaataac tctggaaacc tgaaaataaa cccttcaact gagataacat | 540 |
| cacttactga ccaggctgca gcaaattggc ttattaatga atgtgggccc ggtccagacc | 600 |
| taataacatt gtctgagcag agaatccttg gaggcactga ggctgaggag ggaagctggc | 660 |
| cgtggcaagt cagtctgcgg ctcataatg cccaccactg tggaggcagc ctgatcaata | 720 |
| acatgtggat cctgacagca gctcactgct tcagaagcaa ctctaattcct cgtgactgga | 780 |
| ttgccacgtc tggatttcc acaacatttc ctaaactaag aatgagagta agaaatattt | 840 |
| taattcataa caattataaa tctgcaactc atgaaaatga cattgcactt gtgagacttg | 900 |
| agaacagtgt caccttacc aaagatatcc atagtgtgt tctccagct gctacccaga | 960 |
| atattccacc tggctctact gcttatgtaa caggatgggg cgctcaagaa tatgctggcc | 1020 |
| acacagttcc agagctaagg caaggacagg tcagaataat aagtaatgat gtatgtaatg | 1080 |
| caccacatag ttataatgga gccatcttgt ctggaatgct gtgtgctgga gtacctcaag | 1140 |
| gtggagtgga cgcatgtcag ggtgactctg gtggccact agtacaagaa gactcacggc | 1200 |
| ggctttgggtt tattgtgggg atagtaagct gggagatca gtgtggcctg ccggataagc | 1260 |
| caggagtgt tactcgagtg acagcctacc ttgactggat taggcaacaa actggatct | 1320 |
| agtcaacaa gtgcattccct gttgcaaagt ctgtatgcag gtgtgcctgt cttaaattcc | 1380 |
| aaagctttac attcaactg aaaaagaaac tagaaatgtc ctaatttaac atcttgttac | 1440 |
| ataaaatatgg tttaacaaac actgttaac cttctttat tattaaaggt tttctatttt | 1500 |
| ctccagagaa ctatatgaat gttcatagt actgtggctg tgtaacagaa gaaacacact | 1560 |
| aaactaatta caaagttaac aatttcatta cagttgtgct aaatgcccgt agtgagaaga | 1620 |
| acaggaacct tgagcatgt tactcgatggaa acctgcacag gtctgatggg tcagaggggt | 1680 |
| cttctctggg tttcactgag gatgagaagt aagcaaactg tgaaaacatg caaaggaaaa | 1740 |
| agtgatagaa taatattcaa gacaaaaaga acagtatgag gcaagagaaa taatatgtat | 1800 |
| ttaaaaatttt tggttactca atatcttata cttatgtatga gtcctaaaat taaaatgtg | 1860 |
| aaactgttgt actatacgta taacctaacc ttaatttattc tgtaagaaca tgcttccata | 1920 |
| ggaaaatagtg gataattttc agctattaa ggcaaaagct aaaatagtcc actcctcaac | 1980 |
| tgagacccaa agaattatag atattttca tgatgaccca tgaaaaat cactcatcta | 2040 |
| cataaaggag agactatatc tattttatag agaagctaag aaatatacct acacaaactt | 2100 |
| gtcaggtgct ttacaactac atagtacttt ttaacaacaa aataataatt ttaagaatga | 2160 |

ES 2 886 958 T3

| | | | | | | |
|-------------|-------------|-------------|------------|-------------|------------|------|
| aaaatttaat | catcggaag | aacgtccac | tacagacttc | ctatcactgg | cagttatatt | 2220 |
| tttgagcgta | aaagggtcgt | caaacgctaa | atctaagtaa | cgaattgaaa | gtttaaagag | 2280 |
| gggaaagagt | tggtttgc当地 | aggaaaaagtt | taaatagctt | aatatcaata | gaatgatcct | 2340 |
| gaagacagaa | aaaactttgt | caactttcct | ctctcatttt | ctttctctct | ctctcccctt | 2400 |
| ctcatacaca | tgcctcccc | accaaagaat | ataatgtaaa | ttaaatccac | taaaatgtaa | 2460 |
| tggcatgaaa | atctctgttag | tctgaatcac | taatattcct | gagtttttat | gagctcctag | 2520 |
| tacagctaaa | gttgcctat | gcatgatcat | ctatgcgtca | gagcttcctc | cttctacaag | 2580 |
| ctaactccct | gcatctgggc | atcaggactg | ctccatacat | ttgctgaaaa | cttcttgtat | 2640 |
| ttcctgatgt | aaaattgtgc | aaacacctac | aataaagcca | tctacttttta | gggaaaggga | 2700 |
| gttggaaaatg | caaccaactc | ttggcgaact | gtacaaacaa | atcttgcta | tactttattt | 2760 |
| caaataaatt | ctttttaaaa | taaaaaaaaaa | aaaaaaaaaa | | | 2800 |

<210> 18

<211> 418

<212> PRT

<213> *Homo sapiens*

<400> 18

5

ES 2 886 958 T3

Met Tyr Arg Pro Ala Arg Val Thr Ser Thr Ser Arg Phe Leu Asn Pro
1 5 10 15

Tyr Val Val Cys Phe Ile Val Val Ala Gly Val Val Ile Leu Ala Val
20 25 30

Thr Ile Ala Leu Leu Val Tyr Phe Leu Ala Phe Asp Gln Lys Ser Tyr
35 40 45

Phe Tyr Arg Ser Ser Phe Gln Leu Leu Asn Val Glu Tyr Asn Ser Gln
50 55 60

Leu Asn Ser Pro Ala Thr Gln Glu Tyr Arg Thr Leu Ser Gly Arg Ile
65 70 75 80

Glu Ser Leu Ile Thr Lys Thr Phe Lys Glu Ser Asn Leu Arg Asn Gln
85 90 95

Phe Ile Arg Ala His Val Ala Lys Leu Arg Gln Asp Gly Ser Gly Val
100 105 110

Arg Ala Asp Val Val Met Lys Phe Gln Phe Thr Arg Asn Asn Asn Gly
115 120 125

Ala Ser Met Lys Ser Arg Ile Glu Ser Val Leu Arg Gln Met Leu Asn

ES 2 886 958 T3

130

135

140

| | | | | | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Asn | Ser | Gly | Asn | Leu | Glu | Ile | Asn | Pro | Ser | Thr | Glu | Ile | Thr | Ser | Leu |
| 145 | | | | | 150 | | | | | | 155 | | | | 160 |

Thr Asp Gln Ala Ala Asn Trp Leu Ile Asn Glu Cys Gly Ala Gly
165 170 175

Pro Asp Leu Ile Thr Leu Ser Glu Gln Arg Ile Leu Gly Gly Thr Glu
180 185 190

Ala Glu Glu Gly Ser Trp Pro Trp Gln Val Ser Leu Arg Leu Asn Asn
195 200 205

Ala His His Cys Gly Gly Ser Leu Ile Asn Asn Met Trp Ile Leu Thr
210 215 220

Ala Ala His Cys Phe Arg Ser Asn Ser Asn Pro Arg Asp Trp Ile Ala
225 230 235 240

Thr Ser Gly Ile Ser Thr Thr Phe Pro Lys Leu Arg Met Arg Val Arg
245 250 255

Asn Ile Leu Ile His Asn Asn Tyr Lys Ser Ala Thr His Glu Asn Asp
260 265 270

Ile Ala Leu Val Arg Leu Glu Asn Ser Val Thr Phe Thr Lys Asp Ile
275 280 285

His Ser Val Cys Leu Pro Ala Ala Thr Gln Asn Ile Pro Pro Gly Ser
290 295 300

Thr Ala Tyr Val Thr Gly Trp Gly Ala Gln Glu Tyr Ala Gly His Thr
305 310 315 320

Val Pro Glu Leu Arg Gln Gly Gln Val Arg Ile Ile Ser Asn Asp Val
325 330 335

Cys Asn Ala Pro His Ser Tyr Asn Gly Ala Ile Leu Ser Gly Met Leu
 340 345 350

Cys Ala Gly Val Pro Gln Gly Gly Val Asp Ala Cys Gln Gly Asp Ser
355 360 365

Gly Gly Pro Leu Val Gln Glu Asp Ser Arg Arg Leu Trp Phe Ile Val
370 375 380

ES 2 886 958 T3

Gly Ile Val Ser Trp Gly Asp Gln Cys Gly Leu Pro Asp Lys Pro Gly
 385 390 395 400

Val Tyr Thr Arg Val Thr Ala Tyr Leu Asp Trp Ile Arg Gln Gln Thr
 405 410 415

Gly Ile

<210> 19

<211> 38992

5 <212> ADN

<213> Secuencia artificial

<220>

<223> Polinucleótido recombinante

10 <400> 19

| | | | |
|------------------------|------------------------|-----------------------|------|
| gagggagggt ggtgcttgc | taatggtgaa ttactaactc | ctcaataaag aatattat | 60 |
| gaaataattt ttgaaaatttc | ataattactt tgggttcttt | cttaatgata aataaataat | 120 |
| agtatattac aaacatacat | taatatttcc tgaatgaata | caccacaaat ctcccttaaa | 180 |
| atatacgcaag aataaaaatt | atactatttc tgacaatttt | taatttctca aataataata | 240 |
| ccactctgat ttttaaacat | ctacaccact ctggcttgc | caatctttt aaaaattgaa | 300 |
| aagataataa ttttatcata | attacactga agcatagaac | tttttcttc aaggaaagca | 360 |
| aattttgaa attctataat | ataacctccc ataatcctga | ataaattaaa ggttcaacaa | 420 |
| cttagtaaag taagactgac | cttccctttt atttctttt | cagataaaa atcttacttt | 480 |
| tataggagca gtttcaact | cctaaatgtt gaatataata | gtcagttaaa ttcaccagct | 540 |
| acacaggaat acaggacttt | gagtggaaaga attgaatctc | tggtaagttt atatttgtct | 600 |
| ttgctctta ttccattata | aatgaatat gataataaac | ctaattttt gtaatataatt | 660 |
| ttcagttgct aagtgctcta | catattttcc ttccttgaat | ggtgaaacat gtgttctct | 720 |
| ctgctttat ccagttagtt | tactcatata ctggttctta | ttcacatctt tgtcatgagt | 780 |
| aaaaagtgtt agaaaggcca | cgagtaataa tgcattttat | ttgtttatga attcaaatac | 840 |
| taaaagtttt ttatttgttt | aattaagcat tgacattgtc | tttttaattt ctttcattt | 900 |
| taccttcttc cctttcctt | atccaactaa agacgcaaag | caggagggtt taaaaaacag | 960 |
| gtttaccata tcagcagtaa | catagttgg acaacattac | actttggttc aatgatagac | 1020 |
| atagaagttt gaacagaaat | atgcaaagca agtttgagct | ctaaactgaa gagagcctct | 1080 |
| gggtgcctgc cagggaaacct | cacgagtggc cccttaacat | tcatgtgtca ccacaaacta | 1140 |
| ggggctgccc ttttagtttg | accagtcata gtgtcactca | cttaccctta cttttcaaa | 1200 |
| aaaaagtccct aagaatataa | agtaattcaa tggttctaca | attttagcat gtaactgagt | 1260 |

ES 2 886 958 T3

| | |
|---|------|
| cacctggcag gggtgctttg gtgagctcaa gataaaattt tatacgattt tctacat | 1320 |
| ctggaatatt ccttaatcca ggcttttaat cccttgggc ttttctgaac cactgcaat | 1380 |
| agcttctaact tggtctcaact gtgtgcaggc tctttccctt ctaatctaatttacacactt | 1440 |
| ctgaacacaa atctctcaca gcctgtttcc ttcatgttac ctccagctca agacttttg | 1500 |
| cctacaaaat aaaattcaaa ctgttagt aagcaccc tcatgtctat gcttggctc | 1560 |
| atatttcagc catcgtgtgc cccacttattt cttatagcca acctgaaaag ccattttta | 1620 |
| taagaaaacta cctctgctcttccatgattgg atataattaa tcccttccatc acatcaccc | 1680 |
| gccacaaaat tgtatctgtg ttgatctcat gccacatacc tgatgttatttataatttata | 1740 |
| aatatttgca gacttggtaa atttgccatg ttagactaag ttccatgaag acagctccat | 1800 |
| atccatttcca ttttatata tccacaacat ttggcgggt tgatgcttaa taaatgttta | 1860 |
| ttgaaggaac aggagtctcc cacttctgac ataatgaact tatttcccc agtgttacc | 1920 |
| ctacatctgg ttccgttcca agagtctttt cccaaatcat tctgattcaa ctgttcatc | 1980 |
| tgtatctcatt aaacatttaa atgatatac taacttcgt tgcttatttcatc | 2040 |
| ctgcagtcctc ctcataactt ggttcaatg atgcttgcattt ctagagaaaa aaatgttata | 2100 |
| aataagctta tgattcagtc ctccagctgt gatggtctc actgaacatt agctcagtgg | 2160 |
| ttttcgaagt atggctctta gcataaccta gaaacttggtt agaaatgcaa attcttggc | 2220 |
| tcaccaagac atactaaatc aaaaattctg acattggggc ctagaaatct gtgttttaac | 2280 |
| aaggctgcca gtgcagcctg gtccctttc ttctcgagc cccactcaaa gcttcaagt | 2340 |
| ctcatctccc accaatgaca gggcctctta tggaaaccgg caggacgggt tccaaactcta | 2400 |
| actacgtttt agagtttgc tcctagggtt atccaggcac caagtatcac aggttagttt | 2460 |
| cccagggaaag cagactctga gacttgcattt cagggagtgt ctctgggtt ctctcaacca | 2520 |
| acacccatcag gaagagaagg aagcagcatt gggcagaggc atagtcacac tacagtgc | 2580 |
| ttggcacaga agactgaagg gagtcagagc caggggttag aggtgggccc ttagcatcca | 2640 |
| tccttcacca ttaggtgtga gttgccccac ctccttgatg gtgtaacctc agtcccaagg | 2700 |
| tgggtggag tgcagcagag cagccctac aaggccaaa ccagagatac accaggcgcc | 2760 |
| agaagtgcgt ccagggaaata gagagggaaag gatggctta aggttaggatc cacagaactt | 2820 |
| ggcaatggat tagaagacag gatgagaagt gacaggttaa cactaacaca gaaatgtcta | 2880 |
| acttcggtag ataatggtgc cattggcttag aagaggaaac cgaaatgaaa gcaggttgg | 2940 |
| cagggagaca aaagttcaact gtggacatct cagcagagtg attcagttgg gaaagggatg | 3000 |
| gatgcccaga ccacccatcaga ggaagatcta agctggagcc agcaataaaag atacaagatg | 3060 |
| aacaatccct aacgaactgc tcctcagccca tgctccccag acacgctgct tcagatttt | 3120 |
| agtccgggtg aggcttaggag gtgcgcctcc ctcagttggag gacagcaaaag caccagttggc | 3180 |

ES 2 886 958 T3

| | | | | | | |
|------------|-------------|-------------|------------|-------------|-------------|------|
| tccagggagt | taaaatctt | tgataatttt | tgttctagca | tctgtctgca | gagctgtctc | 3240 |
| tcagccattg | cctgccttta | cacaggagt | gagtccgaaa | ttggggagatg | agtgaardttt | 3300 |
| attatgccta | gagatctgga | tccccagtt | tttgggagta | tattttctga | accacttgg | 3360 |
| ggttaagta | atgcagattt | attgatgcc | cttctcttga | atctgtact | ctggacccac | 3420 |
| catctaagt | aatgtgcaga | gggaacggaa | tggctgcaat | agatctccat | taaaaccagt | 3480 |
| gcacccccc | agacacatac | agtagtaggg | aggtgagtca | atgtcaggac | agcaccagct | 3540 |
| cccgcttcgg | tacatttcca | aagttctcag | tctgtgtaca | aaggtttgct | ctggggcagc | 3600 |
| agaaatagcc | ctggggcaggt | agtcaaaggc | ctgggttgat | ttcctccact | tccaggcaag | 3660 |
| tcactcgaag | gctcacaggg | ttttcctca | cctgccacat | gggtccagt | agatctactg | 3720 |
| agctgtaaat | aatgaaatga | gtgtgtgtgc | agtcatctat | aagttgtaaa | gtactagaaaa | 3780 |
| atggtaaac | tttgggattt | gggctattta | aggctgaatg | ctaaaaatgt | caggcattgt | 3840 |
| ggagaaagga | attnaaatat | aagattgatt | gactgggatt | taaagacaaa | tgaaggcaca | 3900 |
| cacgcaagt | cacacccaca | ctgacactgc | acagctcccg | ttggaggcat | atcctgacca | 3960 |
| tgcagacctg | gggctctgccc | tgtccaagt | cactccttta | ctacataaac | cctccttctc | 4020 |
| ttttgggct | gtcacccac | cagagctggc | accgagccct | tgctgctg | cttccctggg | 4080 |
| gtgtcagctt | ttgacaggg | gtttcctccc | tctgcaggag | ccttaacatc | ccttggactt | 4140 |
| ccttcccccc | acccacccccc | agcagtttta | tctttccta | actcgggacc | cttttttcc | 4200 |
| cacacaaagt | ttattgtcag | ttgctggttt | catctgttt | agcggctgca | acaaaatacc | 4260 |
| atagactggg | tggcatatgc | acgacaaaaaa | tttatttctc | acaggagaag | tcaaagatta | 4320 |
| atgcaccagc | agatctgg | tctgaggggc | cacccctgg | ttttagatg | atgctttcta | 4380 |
| gttaaaacac | ctatttaaca | cactattaa | cactaagt | gttaaatagt | gcagttgatg | 4440 |
| tatttgcatt | gtcaccttta | tcatacacta | aatccttctt | tgtcttttt | tctgtactct | 4500 |
| aatcttttc | tgtaagtaat | cttgcttgc | agcagtagga | tatttagagt | actgtggctt | 4560 |
| gacaatat | tttagtattt | aagatttcca | tgaatttctt | ctgatgtatg | agttccctag | 4620 |
| ttaatcttac | atatgttatcc | cttggtaaaa | acactttgaa | catttaaat | gatacatgaa | 4680 |
| tagtactcta | atacaatgcc | ataaaaatta | taaatcattt | gtatagactg | gtaagtaaag | 4740 |
| attgtgagat | taagaaacgc | atcaaaggcc | attgagctgg | aaagtggat | aatgagaatt | 4800 |
| caaaccagg | tctcttgc | caaaatctaa | ggatcatacc | atttctcatg | ataatatgag | 4860 |
| tattattgtt | atctctatcc | catagacaaa | gtgttaacac | tgaatgagca | gtgaaatagt | 4920 |
| ctcagaattt | tttattttat | ttagcaattc | acttgcatt | tctggcctc | agtttattca | 4980 |
| cgagtaaaat | aaaatagtt | gactagataa | tttctatagt | acattcttac | acaaaaaaatc | 5040 |

ES 2 886 958 T3

| | |
|--|------|
| tatgattttg ttatTTTaa tgtgatatac tcatggact cattcacctc atTTTcccag | 5100 |
| cctgcctcac tggtcattac ttctctgtgt tctttacagg ctccccctcc tctacactgc | 5160 |
| cattaaatat tgaAACACCT caaAGCTTA ctTATgtCCA CCTCTCCTCT gacactatca | 5220 |
| ttctgtctag atgatcccac acatacatgc ccattacttc aacctgtatt tatacgccaa | 5280 |
| tgattcacta tatttccagc cttagacattc ttttgtactc tagttaccag ctTgatATCC | 5340 |
| ttacatggct gtttcaAAAC aactcaaATA tattatCTCT caAAATCAAa CTCATGATGT | 5400 |
| ccccacacca tcctAGCTTT ccaccaacaa tacctatccc tattaatAGC aataccATTt | 5460 |
| attcagttat ccaaATCAAa aacCTAGAAT tcATCCTAA aattCTACTA tcATTCCAAA | 5520 |
| tatCCTATCC atCAGCAGCC actgtattct taatCCCTG tattTCCTTC aaATCCATTc | 5580 |
| acCTCTCTCC ATATCCATTG CTGCTGACT ATCCAAGCCA TCGCCTCTAC CCTAGGGTAC | 5640 |
| caaAAATAGCA acaaACCTAA tctgttcatt tgcattattt tttctccAAA actgattATC | 5700 |
| tatATGTagC aagACAGATT gttCTCAAAT tgCAAATCCC actatATTAT CCTCTTGCTT | 5760 |
| caaACACTC catggTTTCC cattgtttat gataAAACCA aatgCTTCAA gttcGAAGAC | 5820 |
| cggcatgatt gggAAATTCC tgcACCCta gcctacttgc tctccatggT acagttgcac | 5880 |
| tggCTTCTT tcattCCTTA agtacaACCT gttcCTCCC acctcaggac tgcATgtG | 5940 |
| ccattcattc tgctgaggAG ccttttCCT tccacttcaa tcagctaAGT ctgattCTTC | 6000 |
| ctgacaatCT cagCTCAATA agcatttCCT ctaAGAAATG tctctaATAT cattaATTGG | 6060 |
| ctcaggTCCC tctactgtat tgctGCACtT ttcacAGTT taattttACT taatttGAA | 6120 |
| tgattatttG attaggTCTA tttccatCCA ttAgacataa gttcatGAT ggCCAGAtta | 6180 |
| ctgTTTCTA tccatCgttG tattCCAATA cctgacAGAA ggAGGGCggg aggtggTggc | 6240 |
| acacaAGAGA tgctcaAAAAA caattGTTGA ataAGTAAT gaATGAGGCC atttagAAAT | 6300 |
| aacGAAAGTA CCTGTTACA aagtACATGT atcAAAActA tgaATGcATT ctacttACAT | 6360 |
| ggTTTCTCC aaataAAACA aaAGACTTCA atcaggAtta AtacCTGGGA taaACTGAGT | 6420 |
| cattAAATCT ctCCtttGCC atcaggAGtG acattGAAAC aaATGCTGC aaACAACAAA | 6480 |
| tactTTTtC CAAAAATA ttaaatGGCA tttccataAA caAAActAGAA catGGGAGGA | 6540 |
| gaaAGAAAGC aatAttaAtt taaaAttaAt tttatCACAT aacttAtAcc atcaggGATT | 6600 |
| tcgggtaaaa ttccTTTCAG gcacatCCAT ttaacaAGAA ttgattGTTA ctgAAAGCCT | 6660 |
| agaAGAGAAT ttggcacATA cttggTgtTC aaatATTGT tgactGAGtG aataAAAtGAT | 6720 |
| gcaAGTGTCT aagAAACACA aaATAAGGAC atgattACAG tcacGGTGA gttcacAGTC | 6780 |
| atctccAAA tgaggatATG catcccAGGG aggACCAACA attcattGGA gtGCTGAAAT | 6840 |
| aaaataCTCA aaggTCATT tacatgtatt tttctctAA attactttC ttaAGACACA | 6900 |
| gaaaacAAAA aaAGAAACtt agctttGTTA ctttctaACA aatAGTTAA tcattAAACA | 6960 |

ES 2 886 958 T3

| | |
|---|------|
| ggattgacac tagcatcctt gtttggtctt atgccttagg ggaacatgaa atgtgtgaag | 7020 |
| acattctgag atctgaggga aggtagaca gtaatacagt gggactgacc aggcttcagc | 7080 |
| acaccttac ctcctctcag cagattcag tgatgagcag tttacaacta gattgaaaga | 7140 |
| ttatattatc tagttctaaa agaaaactaa gcctcccaa agcaacaagg gaactgagag | 7200 |
| gaatcctgca aaacaaaaac aaatttaaa acttgcactt tgtaataacc ctaatatgt | 7260 |
| atcacagtaa tgaacagtaa gataatgaca gaactgacat atttccttat ctattaaagc | 7320 |
| catattaaca ggtaaagcaa tgccagtcag tggtagactt cttagaagat atttaataca | 7380 |
| tactagacac atacacacac acaacatccc cttcaaggt gtatgtatca gaaaatcact | 7440 |
| tttttaaggcc ggtatgcagtg gtcaggcct gtaatcccag cactttggg ggcgcacgtg | 7500 |
| ggcggatcat ctgaggtcag gagttcaaga ccagcctgcc caacatggcg aaacccccatc | 7560 |
| tctacaaaaaa tacaaaaatt agccagggat gatggtgat gctttagtc ccagctactc | 7620 |
| aagaggcaga ggcaggagaa tcacttgaac ctgggaggca gaggttgcag tgagccaaga | 7680 |
| tcacccattt cactccagcc tggcaacag agtgagactc tgtctcaaaa aaaaaaaaaat | 7740 |
| cacttttag ataaaattca tgctatagag agaagactat gaaaatatgt ttagcaatgt | 7800 |
| gtccatcatt aggtgattga gtttccttt gttttgttt actgaaaatc atataaagta | 7860 |
| tgttatctgt aaaagttctc tgacatgcac acataaaaat ttgggagaaa agattaacta | 7920 |
| taatgtttaa tagattttgt acacattctt taaaaaatat ataaaacaca acacctttca | 7980 |
| atgggttgc aagaataacc aattgacatc atggaaaatg gaaattcact tgctgaat | 8040 |
| taacaaaaat ttgcattatg agtgagactg acaacttagt gtcattttt aatgaattat | 8100 |
| gccaatggta aacttcatgc acatggggcc aggttaattat gtggaaactt tttcaatgct | 8160 |
| taaagccaag tattgaaatt aaacttagaa tcagaccttt gaaccatttt atgacaatgt | 8220 |
| tcaaaaatta taaattctat ccacttatata tataatatta aaaatatcat tacaaaaaaa | 8280 |
| acctgtgttt attttataac tcagccttt taatttctaa tttcataat atattataat | 8340 |
| ggatattgtt agtaatgttag tattattaca tgtatataat ttataagtaa atatacatgt | 8400 |
| tttggctact catgcataaa atgtttcacc cataggagca cataatcaga aatgtctgga | 8460 |
| gaccattata gtaatagata gatcatattt ccacatattt tatctcctcc ttgacaactg | 8520 |
| agctttccag atcttctggt gaaacgaaag agaaagttgt aacagaagag tgattaaaat | 8580 |
| gacaaaagca ttacttctat tacttctatt ctaataatat gagcaaagct ataactatca | 8640 |
| agtaataatg cactaaagaa ggtgatataat ctgatataatt cacaggcaac taataagacc | 8700 |
| tttctattgc agccatgaaa aatatgtgac aattatagat atcctgtgtg cagtgtttca | 8760 |
| acctttatgt gacctgttct actaacagat tttagttagt tcactttgtt agaattttct | 8820 |

ES 2 886 958 T3

| | |
|--|-------|
| tacacatgcc ataacttgct tcagtctttt gattatgaat attatggata ttaaggattc | 8880 |
| tagactattc tagatttaaa aaataatatt gtcaccta tcagaaggga aatattaaat | 8940 |
| agttctcatt tttcaatgt ttactcagtt tttgtccaat gtaatgaaag tgtcagcagt | 9000 |
| acaggttaca aaataaaaatg tgtattaaag taaactcatt tgaacaggtt aataattgta | 9060 |
| gagggagggaa aaaggctaaa agattgaatg taaaacttat gaaaagtaga tacatcgct | 9120 |
| ctatgattt cagtagtcaa ctgcatacag atgaatcatt ttaatacagc ttaactactt | 9180 |
| tcctttaca gatggagaaa ctgagaggaa gaaagtttat atggttcatt aaactttgt | 9240 |
| atgcaagcta aactaacctg tctctgtatt ttccatctac tgcccttac actatctcat | 9300 |
| tagaatactc ttcaagcatc tccttactga ttttcttacc aagcatttgt taagttctaa | 9360 |
| tgagagttgg tagtaacatt ttcacccact ctgtgaaata tgaatctta ttcataggcc | 9420 |
| tcttctttta ttcttgtatt tgcatatcaa ccaattaatc aacttgctt cttatgtt | 9480 |
| cttattatct tagtccttac taaattgcct cttaatgtt tccacataac agaaatgtt | 9540 |
| aggtggatac ttaacatttt agtccagtct agccggtgcc agtgcaatgc caaatcatga | 9600 |
| attaaaaat aattacaaga accacttac aaattttAAC aattccttca gctttgtgac | 9660 |
| agtttttctt acttcgatta aagtcaagta aaattaaagt taaatatttt tattaaaata | 9720 |
| tctcctttaa cattccatat taataaacat attaaagctc atgcttctaa gtagattact | 9780 |
| agaagttact ttatcgaatt acagcaatgg ttaattctag atcatagaat tttagaatgac | 9840 |
| tttttgcctt ctctttttt ttcctttttt ttaaacagag tcttgctctg ttgtccaggc | 9900 |
| tggagtgtac tggcgcgatc ttgactcaact gcacacctg ccctgcaggt tcaagtgtt | 9960 |
| ctcctgcccc agcctcttaa gtagttggga ttacaggtgc ctgccaccac acctggctaa | 10020 |
| ttttttttt gtatTTTtag gagagacagg gtttcaccat gttggccaga ctggtctcga | 10080 |
| actcctgacc tcaagtgtac cacttgcctc agcctccaa agtgctggga ttacaggtgt | 10140 |
| gagccactgt gcctggcctg acttttgct ttcttcttaa tacttactag tatttcttga | 10200 |
| atTTTaaaa aagaaacata aagtactttg ataaaaccaa cagtctcatt gttctaaaa | 10260 |
| ttgttcaaag gttctctgga aaaaaaaaaaag aaaattatca tttggtaag aatcatgtt | 10320 |
| gtctgacatc aatcatccta taggagtcaa tattgaaaaa gtaagatata ttgtggtata | 10380 |
| atcgagattt cataaattttt accattttt agaagaatct gctccaaatc ctggcttaat | 10440 |
| gtaatatcca gcatgctact taatttctt gtcttcaccc tttcatatcc acatccaccc | 10500 |
| aggtgccacc tcacagtata agccagcata atccattctt ctcaatgaaa ccacaataca | 10560 |
| tctgaccctg catctcagga gaactgtatc agccacagca cttccagttt actatgaatc | 10620 |
| tgaatgttat gcctcaggag aaacatcctt gctggactg agtagtgatt caaggagata | 10680 |
| gttatgattc agtcaagaaa ttaataatta gtgttatttt tattattgag acagagtctc | 10740 |

ES 2 886 958 T3

| | |
|--|-------|
| gttctgtgc ccaggctgga gtacagtggc atgatctcg ctcactgcaa cctctacctc | 10800 |
| cccggttcaa gtgattctcc tgcctcagcc tcccaaataa ctgggacagc aggcacttgc | 10860 |
| caccacgcct agctaatttt ttgtattttt agtagagacg gagtttcacc gtgttagcca | 10920 |
| ggatggtctc gatctcctga cctcaaggc cacctgcctc agcctccaa agtgctggg | 10980 |
| ttacaggcgt gagccactgc gcccggccat aaattattaa ctgagccagg cacagtggta | 11040 |
| cacacttata gtcccagata ctcaggagac tgagggttga gtatcctttt ttatgttatt | 11100 |
| ttatTTTAA ttattatggg tacataatag gtgtacatac ccatggagta caagtcatgt | 11160 |
| tctgatacag acacataatg tttaataatc acatcagggt aattgggata tccatcacct | 11220 |
| caagcattta tctttctttg tgtaggaac attccacctc cactcttggaa ataggcaccc | 11280 |
| tgttgtgcta ttaaatacga ggtcttattc atttcatactc actatatttt tctaccatt | 11340 |
| aaccatcacc tctttcccc tcttccccac taccttcct gtgaggctgc aggattctta | 11400 |
| agcacaacag ttagaggcca gcctggacaa catagtgaga ctcaatttct aaaaaataaa | 11460 |
| aaagaaaatta ccaactaatg ctaaaaaaat agtctctgat gcttaggtat gaattagaaa | 11520 |
| tgaccaaaaa aaaaaaaaaa aaaaagactg ccctttgctt ctttctcccc ttctcttcaa | 11580 |
| gtttccatt gctactcatt ttagtctggt ttaatcaggt ttcatccatt aaaagcaatt | 11640 |
| gttggatca cacatttga gttgtgtcag tggacttccc tcatactggc atgattcctg | 11700 |
| ccccaaagccc ttagaaaaag ccaccaagcc atataacata atctctcatt gagtaaaaca | 11760 |
| tctgatgtgt ttagaatgac ttctagcaaa aaaccagcct gtccagcatc atctctgtat | 11820 |
| aacagataaa ggaataggta ctgcatcaaa aggttataga acctgccccaa atcaatcccc | 11880 |
| tgtgtttgc aatggaatta ggttgaacta aagtgaaaat tcagtttct actcctcatt | 11940 |
| aacatgtctc atgttgcag gttgagagga aggagaagaa gaactgtatt tacagagaga | 12000 |
| ttccccctct ctttcttct acagattact aaaacattca aagaatcaaa tttaagaaaat | 12060 |
| cagttcatca gagctcatgt tgccaaactg aggtgagtgg aactgtagaa aaaatattta | 12120 |
| agtatagata caatgtggca tacttgactt tttgtcacag aatgaatagt aaatgacatg | 12180 |
| ttcagataag ttgttgaat attatgaaaa tagtattttt gtcagctaa aaaccaatgc | 12240 |
| aaaaaaagcc aaacatatga tctatTTGC tactaatgt aataaccata ttatATCTAT | 12300 |
| tcttattggg aagaggaaga aggggtggag agagagttgg ggtgaaggta cagtaacaag | 12360 |
| gccatcctat tgtaaaactc cagtggatat cattcacagt gcagcctatg taaacagtcc | 12420 |
| ctcctggagt tgtacaatgc tgggtttgg gtgtatccat ccaagatcaa gacactatga | 12480 |
| ccaacatcaa aagtggcttt ttggTTTAT ctgcctgatg tgctataata aaagggtatt | 12540 |
| atggccaaat ccaaggcatg tctatcatga attaataata ggaggagtag cagcatgcat | 12600 |

ES 2 886 958 T3

| | |
|---|-------|
| gctagttatt tgccattcct gccttagtta aatatgatgt gataaaacca gccttc当地 | 12660 |
| ctgaaatagt caccttact gactctccg caaatgtctc aaatgaccac attgctctag | 12720 |
| tctttaaata atatgc当地 aataactaa gttcttggt agaagaggaa ttataactaa tcttctcaa | 12780 |
| atactagcat cacaagaaaa ttaattcttgc ttctctggag agtcacccatg taagtatctg | 12840 |
| gagcacagat gtctggtc当地 gtaagtttg atgaggagtt aaaggataa gaagagtcca | 12900 |
| tgagaagggt atttccaaa acaccttgc gtcaattc当地 tgcacattca cttagactt | 12960 |
| tcttgc当地 atctgtatca gccactaatg ttcaaaatg agtaaggccct gaaaacctgt | 13020 |
| aggactacat gagccttctg cttttctct cttttgttc acttccact tatcactcaa | 13080 |
| tcctctgcaa cctggcttca ataccaccaaaa atccatcaa ctgctctgc cgattcaaca | 13140 |
| atgacatcca gataacaaaa tccaaagaaaa ccacatc当地 cctattcttgc gaccttcaa | 13200 |
| cagtatttgg tcctgttggc ctgtcactcc ttgaaatagg actatccctt ggttgc当地 | 13260 |
| gccttgata ccctgatccc ccccttaccc cttagctat tccttcttag tttccttac | 13320 |
| taggtcttac ttcttgc当地 attccttaaa tggctgaa catcaggctg tgctctaggc | 13380 |
| ctctcatctt ctcaaggctc actctcttgc ttcccttggc ttcaactgcca cccatatgct | 13440 |
| gagtgc当地 aaagttgtat ctctaggc当地 gtccttgc当地 gctccaaac atgaatata | 13500 |
| gcagccatct acttggtaacc atcacatgga taattctcat gatcttcc agtataactg | 13560 |
| cttcttatt ttttctggc ctctttta gcattgc当地 acatggaact ttatcatgc | 13620 |
| tctcaacctc tattttatct ttatctatg tatgtagact ctgtgtaatt tcttcatctc | 13680 |
| ttttagataa ctaatatctc ttcaagcttgc acttgc当地 tggtaaccc atttattgc当地 | 13740 |
| tttcaattt caatgactat gtttcttat ctgcaagttc tattttttc ttttggaaat | 13800 |
| cttcctggc tttaaacac atttcttatt ttaattttgc ggggtaccta gtagttgtat | 13860 |
| gtattttgg agtacatgag atgtttgtat acaagcaac aatgc当地 aatcacattg | 13920 |
| tgtaaaatgg ggtatccatc ccctcaagca ttatccttgc tggtaaccc caatccaatt | 13980 |
| atattcttgc当地 agtattttttaa atatgtacaa ttaatttttgc attgaccata gtgactctgt | 14040 |
| tgtgctatca gatacttagt gatctttaa aaataatgtt ttctacttgc tctcatttttgc | 14100 |
| atgattccct ctttacgtc atttgc当地 tcaaatacag tcacttgc当地 gttgattcttgc | 14160 |
| ttatgtgaaat gttttggaaat taatcttgc当地 gttactttgc ttccaccc tggatggtttgc | 14220 |
| gctgtggccca cactaaaatc tc当地tgc当地 ctctggcc当地 cataataaccc acatgttgc当地 | 14280 |
| ggagggaccc tggggaggt gattagatta tagggacgtt tcccccttgc gctctgttct | 14340 |
| ttttccctgccc accatgtgaaat aaagatgtgt ttgcttcccc ttctgc当地 attgtaaatt | 14400 |
| tcctgaggcc tccgc当地tgc当地 tgcaggaccc ctttcttgc当地 taaattaccc agtctccggc | 14460 |
| ggttcttgc当地 agtccgtgaaat gaaaaacta atacacaccc catgatgtat tggattaccac | 14520 |

ES 2 886 958 T3

ES 2 886 958 T3

| | |
|---|-------|
| caatggagca tcaatgaaaa gcagaattga gtctgttta cgacaaatgc tgaataactc | 16440 |
| tggaaacctg gaaataaaacc cttcaactga gataacatgt aagtataatt ttccataaac | 16500 |
| aattttatcc caatatatcc ctcaagtttccaattcaaa ttcatatccc aattgagagg | 16560 |
| ctgacttttc ttctttgaa actaaactgt gaaaacaatc cattaaaaag ctaaatatac | 16620 |
| catatagctc cctaacgtaa atcattctaa gacttaaaga atcatttggc atttatatag | 16680 |
| taaattttat ttgctaaaaa ttctcattaa ttatccctgc aacattcctt atgagtgtat | 16740 |
| ttactgtcag atgtcattag tggataggcc ataggagggg tacatagatg ctcaaggtca | 16800 |
| gagaactatt taattaatga tccaccccgaggcttccatttttttt gtaacattta | 16860 |
| tcacaattga aattacaaag ttatctgtgt aaattttgtt tggttgct tcacccctaca | 16920 |
| ctgtaatcat cctaaaagaa agaaccagtc aaccccttccatcctactac ccccttacca | 16980 |
| cccaagtcctcc atcatataac acatattcaa taaataattc ttgcattgact gaaagaaaaag | 17040 |
| aaataatata tgcatacat ttaaggacat tcctccaaatggttacatt ctgcttagtt | 17100 |
| aataagccat tatttttttctt cgtgagctc aagattaaaa ggattttgtt gattccata | 17160 |
| ctagactggt aggtaccagt tacagatgtt ctaactgttta aatattgaaa tggttccata | 17220 |
| tttgggttta aacaattact gcatcaggcc cacaaagtttgc ttccggaga tggttcaat | 17280 |
| ccactgcccc tgctgctaaa gagttatgtt tagcaaagca aagcactcta agacactgct | 17340 |
| ccaaactccat ggcctgatttgcatccat gactggccaa tgctcacgca ctgcagtttgc | 17400 |
| tttaggttagttt gaatattacc tctgcttcca cacattaagg aatgctcccg aacgcacttc | 17460 |
| ccaaagtgtttt atttattttt cattatacta gacaatatgg tgatacgatg gtcacagaat | 17520 |
| agcggtttcc acctccagag cccataatct agtgaaggg aaagatatttca aacacacaaga | 17580 |
| gtgttgacaa tcaagataga atatgatcaa gggccagtg tgaggccag gcaatgatca | 17640 |
| ctgcaggaat ctggggaaaga aagagaccag cgtgcttggg atatctagca aaagtttcat | 17700 |
| gaaggagaat ggactttgac tttgaaatat gggtaggatt tacatattttt gagatgagaa | 17760 |
| aaagaaagtt cccagagaag gaaagcatga aaaggcaaac agtctgtact gaacgcgtat | 17820 |
| ctttgacaga ataatgaaga aaggacccctg ctggaaatgtatgatcgttgcattca | 17880 |
| caccatcatc atcaaaacac ttatccatg agaacttact gttttttagg catggcttta | 17940 |
| atgccctata tgaattttttt tcttgattaa tccttacaac aaacatatacc catagatgt | 18000 |
| tttattgtcc cccttagaaaa agataaatttgcctaggctga cacagtcaatgatgaggca | 18060 |
| gtcaggattc aaactaagtc tggttgcattca aaaaattaag aatggccagc tttttaaaat | 18120 |
| tttctgtctc cagaagtatg atttggctcc actgaagttt gcaaaacaaa tgtgatacc | 18180 |
| aaaccttgcgaaacttttag tggaaataa cttgcataa gtcgggttga gagagcgtgg | 18240 |
| aaacctgtct tggaaatgtttt taatccatg tgcaggaaat aaaaatgtatg ggtttctcaa | 18300 |

ES 2 886 958 T3

| | |
|---|-------|
| ttaaaaattt caatcaagga aggatatgag ctaacataac attttttaa aaagatcagt | 18360 |
| ctggtaaggt agaggtgcat aaactgaaaa ggagcaaaag tggtgaaatt cagttagaaa | 18420 |
| attattgtaa ctgtactgat gtcaaattgat gaaaccatga actaaagtag tacaaaagg | 18480 |
| agtgaggagg atggaataat tcaaaagata gaggacagat gtgcagaacc tggagattat | 18540 |
| aagatgtgaa aggaggagtt tgagaaaatt tcagatttg gaagtggtgt cattttacta | 18600 |
| aaaggatata ataagtagca aattttggat aaagttgggt cccactgagt ttgagatggc | 18660 |
| tgttggacat gcagagaaaa ctgtcttgc tgctgttctt aaattgaaat agacagacct | 18720 |
| ttaccctctg atactgacat atttccttt ccaggctcac cctccatttc cctaaacaca | 18780 |
| acacatgcac tagctctcct tactttattg ctccacaaac atcttacacc tccaagcatt | 18840 |
| tgtgccact gtaccttcta tctggaatct ctttgcctt cttgtgtgcc tgaaaaattc | 18900 |
| ctttcagatc ttcaaaatac agtgcagatg ctatttcttc tagctcaaattt attatctcct | 18960 |
| ccatataatt taattactct ctttttctt ttctctactt tgcaattaca ttatattgaa | 19020 |
| tgattgcttg attaatttctt acctgtaaat tatgtgaggg caggtcctctt atattttgtt | 19080 |
| cgcagttaaa tctgcagcac ttattataga gtggtatcat tagagtaata tacatataatt | 19140 |
| tgaggacatg ataaattaac ttcccctata gtatttatca cattgcattt caatgacttg | 19200 |
| cttatgttcc tggggccca tataaattga gtaacttgaa aaaagagata tctattaat | 19260 |
| attnaatgag aaattaaagt acaaacttta gtatgcataa caacaaattt gggaaaaggaa | 19320 |
| gtaaaacaaag agattttagt ggcggcatgag ttagagatcg tttcagcagg tctgaaagga | 19380 |
| agcctaggaa tctgcattttt agaggaccac ctcccaaccc caacaagtaa ttctgcttct | 19440 |
| tgttgctgg gtactgtact ttaagaaattt atggtaat gatatcagcc ttatattgtat | 19500 |
| ttatcttattt ctcattttt aataactgca cttactgacc aggctgcagc aaattggctt | 19560 |
| attnaatgta agtttaataa ttatattgtt actgttaattt gccaatcat aaagagtaaa | 19620 |
| agtgcagtc ttttgggttac ttttggccaa ggcagttatctt atcaagttga tgtctttgtt | 19680 |
| cttagttcgc tcaggtggtg ttgaaacaag acagtgcgttga tcccaagttgtt cccatggagt | 19740 |
| ggacttttagg ttccccctt ctttttagaa aaaggaagaa gttgttagtgg aggactaccc | 19800 |
| actctgcact caaaatttgc ctcattttttt tttcttggc agctttgaga accttttact | 19860 |
| gccctgggttca taaggtggca tttcttgcata cttacaaattt atgtttgtt acaccgttta | 19920 |
| tgttagcttctt ctcattttttt cttttttttt gttttttttt gttttttttt gttttttttt | 19980 |
| gtataataaa aaagaattgtt cagaagtctt cccagctttt ggtctataac ctgtttttttttt | 20040 |
| agtcaactact cttcaacatc atcctatgtt ctctcaggctt aggtttttttttt gttttttttt | 20100 |
| ccttagaaaaac agcaacttac ttctctgacc aaaaaaaaaatgc agttttttttt tagttcaatg | 20160 |

| | |
|--|-------|
| tacctggtag ctggcctatc ttaggtactt cagtgatttt acaaagtat ggtgccta | 20220 |
| tgggtttt tcagctcac tacgtattta attcatgctt attgttaatg aaactgtat | 20280 |
| aagcaattta ctagggattt tggttggag atgccacaaa ggaacacatg tatctctaa | 20340 |
| tggaagcctg gtcctcctt atccaggaaa ttgtcttaga aaaaaaagcc tttaggttgt | 20400 |
| tgtgttata aaccaggca ctactaaaa gccagcccag caatagttgt gtgatttacc | 20460 |
| attaatttct tagtaataga ccacacaaaa gaagaaaatt atggaatgc gagttgagag | 20520 |
| gaattgggtg atcagcctac cccagccgt ttcagctctg gccagtagac tattcacgag | 20580 |
| ctcttgaaa acattnaat aaaccttatt tagatactag aaaccctctg tcaccctcaa | 20640 |
| gaatattctg tggtatagcg actcccttat gagggcatgt ttggtaatac agcatcagtc | 20700 |
| ttggaggtgg actggattct acaaggtgaa ctgcagtcac taaggagtct ttggatgag | 20760 |
| accagtttc ctccaacttc aatgtgtgca tgaacctcac atcaaaatgt agctttagat | 20820 |
| ttgtccccatg atgtggttcc aagaatcagc acttctaata agttccagg ggatgcccatt | 20880 |
| gctgcaggcc cacaaaccac actgagcata gcaagactat tgagaaaaag gaaatttccc | 20940 |
| aggagtctgt ggcctgagct ggcacatcca ataatgaccc atcttaacct caactcatga | 21000 |
| ggaattccag ggaactctga agctgctcaa aatttgaagc ctatatgcca actaaattca | 21060 |
| gaaatgttct ccaaaatgt atctataagc aacagtagtc acaaattgcat tggatgaaata | 21120 |
| tatcgatcat gcttttgga aaatccagca tgtcctgagg aagaatgtat aagacataaa | 21180 |
| agtcatataat tatggaaaga ctcttcagct tcttccaaat gtggaaat catgatcttc | 21240 |
| ccagcacatt aatgcccttt ctcattagaa tgtggggccg gtccagaccc aataacattg | 21300 |
| tctgagcaga gaatccttgg aggcaactgag gctgaggagg gaagctggcc gtggcaagtc | 21360 |
| agtctgcggc tcaataatgc ccaccactgt ggaggcagcc tgatcaataa catgtggatc | 21420 |
| ctgacagcag ctcactgctt cagaaggtga ggccaccact acctacccat ctggaaacaa | 21480 |
| ttagaataga caggtcatga agactgcacc ctctacccta ggattgaatt gagccagaaa | 21540 |
| taattcaatg caaaaaatc agtaagaatt ttcttcctat tcatgaaagg aaaaggattt | 21600 |
| ttccccctta gcatgctaatt tagtgctat ttctctgttt caggtataaa tatatttagca | 21660 |
| cagtaaagaa caaagattta tatgtcagaa tgtttttaa atcctagcta taaaagctta | 21720 |
| agaaatttac taaatctcca taagctttat ttttttcca aattaaggga caacactgtt | 21780 |
| atctgtgact tagtgtaact ggtgcattt agtacactaa tgtaaacata cgtaaatgt | 21840 |
| tagcgaacg aattgctgtg gaagatttgc acattatatc atggagctg atggctaacc | 21900 |
| tagagactgc cccatgccat taatttatttcc attcataaaag attattgagt atcttagatg | 21960 |
| agcacagtgt tatatatgtt agaagctact agtataaaca aagtattgcc tctgccttca | 22020 |
| aagagcttac actcgaatgt tggaatcaga atgcacaaaa ataatgatca attacaatga | 22080 |

ES 2 886 958 T3

| | |
|--|-------|
| gtagcataaa taaaattaat gtaggcaact tacaagaatt cttaatggag gtgactaaac | 22140 |
| tattgccaac actagggtga tatgctacca gtggcgagta gggtgcataa acttaccta | 22200 |
| ttggtaaaaa gaaaagttca cattgctcat aaaagaagga ttttagattt cagcataact | 22260 |
| aaaatctgtt tcaaacctgc cttgttactg gggcatcgca gaccacaaca gttgttggga | 22320 |
| acttaactca aaaagttcac ccagaaaaat aatggagatt tgaactcgtg tgccctgac | 22380 |
| cataatcaatt ttcttctcag actcttactc taaactggac ctccttatca cacacacaaa | 22440 |
| gccttccata ggcagatcaa tccagtcctt tttctcaaag catgtacctt gagttcaga | 22500 |
| taaacagcat tggctcttc ccctggactc ttccctacatt tccctaccta ttagtatctg | 22560 |
| atcaatctgc ttatccttga aatgttaata tatttaccac atctctattt gaattttatg | 22620 |
| aaatttttga taatttctaa gtagttttt cagattata ggcactactt catggtacag | 22680 |
| tgactgttac aaacgtattt gttaaattta gaaggaataa agataaaaa gactaggta | 22740 |
| gttactgaac taaagttta ggaaatccc aattatttca aatttttctt atggtaattt | 22800 |
| tatgacttaa tattttata tgcagtgaac aaatttgaaa cttaaaaaga tactcccaga | 22860 |
| attatcagtt ttctgatgta gattggcaaa tttattacta tatcccaat aacccaagag | 22920 |
| acaaaattca caaaaacatt tcaatttca ttgccacttg aaaggccaaa aagcagaaat | 22980 |
| ggcacgcatt gattcaatc gtactcttga gtgtggAAC caggaattaa aatacctgga | 23040 |
| cttacaggc acttagcata accaagaacg gaatagaaac ctccctggat tctaagccct | 23100 |
| attcagtcac aatcaccaaa aaccaagtaa acgatatac tataatgaaa gccacagtta | 23160 |
| taaatatcga caacgattac caaaggaatc catggaactt tgaattttgc cacccacat | 23220 |
| ccttctattt attaccatga ttgatccact aaagctaaca gactctgtga accttgtatt | 23280 |
| ggacccctcc ctaaagacct gattgtcact gagaaccatc agtgaggatt tgtttgggc | 23340 |
| atgaccagcc ttacatcaaa gtacatagaa gtgatgaggt cttatcaaag aggattattg | 23400 |
| aattatcacc tcttctatgt agctttccct gatactctt ttcctctcca ttgagttcca | 23460 |
| cagaaatttt tttatctgcc tttaacagtt gtcctcatga tttgtgatat ttgacttacc | 23520 |
| tcttgtcagt ttccctcact agttagatg tcctcaaaga aagagaccat aattacttat | 23580 |
| attttattt ctggagactc atactattcc ttatacaaag tagacactta acaatggctt | 23640 |
| gttgaactat aattaatgaa aataatagct accttcatga aagttcactt tggccaaac | 23700 |
| actatagtt acataataca tttgtctcat taatacttaa caattgtgtg agaaggtatc | 23760 |
| accaatcaca ttttatatgt aaataaacc cagagctatt aattaacttgc tcataaataa | 23820 |
| cacttttcat atgtggcata gccaagattt aaatataaaat gttactgggtt ccaaaatgtat | 23880 |
| gctctaattt cttgtgttga aagaaggaaa ggaagaaaat aaacgagtgg aaggaagaga | 23940 |

ES 2 886 958 T3

| | |
|---|-------|
| gggagggaaag agagaaaagg aaggaaagaa aaaagagtct cttcagaacc ttcactgtaa | 24000 |
| agactccgag caaaagaagt tgaatataaa aacaacatag gtttgttgt tttctaataat | 24060 |
| tttttcttca aaatttttaa ctcaggttca ctcttacaca aactactgtg tcttataaaa | 24120 |
| gtatcccgg tcatagaatt tttattttct gtattaactc cactatctaa tctccataaa | 24180 |
| actcctaaat tggttattatc ggtaacattt tgttttact caacccttag gaacaatgtt | 24240 |
| aagttaatca gccctccaca tcacagatcc ttatttcat cagtctgtac aaggcatttc | 24300 |
| tctcatttta attttttttc ctccgtcat ccctggattt cacttcaact gccctccccc | 24360 |
| cacccatatg cctcataacta atatattcga aatatacatg tcttaaaggt acatgcacgc | 24420 |
| acctacaaaaa cctatagtgt tttttgtat gtatatgtct ttaatttaaa taagtagcat | 24480 |
| tgtgtaaaag tctaataattt tttcttactg ttttcaactca attcttgaa ttttcatctg | 24540 |
| atgcactgct gcatacgacc ccatggatg cagccaccat atttccttca tccaattagg | 24600 |
| ttgcatgacc taccttcccc ttgccacaaa gagtacacac aaaatattt tacttatctt | 24660 |
| tctgtaaacc ttcaggaatt tcagaagcac acatgcaggc tgctaaatat accagaatac | 24720 |
| tttccagcca cttaaatctt taccagtatt gcaaaagagg ccccatttcc ctccacatca | 24780 |
| acatttagta ttattctttt gtttaagttt tatcaatctt ttaaatgtac acaagatgct | 24840 |
| cattttata atttaattt ctcagattac tagtttgagt atctttcat atatctaaga | 24900 |
| gctgtttga tctccctac catgaactgc cactaatatt ctgccttat tttacaatgg | 24960 |
| ttttctgct tatttattac tggtttacag acttttaaaa tatattctac aaaaatttta | 25020 |
| gacattaaac attaccaata tttcccatg gttcctcattc catctggtaa acttgtctat | 25080 |
| ggatatatcta attttgattt aatagaattt attctatttt taccttttag tttgtgtttt | 25140 |
| tgttgtttag ccaaaaagtc cccattccta ggtcataaag gtaatgtcct tttttttttt | 25200 |
| ttaacgctac tttctctct ctgtctcccc ctatgtatat aggtgcacat atacttgta | 25260 |
| acacatacat atacctataat atgaggggg ttcgataagt ttatggaaaa taaaattaaa | 25320 |
| agataaaaata aaaaattata aactttattt ctcaacataa gtcctcaa gttcaagaca | 25380 |
| ctttgttaag caataatacc agccatatcg tccatcccta aagaactgag ggtcctgaga | 25440 |
| attnaactat gtcaatgcag tctttttac attactttt tacagtaactt attgtatgaaa | 25500 |
| aatgggtgcc tttaaagat tggtaaga tttagggaca aaaataagtc agaggaagtc | 25560 |
| aaatcaggac tggaaagggtgg atgccttagt atttattgct gaaactttca taaaactaac | 25620 |
| cttatttgat gagaggaatg agcatgagca tggttgtat ggagaagaac tctggtgag | 25680 |
| cttccttggactaaagctt tggtaactt tcttactctc ataagaagaa | 25740 |
| gatgttattt ttcaactgacc cttagaagg tcaacaagca aaatgccttc agcatccaa | 25800 |
| atgtctgttg tcatgacttt tggtttgac tagtctgggt ttgcttgac tggaccactt | 25860 |

ES 2 886 958 T3

| | |
|--|-------|
| ctaccccttt atagccattt ctttgatggt gctttgtctt caagattgta ttagtaaagg | 25920 |
| catatttcat cttctgttac aattcttcaa agaaataactt cagaatcttgc atctgacatg | 25980 |
| ttaaaaattt ctattggaag ctctgaccctt gggtgcagct gatctggcg aaacagtttt | 26040 |
| ggcatccatc aagtagaaaag tttgctcaac ttttagtttt cagtcagaat tgtataagct | 26100 |
| gaaccagttt agatgtctat ggtgttgtctt attgtttctc acagtttattt gttggccctc | 26160 |
| ttttagacat gaacaagatg aaattttcc tagcaaactg atgtggatga tctgttgctg | 26220 |
| cgggcttcac cctcaacaac atctcttctt ttcttggaaac aaatttatcca ttagtaaact | 26280 |
| gatgatttttggg ggagatgctg tccccataaaa cttttgtaa ggcataaata atttcaccat | 26340 |
| tcttccagtt tcaccataaaa tttgacgttt ttttgcttca atttttagcag cattcatgtt | 26400 |
| gctttgataa gagctttttt caaattcatg tcttattcctt ctttagtgcctt caaactagat | 26460 |
| cttggcagt atgacaagtt agttagatgtt tatctgcattt caaaaatctt tgaaatccat | 26520 |
| gcatagttttggg tttataatat acattttcaa tgaacttttg aagaccccat acatacatat | 26580 |
| gtatatatat gcacacacac acacacacac acacaaaaat cttcaaccat tatcagactt | 26640 |
| agtgcagaaaa aattattcat ccattaacaa gataagaatg ccccttatca tcactactat | 26700 |
| ttaaatggag ctccctggcta aaggaaaaga cagggattga aaaaaatttag ttaaatctaa | 26760 |
| aatgtttattt atttcaggtt tcttagttgc ttaaatggga agggaggtat ggacaaaaga | 26820 |
| gaaatcaaag atatttgtgt tatgctactt atcattaaag tatcagaata acttcattgg | 26880 |
| aatagaaaaa caccaagatc accccacgat atgtttctttaa atatcttctc catttcttta | 26940 |
| gacaagtgcac catgttattcg gccagtgaag aatcaaactc acttgcacgc ttataatgca | 27000 |
| ggaaaatata gcaaagagat gtggatccaa tagttcttag atagtggatc aggtggctt | 27060 |
| agatgaattt atatatctga aatgttcaca aattccctac tcataatagca tgttttcata | 27120 |
| atgttttagc aactctaattc ctcgtgactt gattgccacg tctggatattt ccacaacatt | 27180 |
| tccttaaacta agaatgagag taagaaatat tttaattcat aacaattata aatctgcac | 27240 |
| tcatgaaaaat gacattgcac ttgtgagact tgagaacagt gtcaccttta ccaaagatata | 27300 |
| ccatagtgatg tgctcccttgc ctgctacccca gaatattccca cctggctctt ctgcttatgt | 27360 |
| aacaggatgg ggcgctcaag aatatgctgg taagtgtctc ggaaaaaaaaa attaacaata | 27420 |
| gaaatgtctt atatttgcata ttaggttattt ttttaatttata ggaaacatctt ggaataggtt | 27480 |
| tttctatttct tctacagaca gaaccattctt atattctgtt cagcccaagc tctggctacc | 27540 |
| cctgagtcctc ctttagcaaaag caaagcaatg ctccagaaac tatggaaattt ctcaaatata | 27600 |
| gtaataggaa aatgtaaaag aaagttatga agacacgagt tcttaataa tccagagatt | 27660 |
| ctataagattt caaatagctt ccctataaaac aataaaaaaaat ttttggatgg tttttttttt | 27720 |

| | |
|--|-------|
| tgcttggttt ttagagacaa agactttctc agactggagt gcagtgggc aatcatggct | 27780 |
| tactgcagcc tcaaactctg gtcttaagaa atccttgc ttcagccccc caagtagcta | 27840 |
| gaattataaa taagtgtgta ccaccatacc cagttttttt ttttttttc tacagacagg | 27900 |
| ttcttgctct gttgcccagg ctggctgga attcctgccc tcaagccatc ctcctgcctt | 27960 |
| gttggccccc caaagcaatg ggaggattta gattagacat tgtatgaggg cttataatac | 28020 |
| cttaaggtat taactgccct ttaaagtatt ctgggatattt gcaaaaactc gatgtgtata | 28080 |
| taaacattgg tcataattgt ttattgaatg aataaaatgg aaactaaaat gaggacaatg | 28140 |
| cacaagagct actagaacca gtaagagtat cagcgaagga gtggagggt agcattgaca | 28200 |
| atttccctgg gcttttaccc atgtttaga ttgtctctcc aaggaataat acaaagcctt | 28260 |
| aatagtcccta gaacacattc tattgtgttc ttatggccca aagtaattt gttgttaga | 28320 |
| taacatttgc accagtcattt aaaaactatt ggtgtcattt tgagagtaca tcaatataaa | 28380 |
| atagactagt tccttagcct tgaaactaga ctggtttctc ttttgcgtc aggttaaagg | 28440 |
| ttattcaata tgtaatcttc caatccaaaa tctgtcagtg gataatttaa aagcttttag | 28500 |
| tcaatttaa gatatttggtt ttcttaaat tttaaggggc actgtgtcac aaagctaaag | 28560 |
| aaaaaaaaaga aaaaaaaaaact gatctgtgaa aggggttaccc ctcatctact tggggaaattt | 28620 |
| tggctgcgaa gaaactccaa agtaaatctt tagaaggcctt cattgttaaa tatgaaataa | 28680 |
| tgtttggagt acatttattt cttctcaaat ttattatagg gtcaataatg tacacatctt | 28740 |
| gaagtccatt ttttcctgc ttttataaca aacaggccac acagttccag agctaaggca | 28800 |
| aggacaggc agaataataa gtaatgatgt atgtaatgca ccacatagtt ataatggagc | 28860 |
| catcttgtct ggaatgctgt gtgctggagt acctcaaggt ggagtggacg catgtcaggt | 28920 |
| aagctcaaga caatctcatc catgtcatca tccaagaagt gtataagcac ttcttagtat | 28980 |
| gtgataatgt gatagacata agtgtaacag ttacaataca cagccctgtt cctctaaaat | 29040 |
| ttataatcta gattttagaa ataaattttt ttatgaatga agtttatcta tcatgaaagc | 29100 |
| attaactctg agaggccaaa ttacagagta gttaaccatc caaagctcaa gaatcagaaa | 29160 |
| gacctcgatt tgaattcattt aacctctatt accaagtctc taactaaaag ctggggataa | 29220 |
| tcataatgc acctaacttt ttgggtacta agaaaagtta aatgaagact aaatatatca | 29280 |
| ggcacatggtaaacaacaaa gaaatctcat ctatccact attattaatg tagaccatgg | 29340 |
| tcactcggt taataacttt aacctaacc tttaactgc tgtgaaggat taaataaaaa | 29400 |
| attaatcact atattataaa attaattga tatataataa atgaattttt agagatacgt | 29460 |
| aataattcat ggactccttg aagatagaaa atttatacaa aatccttagta atttgagtca | 29520 |
| caaaagctcc tacaataatg aaacagtatg aatgaaaaag aaaagaaata actattat | 29580 |
| ttggatctag cccataattt ttaaccaaattt gcacaaaaac aaacaacaaa tatgaaattc | 29640 |

ES 2 886 958 T3

| | | | | | | |
|-------------|-------------|------------|--------------|--------------|-------------|-------|
| tcactgtaaa | gtgattaaaa | tcaaatttga | attctaaaat | tttaaattaa | attatctaaa | 29700 |
| cataattgat | gcagttatat | gttttaatag | gttttggca | cataatcgaa | atccaaactcc | 29760 |
| acacagttagc | aggaacagct | ggtgtcagaa | attaaatatt | cttttagtct | ggagttttaa | 29820 |
| aaaatcaatc | tgtttacttg | agtaatttgc | tgctgtttc | atgggtgaat | tgtatacaga | 29880 |
| aggataagaa | ttattcttcg | catcaaagg | tcactgactt | tcatatttag | tgctcatgg | 29940 |
| ctttaaaaag | tggataaaaa | gtagttctca | catttcatgg | aaagccccca | atccatgagc | 30000 |
| acatttccca | aatgaaaaca | tttttatcaa | ctgcaagtttgc | tgtgttaggt | gagatttgg | 30060 |
| tttcaattgt | caagatactg | ttaattaccc | agtcctttat | ctcctttgg | tggagatgtc | 30120 |
| tctgtgctag | gaaacccttc | ttgctctcct | tcctgtttct | cttttactac | tggccctgaa | 30180 |
| acaacaaatt | ctcaagtttc | atgacagctt | tccaaagaat | ccatcaatca | aataagcaac | 30240 |
| acaactcgac | actgacaatt | ccagacctac | taagagcatt | aattaagact | taaaaataaa | 30300 |
| catgagtttt | aaaagggtgt | tattcattat | tttcccattt | ataacgtccc | ttaccttctg | 30360 |
| tccttcagtg | catacaaatt | attatcttc | ttgaagcccc | gttcaagccg | tacctcacca | 30420 |
| tgataccttc | catgtatatt | ccactctagg | cctcaactgat | ttttaactga | aatactataa | 30480 |
| tgcatagttc | acactttaaa | aaaaaaaaaa | aacacagcac | tttacataag | agcttacagg | 30540 |
| atcctatttg | ttttatccat | tctttgttc | attttacaa | tcattaattc | aaaggaattt | 30600 |
| tattaattac | tttctatgca | cccgacgttg | tgttaacaca | acaatactat | ccctgcattc | 30660 |
| agcaagtcta | tggctacaa | gagaggacac | aaattcaaat | gtctgttagtc | aagcagtgaa | 30720 |
| gctggctaga | tatggaaaaaa | ttacaagtcc | ctcttgctt | aacatttgct | tgcccacatt | 30780 |
| tggtcagaca | tcatgcaaaa | taatttctca | ctatagaaaa | aaaaacacta | caaaaacaat | 30840 |
| aatataaaga | actgagaact | ggtaactga | agcatgcata | tgtcatctaa | aagaagcagg | 30900 |
| tgacgaccag | cttcatgaag | tacttgcatt | gcatattggc | acttcacacaca | ctgacccttc | 30960 |
| tccccaccta | gaccagtaat | taaacaggtt | tggatgagct | agctactaag | agcagccaaac | 31020 |
| tgaatagctg | actaacttag | aagcacactt | ggttaataata | gctgactttt | attagtactg | 31080 |
| actatactat | atgctaagct | gtactcaaag | tgctttgagt | tttaaactga | tacaaacatt | 31140 |
| atatgaggaa | acagaggtac | agagagctat | tcaccagctt | accaaaggc | acatagctgg | 31200 |
| taagtggagg | acttaaacc | agactatcta | gttccagaac | ccacagactt | aatccatcg | 31260 |
| gcagaacata | agacatactc | catctgtctc | cccaactagg | ttattatgt | cacaatatt | 31320 |
| tattgggtgg | ttgggttcatt | attatgactg | ggtggtaagt | atgtcattag | gagtgtttg | 31380 |
| cttatgacta | tataaatttc | ttcacaaaaa | gaagactttc | tgtatgatata | ctatgcatca | 31440 |
| gacaccacgc | agggtgctaa | ggttaggaag | ataagtgaga | cttctagaaa | ctcattcatt | 31500 |

ES 2 886 958 T3

| | |
|---|-------|
| caacaatat ctcctaaggg ctagaagctt aggtttcagc agtgaacaga ataggtatgt | 31560 |
| tctcttcgt gttggacctt atagtatatac tggaaaaca gacattgaat aaatatcaca | 31620 |
| aatgcaagtg agtgtttcag agacatgcag ctgctacatc aaaacaaaac agaacaaaac | 31680 |
| aaacaaacaa aaactgacca gtgggattaa gtgtaaatag gcacacaaat gcacaaatata | 31740 |
| gctttataa aatagtgaag cagtgacaga gacacacaca agatataaag acacaatgaa | 31800 |
| gaacaattga gcccaaagct ggaaagggtg agagtgtgaa ggaaaaaggt tgatcagaga | 31860 |
| agtttcccg aaggagagaa agcctggatg attaggaggc aaccactcgg tgactgaggg | 31920 |
| aatctgaaa aatgtatttg tcattttctc agacttgctg aaggaatgac ttgggtactt | 31980 |
| tgaggatttc agtaattttt ccatgacttg gtataatatt tcaaaaggaa ataggctgac | 32040 |
| tttatttcta taatgaatgt gactccttcc tcgactgcc a tagaaataaa ctcctaata | 32100 |
| ttttgggttt gtcttgcac ttaagtaatc agtcattctg ttttttaca gggtaactct | 32160 |
| ggccccac tagtacaaga agactcacgg cggttttgtt ttattgtggg gatagtaagc | 32220 |
| tggggagatc agtgtggcct gccggataag ccaggagtgt atactcgagt gacagcctac | 32280 |
| cttgactgga ttaggcaaca aactggatc tagtgcacca agtgcattttc tggtgaaag | 32340 |
| tctgtatgca ggtgtgcctg tcttaattt caaagcttta catttcaact gaaaaagaaaa | 32400 |
| ctagaaatgt cctaatttaa catctgtta cataaatatg gtttaacaaa cactgtttaa | 32460 |
| cctttcttta ttattaaagg ttttctattt tctccagaga actatatgaa tggtgcata | 32520 |
| tactgtggct gtgtacaga agaaacacac taaactaatt acaaagttaa caatttcatt | 32580 |
| acagttgtgc taaatgccc tagtgaaag aacaggaacc ttgagcatgt atagtagagg | 32640 |
| aacctgcaca ggtctgatgg gtcagagggg ttttctctgg gtttcaactga ggatgagaag | 32700 |
| taagcaaact gtggaaacat gcaaaggaaa aagtgtataga ataatattca agacaaaaag | 32760 |
| aacagtatga ggcaagagaa ataatatgtt tttttttttt ttggttactc aatatcttat | 32820 |
| acttagtatg agtcctaaaa ttaaaaatgt gaaactgttg tactatacgta ataacctaac | 32880 |
| cttaattatt ctgtacaaac atgcttccat agggaaatagt ggataatttt cagctattta | 32940 |
| aggcaaaagc taaaatagtt cactcctcaa ctgagaccca aagaattata gatatttttc | 33000 |
| atgatgaccc ataaaaataa tcactcatct acataaagga gagactatat ctatttata | 33060 |
| gagaagctaa gaaatataacc tacacaaact tgcagggtgc tttacaacta catagtactt | 33120 |
| tttaacaaca aaataataat tttaagaatg aaaaatttaa tcattggaa gaacgtcccc | 33180 |
| ctacagactt cctatcactg gcagttatat ttttgcgtt aaaagggtcg tcaaacgcta | 33240 |
| aatctaagta acgaattgaa agtttaaaga gggggaaagag ttggttgc aaggaaaaagt | 33300 |
| ttaaatagct taatatacat agaatgatcc tgaagacaga aaaaacttttgc tcaactttcc | 33360 |
| tctctcattt tttttcttc tctctccct ttcatacac atgcctcccc cacaaagaa | 33420 |

| | | | | | | |
|-------------|-------------|-------------|-------------|-------------|-------------|-------|
| tataatgtaa | attaaatcca | ctaaaatgtaa | atggcatgaa | aatctctgtaa | gtctgaatca | 33480 |
| ctaataattcc | tgagtttta | tgagctccta | gtacagctaa | agttgccta | tgcatgatca | 33540 |
| tctatgcgtc | agagcttcct | ccttctacaa | gctaactccc | tgcatactggg | catcaggact | 33600 |
| gctccataaca | tttgctgaaa | acttctgtaa | tttcctgatg | taaaattgtg | caaacaccta | 33660 |
| caataaaagcc | atctactttt | agggaaaggg | agttgaaaat | gcaaccaact | cttggcgaac | 33720 |
| tgtacaaaca | aatctttgct | atactttatt | tcaaataaaat | tcttttaaaat | ataatttccc | 33780 |
| tgcctaatta | tttatggaag | ttatgacttt | tgaaggacaa | ttcaaaacca | tttatttaat | 33840 |
| tggttctgca | atgaaagaac | tgccccatat | actctactaa | aggcttggca | ctttctgctg | 33900 |
| ccttttaatc | cagcgctata | attgaggcaa | gcgtccagct | tgacacctcg | agataacttc | 33960 |
| gtataatgtaa | tgctatacga | agttatatgc | atggcctccg | cgcggggttt | tgccgcctcc | 34020 |
| cgcgggcgcc | cccctcctca | cggcgagcgc | tgccacgtca | gacgaagggc | gcagcgagcg | 34080 |
| tcctgatcct | tccgccccga | cgctcaggac | agcggcccg | tgctcataag | actcggcctt | 34140 |
| agaaccccaag | tatcagcaga | aggacatttt | aggacgggac | ttgggtgact | ctagggcact | 34200 |
| ggttttcttt | ccagagagcg | gaacaggcga | ggaaaagtag | tcccttctcg | gcgattctgc | 34260 |
| ggagggatct | ccgtggggcg | gtgaacgccc | atgattata | aaggacgcgc | cgggtgtggc | 34320 |
| acagctagtt | ccgtcgcagc | cgggatttgg | gtcgccgttc | ttgtttgtgg | atcgctgtga | 34380 |
| tcgtcacttg | gtgagtagcg | ggctgctggg | ctggccgggg | cttcgtggc | cgcggcccg | 34440 |
| ctcggtggga | cggaaagcgtg | tggagagacc | gc当地ggct | gtagtctgg | tccgcgagca | 34500 |
| aggttgcctt | gaactgggggg | ttggggggag | cgcagcaaaa | tggcggctgt | tcccgagtct | 34560 |
| tgaatggaaag | acgcttgtga | ggcgggctgt | gaggtcggtt | aaacaagggt | gggggcattgg | 34620 |
| tgggcggcaa | gaacccaagg | tcttggggcc | ttcgctaatg | cggaaagct | cttattcggg | 34680 |
| ttagatgggc | tggggcacca | tcttggggacc | ctgacgtgaa | gtttgtca | gactggagaa | 34740 |
| ctcggtttgt | cgtctgttgc | gggggcggca | gttatggcgg | tgccgttggg | cagtgcaccc | 34800 |
| gtacctttgg | gagcgcgcgc | cctcgctgt | tcgtgacgtc | accgggtctg | ttggcttata | 34860 |
| atgcagggtg | ggcccacctg | ccggtaggtg | tgcggtaggc | tttctccgt | cgcaggacgc | 34920 |
| agggttcggg | ccttaggtag | gctctcctga | atcgacaggc | gc当地ggct | tggtagggg | 34980 |
| agggataagt | gaggcgtcag | tttcttgggt | cgtttttatg | tacctatctt | cttaagttagc | 35040 |
| tgaagctccg | gttttgaact | atgcgctcgg | ggttggcggag | tgtgtttgt | gaagttttt | 35100 |
| aggcaccttt | tgaaatgtaa | tcatttgggt | caatatgtaa | tttcagtgt | tagacttagta | 35160 |
| aattgtccgc | taaattctgg | cgtttttgg | ctttttgtt | agacgtgttgc | acaattaatc | 35220 |
| atcggcatag | tatatcgca | tagtataata | cgacaagggt | aggaactaaa | ccatgggatc | 35280 |

ES 2 886 958 T3

| | |
|---|-------|
| ggccattgaa caagatggat tgcacgcagg ttctccggcc gttgggtgg agaggctatt | 35340 |
| cggctatgac tgggcacaac agacaatcg ctgctctgat gcccgcgtgt tccggctgtc | 35400 |
| agcgcagggg cgccccggttc ttttgcataa gaccgacctg tccggtgccc tgaatgaact | 35460 |
| gcaggacgag gcagcgcggc tatcgtggct ggccacgacg ggcgttcctt ggcagctgt | 35520 |
| gctcgacgtt gtcactgaag cgggaaggga ctggctgcta ttggcgaag tgccggggca | 35580 |
| ggatctcctg tcatctcacc ttgctcctgc cgagaaaagta tccatcatgg ctgatgcaat | 35640 |
| gcggcggctg catacgcttg atccggctac ctgcccattc gaccaccaag cgaaacatcg | 35700 |
| catcgagcga gcacgtactc gnatggaagc cggcttggtc gatcaggatg atctggacga | 35760 |
| agagcatcag gggctcgcbc cagccgaact gttcgccagg ctcaaggcgc gcatgcccga | 35820 |
| cggcgatgat ctcgtcgtga cccatggcga tgcctgcttg ccgaatatca tggtgaaaaa | 35880 |
| tggccgctt tctggattca tcgactgtgg cccgctgggt gtggcggacc gctatcagga | 35940 |
| catagcgttg gctacccgtg atattgctga agagcttggc ggcgaatggg ctgaccgctt | 36000 |
| cctcgtgctt tacggtatcg ccgctcccgta ttgcagcgc atgccttct atgccttct | 36060 |
| tgacgagttc ttctgagggg atccgctgta agtctgcaga aattgatgat ctattaaaca | 36120 |
| ataaaagatgt ccactaaaat ggaagtttt cctgtcatac tttgttaaga agggtgagaa | 36180 |
| cagagtacct acattttgaa tggaggatt ggagctacgg ggggtgggggt ggggtgggat | 36240 |
| tagataaaatg cctgctcttt actgaaggct ctttactatt gtttatgat aatgtttcat | 36300 |
| agttggatat cataatttaa acaagaaaa ccaaattaag ggccagctca ttcccccac | 36360 |
| tcatgatcta tagatctata gatctctcggtt gggatcattt ttttctctt gattccact | 36420 |
| ttgtggttct aagtactgtg gttccaaat gtgtcagttt catagcctga agaacgagat | 36480 |
| cagcagccctc tggccacat acacttcatt ctcagttt gtttgcctaa ttctaaattcc | 36540 |
| atcagacctc gacctgcagc ccctagcccg ggccgcgtta gcagcaccca cgtccaccc | 36600 |
| ctgtcttagta atgtccaaaca cctccctcag tccaaacact gctctgcata catgtggctc | 36660 |
| ccatttatac ctgaagcact tggatggcc tcaatgtttt actagagccc accccctgc | 36720 |
| aactctgaga ccctctggat ttgtctgtca gtgcctcaact gggcggtgg ataatttctt | 36780 |
| aaaaggtaaa gttccctcag cagcattctc tgagcagtct gaagatgtgt gctttcaca | 36840 |
| gttcaaatcc atgtggctgt ttcacccacc tgcctggcc tgggttatct atcaggaccc | 36900 |
| agcctagaag caggtgtgtg gcacttaaca cctaagctga gtgactaact gaacactcaa | 36960 |
| gtggatgcca tctttgtcac ttcttgactg tgacacaagc aactcctgat gccaaagccc | 37020 |
| tgccccacccc tctcatgccc atatttggac atggtacagg tcctcactgg ccatggctcg | 37080 |
| tgaggtcctg gtcctctttg acttcataat tcctaggggc cactagtatc tataagagga | 37140 |
| agagggtgct ggctcccagg ccacagccca caaaattcca cctgctcaca gttggctgg | 37200 |

| | |
|---|-------|
| ctcgacccag gtgggtgtccc ctgctctgag ccagctcccg gccaaggccag caccatgggt | 37260 |
| acccccaaga agaagaggaa ggtgcgtacc gatttaaatt ccaatttact gaccgtacac | 37320 |
| caaaatttgc ctgcattacc ggtcgatgca acgagtgatg aggttcgcaa gaacctgatg | 37380 |
| gacatgttca gggatcgcca ggcgtttct gagcatacct ggaaaatgct tctgtccgtt | 37440 |
| tgccggtcgt gggcggcatg gtgcaagttg aataaccgga aatggttcc cgacagaacct | 37500 |
| gaagatgttc gcgattatct tctatatctt caggcgcgcg gtctggcagt aaaaactatc | 37560 |
| cagcaacatt tgggccagct aaacatgctt catcgtcggt ccgggctgcc acgaccaagt | 37620 |
| gacagcaatg ctgtttcact gtttatgcgg cgatccgaa aagaaaacgt tgatgccggt | 37680 |
| gaacgtgcaa aacaggctct acggttcgaa cgcaactgatt tcgaccaggt tcgttcactc | 37740 |
| atggaaaata gtgatcgctg ccaggatata cgtaatctgg catttctggg gattgcttat | 37800 |
| aacaccctgt tacgtatagc cgaaattgcc aggatcaggg tttaagatat ctcacgtact | 37860 |
| gacggtgaaa gaatgttaat ccataattggc agaacgaaaa cgctggtag caccgcaggt | 37920 |
| gtagagaagg cacttagcct ggggtaact aaactggtcg agcgatggat ttccgtctct | 37980 |
| ggtagctg atgatccgaa taactacctg tttgccggg tcagaaaaaa tggtgttgc | 38040 |
| gcgccatctg ccaccagcca gctatcaact cgccctgg aagggatttt tgaagcaact | 38100 |
| catcgattga ttacggcgc taaggtaaat ataaaatttt taagtgtata atgtgttaaa | 38160 |
| ctactgattc taattgtttg tgtatTTTg gatgactctg gtcagagata cctggcctgg | 38220 |
| tctggacaca gtgcccgtgt cgagccgcg cgagatatgg cccgcgtgg agttcaata | 38280 |
| ccggagatca tgcaagctgg tggctggacc aatgtaaata ttgtcatgaa ctatatccgt | 38340 |
| aacctggata gtgaaacagg ggcaatggtg cgccgtgg aagatggcga ttgatctaga | 38400 |
| taagtaatga tcataatcag ccatacaca tctgttaggg ttttacttgc tttaaaaaac | 38460 |
| ctccccacacc tccccctgaa cctgaaacat aaaatgaatg caattgtgt tgttaaacct | 38520 |
| gccctagttt cggccaattc cagctgagcg tgccctcgca ccattaccag ttggctgg | 38580 |
| gtcaaaaata ataataaccg ggcagggggg atctaagctc tagataagta atgatcataa | 38640 |
| tcagccatat cacatctgtt gaggtttac ttgtttaaa aaacccccca caccctcccc | 38700 |
| tgaacctgaa acataaaatg aatgcaattt ttgttttttatt cttgtttatt gcagttata | 38760 |
| atggttacaa ataaagcaat agcatcacaa atttcacaaa taaagcattt ttttacttgc | 38820 |
| attctagttt tggtttgtcc aaactcatca atgtatctt tcatgtctgg aataacttcg | 38880 |
| tataatgtat gctatacgaa gttatgttag taactataac ggtcctaagg tagcgagcta | 38940 |
| gctgcaacccg agaaaaaaac gtgccatgag gtctctgtat ccaagtgtga ct | 38992 |

<210> 20

<211> 34073

<212> ADN

<213> Secuencia artificial

<220>

ES 2 886 958 T3

<223> Polinucleótido recombinante

<400> 20

| | |
|---|------|
| gagggagggt ggtgcttgc taatggtaaa ttactaactc ctcaataaag aatattat | 60 |
| gaaataattt ttgaaatttc ataattactt tgggttcttt cttaatgata aataaaataat | 120 |
| agtatattac aaacatacat taatattcc tgaatgaata caccacaat ctcccttaaa | 180 |
| atatacgaa aataaaaaatt atactatttc tgacaatttt taatttctca aataataata | 240 |
| ccactctgat ttttaaacat ctacaccact ctggcttgc caatctttt aaaaattgaa | 300 |
| aagataataa ttttatcata attacactga agcatagaac ttttcttca aaggaaagca | 360 |
| aattttgaa attctataat ataacctccc ataattcctga ataaataaa gttcaacaa | 420 |
| cttagtaaag taagactgac cttccctttt atttctttt cagatcaaaa atcttacttt | 480 |
| tataggagca gtttcaact cctaaatgtt gaatataata gtcagttaaa ttcaccagct | 540 |
| acacaggaat acaggacttt gagtgaaaga attgaatctc tggtaagttt atatttgct | 600 |
| ttgctcttta ttccattata aatgaatat gataataaac ctaatgttt gtaatataatt | 660 |
| ttcagttgct aagtgcctca catatttcc ttcccttgaat ggtgaaacat gtgtttctct | 720 |
| ctgctttat ccagtttagtt tactcatata ctggttctta ttcacatctt tgtcatgagt | 780 |
| aaaaagtgtt agaaaggcca cgagtaataa tgcattttat ttgttatga attcaaatac | 840 |
| taaaagttt ttatggttt aattaagcat tgacattgtc tttttaaatt ctttcattt | 900 |
| taccttcttc cctttccctt atccaaactaa agacgcaaaag caggaggtgt taaaaaacag | 960 |
| gtttaccata tcagcagtaa catagttgg acaacattac actttggttc aatgatagac | 1020 |
| atagaagttt gaacagaaat atgcaagca agtttgagct ctaactgaa gagagcctct | 1080 |
| gggtgcctgc cagaaacct cacgagtgg cccttaacat tcatgtgtca ccacaaacta | 1140 |
| ggggctgccc tttagtttg accagtcacta gtgtcactca cttaccctta cttttcaaa | 1200 |
| aaaaagtccct aagaatataa agtaattcaa tggttctaca attttagcat gtaactgagt | 1260 |
| cacctggcag ggttgccttgc gtgagctaa gataaaattt ttcacatctt tctacat | 1320 |
| ctggaaatatt ccttaatcca ggcttttaat cccttggc ttttctgaac cactgcaatg | 1380 |
| agcttctaact ttttctact gtgtgcaggc tttttccctt ctaatctaattt ttacacactt | 1440 |
| ctgaacacaa atctctcaca gcctgtttcc ttcatgttac ctccagctca agacttttg | 1500 |
| cctacaaaat aaaattcaaa cttgttagct aagcaccttc tcatgtctat gctttggctc | 1560 |
| atatttcagc catcgtgtgc cccacttattt cttatagcca acctgaaaag ccatctttta | 1620 |
| taagaaacta cctctgctct ccatgattgg atataattaa tcctcccttcc acatcacctc | 1680 |

ES 2 886 958 T3

| | |
|---|------|
| gccacaaaat tgtatctgtg ttgatctcat gccacatacc tgtatgtatt ttatattata | 1740 |
| aatatttgca gacttggta atttgccatg ttagactaag ttccatgaag acagctccat | 1800 |
| atccattcca ttttatata tccacaacat ttggcggt tgatgctaa taaatgtta | 1860 |
| ttgaaggaac aggagtctcc cacttctgac ataatgaact tatttcccc agtgttaacc | 1920 |
| ctacatctgg ttcctgtcca agagtcctt cccaaatcat tctgattcaa ctgttcattc | 1980 |
| tgatctcatt aaacatttaa atgatatac taacttcgct tgcttatttc tatgctcatc | 2040 |
| ctgcagtctc ctcataactt ggttcaatg atgcttgctt ctagagaaaa aaatgtatta | 2100 |
| aataagctta tgattcagtc ctccagctgt gatggttctc actgaacatt agtcagtg | 2160 |
| ttttcgaagt atggctctta gcataaccta gaaacttgtt agaaatgcaa attcttggc | 2220 |
| tcaccaagac atactaaatc aaaaattctg acattggggc ctagaaatct gtgtttaac | 2280 |
| aaggctgcca gtgcagcctg gtccctttc ttctcgagc cccactcaaa gcttcagtg | 2340 |
| ctcatctccc accaatgaca gggcctcta tggaaaccgg caggacggtt tccaaactcta | 2400 |
| actacgtttt agagtttgct tccttagggct atccaggcac caagtatcac aggttagtt | 2460 |
| cccaaggaag cagactctga gacttgcatt cagggagtgt ctctgggtg ctctcaacca | 2520 |
| acaccttcag gaagagaagg aagcagcatt gggcagagggc atagtcaaac tacagtgc | 2580 |
| ttggcacaga agactgaagg gagtcagagc cagggggtag aggtgggccc ttagcatcca | 2640 |
| tccttcacca ttaggtgtga gttgccccac ctccttgatg gtgtaacctc agtcccaagg | 2700 |
| tgggtggag tgcagcagag cagccctac aaggccaaa ccagagatac accaggcgcc | 2760 |
| agaagtgcg ccagggata gagaggaaag gatgggctta aggtaggatc cacagaactt | 2820 |
| ggcaatggat tagaagacag gatgagaagt gacaggttaa cactaacaca gaaatgtcta | 2880 |
| acttcggtag ataatggtgc cattggctag aagaggaaac cgaaatgaaa gcaggttgg | 2940 |
| cagggagaca aaagttcact gtggacatct cagcagagtg attcagtggg gaaaggaatg | 3000 |
| gatgccaga ccacctcaga ggaagatcta agctggagcc agcaataaag atacaagatg | 3060 |
| aacaatccct aacgaactgc tcctcagcca tgctccccag acacgctgct tcagatttat | 3120 |
| agtccgggtg aggctaggag gtgcgcctcc ctcagtggag gacagcaaag caccagtggc | 3180 |
| tccagggagt taaaatctt tgataatttt tggtctagca tctgtctgca gagctgtctc | 3240 |
| tcagccattt cctgccttta cacaggagtg cagtccgaaa ttgggagatg agtcaaattt | 3300 |
| attatgccta gagatctgga tccccagttg tttgggagta tattttctga accacttgg | 3360 |
| ggtttaagta atgcagattt attgatgcca cttcttttga atctgtgact ctggaccac | 3420 |
| catctaagtg aatgtgcaga gggAACGGAA tggctgcaat agatctccat taaaaccagt | 3480 |
| gcacccatcc agacacatac agtagtaggg aggtgagtca atgtcaggac agcaccagct | 3540 |

ES 2 886 958 T3

| | |
|--|------|
| cccgcttcgg tacatccc aagttctcag tctgtgtaca aaggttgct ctggggcagc | 3600 |
| agaaaatagcc ctgggcaggt agtcaaaggc ctgggttgat ttccctccact tccaggcaag | 3660 |
| tcactcgaag gctcacaggc ttttcctca cctgccacat gggtccagtg agatctactg | 3720 |
| agctgtaaat aatgaaatga gtgtgtgtgc agtcatctat aagttgtaaa gtactagaaa | 3780 |
| atggtaaac tttgggattt gggctattta aggctgaatg ctaaaaatgt caggcattgt | 3840 |
| ggagaaaagga atttaaatat aagattgatt gactgggatt taaagacaaa tgaaggcaca | 3900 |
| cacgcaagtg cacacccaca ctgacactgc acagctcccg ttggaggcat atcctgacca | 3960 |
| tgcagacctg gggctctgcc tgtccaagtg cactccttta ctacataaac cctccttctc | 4020 |
| ttttgggct gtcacccccac cagagctggc accgagccct tgctgctgctg cttccctggg | 4080 |
| gtgtcagctt ttgacaggggt gtttcctccc tctgcaggag ccttaacatc ccttggactt | 4140 |
| ccttcccccc acccaccmmm agcagttta tctcttctta actcgggacc cttttttcc | 4200 |
| cacacaaagt ttattgtcag ttgctggttt catctgtttt agcggctgca acaaaaatacc | 4260 |
| atagactggg tggcatatgc acgacaaaaa ttatattctc acaggagaag tcaaagatta | 4320 |
| atgcaccagc agatctggtg tctgaggggc caccttctgg tttgtagatg atgctttcta | 4380 |
| gttaaaacac ctatthaaca cactattaaa cactaagtgt gttaaatagt gcagttgatg | 4440 |
| tattngtcat gtcaccttta tcatacacta aatccttctt tgtcttttt tctgtactct | 4500 |
| aatctcttcc tptaagtaat ctttgcttgc agcagtagga tattnagatg actgtggctt | 4560 |
| gacaatatat ttagtatttc aagatttcca tgaatttctt ctgatgtatg agttccctag | 4620 |
| ttaatcttac atatgtatcc ctttgtaaaa acactttgaa cattttaaat gatacatgaa | 4680 |
| tagtactcta atacaatgcc ataaaaattttaaaatcattt gtatagactg gtaagtaaag | 4740 |
| attgtgagat taagaaacgc atcaaaggcc attgagctgg aaagtggat aatgagaatt | 4800 |
| caaaccaggc tctcttgact caaaatctaa ggatcataacc atttctcatg ataataatgag | 4860 |
| tattattgtt atctcttatcc catagacaaa gtgttaacac tgaatgagca gtgaaatagt | 4920 |
| ctcagaattt ttatTTTtat ttagcaatttcc acttgcattt tctggccttc agtttattca | 4980 |
| cgagttttttt aaaaatgttg gactagataa ttcttatagt acatttttac acaaaaaatc | 5040 |
| tatgatTTTt ttatTTTaa tttgtatatac tcatggcact cattcacccat atttccctag | 5100 |
| cctgcctcac tggcatttac ttctctgtgt tctttacagg ctccccctcc tctacactgc | 5160 |
| cattaaatat tgaacacacctt caaagttta cttatgtcca ccttcctct gacactatca | 5220 |
| ttctgtctag atgatcccat acatacatgc ccattacttc aacctgtatt tatacgccaa | 5280 |
| tgattcacta tattttccaggc ctagacatttcc ttttgcatttgc tagttaccag cttgtatcc | 5340 |
| ttacatggct gtttcaaaac aactcaaata tattatctct caaaatcaaa ctcatgtatgt | 5400 |
| ccccacacca tccttagctt ccaccaacaa tacctatccc tattaatgc aataccattt | 5460 |

ES 2 886 958 T3

| | |
|--|------|
| attcagttat ccaaataaaa aacctagaat tcatccttaa aattctacta tcattccaaa | 5520 |
| tatccatatcc atcagcagcc actgtattct taatcccctg tatttccttc aaatccattc | 5580 |
| acctctctcc atatccatttgc ctgcatttgc actccaagcc tcgcctctac cctagggtac | 5640 |
| caaaatagca acaaaccctaa tctgttcatt tgcatcattttt tttctccaaa actgattatc | 5700 |
| tatatgttagc aagacagatt gttctcaaat tgcaaataccc actatattat cctcttgctt | 5760 |
| caaacacttc catggtttcc cattgtttat gataaaacca aatgcttcaa gttcgaagac | 5820 |
| cggcatgatt gggaaatttcc tgtcaccccta gcctacttgc tctccatggc acagttgcac | 5880 |
| tggctttctt tcattcctta agtacaacct gtttcctccc acctcaggac tgcgtatgt | 5940 |
| ccattcattc tgctgaggag ccttttcct tccacttcaa tcagctaagt ctgattcttc | 6000 |
| ctgacaatct cagctcaata agcatttcct ctaagaaatg tctctaataat cattaattgg | 6060 |
| ctcaggtccc tctactgtat tgctgcactt ttcacagttta taattttact taatttatgaa | 6120 |
| tgattatttg attaggtcta ttccatcca tttagacataa gcttcatgat ggcagatta | 6180 |
| ctgaaaaatcta tccatcggtt tattccaata cctgacagaa ggagggcggg aggtggtggc | 6240 |
| acacaagaga tgctcaaaaaa caatttgtga ataagtaat gaatgaggcc atttagaaat | 6300 |
| aacgaaagta cctgtttaca aagtacatgt atcaaaaacta tgaatgcatt ctacttacat | 6360 |
| ggttttctcc aaataaaaaca aaagacttca atcaggatta atacctggaa taaactgagt | 6420 |
| cattaaatct ctccttgc acaggagtg acattgaaac aaatgtctgc aaacaacaaa | 6480 |
| tactttttcc caaaaatata ttgaatggca ttccataaa caaactagaa catgggagga | 6540 |
| gaaagaaaagc aatattaatt taaaattaat cttatcacat aacttataacc atcagggatt | 6600 |
| tcgggtaaaa ttcccttcag gcacatccat ttaacaagaa ttgattgtta ctgaaagcct | 6660 |
| agaagagaat ttggcacata cttgggttcc aaatatttg tgactgagtg aataaatgat | 6720 |
| gcaagtgtct aagaaacaca aaataaggac atgattacag tcacggtgaa gttcacagtc | 6780 |
| atctccaaaa tgaggatatg catcccagg aggaccaaca attcattggaa gtgctgaaat | 6840 |
| aaaatactca aaggtcattt tacatgtatt tttctctaa attactttc ttaagacaca | 6900 |
| gaaaacaaaa aaagaaaactt agctttgtta ctttctaaaca aatagttaaa tcattaaaca | 6960 |
| ggattgacac tagcatcctt gtttggtctt atgccttagg ggaacatgaa atgtgtgaag | 7020 |
| acattctgag atctgaggga aggtagaca gtaatacagt gggactgacc aggcttcagc | 7080 |
| acacccttac ctcctctcag cagatttcag tgatgagcag tttacaacta gattgaaaga | 7140 |
| ttatattatc tagttctaaa agaaaactaa gcctccaaa agcaacaagg gaactgagag | 7200 |
| gaatcctgca aaacaaaaac aaattttaaa acttgcactt tgtaataacc ctaatatgt | 7260 |
| atcacagtaa tgaacagtaa gataatgaca gaactgacat atttccttat ctattaaagc | 7320 |

ES 2 886 958 T3

| | |
|--|------|
| catattaaca ggtaaagcaa tgccagtcag tggcacactt cttagaagat atttaataca | 7380 |
| tactagacac atacacacac acaacatttt ccttcaaggt gtatgtatca gaaaatcact | 7440 |
| ttttaaggcc ggatgcagtg gctcaggcct gtaatcccag cactttggga ggccgacgtg | 7500 |
| ggcggatcat ctgaggtcag gagtcaaga ccagcctgcc caacatggcg aaaccccatc | 7560 |
| tctacaaaaa tacaaaaaatt agccagggat gatggtggat gctttagtgc ccagctactc | 7620 |
| aagaggcaga ggcaggagaa tcacttgaac ctgggaggca gaggttgcag tgagccaaga | 7680 |
| tcaccattg cactccagcc tgggcaacag agtgagactc tgtctaaaaa aaaaaaaaaat | 7740 |
| cacttttag ataaaattca tgctatagag agaagactat gaaaatatgt ttagcaatgt | 7800 |
| gtccatcatt aggtgattga gtttccttt gttttgtttt actgaaaatc atataaagta | 7860 |
| tgttatctgt aaaagttctc tgacatgcac acataaaaat ttgggagaaa agattaacta | 7920 |
| taatgttaa tagattttgt acacatttct taaaaaatat ataaaacaca acaccttca | 7980 |
| attggttgc aagaataacc aattgacatc atggaaaatg gaaattcaact tgctgaattt | 8040 |
| taacaaaaat ttgcatgatg agtgagactg acaacttagt gtcatgattt aatgaattat | 8100 |
| gccaatggta aacttcatgc acatggggcc aggttaattat gtggaaactt tttcaatgct | 8160 |
| taaagccaag tattgaaatt aaacttagaa tcagaccttt gaaccatttt atgacaatgt | 8220 |
| tcaaaaatta taaattctat ccacttatata tataatatta aaaatatcat tacaaaaaaa | 8280 |
| acctgtgttt atttataac tcagccttt taatttctaa tttcataat atattataat | 8340 |
| ggatattgtt agtaatgttag tattattaca tttatataat ttataagtaa atatacatgt | 8400 |
| tttggctact catgcataaa atgttcacc cataggagca cataatcaga aatgtctgga | 8460 |
| gaccattata gtaatagata gatcatattt ccacatattt tatctcctcc ttgacaactg | 8520 |
| agctttccag atcttctggt gaaacgaaag agaaagtgtt aacagaagag tgattaaaat | 8580 |
| gacaaaagca ttacttctat tacttctatt ctaataatat gagcaaagct ataactatca | 8640 |
| agtaataatg cactaaagaa ggtgattaat ctgatataatt cacaggcaac taataagacc | 8700 |
| tttctattgc agccatgaaa aatatgtgac aattatacat atcctgtgt cagtgtttca | 8760 |
| acctttatgt gacctgttct actaacagat ttatgtatgt tcactttgtt agaattttct | 8820 |
| tacacatgcc ataacttgct tcagtcattt gattatgaat attatggata ttaaggattc | 8880 |
| tagactattc tagatttaaa aaataatatt gtcacctcaa tcagaaggaa aatattaaat | 8940 |
| agttctcatt tttcaatgt ttactcagtt ttgtccat gtaatgaaag tgtcagcagt | 9000 |
| acaggttaca aaataaaaatg tgtattaaag taaactcatt tgaacagggtt aataattgt | 9060 |
| gagggaggga aaaggctaaa agattgaatg taaaacttat gaaaagttaga tacatgtct | 9120 |
| ctatgatttg cagtagtcaa ctgcatacag atgaatcatt ttaatacagc ttaactactt | 9180 |
| tcctttaca gatggagaaa ctgagaggaa gaaagtttat atggttcatt aaactttgt | 9240 |

ES 2 886 958 T3

| | |
|---|-------|
| atgcaagcta aactaacctg tctctgtatt ttccatctac tgcccttatac actatctcat | 9300 |
| tagaataactc ttcaaggcatc tccttactga ttttcttacc aagcatttgt taagttctaa | 9360 |
| ttagagttgg tagtaacatt ttcacccact ctgtgaaata tgaaatctta ttcataggcc | 9420 |
| tcttctttta ttcttgatt tgcataatcaa ccaattaatc aacttgcttt ctttatgttg | 9480 |
| cttattatct tagtccttac taaattgcct cttaatgttg tccacataac agaaaatgtta | 9540 |
| aggtggatac ttaacatttt agtccagtct agccggtgcc agtgcataatgc caaatcatga | 9600 |
| attaaaatat aattacaaga accacttatac aaattttaaac aattccttca gctttgtgac | 9660 |
| agttttttct acttcgatta aagtcaagta aaattaaagt taaatatttt tattaaaata | 9720 |
| tctcctttaa cattccatat taataaacat attaaagctc atgcttctaa gtagattact | 9780 |
| agaagttact ttatcgaatt acagcaatgg ttaattctag atcatagaat tttagaatgac | 9840 |
| tttttgcctt cttctttttt ttcccttttt ttaaacagag tcttgctctg ttgtccaggc | 9900 |
| tggagtgtac tggcgcgatc ttgactcaact gcgcacctctg ccctgcaggt tcaagtgatt | 9960 |
| ctcctgcccc agcctcttaa gtagttggga ttacaggtgc ctgccaccac acctggctaa | 10020 |
| tttttttttt gtattttag gagagacagg gtttcaccat gttggccaga ctggtctcga | 10080 |
| actcctgacc tcaagtgatc cacttcctc agcctccaa agtgcgtggaa ttacaggtgt | 10140 |
| gagccactgt gcctggcctg acttttgct ttcttcttaa tacttactag tatttcttga | 10200 |
| atttttaaaa aagaaacata aagtactttg ataaaaccaa cagtcattt gttctaaaa | 10260 |
| ttgttcaaag gttctctgga aaaaaaaaaaag aaaattatca tttggtaag aatcatgttg | 10320 |
| gtctgacatc aatcatccta taggagtgaa tattgaaaaa gtaagatata ttgtggata | 10380 |
| atcgagattt cataaatttt accattttt agaagaatct gctccaaatc ctggcttaat | 10440 |
| gtaatatcca gcatgctact taattttctt gtcttcaccc tttcatatcc acatccaccc | 10500 |
| aggtgccacc tcacagtata agccagcata atccattttt ctcaatgaaa ccacaataca | 10560 |
| tctgaccctg catctcagga gaactgtatc agccacagca cttccagttt actatgaatc | 10620 |
| tgaatgttat gcctcaggag aaacatcctt gctggactg agtagtgatt caaggagata | 10680 |
| gttatgattc agtcaagaaa ttaataattt gtgttatttt tattattgag acagagtctc | 10740 |
| gttctgttagc ccaggctgga gtacagtggc atgatctcg ctcactgcaa cctctaccc | 10800 |
| cccggttcaa gtgattctcc tgcctcagcc tcccaaataa ctgggacagc aggcaattgc | 10860 |
| caccacgcct agctaatttt ttgttattttt agtagagacg gagtttcacc gtgttagcca | 10920 |
| ggatggtctc gatctcctga cctcaaggatc cacctgcctc agcctccaa agtgcgtggaa | 10980 |
| ttacaggcgt gagccactgc gcccgccat aaattattaa ctgagccagg cacagtggta | 11040 |
| cacacttata gtcccaagata ctcaggagac tgaggttggaa gtatcctttt ttatgttatt | 11100 |

ES 2 886 958 T3

| | |
|---|-------|
| ttatTTTaa ttattatggg tacataatag gtgtacatac ccatggagta caagtcatgt | 11160 |
| tctgatacag acacataatg ttaataatc acatcagggt aattgggata tccatcacct | 11220 |
| caagcattta tcttctttg tgtaggaac attccacctc cactcttgg aatggcaccc | 11280 |
| tgttgctta ttaaatacga ggtcttattc atttcatcta actatattt tctaccatt | 11340 |
| aaccatcacc tctttcccc taccccttgc gtgaggctgc aggattctta | 11400 |
| agcacaacag ttagaggcca gcctggacaa catagtgaga ctcaatttct aaaaaataaa | 11460 |
| aaagaaatta ccaactaatg ctaaaaaat agtctctgat gcttaggtat gaattagaaa | 11520 |
| tgacaaaaaa aaaaaaaaaaaa aaaaagactg cccttgctt ccttctcccc ttcttcaa | 11580 |
| gttttccatt gctactcatt ttagtctgg ttaatcaggt ttcatccatt aaaagcaatt | 11640 |
| gttgggatca cacatTTGA gttgtgtcag tggacttccc tcattgtggc atgattcctg | 11700 |
| ccccaaagccc tttagaaaaag ccaccaagcc atataacata atctctcatt gagaaaaaca | 11760 |
| tctgatgtgt ttagaatgac ttctagcaaa aaaccagcct gtccagcatc atctctgtat | 11820 |
| aacagataaa ggaataggtt ctgcatcaaa aggttataga acctgccc aaatccca | 11880 |
| tgtgtttgc aatggaaatta ggttgaacta aagtggaaat tcagtttct actcctcatt | 11940 |
| aacatgtctc atgttgcaag gttgagagga aggagaagaa gaactgtatt tacagagaga | 12000 |
| ttccccctct ctTTCTTCT acagattact aaaacattca aagaatcaaa tttaagaaaat | 12060 |
| cagttcatca gagctcatgt tgccaaactg aggtgagtgg aactgttagaa aaaatattta | 12120 |
| agtatacgata caatgtggca tacttgactt ttgtcacag aatgaatagt aaatgacatg | 12180 |
| ttcagataag ttgttgtaat attatgaaaa tagtattttt gtcagctaa aaaccaatgc | 12240 |
| aaaaaaAGCC aaacatatga tctatttagc tactaatgtt aataaccata ttatATCTAT | 12300 |
| tcttattggg aagaggaaga aggggtggag agagagttgg ggtgaaggta cagtaacaag | 12360 |
| gccatctat tgtAAAactc cagtggatat cattcacagt gcagcctatg taaacagtcc | 12420 |
| ctcctggagt tgtacaatgc tgggtttgg gtgtatccat ccaagatcaa gacactatga | 12480 |
| ccaacatcaa aagtggctt ttggTTTCT ctgcctgtat tgctataata aaagggtatt | 12540 |
| atggccaaat ccaaggcatg tctatcatga attaataata ggaggagtag cagcatgcat | 12600 |
| gctagttatt tgccattcc GCCTTAGTTTAAATATGATGT gataaaacca GCTTTCCAA | 12660 |
| ctgaaatagt caccttact gactctcccg caaatgtctc aaatgaccac attgctctag | 12720 |
| tctttaaata atatgcaata gttcttgg tggatggaa ttataactaat tctttctcaa | 12780 |
| atactagcat cacaagaaaa ttaattcttgc ttctctggag agtcacctag taagtatctg | 12840 |
| gagcacagat gtctggtcag gtaagtttg atgaggagtt aaaggataa gaagagtcca | 12900 |
| tgagaagggt atttccaaa acaccttgc gtcaattcag tgcacattca cttagtactt | 12960 |
| tcttgcagt atctgtatca gccactaatg ttcaaaaatg agtaagccct gaaaacctgt | 13020 |

ES 2 886 958 T3

| | |
|--|-------|
| aggactacat gaggccttctg ccttttctct cctttgttc acttcccact tatcactcaa | 13080 |
| tcctctgcaa cctggcttca ataccaccat aaaatatcaa ctgctcttgc cgattcaaca | 13140 |
| atgacatcca gataacaaaa tccaaagaaa ccacatcagt cctattcttgc gaccttcaa | 13200 |
| cagtatttgg tcctgttggc ctgtcactcc ttgaaatagg actatccctt ggtttgcattg | 13260 |
| gccttgtata ccctgatttt ccccttacctt cccttagctat tccttcttag tttcctttac | 13320 |
| taggtcttac ttctttgtat attccttaaa tggtgctgaa catcaggctg tgctcttaggc | 13380 |
| ctctcatctt ctcaggtcac actctctcct ttcccttggcc ttcaactgcca cccatatgct | 13440 |
| gagtgccttc aaagttgtat ctctaggcca gtcctctttt gcctccaaac atgaatata | 13500 |
| gcagccatct acttggtacc atcacatgga taattctcat gatctctcc agtatgactg | 13560 |
| cttctttatt ttttctggg ctcttttta gcattgcttt acatggaact ttatcatgtc | 13620 |
| tctcaacctc tattttatct tttatctatg tatgttagagt ctgtgttaatt tcttcatctc | 13680 |
| tttagataa ctaatatctc ttcaagcttg acttgtattc tgtgtAACCC atttattgCG | 13740 |
| tttcaattt caatgagttt gtttccatct ctgcaagttc tatttgcattt ttttgagaat | 13800 |
| cttcctggc ttttaaacac atttcttatt ttaatttttgc ggggtaccta gtatgttat | 13860 |
| gtatTTTGG agtacatgag atgtttgtat acaagcaaac aatgcataat aatcacattg | 13920 |
| tgtaaaatgg ggtatccatc ccctcaagca tttatcctt gtgttacaaa caatccaatt | 13980 |
| atattctttt agtattttttaa aatgtacaa ttaaattattt attgaccata gtgactctgt | 14040 |
| tgtgctatca gatacttagt gatctttaa aaataatgtt ttctacttaa tctcattttt | 14100 |
| atgattccctt cttttacgtc atttgcattt tcaaatacag tcacttgtct gttgattcta | 14160 |
| ttatgtgaag ttttgagga taatctttt gttactttga ttccacccctt gtatggtttgc | 14220 |
| gctgtcccc cactaaaatc tcatacttgcata ctctggttcc cataataccc acatgttgc | 14280 |
| ggagggaccc tggggaggt gattagatta tagggacgtt tcccccttt gctctgttct | 14340 |
| ttttcctgcc accatgtaaag aaagatgtgt ttgcttcccc ttctgccatg attgtaaatt | 14400 |
| tcctgaggcc tccgcagcca tgcaggaccc ttttcttttgcataattaccc agtctccggc | 14460 |
| ggttctttat agtcccgta gaaaaaacta atacacacccatgtat ttttaccac | 14520 |
| tgaaattgtat tgcttaaattttaa taatctcact tggacccttgcataacccatgacttacata | 14580 |
| tctacccatca gagcagttac atctgtcaga cattcttagag gaatcagcag cacatggact | 14640 |
| ttgttgttgtttaaatttgcata tggggggagg ggggagggat agcattagga gatacaccta | 14700 |
| atgctaaatg acgagttaat ggggtgcagca caccacatg gcacatgtat acatatgtaa | 14760 |
| caaaccctgca cgttgcac atatacccta aaacttaaag tataataata ataaaattaa | 14820 |
| aaaaaaaaaaag gttctggag tattcaggtatgtat agattcagac atcgtgcagc | 14880 |

| | |
|---|-------|
| caggccccatg cttatgaatt ttcaggtgat acttcttttt cttnnnnctt aatttaaagc | 14940 |
| tggatctcg aaacagataa atttattttt ttatgacatg acgagcattt ttttcattct | 15000 |
| agttcatgct gttattgggt gtttagttct ttgagactcc tggcctttt ctaaaacctc | 15060 |
| aagttcaact tcctatttt cactggccca aggtcccata tccagttct atgtaaatgc | 15120 |
| taaacataag cctgtggaat attctagtct caccacatac tattcacatt ctctttgtt | 15180 |
| tttggcttc caggatttc cttaactttc tatgaaccca gtcttcatt tgaatggaa | 15240 |
| tttattatat attatctatc ctttctattt gtttatgca gaaagtgtt tctaaaatta | 15300 |
| tttaggcttc catattgcta gacatggaag ttgtaattat ttgttcagtg cctgtttcta | 15360 |
| catctaaact gcaagaccca tatggcaact gtgaatctt gtcccaagctt atttctgaag | 15420 |
| cttagaatag tgccctagcac aagaagtgt ttatctaaca tttttaaaaa taaatattaa | 15480 |
| attcatatct ggaatgaata ttaagttaga gctggtcatt gaggtgagag gaggaagcca | 15540 |
| agagagaata tgagagcctc aaagccaaat atcttaatg tacttttca gaaaagaaga | 15600 |
| cagccaatgt caggtggagg aactggttt tgaggttaact ttccctggaag aaaatagaaa | 15660 |
| ttactgaggt tttagataat ccaaataattt aatcaagtca ccaaggttt ttgtgggaa | 15720 |
| tctttattat taattaaaaat gagtgtgaa atcttaatat acgacaaaaag ttaaaaattt | 15780 |
| ctttgcagg cagatgaatg gtcttaggtat caaaaaatta agttgagtct ctaactcaca | 15840 |
| caaatttaca accctatcac tttatgaatt ttttaggag attatttttataacactgg | 15900 |
| tgaagtctaa gaatagctaa aatttatagt acacttattt tttgttattt actttttttt | 15960 |
| gaagtttgc atatagtgtat tcataatc ttccataaccc attttacatg tgaagaaact | 16020 |
| tagatataga aagattaaga aacttacata acttatccaa agttacacag taaaactctg | 16080 |
| gcattataac ttcaaaatca gctatcctac agtgagtaca gtgttctgtt cattgaaatc | 16140 |
| aaataagtga gatagcatcg tgatatagtt ttacgtatgc aaacactgtt acagagatct | 16200 |
| gtctaaagtt aaattccaca aatgaattct ttaaaagggt ttaatcaaga agaatatata | 16260 |
| aacaggatgg tgaaaaattt tcataattttt tttttttttt aatatcttta tgatttacag | 16320 |
| gcaagatggt agtggtgtga gagcggatgt tttcatgaaa ttcaattca ctagaaataaa | 16380 |
| caatggagca tcaatgaaaaa gcagaattga gtctgtttt ccacaaatgc tgaataactc | 16440 |
| tggaaacctg gaaataaacc cttcaactga gataacatgt aagtataatt ttccataaaac | 16500 |
| aattttattt caatataatcc ctcaaggttt ccaattcaaa ttcatatttt aatttgagagg | 16560 |
| ctgacttttc ttctttgaa actaaactgt gaaaacaatc cattaaaaag ctaaatataac | 16620 |
| catatagctc cctaacgtaa atcattctaa gacttaaaga atcatttggc atttatataag | 16680 |
| taaattttat ttgctaaaaa ttctcattaa ttatccctgc aacattcctt atgagtgtatg | 16740 |
| ttactgtcag atgtcattag tggataggcc ataggaggg tacatagatg ctcaaggta | 16800 |

ES 2 886 958 T3

| | | | | | | |
|-------------|-------------|------------|-------------|------------|------------|-------|
| gagaactatt | taattaatga | tccacctcag | aggcttcttc | attttcttt | gtaacattta | 16860 |
| tcacaattga | aattacaaag | ttatctgtgt | aaattttgt | ttgtttggct | tcatcctaca | 16920 |
| ctgtaatcat | cctaaaagaa | agaaccagtc | aaccttcttc | atcctactac | cctcctacca | 16980 |
| cccagtctcc | atcatataac | acatattcaa | taaataattc | ttgcatgact | gaaagaaaag | 17040 |
| aaataatata | tgcatalogat | ttaaggacat | tcctccaagt | tggttacatt | ctgctagttt | 17100 |
| aataagccat | tatttcttct | cgatgagctc | aagattaaaa | ggattttgat | gattccata | 17160 |
| ctagactgg | aggtaaccgt | tacagatgt | ctaactgtt | aatattgaaa | tgctttccta | 17220 |
| tttgttgta | aacaattact | gcatcaggcc | cacaaagtt | tcttccgaga | tgtttcaa | 17280 |
| ccactgcccc | tgctgctaaa | gagttatgct | tagcaaagca | aagcactcta | agacactgct | 17340 |
| ccaactccat | ggcctgattt | catctttat | gactggccaa | tgctcacgca | ctgcagttt | 17400 |
| tttagtagtt | gaatattacc | tctgcttcca | cacattaagg | aatgctcccg | aacgcacttc | 17460 |
| ccaagtgtt | atttatttat | cattatacta | gacaatatgg | tgatacgt | gtcacagaat | 17520 |
| agcggttcc | acctccagag | cccataatct | agttgaaggg | aaagatattc | caacacaaga | 17580 |
| gtgttgacaa | tcaagataga | atatgatcaa | gggccagtg | tgaggcccag | gcaatgatca | 17640 |
| ctgcaggaat | ctgggagaaga | aagagaccag | cgtgcttgg | atatctagca | aaagtttcat | 17700 |
| gaaggagaat | ggactttgac | tttgaatata | gggttaggatt | tacatattt | gagatgagaa | 17760 |
| aaagaaaagtt | cccagagaag | gaaagcatga | aaaggcaa | agtctgtact | gaacgcgatg | 17820 |
| ctttgacaga | ataatgaaga | aaggacactg | ctggaatgat | tgatcagt | tcatcattca | 17880 |
| caccatcatc | atcaaaacac | ttatttaatg | agaacttact | gttttttagg | catggcttta | 17940 |
| atgccctata | tgaattttt | tcttgattaa | tccttacaac | aaacatatcc | catagatagt | 18000 |
| tttattgtcc | cccttagaaa | agataaattt | cctaggctga | cacagt | atatgaggca | 18060 |
| gtcaggattc | aaactaagtc | tgtttgtca | aaaaattaag | aatggccagc | tttttaaaat | 18120 |
| tttctgtctc | cagaagtatg | atttggctcc | actgaagttt | gcaaaacaaa | tgtgataccc | 18180 |
| aaaccttgtg | aaacttttag | tggaaataa | cttgcataa | gtcggttga | gagagcgtgg | 18240 |
| aaacctgtct | tgaaaagttt | taatttaact | tgcaaggaaat | aaaaatgt | ggtttctcaa | 18300 |
| ttaaaaattt | caatcaagga | aggatatgag | ctaacataac | attttttaa | aaagatcagt | 18360 |
| ctggtaaggt | agaggtgc | aaactgaaaa | ggagcaaaag | tggtggatt | cagttagaaa | 18420 |
| attattgtaa | ctgtactgat | gtcaaatgat | gaaaccatga | actaaagt | tacaaaagg | 18480 |
| agtgaggagg | atgaaataat | tcaaaagata | gaggacagat | gtcagaacc | tggagattat | 18540 |
| aagatgtgaa | aggaggagtt | tgagaaaatt | tcaagatttt | gaagtgg | tgtgttcat | 18600 |
| aaaggatata | ataagtagca | aattttggat | aaagttgggt | cccactgagt | ttgagatggc | 18660 |

ES 2 886 958 T3

| | |
|---|-------|
| tgttggacat gcagagaaaa ctgtcttgta tgctgttctt aaattgaaat agacagaccc | 18720 |
| ttaccctctg atactgacat atttccctt ccaggctcac cctccatttc cctaaacaca | 18780 |
| acacatgcac tagctctcct tactttattg ctccacaaac atcttacacc tccaaggcatt | 18840 |
| tgtgcccaact gtaccttcta tctggaatct ctttgcctt cttgtgtgcc tgaaaaattc | 18900 |
| ctttcagatc ttcaaaatac agtgcagatg ctatttcttc tagctcaaattt attatctcct | 18960 |
| ccatataatt taattactct ctttttctt ttctctactt tgcacttaca tttatggaa | 19020 |
| tgattgcttg attaatttct acctgtaaat tatgtgaggg caggtcctct atatttgct | 19080 |
| cgcagttaaa tctgcagcac ttattataga gtggtatcat tagagtaata tacatatatt | 19140 |
| tgaggacatg ataaattaac ttcccctata gtatttatca cattgcacatc caatgacttg | 19200 |
| cttatgttcc tggggccca tataaattga gtaacttggaa aaaagagata tctattaagt | 19260 |
| attnaatgag aaattaaagt acaaacttta gtatgcataa caacaaattt ggaaaagggtt | 19320 |
| gtaaacaaag agattttag ggcgcatttag ttagagatcg tttcagcagg tctgaaagga | 19380 |
| agcctaggaa tctgcatttt agaggaccac ctcccaaccc caacaagtaa ttctgcttct | 19440 |
| tgttgtctgg gtactgtact ttaagaaattt atggtgaaat gatatcagcc tttattgtat | 19500 |
| ttatcttatt ctcattttt aatactagca cttactgacc aggctgcagc aaattggctt | 19560 |
| attaatggta agtttaata ttatggta actgtaattt gccaaatcat aaagagtaaa | 19620 |
| agtgcagtc ttttgttac tttggccaa ggcagttatct atcaagttga tgtctttgtt | 19680 |
| cttagttcgc tcaggtgggtg ttgaaacaag acagtgcgttgc tcccaagtgt cccatggagt | 19740 |
| ggacttttagg tttccccctt ctttttagaa aaaggaagaa gttgttagtgg aggactaccc | 19800 |
| actctgcact caaaattgcc ctcatgaaaa ttctttggc agcttggaa acctttact | 19860 |
| gccctggttc taaggtggca ttctgttaca cttacaaattt atggttgcgtt acaccgttt | 19920 |
| tgttagcttct cctaaccacc agatgtactt gcttgcgtt gaattcaggt taatcacaaa | 19980 |
| gtataataaa aaagaattgt cagaagtctt cccagtttgc ggtctataac ctgaaggaaa | 20040 |
| agtcaactact cttcaacatc atcctatgtt ctctcaggctt agatgtcagaaatgc | 20100 |
| cctagaaaaac agcaacttac ttctctgacc aaaaaatgc agttttttt tagttcaatg | 20160 |
| tacctggtag ctggcttac ttaggtactt cagtgtttt acaaaatgtt ggtgttcaatg | 20220 |
| tgggtgtttt tcagttcac tacgtatttta attcatgtt attgttaatg aaactgtgt | 20280 |
| aagcaattta ctagggattt tggttggag atgccacaaa ggaacacatg tatctctttaa | 20340 |
| tggaaaggctg gtctccctt atccaggaaa ttgcgttgc aaaaaagcc tttaggtgg | 20400 |
| tgtgttattt aaccaggca ctactttttt gccagccccag caatgttgc gtgatttacc | 20460 |
| attaaatttct tagtaataga ccacacaaaaa gaagaaaattt atggaaatgc gagttggag | 20520 |
| gaattgggtt atcagcttac cccagccccgt ttcaagcttgc qccagttttttt tatttcacgg | 20580 |

ES 2 886 958 T3

| | |
|--|-------|
| ctctttgaaa acattnaat aaacccatt tagatactag aaaccctctg tcaccctcaa | 20640 |
| gaatattctg tggtagatcg actcctttat gagggcatgt ttggtaatac agcatcagtc | 20700 |
| ttggagggtgg actggattct acaaggtgaa ctgcagtcac taaggagtct tttggatgag | 20760 |
| accagtttc ctccaacttc aatgttgca tgaacctcac atcaaaatgt agctttagat | 20820 |
| ttgtccccatg atgtggttcc aagaatcagc acttctaata agttccagg ggatgccat | 20880 |
| gctgcaggcc cacaaccac actgagcata gcaagactat tgaaaaaag gaaatttccc | 20940 |
| aggagtctgt ggcctgagct ggcacatcca ataatgacct atcttaacct caactcatga | 21000 |
| ggaattccag ggaactctga agctgctcaa aatttgaagc ctatatgccca actaaattca | 21060 |
| gaaatgttct ccaaaatgct atctataagc aacagtagtc acaaatgcat tgttagaaata | 21120 |
| tatcgatcat gcttttggaa aatccagca tgcctgagg aagaatgtat aagacataaa | 21180 |
| agtcataaat tatggaaaga ctcttcagct tcttccaaat gtaaaggaat catgatctc | 21240 |
| ccagcacatt aatgccttt ctcattagaa tggggcccg gtccagacct aataacattg | 21300 |
| tctgagcaga gaatccttgg aggcaactgag gctgaggagg gaagctggcc gtggcaagtc | 21360 |
| agtctgcggc tcaataatgc ccaccactgt ggaggcagcc tgatcaataa catgtggatc | 21420 |
| ctgacagcag ctcactgctt cagaaggtga ggcaccact acctacccat ctggaaacaa | 21480 |
| ttagaataga caggtcatga agactgcacc ctctacccta ggattgaatt gagccagaaa | 21540 |
| taattcaatg caaaaaaaatc agtaagaatt ttcttcctat tcatgaaagg aaaaggattt | 21600 |
| ttccccctta gcatgctaattt tagtgctat ttctctgttt caggtataaa tatattagca | 21660 |
| cagtaaagaa caaagattta tatgtcagaa tgtttttaa atcctagcta taaaagctta | 21720 |
| agaaatttac taaatctcca taagctttat ttttttcca aattaaggga caacactgtt | 21780 |
| atctgtgact tagtgtaact ggtgcattt agtacactaa tgtaaacata cgtaaatgt | 21840 |
| tagcgaaacg aattgctgtg gaagattgc acattataatc atggagctg atggctaacc | 21900 |
| tagagactgc cccatgccat taatttatttcc attcataaaag attattgagt atcttagatg | 21960 |
| agcacagtgt tatatatattgt agaagctact agtataaaca aagtattgcc tctgccttca | 22020 |
| aagagcttac actcgaatgt tggaatcaga atgcacaaaaa ataatgatca attacaatga | 22080 |
| gtagcataaa taaaattaaat gtaggcaact tacaagaatt cttaattgag gtgactaaac | 22140 |
| tattgccaac actagggtga tatgctacca gtggcgagta ggtgcataa acttacctta | 22200 |
| ttggtaaaaaa gaaaagttca cattgctcat aaaagaagga ttttagattt cagcataact | 22260 |
| aaaatctgtt tcaaaccctgc cttgttactg gggcatcgca gaccacaaca gttgttggga | 22320 |
| acttaactca aaaagttcac ccagaaaaat aatggagatt tgaactcggt tgccctgac | 22380 |
| cataatcaatt ttcttctcag actcttactc taaaactggac ctccttatca cacacacaaa | 22440 |

ES 2 886 958 T3

| | |
|---|-------|
| gccttcata ggcagatcaa tccagtctta tttctcaaag catgtacctt gagttcaga | 22500 |
| taaacagcat tgttctttc ccctggactc ttccctacatt tccctaccta tgagtatctg | 22560 |
| atcaatctgc ttatccttga aatgttaata tatttaccac atctctattt gaattttatg | 22620 |
| aaattttga taatttctaa gtagttttt cagattata ggcactactt catggcacag | 22680 |
| tgactgttac aaacgtattt gttaaattta gaaggaataa agataaaaa gactaggta | 22740 |
| gttactgaac taaagttta ggaaatccc aattatttca aattttctt atggtat | 22800 |
| tatgacttaa tattttata tgcagtgaac aaatttggaaa cttaaaaaga tactcccaga | 22860 |
| attatcagtt ttctgatgtt gattggcaaa tttattacta tatccccaaat aacccaagag | 22920 |
| acaattca caaaaacatt tcaatttca ttgccacttggaaaggccaaa aagcagaaat | 22980 |
| ggcacgcatt gatttcaatc gtactcttga gtgtgggaac caggaattaa aataccttga | 23040 |
| cttatcaggc acttagcata accaagaacg gaatagaaac ctccctggat tctaagccct | 23100 |
| attcagttccc aatcaccaaa aaccaagtaa acgatatcac tataatggaaa gccacagtt | 23160 |
| taaatatcga caacgattac caaaggaatc catggactt tgaattttgc caccacat | 23220 |
| ccttctatttccattaccatga ttgatccact aaagctaaca gactctgtga accttgttatt | 23280 |
| ggaccctcc ctaaagacctt gattgtcact gagaaccatc agtgaggatt tggttgggc | 23340 |
| atgaccagcc ttacatcaaa gtacatagaa gtgtgaggt cttatcaaag aggattattt | 23400 |
| aattatcacc tcttctatgt agctttccct gatactctt ttcctctcca ttgagttcca | 23460 |
| cagaaatttt tttatctgcc tttaacagtt gtctcatga tttgtgatat ttgacttacc | 23520 |
| tcttgcgtt cttcttcact agttagagt tcctcaaaga aagagaccat aattacttat | 23580 |
| attttatttcccttgcgtt cttcttcact agttagagt tcctcaaaga aagagaccat aattacttat | 23640 |
| gttgaactat aattaatgaa aataatagct accttcatga aagttcaattt tgccttgcgtt | 23700 |
| actatacgatc acataataca ttgtctcat taatacttaa caattgtgtg agaaggatc | 23760 |
| accaatcaca ttttatatgt aaataaacc cagagctatt aattaacttgc tctataataa | 23820 |
| cactttcat atgtggcata gccaagattt aaatataat gttactggtt cccaaatgt | 23880 |
| gctctaaatttccat atgtggcata gccaagattt aaatataat gttactggtt cccaaatgt | 23940 |
| gggagggaaag agagaaaagg aaggaaagaa aaaagagtct cttcagaacc ttctgttac | 24000 |
| agactccgag caaaagaagt tgaatataaa aacaacatag gtttgcgtt tttctaaat | 24060 |
| tttttcttca aaatttttaa ctcaggttca ctcttacaca aactactgtg tcttataaaa | 24120 |
| gtatccgg tcatagaatt ttatccgttct gtattactc cactatctaa tctccataaa | 24180 |
| actcctaaat tggattatc ggtaacattt tggtttact caacccttag gaacaatgtt | 24240 |
| aagttaatca gcccctccaca tcacagatcc ttatccat cagtctgtac aaggcatttc | 24300 |
| tctcatttttta attttttttc ctccgttcat ccctggattt cactttcaattt qccctccctt | 24360 |

ES 2 886 958 T3

| | |
|--|-------|
| cacccatatg cctcatacta atatattcga aatatacatg tcttaaaggt acatgcacgc | 24420 |
| acctacaaaa cctatagtgt tttttgtat gtatatgtct ttaatttaaa taagtagcat | 24480 |
| tgtgtaaaag tctaataattt tttcttactg ttttcactca attcttggaa ttttcatctg | 24540 |
| atgcactgct gcatagcacc ccatggtatg cagccaccat atttccttca tccaattagg | 24600 |
| ttgcatgacc taccttccc ttgccacaaa gagtacacac aaaatattt tacttatctt | 24660 |
| tctgtaaacc ttcaggaatt tcagaagcac acatgcaggg tgctaaatat accagaatac | 24720 |
| tttccagcca cttaaatctt taccagtatt gcaaaagagg ccccattcc ctccacatca | 24780 |
| acatttagta ttattctttt gtttaagttt tatcaatctt tttaatgtac acaagatgtct | 24840 |
| cattttata attttaattt ctcagattac tagtttgagt atctttcat atatctaaga | 24900 |
| gctgtttga tctccctac catgaactgc cactaatatt ctggcttat ttacaatgg | 24960 |
| ttttctgct tatttattac tggtttacag acttttaaaa tatattctac aaaaatttta | 25020 |
| gacattaaac attaccaata ttttccatg gttcctcatc catctggtaa acttgtctat | 25080 |
| ggtatatcta attttgattt aatagaattc attctatttt taccttttag tttgtgtttt | 25140 |
| tgttgttttag ccaaaaagtc cccattccta ggtcataaag gtaatgtcct ttttttttt | 25200 |
| ttaacgctac tggctctctct ctgtctcccc ctatgtatat aggtgcacat atacttgtac | 25260 |
| acacatacat atacctataat atgaggggag ttcgataagt ttatggaaaa taaaattaaa | 25320 |
| agataaaaata aaaaattata aactttattt ctcaacataa gtccttcaa gttcaagaca | 25380 |
| cttttgaag caataataacc agccatatcg tccatcccta aagaactgag ggtcctgaga | 25440 |
| attnaactat gtcaatgcag tctttttac attactttt tacagtactt attgatgaaa | 25500 |
| aatgggtgcc ttttaaagat tggtaaaga ttaggaaaca aaaataagtc agaggaagtc | 25560 |
| aaatcaaggac tggaaagggtgg atgccttagtg atttattgt gaaactttca taaaactaac | 25620 |
| cttatttgat gagaggaatg agcatgagca tgggtgtat ggagaagaac tctggtggag | 25680 |
| cttccttggacttttct actaaagctt tggcttaactt tcttactctc ataagaagaa | 25740 |
| gatgttattt ttcactgacc cttagaagg tcaacaagca aaatgccttc agcatcccaa | 25800 |
| atgtctgttg tcatgacttt tggcttgac tagtctgggtt ttgctttgac tggaccactt | 25860 |
| ctacctcttt atagccattt gttgtatggt gctttgtctt caagattgtt ttagtaaagc | 25920 |
| catatttcat ctctgtttac aattttcaaa agaaataactt cagaatctt atctgacatg | 25980 |
| tttaaaattt ctatttggaaag ctctgacctt ggggtgcagct gatctggcg aaacagtttt | 26040 |
| ggcatccatc aagttagaaag tttgctcaac ttttagtttt cagtcagaat tgtataagct | 26100 |
| gaaccagttg agatgtctat ggtgtgtctt attgtttctc acagtttaatt gttggcctc | 26160 |
| tttgagacat gaacaagatg aaattttcc tagcaaactg atgtggatga tctgttgctg | 26220 |

| | |
|---|-------|
| cgggcttcac cctcaacaac atctctttct ttcttgaacaa aaattatcca ttagtaaact | 26280 |
| gatgattggg ggagatgctg tccccataaa cttttgtaa ggcataaata atttcaccat | 26340 |
| tcttccagtt tcaccataaa tttgacgaaa ttttgcttca attttagcag cattcatgtt | 26400 |
| gctttgataa gagctcttt caaattcatg tcttattcct cttagtgccc caaactagat | 26460 |
| cttggtcagt atgacaagtt agtatgagtt tatctgcatg caaaaatctt taaaatccat | 26520 |
| gcatagtttgc ttataatat acatttcaa tgaactttt aagaccat acatacatat | 26580 |
| gtatatatat gcacacacac acacacacac acacaaaaat cttcaaccat tatcagactt | 26640 |
| agtgcagaaa aattattcat ccattaacaa gataagaatg ccccttatca tcactactat | 26700 |
| ttaaatggag ctcctggcta aaggaaaaga cagggattga aaaaaattag ttaaatctaa | 26760 |
| aatgtttatt atttcaggtt tcttagttgc ttaaatggga agggaggtat ggacaaaaga | 26820 |
| gaaatcaaag atatttgtgt tatgctactt atcattaaag tatcagaata acttcattgg | 26880 |
| aatagaaaaa caccaagatc accccacatc atgtttctaa aatcttc catttttttca | 26940 |
| gacaagtgc catgtattcg gccagtgaag aattaaactc acttgccagc ttataatgca | 27000 |
| gaaaaatata gcaaagagat gtggatccaa tagttctag atagtgtac aggatggcta | 27060 |
| agatgaattt atatatctga aatgttcaca aattccctac tcataatgca ttttttctata | 27120 |
| atgttttagc aactctaatac ctcgtactg gattgccacg tctggattt ccacaacatt | 27180 |
| tcctaaacta agaatgagag taagaaatat tttaattcat aacaattata aatctgcaac | 27240 |
| tcatgaaaat gacattgcac ttgtgagact tgagaacagt gtcacccctt cccaaagat | 27300 |
| ccatagtgtg tgtctccag ctgctaccca gaatattcca cctggctcta ctgcttatgt | 27360 |
| aacaggatgg ggcgctcaag aatatgtgg taagtgtctc ggaaaaaaaaa attaacaata | 27420 |
| gaaatgtctt atatttgcta ttaggttaatt ttttaatttta ggaaacatct ggaataggtg | 27480 |
| tttctattct tctacagaca gaaccattct atattctgct cagcccaagc tctggctacc | 27540 |
| cctgagtctc cttagcaaag caaagcaatg ctccagaaac tatggaaatt ctcaaataata | 27600 |
| gtaataggaa aatgtaaaag aaagttatga agacacgagt tcttaataa tccagagatt | 27660 |
| ctataagatt caaatagctt ccctataaac aataaaaaag atttgtttt tttgtttgtt | 27720 |
| tgcttggttt ttagagacaa agactttctc agactggagt gcagtgggtgc aatcatggct | 27780 |
| tactgcagcc tcaaactctg gtcttaagaa atcctttgc ttcaagccatc ctcctgcctt | 27840 |
| gaattataaa taagtgtgtt ccaccatacc cagctttttt tttttttt tacagacagg | 27900 |
| ttcttgctct gttggccagg ctggctggaa attcctgccc tcaagccatc ctcctgcctt | 27960 |
| gttggcctcc caaagcaatg ggaggattta gattagacat tgtatgaggg cttataatac | 28020 |
| cttaaggtat taactgcct ttaaagtatt ctggatatg gcaaaaactc gatgtgtata | 28080 |
| taaacattgg tcatattgtt ttattgaatg aataaaatgg aaactaaaat gaggacaatg | 28140 |

ES 2 886 958 T3

| | |
|---|-------|
| cacaagagct actagaacca gtaagagtat cagcgaagga gtggaagggt agcattgaca | 28200 |
| atttccctgg gcttttaccc atgttgtaga ttgtctctcc aaggaataat acaaagcctt | 28260 |
| aatagtccta gaacacattc tattgtgttc ttatggccca aagtaaattg gtgttaga | 28320 |
| taacatttgc accagtcatg aaaaactatt ggtgtcattc tgagagtaca tcaatataaa | 28380 |
| atagactagt tctttagcct tgaaactaga ctgggttctc ttttgctgct aggttaaagg | 28440 |
| ttattcaata tgtaatcttc caatccaaaa tctgtcagtg gataattaa aagcttttag | 28500 |
| tcaattttaa gatatttggtt ttcttaaat ttaaggggc actgtgtcac aaagctaaag | 28560 |
| aaaaaaaaaga aaaaaaaaaact gatctgtgaa aggggttatac ctcatctact tgggaaattt | 28620 |
| tggctgcgaa gaaactccaa agtaaatctt tagaagcctt cattgttaaa tatgaaataa | 28680 |
| tgttggagt acatttattt cttctcaaat ttattatagg gtcataatg tacacatctt | 28740 |
| gaagtccatt ttttcctgc ttttataaca aacaggccac acagttccag agctaaggca | 28800 |
| aggacaggc tc agaataataa gtaatgatgt atgtaatgca ccacatagtt ataatggagc | 28860 |
| catcttgc t ggaatgctgt gtgctggagt acctaagggt ggagtggacg catgtcagg | 28920 |
| aagctcaaga caatctcatc catgtcatca tccaagaagt gtataagcac ttccatgtat | 28980 |
| gtgataatgt gatagacata agtgtacag ttacaataca cagccctgtt cctctaaaat | 29040 |
| ttataatcta gattttagaa ataaattttt ttatgaatga agtttatcta tcatgaaagc | 29100 |
| attaactctg agaggccaaa ttacagagta gttaccatc caaagctcaa gaatcagaaa | 29160 |
| gacctcgatt tgaattcctt aacctctatt accaagtctc taactaaaag ctggggataa | 29220 |
| tcataatgc acctaacttt ttgggtacta agaaaagttt aatgaagact aaatatatca | 29280 |
| ggcacatggt aaacaacaaa gaaatctcat ctatctact attattaatg tagaccatgg | 29340 |
| tcactcgtgt taataacttt aacctaacc tttaactgc tgtgaaggat taaataaaaa | 29400 |
| attaatcact atattataaa aattaattga tatataataa atgaattttt agagatacgt | 29460 |
| aataattcat ggactccttg aagatagaaa atttatacaa aatcctagta atttgagtca | 29520 |
| caaaagctcc tacaataatg aaacagtatg aatgaaaaag aaaagaaata actattat | 29580 |
| ttggatctag cccataattt ttaaccaat gcacaaaaac aaacaacaaa tatgaaattc | 29640 |
| tcactgtaaa gtgattaaaa tcaaatttga attctaaaat tttaattaa attatctaaa | 29700 |
| cataattgat gcagttat gtttaatag gttttgtca catatctgaa atccaaactcc | 29760 |
| acacagtgc aggaacagct ggtgtcagaa attaaatatt cttttagtct ggagttttaa | 29820 |
| aaaatcaatc tgtttacttg agtaatttgc tgctgtttc atgggtgaat tgtatacaga | 29880 |
| aggataagaa ttattcttcg catcaaagg tcactgactt tcataattag tgctcatgg | 29940 |
| ctttaaaaag tggataaaaaa gtagtctca catttcattgg aaagccccca atccatgagc | 30000 |

| | |
|---|-------|
| acatttccca aaatgaaaca ttttatcaa ctgcaagttg tgtgttaggtg gagatttgtt | 30060 |
| tttcaattgt caagatactg ttaattaccc agtcctttat ctcctttgg tggagatgtc | 30120 |
| tctgtgctag gaaacccttc ttgctctcct tcctgtttct ctttactac tgcccctgaa | 30180 |
| acaacaaatt ctcagaattt atgacagctt tccaaagaat ccatcaatca aataagcaac | 30240 |
| acaactcgac actgacaatt ccagacctac taagagcatt aattaagact taaaaataaa | 30300 |
| catgagttt aaaagggtgt tattcattat tttccattt ataacgtccc ttaccttctg | 30360 |
| tccttcagtg catacaaatt attatcttcc ttgaagccc gttcaagccg tacctcacca | 30420 |
| tgatacctc catgtatatt ccactctagg cctcactgat ttttaactga aatactataa | 30480 |
| tgcatagttc acacttaaaa aaaaaaaaaaacacagcac tttacataag agcttacagg | 30540 |
| atcctatTTG ttttatccat tctttgttc attttacaa tcattaattc aaaggaatta | 30600 |
| tattaattac ttctatgca cccgacgttg tgtaacaca acaatactat ccctgcattc | 30660 |
| agcaagtcta tggtctacaa gagaggacac aaattcaa at gtctgttagtc aagcagtgaa | 30720 |
| gctggctaga tatggaaaaa ttacaagtcc ctcttgctt aacatttgct tgcccacatt | 30780 |
| tggtcagaca tcatgcaaaa taatttctca ctatagaaaa aaaaacacta caaaaacaat | 30840 |
| aatataaaga actgagaact ggttaactga agcatgcata tgtcatctaa aagaagcagg | 30900 |
| tgacgaccag ctcatgaag tacttgcatt gcatttggc acttcacacaca ctgacccttc | 30960 |
| tccccaccta gaccagtaat taaacaggtt tggatgagct agctactaag agcagccaac | 31020 |
| tgaatagctg actaacttag aagcacactt ggttaataata gctgactttt attagtaactg | 31080 |
| actataactat atgctaagct gtactcaaag tgcttgagt tttaaactga tacaaacatt | 31140 |
| atatgagggaa acagaggtac agagagctat tcaccagctt accaaaggc acatagctgg | 31200 |
| taagtggagg acttaaacc agactatcta gttcagaac ccacagactt aatccatcgt | 31260 |
| gcagaacata agacatactc catctgtctc cccaaactagg ttattatgtg cacaatatt | 31320 |
| tattgggtgg ttggttcatt attatgactg ggtggtaagt atgtcattag gagtgttttg | 31380 |
| cttatgacta tataaatttc ttccaaaaaa gaagactttc tgatgatata ctatgcatca | 31440 |
| gacaccacgc agggtgctaa ggtaggaag ataagtgaga cttctagaaa ctcattcatt | 31500 |
| caacaaatat ctcttaaggg cttagaagctt aggtttcagc agtgaacaga ataggtatgt | 31560 |
| tctcttcgt gttggacctt atagtatatc tggaaaaaca gacattgaat aaatatcaca | 31620 |
| aatgcaagtg agtgtttcag agacatgcag ctgctacatc aaaacaaaac agaacaacaaac | 31680 |
| aaacaaacaa aaactgacca gtgggattaa gtgtaaatag gcacacaaat gcacaaatat | 31740 |
| gctttataa aatagtgaag cagtgcacaga gacacacaca agatataaag acacaatgaa | 31800 |
| gaacaattga gcccaaagct ggaaagggtg agagtgtgaa ggaaaaaggt tgatcagaga | 31860 |
| agtttcccg aaggagagaa agcctggatg attaggaggc aaccactcgg tgactgaggg | 31920 |

ES 2 886 958 T3

| | |
|--|-------|
| aaatctgaaa aatgtatttg tcatcttctc agacttgctg aaggaatgac ttgggtactt | 31980 |
| tgaggatttc agtaattttt ccatgacttg gtataatatt tcaaaaggaa ataggctgac | 32040 |
| tttatttgta taatgaatgt gactccttcc tcgactgcca tagaaataaa ctccctaata | 32100 |
| ttttgggttt gtctttgcac ttaagtaatc agtcattctg ttttttaca gggtaactct | 32160 |
| ggtgccccac tagtacaaga agactcacgg cggtttggg ttattgtggg gatagtaagc | 32220 |
| tggggagatc agtgtggcct gccggataag ccaggagtgt atactcgagt gacagcctac | 32280 |
| cttgactgga ttaggcaaca aactggatc tagtgcaca agtgcacccc ttttgc当地 | 32340 |
| tctgtatgca ggtgtgcctg tcttaattc caaagctta cattcaact gaaaaagaaaa | 32400 |
| ctagaaatgt cctaatttaa catcttgta cataaatatg gtttacaaa cactgtttaa | 32460 |
| cctttcttta ttattaaagg ttttcttattt tctccagaga actatatgaa ttttgc当地 | 32520 |
| tactgtggct gtgtacacaga agaaacacac taaactaatt acaaagttaa caatttcatt | 32580 |
| acagttgtgc taaatgcccc tagtgagaag aacaggaacc ttgagcatgt atagtagagg | 32640 |
| aacctgcaca ggtctgatgg gtcagagggg tcttctctgg gtttcaactga ggatgagaag | 32700 |
| taagcaaact gtggaaacat gcaaaggaaa aagtgtataga ataatattca agacaaaaag | 32760 |
| aacagtatga ggcaagagaa ataatatgta tttaaaattt ttggttactc aatatcttat | 32820 |
| acttagtatg agtcctaaaa tttaaaatgt gaaactgttg tactatactg ataacctaacc | 32880 |
| cttaattatt ctgttacaaac atgcttccat aggaaatagt ggataattttt cagctttaa | 32940 |
| aggcaaaagc taaaatagt cactcctcaa ctgagaccca aagaattata gatattttc | 33000 |
| atgatgaccc atgaaaaata tcactcatct acataaagga gagactatcttata | 33060 |
| gagaagctaa gaaatataacc tacacaaaact tgtcaggtgc ttacaacta catagtactt | 33120 |
| tttaacaaca aaataataat tttaagaatg aaaaatttaa tcattggaa gaacgtcccc | 33180 |
| ctacagactt cctatcactg gcagttatatttttgcgt aaaagggtcg tcaaacgcta | 33240 |
| aatctaagta acgaattgaa agtttaaaga gggggaaagag ttgggttgca aaggaaaagt | 33300 |
| tttaatagct taatatcaat agaatgatcc tgaagacaga aaaaacttttgcacttcttcc | 33360 |
| tctctcattt tctttctctc tctctccct tctcatacac atgcctcccc caccaaaagaa | 33420 |
| tataatgtaa attaaatcca ctaaaatgtt atggcatgaa aatctctgtt gctgtatca | 33480 |
| ctaataattcc tgagttttta tgagctccta gtacagctaa agtttgccta tgcatactg | 33540 |
| tctatgcgtc agagcttcct ctttctacaa gctaactccc tgcatactggg catcaggact | 33600 |
| gctccataaca ttgtgtgaaa acttcttgta ttgcgtatg taaaattgtt gaaacaccta | 33660 |
| caataaagcc atctactttt agggaaaggg agttgaaaat gcaaccaact ctggcgaac | 33720 |
| tgtacaaaca aatctttgct atacttttattt tcaaataat tctttttaaa ataatttccc | 33780 |

ES 2 886 958 T3

| | |
|--|-------|
| tgcctaatta tttatggaag ttatgacttt tgaaggacaa ttcaaaacca tttatttaat | 33840 |
| tggttctgca atgaaaagaac tgccccatat actctactaa aggcttggca ctttctgctg | 33900 |
| ccttttaatc cagcgctata attgaggcaa gcgtccagct tgacacctcg agataacttc | 33960 |
| gtataatgtt tgctatacga agttatgcta gtaactataa cggtcctaag gttagcgagct | 34020 |
| agctgcaacc gagaaaaaaa cgtgccatga ggtctctgta tccaagtgtg act | 34073 |

5 <210> 21
<211> 418
<212> PRT
<213> Secuencia artificial

10 <220>
<223> Proteína recombinante
<400> 21

ES 2 886 958 T3

Met Tyr Arg Pro Arg Pro Met Leu Ser Pro Ser Arg Phe Phe Thr Pro
1 5 10 15

Phe Ala Val Ala Phe Val Val Ile Ile Thr Val Gly Leu Leu Ala Met
20 25 30

Met Ala Gly Leu Leu Ile His Phe Leu Ala Phe Asp Gln Lys Ser Tyr
35 40 45

Phe Tyr Arg Ser Ser Phe Gln Leu Leu Asn Val Glu Tyr Asn Ser Gln
50 55 60

Leu Asn Ser Pro Ala Thr Gln Glu Tyr Arg Thr Leu Ser Gly Arg Ile
65 70 75 80

Glu Ser Leu Ile Thr Lys Thr Phe Lys Glu Ser Asn Leu Arg Asn Gln
85 90 95

Phe Ile Arg Ala His Val Ala Lys Leu Arg Gln Asp Gly Ser Gly Val
100 105 110

Arg Ala Asp Val Val Met Lys Phe Gln Phe Thr Arg Asn Asn Asn Gly
115 120 125

Ala Ser Met Lys Ser Arg Ile Glu Ser Val Leu Arg Gln Met Leu Asn
130 135 140

Asn Ser Gly Asn Leu Glu Ile Asn Pro Ser Thr Glu Ile Thr Ser Leu
145 150 155 160

Thr Asp Gln Ala Ala Ala Asn Trp Leu Ile Asn Glu Cys Gly Ala Gly
165 170 175

ES 2 886 958 T3

Pro Asp Leu Ile Thr Leu Ser Glu Gln Arg Ile Leu Gly Gly Thr Glu
180 185 190

Ala Glu Glu Gly Ser Trp Pro Trp Gln Val Ser Leu Arg Leu Asn Asn
195 200 205

Ala His His Cys Gly Gly Ser Leu Ile Asn Asn Met Trp Ile Leu Thr
210 215 220

Ala Ala His Cys Phe Arg Ser Asn Ser Asn Pro Arg Asp Trp Ile Ala
225 230 235 240

Thr Ser Gly Ile Ser Thr Thr Phe Pro Lys Leu Arg Met Arg Val Arg
245 250 255

Asn Ile Leu Ile His Asn Asn Tyr Lys Ser Ala Thr His Glu Asn Asp
260 265 270

Ile Ala Leu Val Arg Leu Glu Asn Ser Val Thr Phe Thr Lys Asp Ile
275 280 285

His Ser Val Cys Leu Pro Ala Ala Thr Gln Asn Ile Pro Pro Gly Ser
290 295 300

Thr Ala Tyr Val Thr Gly Trp Gly Ala Gln Glu Tyr Ala Gly His Thr
305 310 315 320

Val Pro Glu Leu Arg Gln Gly Gln Val Arg Ile Ile Ser Asn Asp Val
325 330 335

Cys Asn Ala Pro His Ser Tyr Asn Gly Ala Ile Leu Ser Gly Met Leu
340 345 350

Cys Ala Gly Val Pro Gln Gly Gly Val Asp Ala Cys Gln Gly Asp Ser
355 360 365

Gly Gly Pro Leu Val Gln Glu Asp Ser Arg Arg Leu Trp Phe Ile Val
370 375 380

Gly Ile Val Ser Trp Gly Asp Gln Cys Gly Leu Pro Asp Lys Pro Gly
385 390 395 400

Val Tyr Thr Arg Val Thr Ala Tyr Leu Asp Trp Ile Arg Gln Gln Thr
405 410 415

Gly Ile

ES 2 886 958 T3

| | | |
|----|--|--|
| | <210> 22 | |
| | <211> 257 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 5 | <220> | |
| | <223> oligonucleótido sintético | |
| 10 | <400> 22 | |
| | agcacccctc tattccgcag agtctaagaa atcgctgtgt ttagccctcg ccctgggac 60 | |
| | tgtcctcactg ggagctgctg tggctgctgt ctgccttgg aagttcagta agtgcaggga 120 | |
| | gcctcgatcc caccatgtgc tcctgcagtc cccagtgcgc tgagccagac cctgctct 180 | |
| | gggctattga gacctctgga ggccctccgt gaggttcctc tcttacataa cgaggctgtc 240 | |
| | tctcttccct tctcttg 257 | |
| 15 | <210> 23 | |
| | <211> 190 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| | <220> | |
| | <223> oligonucleótido sintético | |
| 20 | <400> 23 | |
| | ggtcagagga ccaaagggtga ggcaaggcca gacttggtgc tcctgtggtt ctcgagataa 60 | |
| | cttcgtataa tgtatgctat acgaagttat atgcattggcc tccgcgccgg gttttggcgc 120 | |
| | ctccccgggg cgccccccctc ctcacggcga gcgctgccac gtcagacgaa gggcgcagcg 180 | |
| | agcgtcctga 190 | |
| 25 | <210> 24 | |
| | <211> 171 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 30 | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 24 | |
| | attgtttgc caagttctaa ttccatcaga cctcgacctg cagcccctag ataacttcgt 60 | |
| | ataatgtatg ctatacgaag ttatgctagt aactataacg gtcctaaggt agcgagctag 120 | |
| | ctccacgtgg ctttgtccca gacttcctt gtcttcaaca accttctgca a 171 | |
| 35 | <210> 25 | |
| | <211> 177 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| | <220> | |
| | <223> oligonucleótido sintético | |
| 40 | <400> 25 | |

| | | |
|----|--|-----|
| | ggtcagagga ccaaaggta ggcaaggcca gacttggtgc tcctgtggtt ctcgagataa | 60 |
| | cttcgtataa tgtatgctat acgaagttat gctagtaact ataacggtcc taaggttagcg | 120 |
| | agctagctcc acgtggctt gtcccagact tccttgcct tcaacaacct tctgcaa | 177 |
| 5 | <210> 26 <211> 20 <212> ADN <213> Secuencia artificial | |
| 10 | <220> <223> oligonucleótido sintético | |
| | <400> 26 gccgtgactg tgacattc | 20 |
| 15 | <210> 27 <211> 22 <212> ADN <213> Secuencia artificial | |
| 20 | <220> <223> oligonucleótido sintético | |
| | <400> 27 tggaggagcc acctgatgcc tc | 22 |
| 25 | <210> 28 <211> 20 <212> ADN <213> Secuencia artificial | |
| 30 | <220> <223> oligonucleótido sintético | |
| | <400> 28 gccttgccct caatggaaac | 20 |
| 35 | <210> 29 <211> 21 <212> ADN <213> Secuencia artificial | |
| 40 | <220> <223> oligonucleótido sintético | |
| | <400> 29 ggttgcacag caaggaagaa g | 21 |
| 45 | <210> 30 <211> 24 <212> ADN <213> Secuencia artificial | |
| 50 | <220> <223> oligonucleótido sintético | |
| | <400> 30 ccaggagttc ctgtgagcct accc | 24 |
| 55 | <210> 31 <211> 20 <212> ADN | |
| 60 | | |

| | | |
|----|---------------------------------|----|
| | <213> Secuencia artificial | |
| 5 | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 31 | |
| | tggaatggaa ggagctggag | 20 |
| 10 | <210> 32 | |
| | <211> 19 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 15 | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 32 | |
| | gtccccaccc tcgtcaactg | 19 |
| 20 | <210> 33 | |
| | <211> 22 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 25 | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 33 | |
| | tgaggcctcc catcagcctg gg | 22 |
| 30 | <210> 34 | |
| | <211> 21 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 35 | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 34 | |
| 40 | ccacaatgc acatgggtct g | 21 |
| | <210> 35 | |
| | <211> 18 | |
| | <212> ADN | |
| 45 | <213> Secuencia artificial | |
| | <220> | |
| | <223> oligonucleótido sintético | |
| 50 | <400> 35 | |
| | ggtgcttgct ccccaaga | 18 |
| | <210> 36 | |
| | <211> 20 | |
| 55 | <212> ADN | |
| | <213> Secuencia artificial | |
| | <220> | |
| | <223> oligonucleótido sintético | |
| 60 | <400> 36 | |
| | cctaaaagggt gttgtaatgg | 20 |
| | <210> 37 | |
| 65 | <211> 24 | |
| | <212> ADN | |

| | | |
|----|--|-----|
| | <213> Secuencia artificial | |
| | <220> | |
| 5 | <223> oligonucleótido sintético | |
| | <400> 37 | |
| | ggcaataaaag aaggaagacg tttt | 24 |
| 10 | <210> 38 | |
| | <211> 120 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 15 | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 38 | |
| | ccagtcaggg acacacatgc tcacacgccc gcccacccgc acacactaca gtcgagataa | 60 |
| | cttcgtataa tgtatgctat acgaagttat atgcattggcc tccgcgccgg gttttggcgc | 120 |
| 20 | <210> 39 | |
| | <211> 198 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 25 | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 39 | |
| 30 | attcttagttg tggtttgtcc aaactcatca atgtatctta tcatgtctgg aataacttcg | 60 |
| | tataatgtat gctatacgaa gttatgctag taactataac ggtcctaagg tagcgagcta | 120 |
| | gccaagtctg tgtgctacca agtagaaaa ctgagcctgg aactcacaca tgcgtgtctg | 180 |
| | agagccccagc actatcgc | 198 |
| 35 | <210> 40 | |
| | <211> 100 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| | <220> | |
| 40 | <223> oligonucleótido sintético | |
| | <400> 40 | |
| | taatctgact ttctcttcat cggctctct tattcttaggc tgagctgtaa cgctgccgtc | 60 |
| | ccccacatcc agaaagctgct tccattcaga cctacacctacg | 100 |
| 45 | <210> 41 | |
| | <211> 177 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 50 | <220> | |
| | <223> oligonucleótido sintético | |

| | | |
|----|--|-----|
| | <400> 41 | |
| | ccagtcaggg acacacatgc tcacacgccc gcccacccgc acacactaca gtcgagataa | 60 |
| | cttcgtataa tgtatgctat acgaagttat gctagtaact ataacggtcc taaggttagcg | 120 |
| | agcttagccaa gtctgtgtgc taccaagtag caaaacttag cctgaaactc acacatg | 177 |
| 5 | <210> 42 | |
| | <211> 19 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 10 | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 42 | |
| | gagcaggccc atgacacat | 19 |
| 15 | <210> 43 | |
| | <211> 24 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| 20 | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 43 | |
| 25 | accattagat cccagcactg gaca | 24 |
| | <210> 44 | |
| | <211> 20 | |
| | <212> ADN | |
| 30 | <213> Secuencia artificial | |
| | <220> | |
| | <223> oligonucleótido sintético | |
| 35 | <400> 44 | |
| | aaacccttcc cgagagagaa | 20 |
| | <210> 45 | |
| | <211> 23 | |
| 40 | <212> ADN | |
| | <213> Secuencia artificial | |
| | <220> | |
| | <223> oligonucleótido sintético | |
| 45 | <400> 45 | |
| | gaggaacact gtgtcaagga ctt | 23 |
| | <210> 46 | |
| 50 | <211> 22 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| | <220> | |
| 55 | <223> oligonucleótido sintético | |
| | <400> 46 | |
| | cctgaaaagc ccggagtgcc ag | 22 |
| 60 | <210> 47 | |

| | | |
|----|---|----|
| | <211> 19 <212> ADN <213> Secuencia artificial | |
| 5 | <220> <223> oligonucleótido sintético | |
| 10 | <400> 47 gggcagagac cacatctga | 19 |
| 15 | <210> 48 <211> 22 <212> ADN <213> Secuencia artificial | |
| 20 | <220> <223> oligonucleótido sintético | |
| 25 | <400> 48 ggaaggccctc tctcgatact tg | 22 |
| 30 | <210> 49 <211> 22 <212> ADN <213> Secuencia artificial | |
| 35 | <220> <223> oligonucleótido sintético | |
| 40 | <400> 49 ttctaccctg agggcatgca gc | 22 |
| 45 | <210> 50 <211> 22 <212> ADN <213> Secuencia artificial | |
| 50 | <220> <223> oligonucleótido sintético | |
| 55 | <400> 50 tgggatgtag aagggtgtca ga | 22 |
| 60 | <210> 51 <211> 22 <212> ADN <213> Secuencia artificial | |
| 65 | <220> <223> oligonucleótido sintético | |
| | <400> 51 ctgagcctgg aactcacaca tg | 22 |
| | <210> 52 <211> 23 <212> ADN <213> Secuencia artificial | |
| | <220> | |
| | <223> oligonucleótido sintético | |
| | <400> 52 tctgagagcc cagcactatc gcc | 23 |
| | <210> 53 | |

| | | |
|----|--|-----|
| | <211> 19 <212> ADN <213> Secuencia artificial | |
| 5 | <220> <223> oligonucleótido sintético | |
| 10 | <400> 53 gctgagggtc aggcttgag | 19 |
| 15 | <210> 54 <211> 21 <212> ADN <213> Secuencia artificial | |
| 20 | <220> <223> oligonucleótido sintético | |
| 25 | <400> 54 tctgcagggt agggagagaa g | 21 |
| 30 | <210> 55 <211> 29 <212> ADN <213> Secuencia artificial | |
| 35 | <220> <223> oligonucleótido sintético | |
| 40 | <400> 55 tgttcagaa aaggaagact cacgttaca | 29 |
| 45 | <210> 56 <211> 24 <212> ADN <213> Secuencia artificial | |
| 50 | <220> <223> oligonucleótido sintético | |
| 55 | <400> 56 gagaccgatg aagagaaaat caga | 24 |
| 60 | <210> 57 <211> 100 <212> ADN <213> Secuencia artificial | |
| 65 | <220> <223> oligonucleótido sintético | |
| | <400> 57 gaccattta aggtttgct tggttgttt ggagggaggg tggtgcttg ctaatggta | 60 |
| | attactaact cctcaataaa gaatattatt taaaataatt | 100 |
| | <210> 58 <211> 190 <212> ADN <213> Secuencia artificial | |
| | <220> <223> oligonucleótido sintético | |
| 65 | <400> 58 | |

| | | |
|----|---|-----|
| | gctgccttt aatccagcgc tataatttag gcaagggtcc agcttgacac ctcgagataa | 60 |
| | cttcgtataa tgtatgctat acgaagttat atgcattggcc tccgcgcgg gttttggcgc | 120 |
| | ctccccggg cgccccccctc ctcacggcga gcgctgccac gtcagacgaa gggcgcagcg | 180 |
| | agcgtcctga | 190 |
| | | |
| | <210> 59 | |
| 5 | <211> 171 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| | | |
| | <220> | |
| 10 | <223> oligonucleótido sintético | |
| | | |
| | <400> 59 | |
| | | |
| | attgtttgc caagttctaa ttccatcaga cctcgaccctg cagcccctag ataacttcgt | 60 |
| | ataatgtatg ctatacgaag ttatgctagt aactataacg gtcctaaggt agcgagctag | 120 |
| | | |
| | ctgcaaccga ggaaaaaaacg tgccatgagg tctctgtatc caagtgtgac t | 171 |
| | | |
| 15 | <210> 60 | |
| | <211> 177 | |
| | <212> ADN | |
| | <213> Secuencia artificial | |
| | | |
| 20 | <220> | |
| | <223> oligonucleótido sintético | |
| | | |
| | <400> 60 | |
| | | |
| | ccagtcaggg acacacatgc tcacacgccc gcccacccgc acacactaca ctcgagataa | 60 |
| | cttcgtataa tgtatgctat acgaagttat gctagtaact ataacggtcc taaggtagcg | 120 |
| | agctagctgc aaccgaggaa aaaacgtgcc atgaggtctc tgtatccaag tgtgact | 177 |
| 25 | | |
| | <210> 61 | |
| | <211> 21 | |
| | <212> ADN | |
| 30 | <213> Secuencia artificial | |
| | | |
| | <220> | |
| | <223> oligonucleótido sintético | |
| | | |
| 35 | <400> 61 | |
| | tcctctccag acaagaaaagc t | 21 |
| | | |
| | <210> 62 | |
| | <211> 30 | |
| 40 | <212> ADN | |
| | <213> Secuencia artificial | |
| | | |
| | <220> | |
| | <223> oligonucleótido sintético | |
| 45 | | |
| | <400> 62 | |

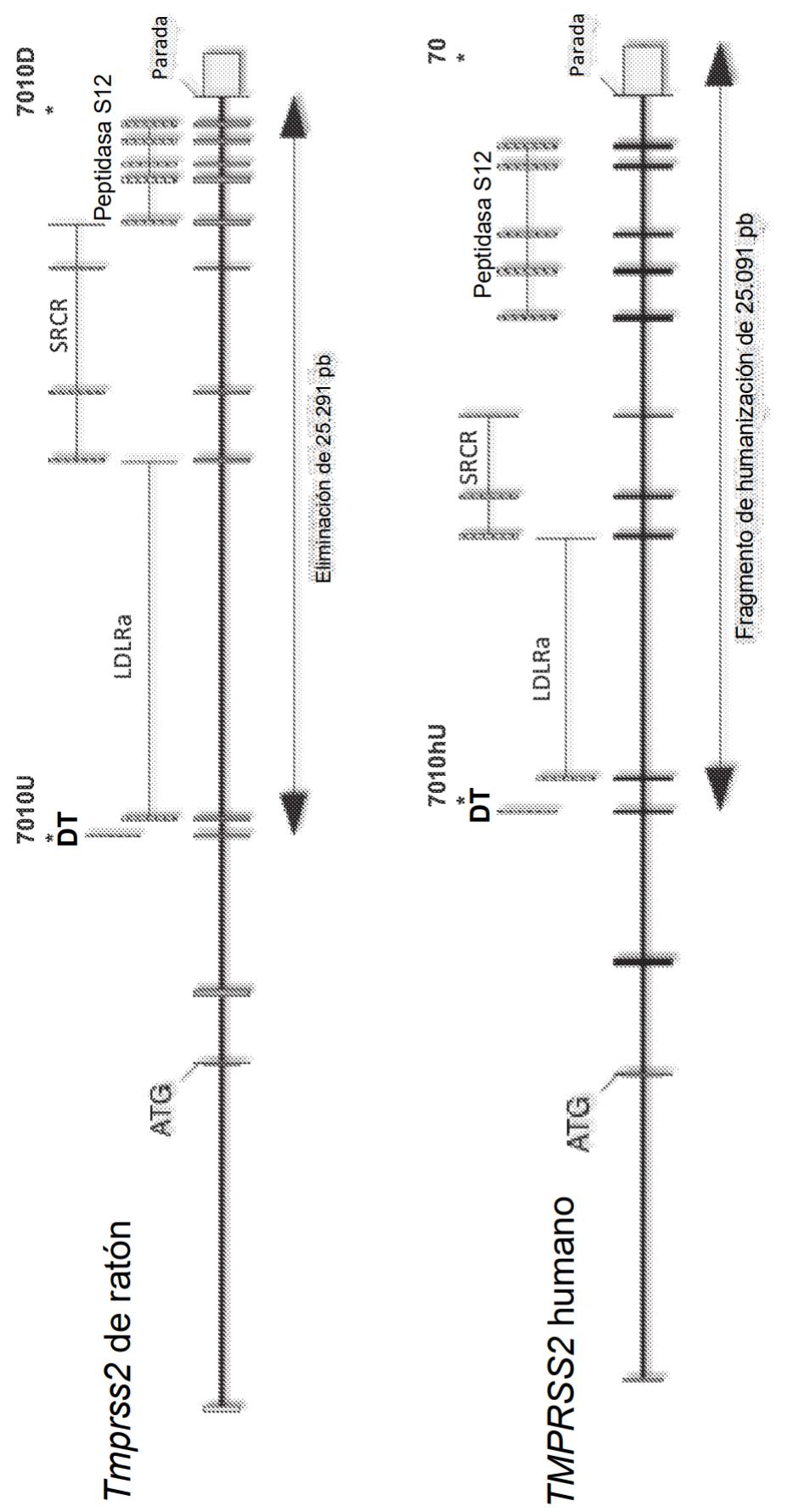
| | | |
|----|---|----|
| | tcatagcagc ttcaaattcc taaacgttga | 30 |
| 5 | <210> 63 <211> 20 <212> ADN <213> Secuencia artificial | |
| 10 | <220> <223> oligonucleótido sintético | |
| 10 | <400> 63 tcgtgttagt ctggtagtt | 20 |
| 15 | <210> 64 <211> 22 <212> ADN <213> Secuencia artificial | |
| 20 | <220> <223> oligonucleótido sintético | |
| 20 | <400> 64 catgcgatca caggaggaga tc | 22 |
| 25 | <210> 65 <211> 22 <212> ADN <213> Secuencia artificial | |
| 30 | <220> <223> oligonucleótido sintético | |
| 30 | <400> 65 aattggggcc gaagccagat gc | 22 |
| 35 | <210> 66 <211> 20 <212> ADN <213> Secuencia artificial | |
| 40 | <220> <223> oligonucleótido sintético | |
| 40 | <400> 66 cggaaggctt ctgtgacttc | 20 |
| 45 | <210> 67 <211> 25 <212> ADN <213> Secuencia artificial | |
| 50 | <220> <223> oligonucleótido sintético | |
| 50 | <400> 67 gtctccccact tctgacataa tgaac | 25 |
| 55 | <210> 68 <211> 27 <212> ADN <213> Secuencia artificial | |
| 60 | <220> <223> oligonucleótido sintético | |
| 65 | <400> 68 | |

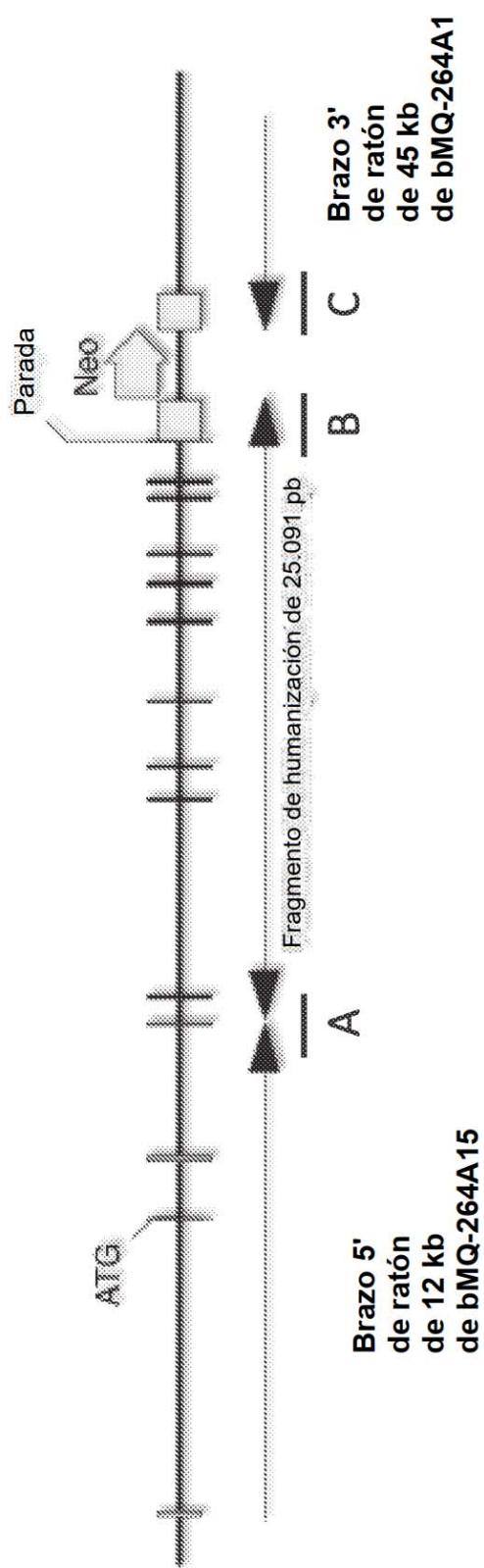
| | | |
|----|---|----|
| | cccagtgta accctacatc tggtcc | 27 |
| 5 | <210> 69 <211> 20 <212> ADN <213> Secuencia artificial | |
| 10 | <220> <223> oligonucleótido sintético | |
| 10 | <400> 69 tggaaagaga ctctggaca | 20 |
| 15 | <210> 70 <211> 25 <212> ADN <213> Secuencia artificial | |
| 20 | <220> <223> oligonucleótido sintético | |
| 20 | <400> 70 atgagtcct agtacagcta aagt | 25 |
| 25 | <210> 71 <211> 26 <212> ADN <213> Secuencia artificial | |
| 30 | <220> <223> oligonucleótido sintético | |
| 30 | <400> 71 atgcatgatc atctatgcgt cagagc | 26 |
| 35 | <210> 72 <211> 21 <212> ADN <213> Secuencia artificial | |
| 40 | <220> <223> oligonucleótido sintético | |
| 40 | <400> 72 tgcccagatg caggagtt g | 21 |

REIVINDICACIONES

1. Un roedor cuyo genoma comprende un gen *Tmprss* humanizado, en donde el gen *Tmprss* humanizado
 - 5 - comprende una secuencia de nucleótidos de un gen *Tmprss* de roedor endógeno y una secuencia de nucleótidos de un gen *TMPRSS* humano afín,
 - se encuentra en un locus *Tmprss* de roedor endógeno, y es el resultado de una sustitución de una secuencia genómica del gen *Tmprss* de roedor endógeno con dicha secuencia de nucleótidos del gen *TMPRSS* humano afín y está bajo el control del promotor del gen *Tmprss* de roedor endógeno, y
 - 10 - codifica una proteína *Tmprss* humanizada que comprende: (i) un ectodominio que tiene al menos un 85 % de identidad de secuencia con el ectodominio de la proteína *TMPRSS* humana codificada por el gen *TMPRSS* humano afín y (ii) una porción citoplasmática y transmembrana que tiene al menos un 85 % de identidad de secuencia con la porción citoplasmática y transmembrana de la proteína *Tmprss* de roedor endógena codificada por el gen *Tmprss* de roedor endógeno.
 - 15
2. El roedor de la reivindicación 1, en donde el gen *Tmprss* humanizado es un gen *Tmprss2* humanizado, el gen *Tmprss* de roedor endógeno es un gen *Tmprss2* de roedor endógeno y el gen *TMPRSS* humano afín es un gen *TMPRSS2* humano.
- 20 3. El roedor de la reivindicación 2, en donde la secuencia de nucleótidos del gen *TMPRSS2* humano comprende el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *TMPRSS2* humano.
4. El roedor de la reivindicación 2, en donde el gen *Tmprss2* humanizado comprende
 - 25 (i) los exones 1-2 codificantes del gen *Tmprss2* de roedor endógeno,
 - (ii) un exón 3 que comprende una porción 5' del exón 3 codificante del gen *Tmprss2* de roedor endógeno y una porción 3' del exón 3 codificante del gen *TMPRSS2* humano, y
 - (iii) el exón 4 codificante hasta el exón 13 codificante del gen *TMPRSS2* humano, y
- 30 5. El roedor de la reivindicación 1, en donde el gen *Tmprss* humanizado codifica una proteína *Tmprss2* humanizada que comprende una porción citoplasmática y transmembrana que tiene al menos un 85 % de identidad de secuencia con la porción citoplasmática y transmembrana de la proteína *Tmprss2* de roedor codificada por dicho gen *Tmprss2* de roedor endógeno y un ectodominio que tiene al menos un 85 % de identidad de secuencia con el ectodominio de la proteína *TMPRSS2* humana codificada por dicho gen *TMPRSS2* humano.
- 35 6. El roedor de la reivindicación 5, en donde el gen *TMPRSS4* humano comprende el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *TMPRSS4* humano.
7. El roedor de la reivindicación 5, en donde el gen *Tmprss4* humanizado comprende el exón 1 codificante hasta el exón 3 codificante del gen *Tmprss4* de roedor endógeno y el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *TMPRSS4* humano.
- 45 8. El roedor de la reivindicación 1, en donde el gen *Tmprss* humanizado es un gen *Tmprss11d* humanizado, el gen *Tmprss* de roedor endógeno es un gen *Tmprss11d* de roedor endógeno y el gen *TMPRSS* humano afín es un gen *TMPRSS11D* humano.
- 50 9. El roedor de la reivindicación 8, en donde la secuencia de nucleótidos del gen *TMPRSS11D* humano comprende el exón 3 codificante hasta el codón de parada en el exón 10 codificante del gen *TMPRSS11D* humano.
10. El roedor de la reivindicación 8, en donde el gen *Tmprss11d* humanizado que comprende los exones 1-2 codificantes del gen *Tmprss11d* de roedor endógeno y el exón 3 codificante hasta el exón 10 codificante del gen *TMPRSS11D* humano.
- 55 11. El roedor de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde el roedor es homocigoto para el gen *Tmprss* humanizado.
- 60 12. Un método para producir un roedor que tiene un gen *Tmprss* humanizado, que comprende:
 - (a) insertar un fragmento genómico en un locus *Tmprss* de roedor endógeno en una célula madre embrionaria de roedor, comprendiendo dicho fragmento genómico una secuencia de nucleótidos de un gen *TMPRSS* humano afín, formando, de este modo, un gen *Tmprss* humanizado,
 - 65 en donde el gen *Tmprss* humanizado está bajo el control del promotor del gen *Tmprss* de roedor en el locus *Tmprss*

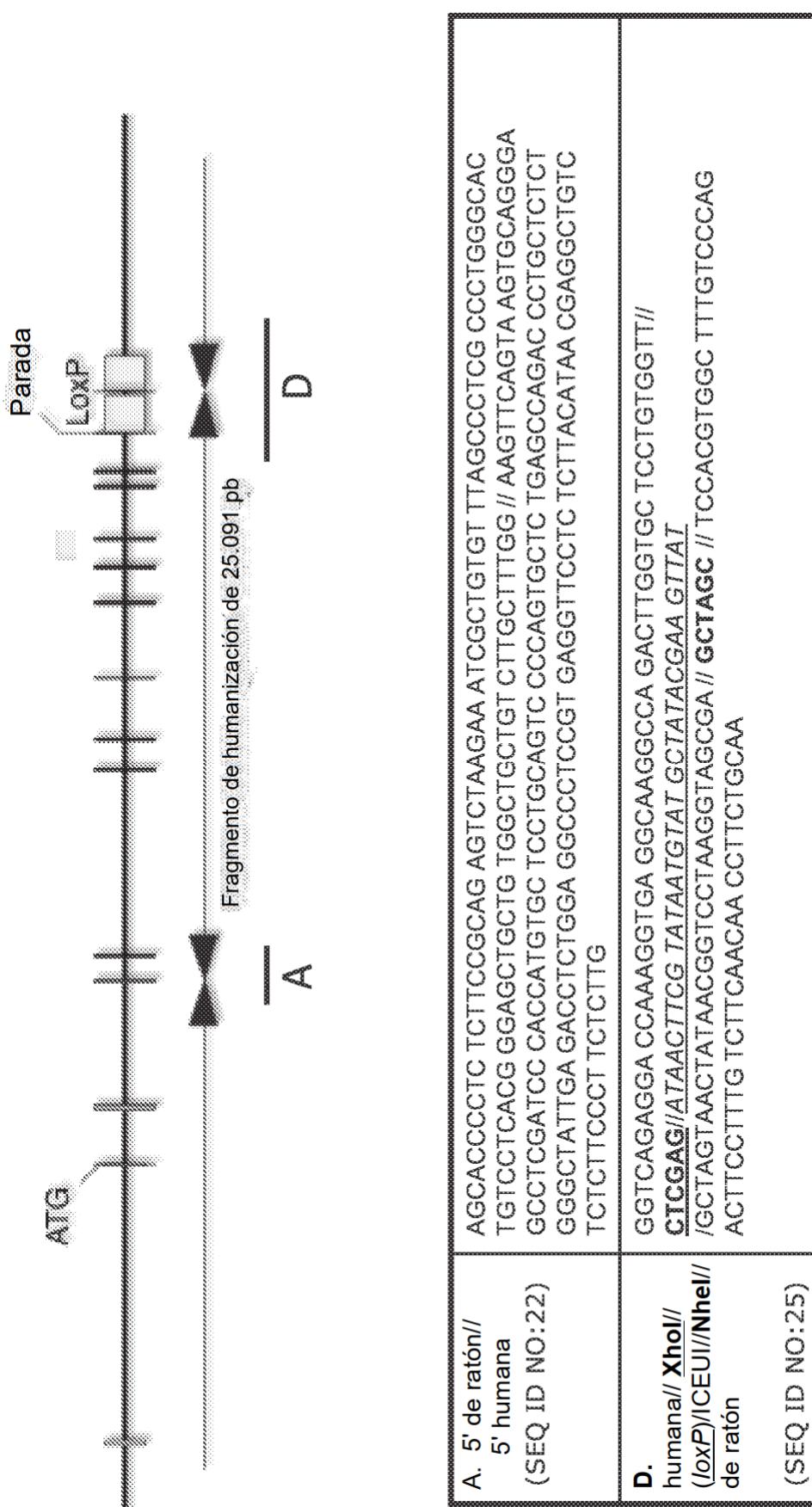
- de roedor endógeno y codifica una proteína Tmprss humanizada, y en donde la proteína Tmprss humanizada comprende (i) un ectodominio que tiene al menos un 85 % de identidad de secuencia con el ectodominio de la proteína TMPRSS humana codificada por el gen *TMPRSS* humano afín y (ii) una porción citoplasmática y transmembrana que tiene al menos un 85 % de identidad de secuencia con la porción citoplasmática y transmembrana de la proteína Tmprss de roedor codificada por el gen *Tmprss* de roedor en dicho locus *Tmprss* de roedor endógeno;
- (b) obtener una célula madre embrionaria de roedor que comprende el gen *Tmprss* humanizado de (a); y,
- (c) crear un roedor utilizando la célula madre embrionaria de roedor de (b).
- 10 13. El método de la reivindicación 11 o 12, en donde el gen *Tmprss* humanizado se selecciona del grupo que consiste en un gen *Tmprss2* humanizado, un gen *Tmprss4* humanizado y un gen *Tmprss11d* humanizado.
14. El método de la reivindicación 13, en donde el gen *Tmprss2* humanizado comprende (i) los exones 1-2 codificantes del gen *Tmprss2* de roedor endógeno, (ii) un exón 3 que comprende una porción 5' del exón 3 codificante del gen *Tmprss2* de roedor endógeno y una porción 3' del exón 3 codificante del gen *TMPRSS2* humano y (iii) el exón 4 codificante hasta el exón 13 codificante del gen *TMPRSS2* humano; y/o en donde el gen *Tmprss4* humanizado comprende el exón 1 codificante hasta el exón 3 codificante del gen *Tmprss4* de roedor endógeno y el exón 4 codificante hasta el codón de parada en el exón 13 codificante del gen *TMPRSS4* humano; y/o en donde el gen *Tmprss11d* humanizado comprende los exones 1-2 codificantes del gen *Tmprss11d* de roedor endógeno y el exón 3 codificante hasta el exón 10 codificante del gen *TMPRSS11D* humano.
- 15 15. El roedor acuerdo con una cualquiera de las reivindicaciones 1-11 o el método de acuerdo con una cualquiera de las reivindicaciones 12-14, en donde el roedor es un ratón o una rata.
- 20 16. Un método para evaluar la eficacia terapéutica de un compuesto en el tratamiento de una infección por el virus de la gripe A, que comprende:
- proporcionar un roedor de acuerdo con una cualquiera de las reivindicaciones 1-11 o 15;
- 30 administrar un virus de la gripe A y un compuesto candidato al roedor, en donde opcionalmente el compuesto candidato es un anticuerpo o fragmento de unión a antígeno del mismo específico para una proteína TMPRSS humana; y monitorizar la presencia y la gravedad de la infección por el virus de la gripe A en el roedor para determinar la eficacia terapéutica del compuesto candidato.

**Figura 1A**



| | |
|---|--|
| A. 5' de ratón//5 humana (SEQ ID NO:22) | AGCACCCCTC TCTTCCGAG AGTCATAAGAA ATCGCTGTGT TTAGCCCTCG CCCTGGGCAC TGTCCCTCACG GGAGCTGCTG TGGCTGCTGT CTTGCTTTGG // AAGTCAGTA AGTCAGGGGA GCCTCGATCC CACCATGTGC TCTTGAGTC CCCAGTGCTC TGAGGCCAGAC CCTGCTCTCT GGGCTATTGA GACCTCTGGA GCCGCTCCGT GAGGTTCCTC TCTTACATAACGAGGCTGTCTC TCTCTTCCCT TCTCTTG |
| B. humana// <u>Xba</u> I// casete (<i>loxP</i>) (SEQ ID NO:23) | GTCAGAGGA CCAAAGGTGA GGCAAGGCCA GACTTGGTGC TCTCTGGTT// <u>CICGAG</u> //ATAACTTCG TATAATGTAI GCTATAACGAA GTTATATGCATGGCC TCCGGCCGG GTTTGGCG CTCCCCGGG CGCCCCCCCCTC CTCACGGGA GGCTGCAC GTCAGACGA GGGCGAGCG AGCGTCTGA |
| C. Casete (<i>loxP</i>)/ICEU//NheI// de ratón (SEQ ID NO:24) | ATTGTTTGC CAAGTTCTAA TTCCATCAGA CCTCGACCTG CAGCCCCCTAG <u>ATAACTTCGT</u> ATAATGTAI GCTATAACGAG <u>TTAT</u> / GCTAGTAACATAACGGCTTAAGTAGCGA // <u>GCTAGC</u> // TCCACGTTGGC TTGTCCTTGC ACTTCCAG ACCTTCCCTTG TCTTCAACAA CCTTCTGCAA |

Figura 1B

**Figura 1C**

Alineación de proteínas Tmprss2

| | | | | | |
|--|-------|-------|-------|-------|-------|
| 10 | 20 | 30 | 40 | 50 | 60 |
| hTMRSS2 | | | | | |
| mTmrss2 | | | | | |
| pro 7010 mutante | ***** | ***** | ***** | ***** | ***** |
| | | | | | |
| 70 | 80 | 90 | 100 | 110 | 120 |
| hTMRSS2 | | | | | |
| mTmrss2 | | | | | |
| pro 7010 mutante | ***** | ***** | ***** | ***** | ***** |
| | | | | | |
| 130 | 140 | 150 | 160 | 170 | 180 |
| DSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQVYSSQRKSWMHPCQDDWNENY | | | | | |
| GSSGTCISSSLWCDGVAHCPNGEDENRCVRLYQSEFILQVYSSQRKAWYYPVCQDDWSESY | | | | | |
| DSSGTCINPSNWCDGVSHCPGGEDENRCVRLYGPNFILQVYSSQRKSWMHPCQDDWNENY | | | | | |
| ***** | ***** | ***** | ***** | ***** | ***** |
| | | | | | |
| 190 | 200 | 210 | 220 | 230 | 240 |
| GRAACRDMGYKNFYYSSQGIVDDSGSTSEFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLR | | | | | |
| GRAACKDMGYKNFYYSSQGIPDQSGATSEFMKLNVSSGNGVNDIYKKLYHSDCSSRMVVSLR | | | | | |
| GRAACRDMGYKNFYYSSQGIVDDSGSTSEFMKLNTSAGNVDIYKKLYHSDACSSKAVVSLR | | | | | |
| ***** | ***** | ***** | ***** | ***** | ***** |
| | | | | | |
| 250 | 260 | 270 | 280 | 290 | 300 |
| CIAAGVNLNSSRQSRRIVGGESALPGAWPWQVSLLHVQNWHVCGGSIITPEWIVTAAHCVVK | | | | | |
| CIECGVRS-VKRQSRIVGLNAPGDWPWQVSLLHVQGVHVCGGSIITPEWIVTAAHCVVE | | | | | |
| CIAAGVNLNSSRQSRRIVGGESALPGAWPWQVSLLHVQNWHVCGGSIITPEWIVTAAHCVK | | | | | |
| ***** | ***** | ***** | ***** | ***** | ***** |

Figura 1D

Alineación de proteínas Tmprss2

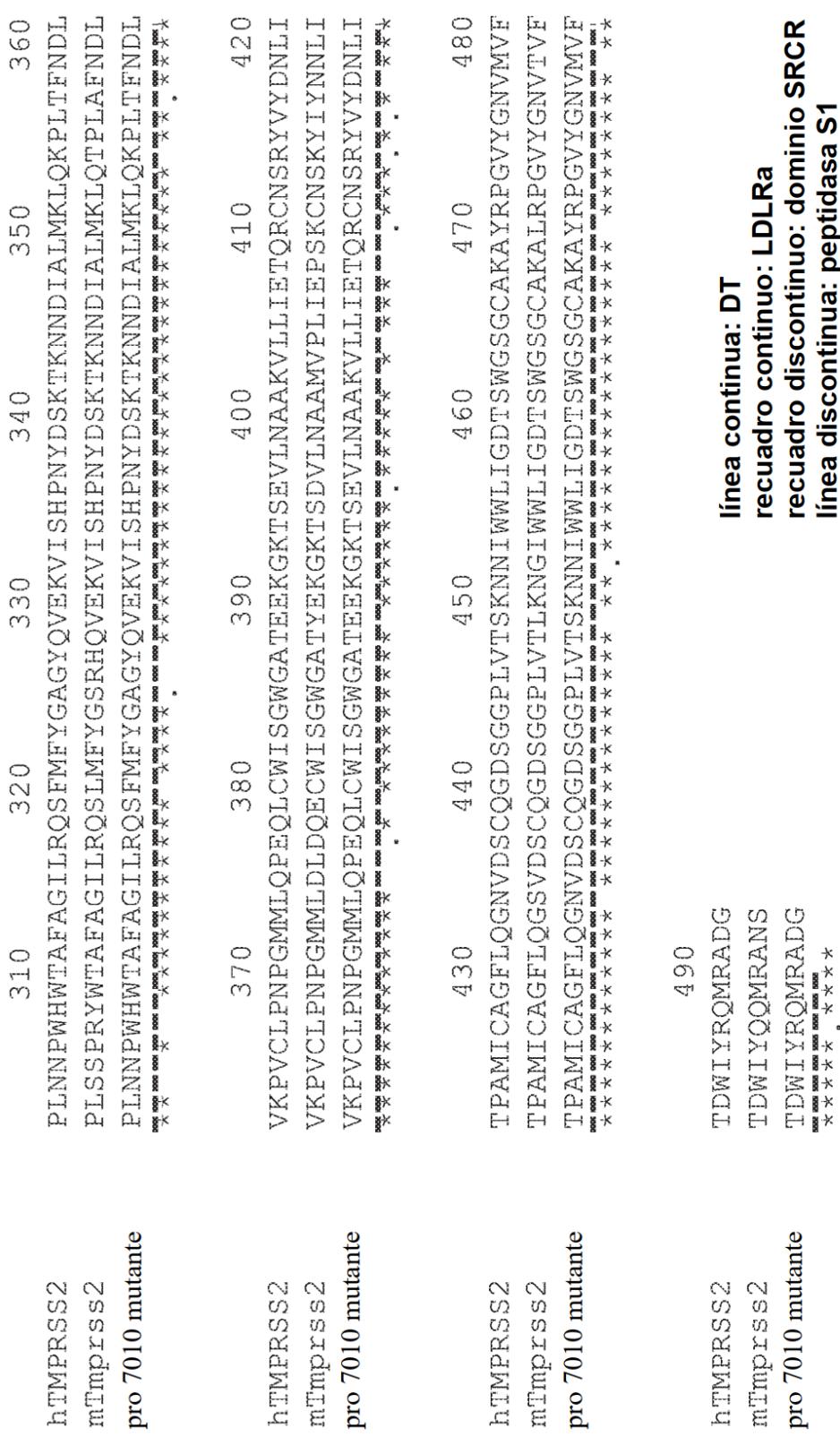
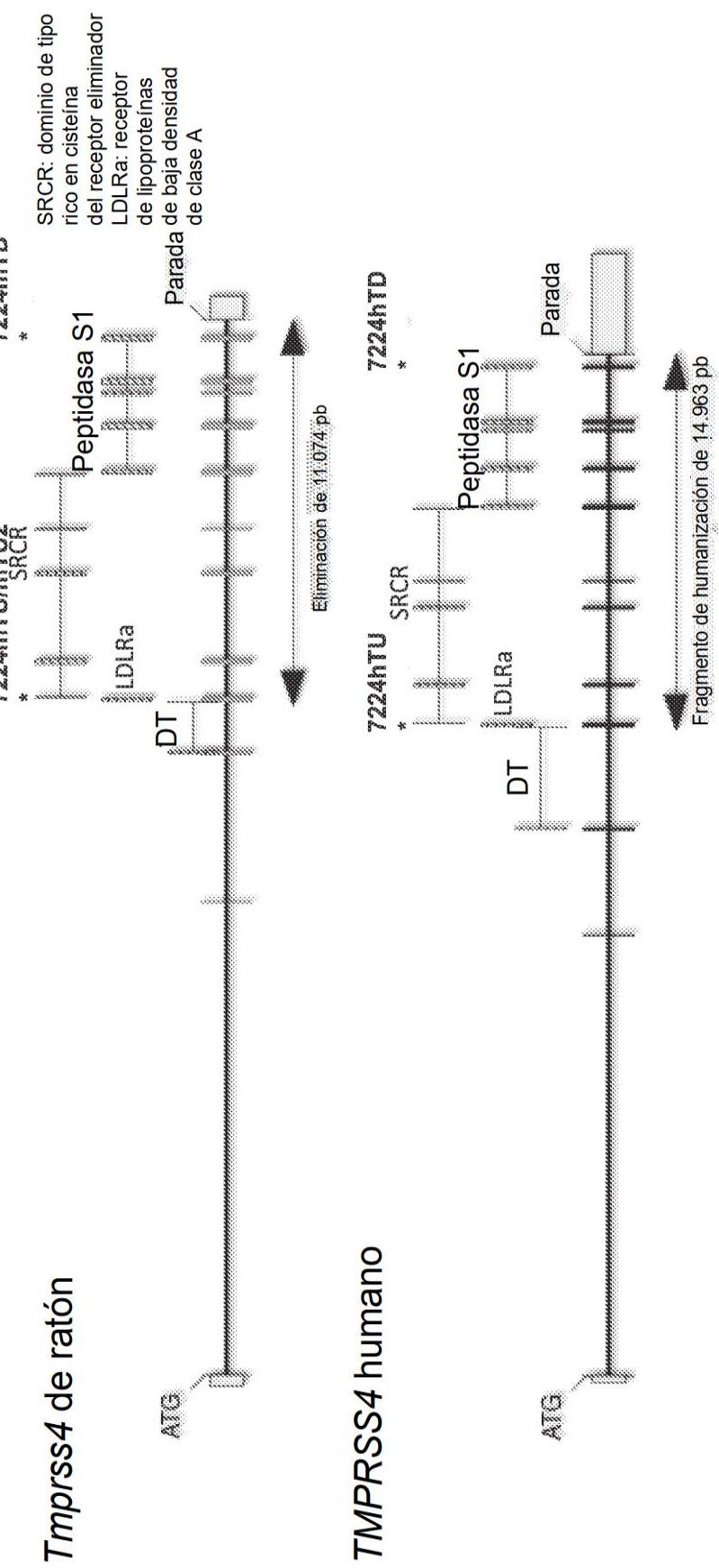
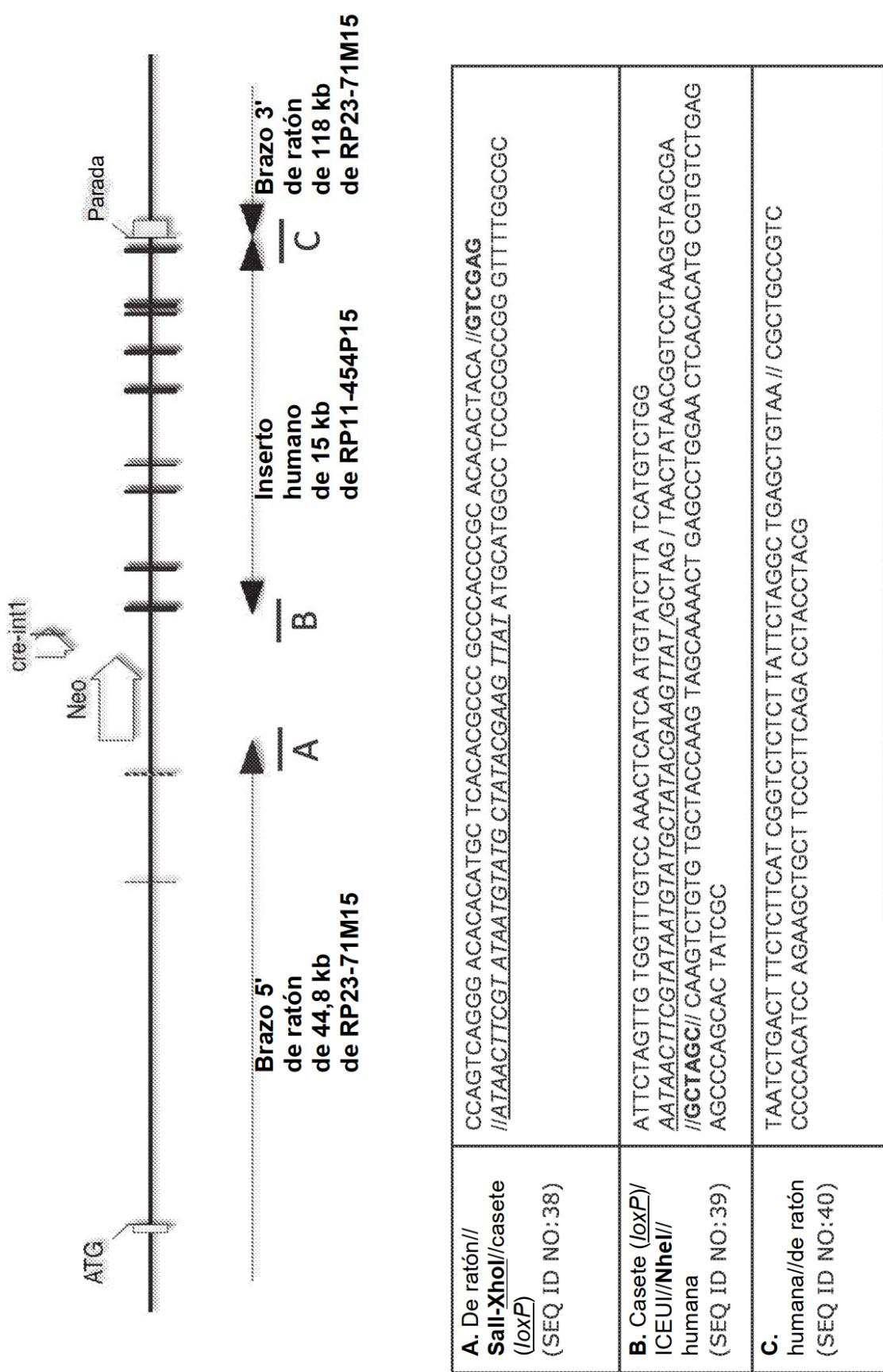


Figura 1D (continuación)

**Figura 2A**

**Figura 2B**

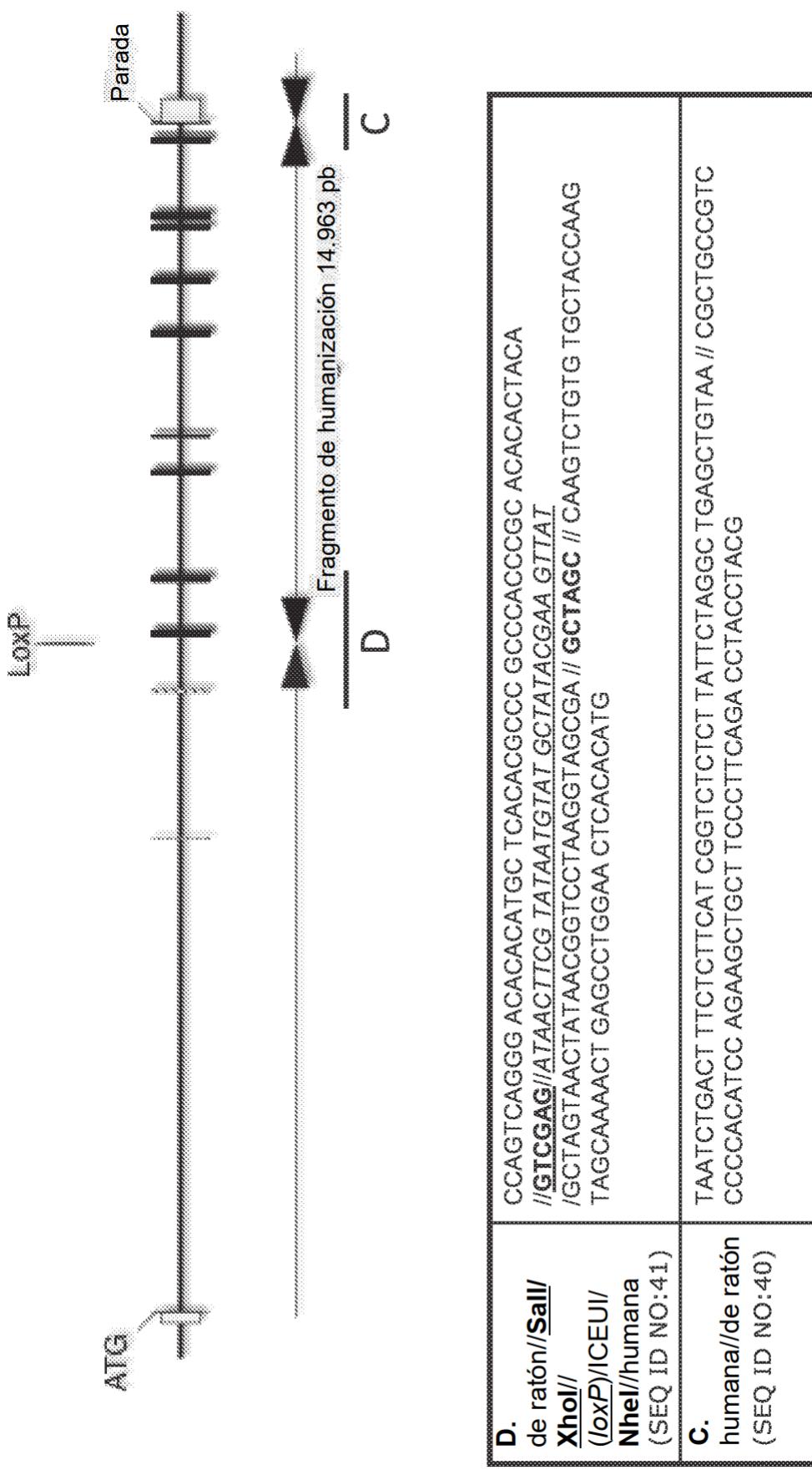


Figura 2C

Alineación de proteínas Tmprrss4

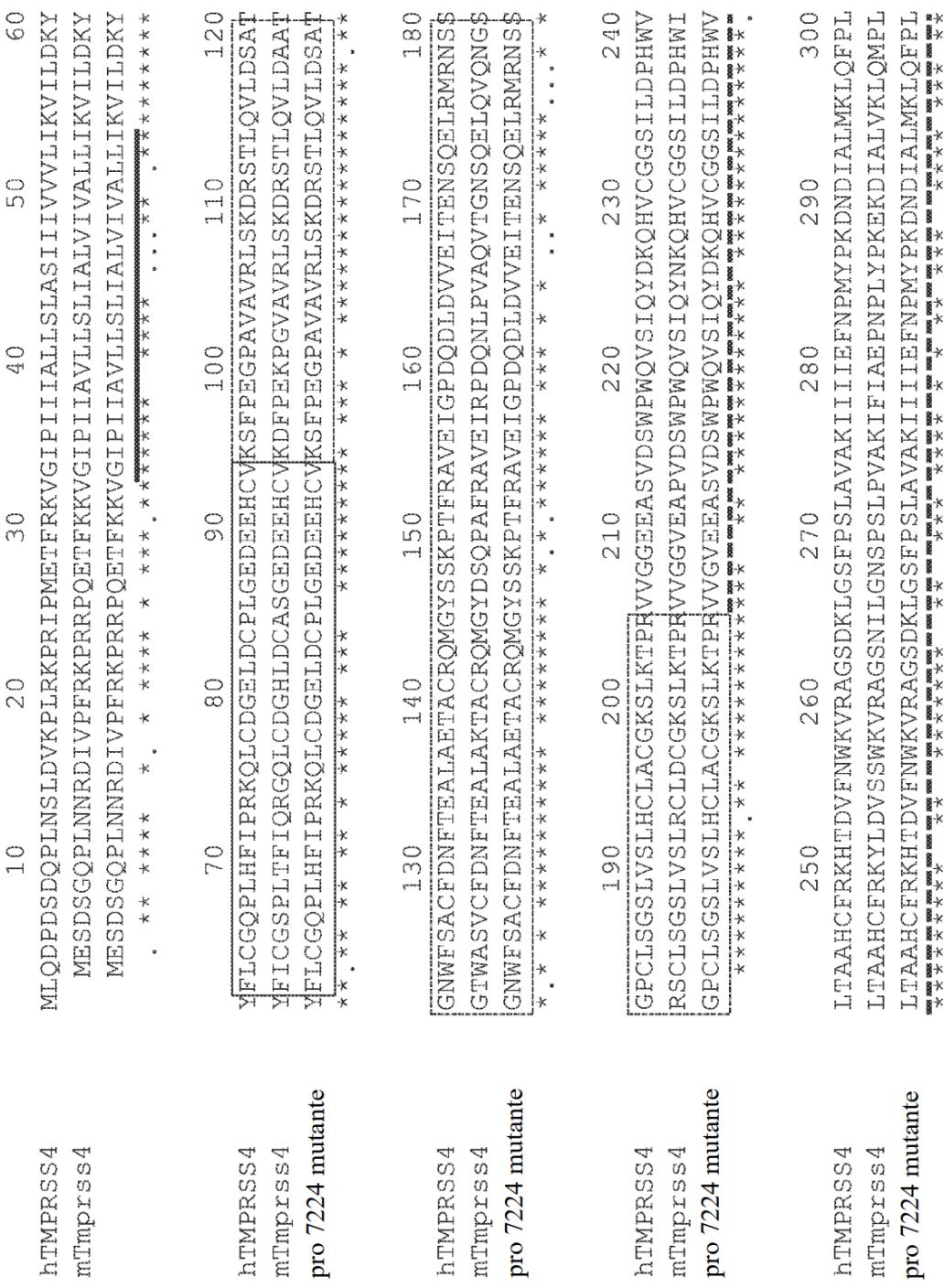


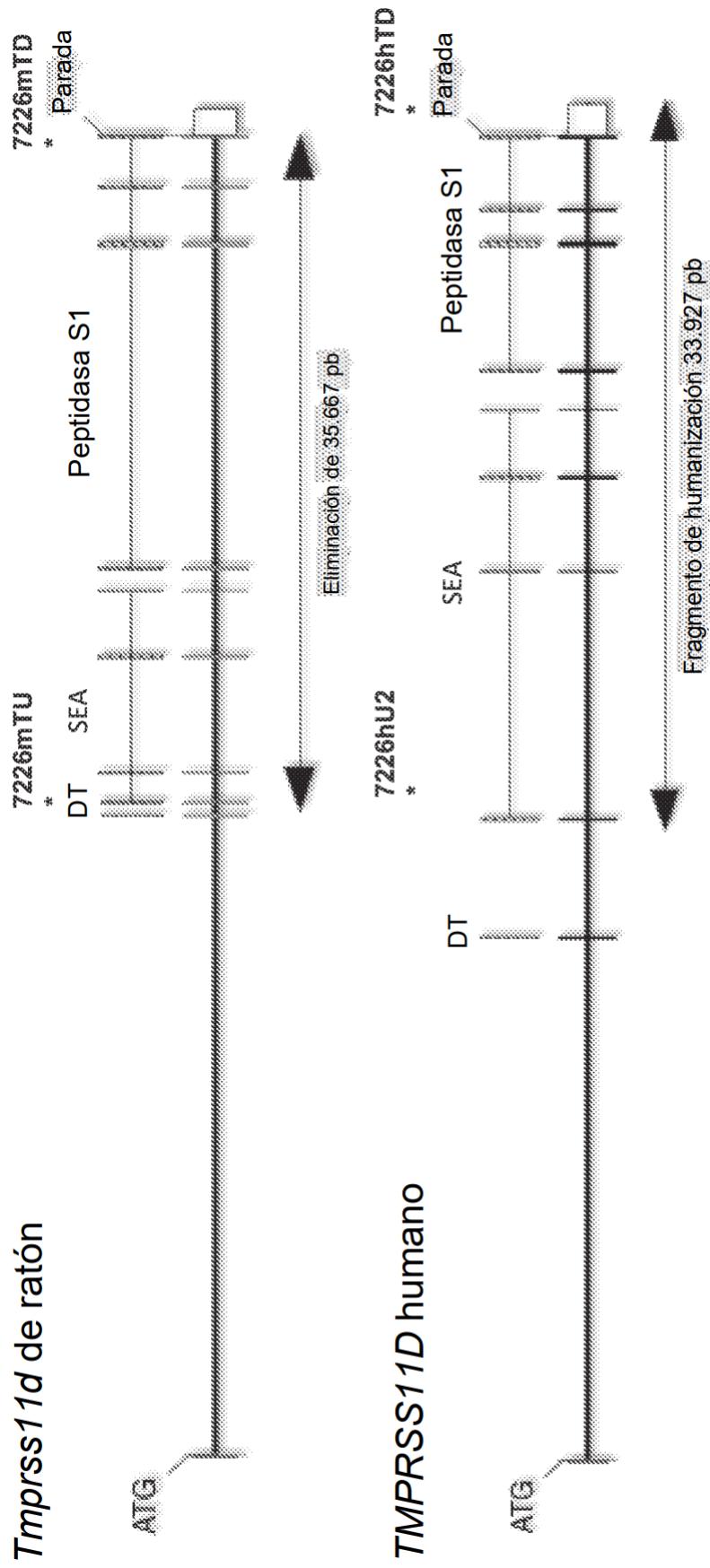
Figura 2D

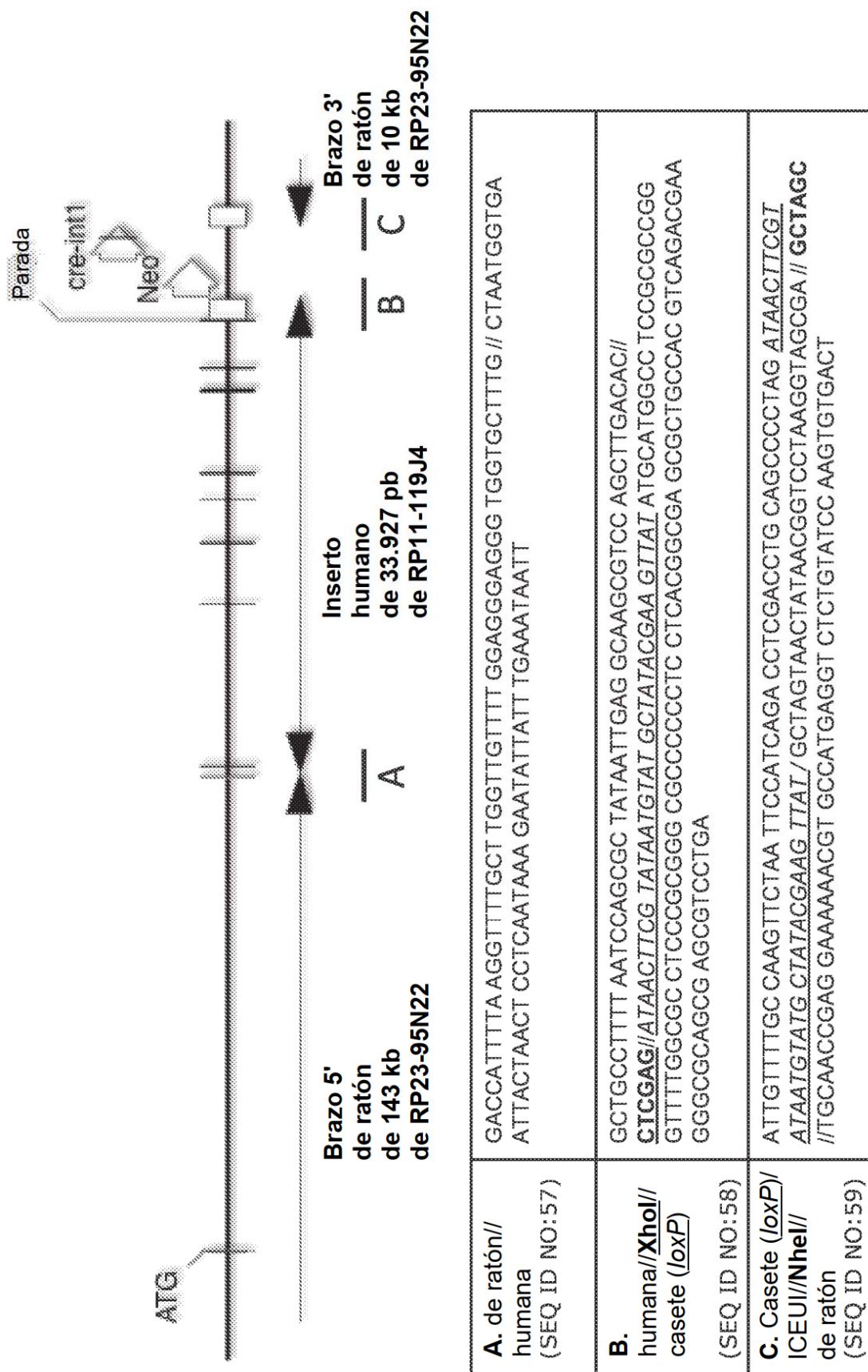
ES 2 886 958 T3

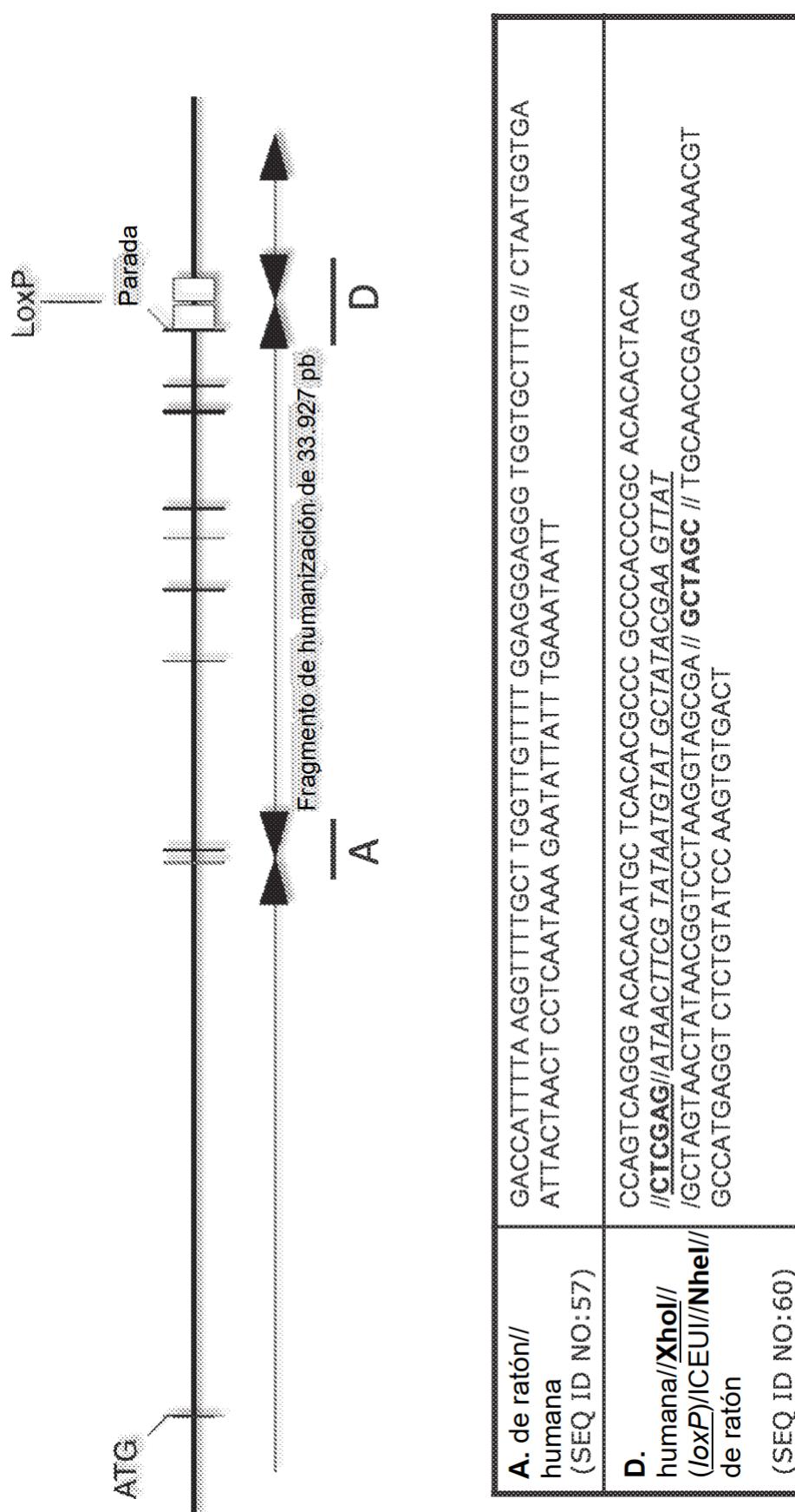
Alineación de proteínas Tmpress4

| | | | | | |
|------------------|-----|-----|-----|-----|-----|
| | | | | | |
| hTmprss4 | 310 | 320 | 330 | 340 | 350 |
| mTmprss4 | | | | | 360 |
| pro 7224 mutante | | | | | |
| | | | | | |
| hTmprss4 | 370 | 380 | 390 | 400 | 410 |
| mTmprss4 | | | | | 420 |
| pro 7224 mutante | | | | | |
| | | | | | |
| hTmprss4 | 430 | | | | |
| mTmprss4 | | | | | |
| pro 7224 mutante | | | | | |

Figura 2D (continuación)

**Figura 3A**

**Figura 3B**

**Figura 3C**

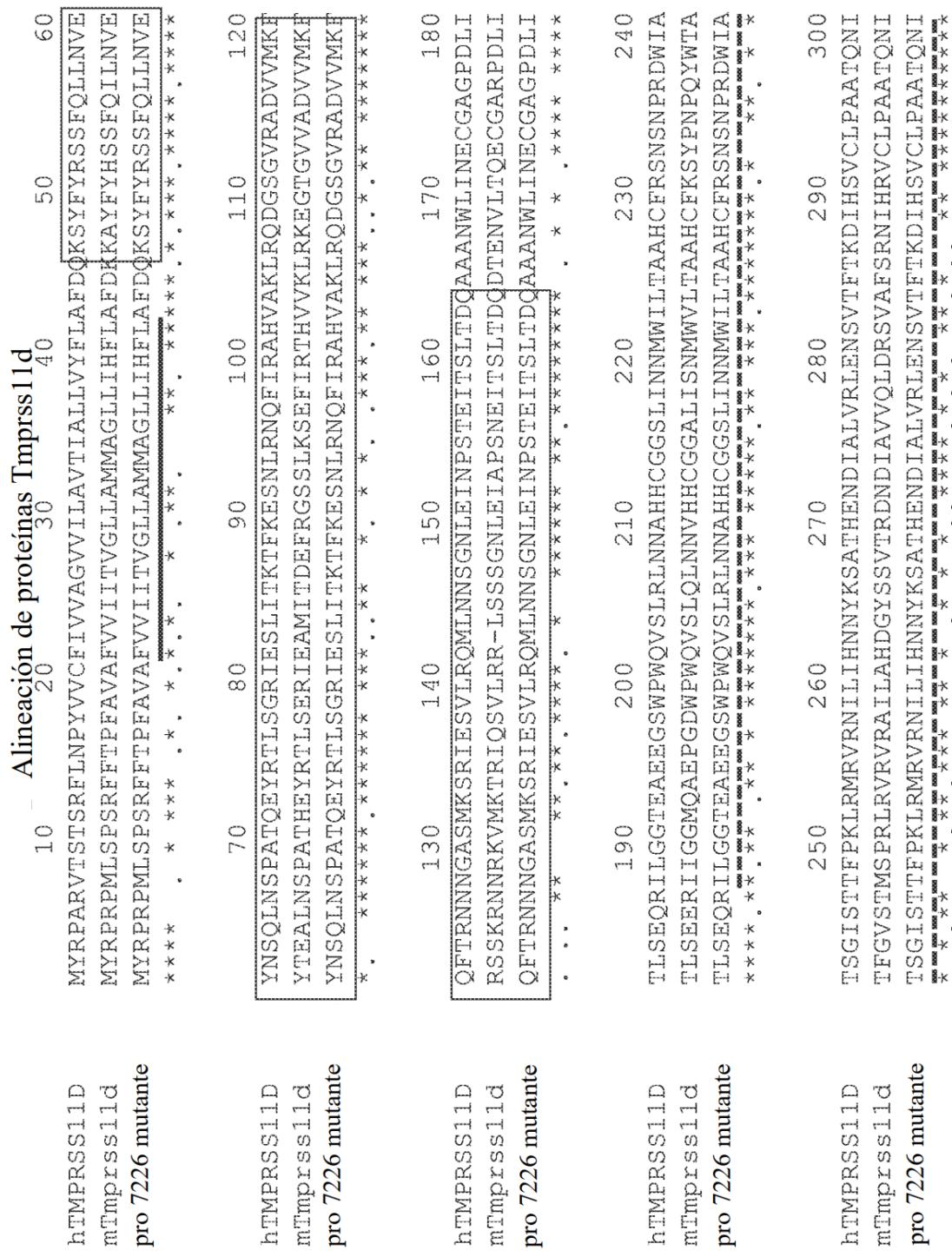


Figura 3D

ES 2 886 958 T3

Alineación de proteínas Tmpress1d

línea continua: DT
recuadro continuo: dominio SEA
línea discontinua: peptidasa S1

Figura 3D (continuación)

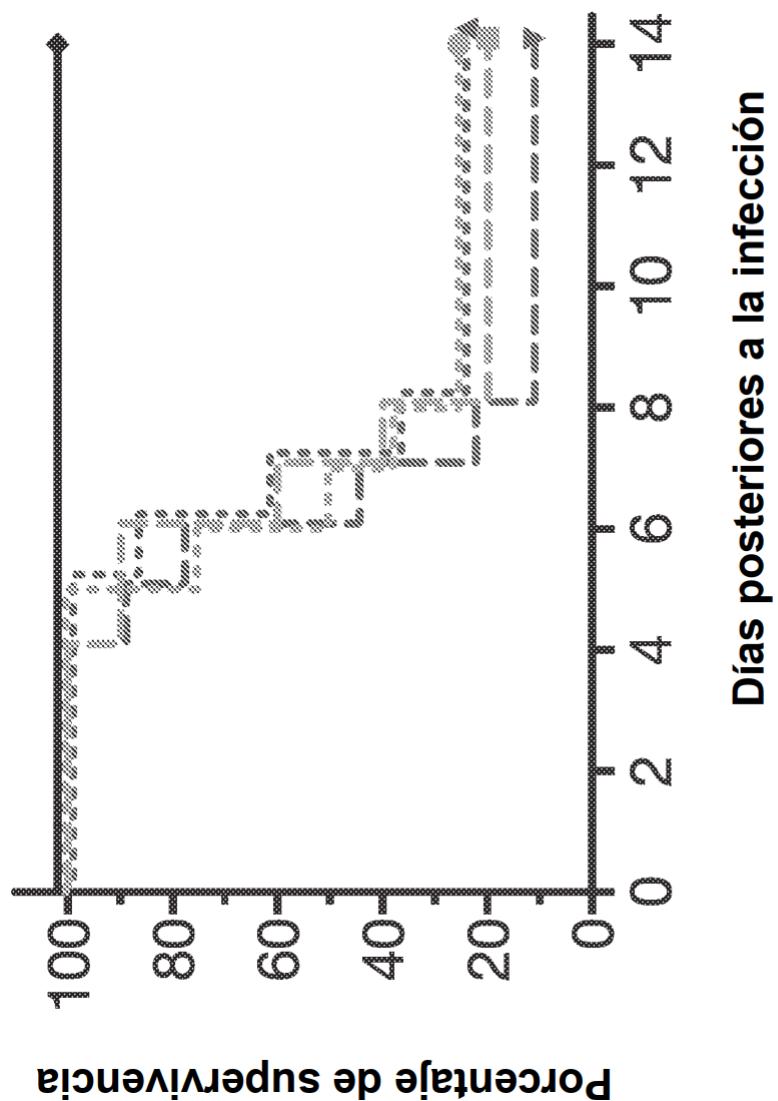


Figura 4