

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-Friebertshauser *et al.* 2010 J. Virol. 84: 5604; Bertam *et al.* 2010 J. Virol. 84:10016; Bertam *et al.* 2010 J. Virol. 84:10016; Böttcher-Friebertshauser *et al.* 2011 J. Virol. 85: 1554; Bahgat *et al.* 2011 Virol. J. 8:27).

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.

También se hace referencia a los siguientes documentos:

- Kuhn *et al.* (2015) consultado de internet: URL: http://elib.tiho.hannover.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:

- 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.

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, 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*. 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 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 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 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 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 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 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 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 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 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 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 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 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.

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.

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.

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.

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.

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.

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.

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.

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:

- (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).

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 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 TMPRSS11D 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.

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 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
 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.

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.

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 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

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.

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.

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).

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).

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.

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.

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).

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).

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.

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.

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).

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).

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).

La **figura 4** muestra los resultados de un experimento que muestra que los ratones MAID7225 HumInTMPRSS4 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 HumIn 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 HumIn 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 HumIn TMPRSS4 no infectado como control negativo y ratones de tipo silvestre (rombos negros, línea continua).

Descripción detallada de determinadas realizaciones

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*")

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 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 *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 erizo de mar), un dominio receptor eliminador del grupo A, un dominio LDLA (receptor de lipoproteínas de *baja* densidad de clase A), un dominio CUB (factor de crecimiento embrionario de erizo *Cls/Clr*, 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*.

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 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)).

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.

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</i> de <i>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</i> de <i>Mus musculus</i> , proteína	Longitud: 490 aa
3	<i>TMPRSS2</i> de <i>Homo sapiens</i> , variante 2 de transcrito, 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</i> de <i>Homo sapiens</i> , variante 2 de transcrito, 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>XhoI</i> -LoxP-Casete-loxP- <i>ICeuI</i> - <i>NheI</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>XhoI</i> -loxP- <i>ICeuI</i> - <i>NheI</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 transcrito, 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 transcrito, 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</i> / <i>XhoI</i> -LoxP- <i>hUbi</i> -EM7-Neo-Pm-Cre-loxP- <i>ICeuI</i> - <i>NheI</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</i> / <i>XhoI</i> -LoxP- <i>ICeuI</i> - <i>NheI</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>XhoI</i> -LoxP- <i>hUbi</i> -EM7-Neo-Pm-Cre-loxP- <i>ICeuI</i> - <i>NheI</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: XhoI-LoxP- <i>ICeuI</i> -NheI (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//XhoI//Casete loxP
24	secuencia de unión casete/3' de ratón para humanización de <i>Tmprss2</i>	Casete (loxP)/ <i>ICEUI</i> //NheI//de ratón
25	unión 3' humana/loxP/3' de ratón para humanización de <i>Tmprss2</i>	3' humana//XhoI//((loxP)/ <i>ICEUI</i> //NheI//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-XhoI//casete(<i>loxP</i>)
39	Secuencia de unión casete/5' humana para humanización de <i>Tmprss4</i>	Casete (<i>loxP</i>)/ <i>ICEUI</i> //NheI//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/XhoI//((loxP)/ <i>ICEUI</i> //NheI//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//XhoI//Casete (<i>loxP</i>)
59	secuencia de unión casete/3' de ratón para humanización de <i>Tmprss11d</i>	casete (loxP)/ <i>ICEUI</i> //NheI//3' de ratón
60	unión 3' humana/loxP/3' de ratón para humanización de <i>Tmprss11d</i>	3' humana//XhoI//((loxP)/ <i>ICEUI</i> //NheI//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

- 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.

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, jerbos, 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).

En algunas realizaciones determinadas, un roedor genéticamente modificado de la presente invención se selecciona de un auténtico ratón o rata (familia Muridae), un jerbo, 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.

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 %, 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. 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.

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.

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.

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.

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.

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.

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.

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.

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

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.

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 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 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 %.

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.

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 es una célula madre embrionaria de rata. 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.

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.

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*.

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.

Roedores con Tmprss2 humanizado

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.

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.

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.

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 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.

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 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 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.

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 el mismo 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 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 igual que, una proteína Tmprss2 de roedor homóloga o una forma soluble de la misma en un roedor de control.

Roedores con Tmprss4 humanizado

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.

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.

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 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.

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.

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.

- 25 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.

- 30 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.

- 35 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.

- 40 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.

- 50 En realizaciones específicas, 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.

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.

- 65 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.

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.

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, 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 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.

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, 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 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.

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.

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.

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.

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.

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.

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

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:

- (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).

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 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 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): doi: 10.1038/nprot.2011.338 describen células ES de rata y métodos para producir una rata genéticamente modificada.

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 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 mediante genotipificación del ADN aislado de cortes de la cola utilizando ensayos de pérdida de alelos de roedor y/o ganancia de alelos humanos.

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

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.

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:

- 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
- 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.

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.

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.

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.

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.

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, 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.

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.

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) inmunoquímica, 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 IFN γ , en comparación con los animales no infectados.

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 desee.

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 genotipan 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)	TGGAGGAGCCACCTGATGCCTC	(SEQ ID NO: 27)
	Inverso	GCCTTGCCCTCAATGGAAAC	(SEQ ID NO: 28)
7010D	Directo	GGTTGCACAGCAAGGAAGAAG	(SEQ ID NO: 29)
	Sonda (BHQ)	CCAGGAGTTCCTGTGAGCCTACCC	(SEQ ID NO: 30)
	Inverso	TGGAATGGAAGGAGCTGGAG	(SEQ ID NO: 31)
7010hU	Directo	GTCCCACCTCCTGCAACTG	(SEQ ID NO: 32)
	Sonda (BHQ)	TGAGCCTTCCCATCAGCCTGGG	(SEQ ID NO: 33)
	Inverso	CCACAATGGCACATGGGTCTG	(SEQ ID NO: 34)
7010hTD	Directo	GGTGCTTGCTCCCCAAGA	(SEQ ID NO: 35)
	Sonda (BHQ)	CCTAAAAGGTGTTGTAATGG	(SEQ ID NO: 36)
	Inverso	GGCAATAAAGAAGGAAGACGTTTT	(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 desee.

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.

TABLA 2

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

(continuación)

Nombre	Cebador	Secuencia (5'-3')	SEQ ID NO
7224mTD	Directo	GGAAGCCCTCTCTCGATACTTG	(SEQ ID NO: 48)
	Sonda (BHQ)	TTCTACCTGAGGGCATGCAGC	(SEQ ID NO: 49)
	Inverso	TGGGATGTAGAAGGTTGTCAGA	(SEQ ID NO: 50)
7224hTU	Directo	CTGAGCCTGGAACACACATG	(SEQ ID NO: 51)
	Sonda (BHQ)	TCTGAGAGCCCAGCACTATCGCC	(SEQ ID NO: 52)
	Inverso	GCTGAGGGTCAGGCTTGAG	(SEQ ID NO: 53)
7224hTD	Directo	TCTGCAGGGTAGGGAGAGAAG	(SEQ ID NO: 54)
	Sonda (BHQ)	TGTTTCAGAAAAGGAAGACTCACGT TACA	(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 desee.
- 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*).

Brevemente, se utilizó el clon RP23-95N22 de cromosoma artificial bacteriano (BAC) de ratón que contiene un gen

15 *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

25 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 aproximadamente 143 kb de ADN genómico de ratón que incluye la 5' UTR de *Tmprss11d* de ratón, los exones 1-2

30 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

35 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.

El clon BAC modificado que contiene el gen *Tmprss11d* humanizado, como se describe anteriormente, se utiliza para

45 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

50 *Tmprss11d* endógeno como se describe anteriormente (figuras 3A-3B). 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 *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.

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. 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 genotipan 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.

TABLA 3

Nombre	Cebador	Secuencia (5'-3')	SEQ ID NO
7226mTU	Directo	TCCTCTCCAGACAAGAAAGCT	(SEQ ID NO: 61)
	Sonda (BHQ)	TCATAGCAGCTTTCAAATCCTAAACGT TGA	(SEQ ID NO: 62)
	Inverso	TCGTGTGTAGCTGGTGAAGT	(SEQ ID NO: 63)
7226mTD	Directo	CATGCGATCACAGGAGGAGATC	(SEQ ID NO: 64)
	Sonda (BHQ)	AATTGGGCCCCGAAGCCAGATGC	(SEQ ID NO: 65)
	Inverso	CGGAAGGCTTCTGTGACTTC	(SEQ ID NO: 66)
7226hTU	Directo	GTCTCCCACTTCTGACATAATGAAC	(SEQ ID NO: 67)
	Sonda (BHQ)	CCCAGTGTTAACCCTACATCTGGTTCC	(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

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.

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).

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)
<i>TMPRSS4</i> TS; infectado con H1_PR34	10	20 (2/10)
Humln <i>TMPRSS4</i> ; infectado con H1_PR34	8	25 (2/8)
<i>TMPRSS4</i> 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> REGENERON PHARMACEUTICALS, INC.

<120> ROEDORES QUE TIENEN UN GEN TMPRSS HUMANIZADO

<130> 33093PCT (10234WO01)

<150> 62/301.023

<151> 29/02/2016

<160> 72

<170> Patentln versión 3.5

<210> 1

<211> 3175

<212> ADN

<213> *Mus musculus*

<400> 1

ES 2 886 958 T3

gcctttcctg gccgttcctt ccttctggcc gaggtgcctg cgtttagggg tgtcacccctg	60
gctcccggga cgccgcctcc ggagatttaa gcgagaactg gagtaggtcg tgtacttgga	120
gcggacgagg aagccaagag ctcgacaga ggcggagagg ggcgggaagc gcaacaggtc	180
acctggagga agccccatac tgacctctc atgctgctga cacaggcagg atggcattga	240
actcagggtc acctccagga atcggacctt gctatgagaa ccacgggtat cagtctgagc	300
acatctgtcc tccgagacca ccagtggctc ccaatggcta caacttgtat ccagcccagt	360
actaccatc tccagtgcct cagtatgctc cgaggattac aacgcaagcc tcaacatctg	420
tcatccacac acatcccaag tcctcaggag cactgtgcac ctcaaagtct aagaaatcgc	480
tgtgttttagc cctcgccctg ggactgtcc tcacgggagc tgctgtggct gctgtcttgc	540
tttgagggtt ctgggacagc aactgttcta cgtctgagat ggagtgtggg tcttcaggca	600
catgcatcag ctcttctctc tgggtgtgacg gggtagcaca ttgtcccaac ggagaagatg	660
agaaccgttg tgttcgtctc tacggacaaa gcttcacctt ccagggtttac tcatctcaga	720
ggaaagcctg gtatcccgtg tgccaggatg attggagtga gagctacggg agagcagcat	780
gtaaagacat gggatacaag aacaattttt attctagcca agggatacca gaccagagcg	840
gggcaacgag ctttatgaag ctgaatgtga gctcaggcaa cgttgacctc tataaaaaac	900
tctaccacag tgactcatgt tcatcccgca tgggtggtttc tttgcgctgt atagaatgcg	960
gggttcgctc agtgaaacgc cagagcagga ttgtgggtgg attgaatgcc tcaccaggag	1020
actggccctg gcaggtcagc ctgcacgtcc aaggcgcca cgtctgcgga ggctccatca	1080
tcacccccga gtggattgtg acggccgccc actgtgtgga agaaccctc agcagccga	1140
ggtactggac ggcatttgcg ggaattctga gacagtctct catgttctat ggaagtagac	1200
accaggtaga aaaagtaatt tcccatcaa attacgactc taagaccaag aataacgaca	1260

ES 2 886 958 T3

ttgctctcat gaagctgcag acaccttttg cttttaatga tctagtgaag ccagtgtgtc	1320
tgccgaaccc aggcattgatg ctagacctag accaggaatg ctggatttcg ggggtggggg	1380
ccacctatga gaaagggaag acctcggacg tgttgaatgc tgccatggta cccttgatcg	1440
agccctccaa atgtaatagt aaatacatat acaacaacct aatcacacca gccatgatct	1500
gtgccggcctt cctccagggg tctgtcgact cttgccaggg agacagtgga gggccgctgg	1560
ttactttgaa gaatgggatc tgggtggctga ttggggacac gagctggggc tcgggctgtg	1620
ccaaggcact cagacctgga gtatacggga acgtgacggg atttacagat tggatctacc	1680
agcaaagtag ggcgaacagc taatccacgt ggctttgtcc cagacttctt ttgtcttcaa	1740
caaccttctg caagaaaacc aagggcctga attttaactt cctgtgcaca atgtaccttt	1800
tgagatgatt cgaagggcct ttcactttta ttaaacagtg acttgtttga ctgtgtctcc	1860
tggtcctgtg agggccttcag tgccccaccc ctggggcact tctgcagctc ccaccagaat	1920
ggatgaccag attctgtttg gtttgggcac atagggccaa aggcagagga ggggtggcact	1980
ctcatgtttg aacttctttt gggctcatgc tcaggccttt tttggatcac taaggactat	2040
gacctctgag taacctgatg acctgagaaa gagtaaggag gccaggcagg gccttggggc	2100
caggaacagc taccttgaga gtgagagcta cccattgcct gtggcctaaa tctgctgtgc	2160
aggttgggct ggtcatactg tcatgatttc attaacagcc tgggtgaaca tggctgggag	2220
taaagggtt gctctcctgc atgttgacat gacggccctt tccaagggtg atggaggctt	2280
tccaagcta agggcctagg cagatctctc agagcaagaa gctaataccg gcatgtccct	2340
tgggtgagct ctacatggtg ttattcagtc tggttcttgg ctccccacta ctgtttctct	2400
cagcctctca gagcctgaaa cttacctctt agctttggct acaggcatgg cctagtacct	2460
gatggagcct gtatagctca gctaatacaa tggaggctca ggtccatcag aatcaggac	2520
ttgtgatttc agtcaccttg cttctgggtt gtgtttcttc tcttactacc tactgcacc	2580
tggacactag agtggatgaa tgtctggagt tcacctgcat ttggactgtg tgattgtgcc	2640
tcagacacta gacctcttcc agatgggttag gttgttctgt agactggcaa tgagattaga	2700
agttcctagc ttcagataaa gatgaaagag aggagatcat tgtcttctgt cttcttctgg	2760
ccctgggttt ataccaggaa agccatgcca gaattaccaa atatgaagta tgaatgtctt	2820
accacgggtg aggtctctgcc tccttctctc tgcttggttc ttcagaaggc agtgaatggg	2880
tcataactgg gactccatct ttgctgggga aagtctccca cctagggaat ggttaccact	2940
ccatgtaaag aaaactccct catgcgtcct ctgggacctt cttagatgct gtaaggtacc	3000
tacatacaga ctaaatgtgc aagcaccttg aagtgtgaga acctgtcccc tccttagctc	3060
tccttgtctt tgctgttggg tgggtatttc ctgctttgtg tctgttctga gctgtgagat	3120
tccactgtga aatatatgaa taaagtatat aattctttta aaaaaaaaaa aaaaa	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

ttgataacag caagatggct ttgaactcag ggtcaccacc agctattgga ctttactatg 180

aaaaccatgg ataccaaccg gaaaaccctt atcccgcaca gccactgtg gtccccactg 240

tctacgaggt gcatccggct cagtactacc cgtccccgt gccccagtac gcccgcaggg 300

tcctgacgca ggcttccaac ccgctcgtct gcacgcagcc caaatccca tccgggacag 360

tgtgcacctc aaagactaag aaagcactgt gcatcacctt gaccctgggg accttcctcg 420

tgggagctgc gctggccgct ggctactctt ggaagtcat gggcagcaag tgctccaact 480

ctgggataga gtgcgactcc tcaggtagct gcatcaacct ctctaactgg tgtgatggcg 540

tgtcacactg ccccgccggg gaggacgaga atcgggtgtg tcgcctctac ggaccaaact 600

tcacacctca ggtgtactca tctcagagga agtcctggca ccctgtgtgc caagacgact 660

ggaacgagaa ctacggggcg gcggcctgca gggacatggg ctataagaat aatttttact 720

ctagccaagg aatagtggat gacagcggat ccaccagctt tatgaaactg aacacaagtg 780

ccggcaatgt cgatatctat aaaaaactgt accacagtga tgctgttct tcaaaagcag 840

tggtttcttt acgctgtata gcctgcgggg tcaacttgaa ctcaagccgc cagagcagga 900

ttgtgggcgg cgagagcgcg ctcccggggg cctggccctg gcaggtcagc ctgcacgtcc 960

agaacgtcca cgtgtgcgga ggctccatca tcacccccga gtggatcgtg acagccgccc 1020

actgcgtgga aaaacctctt aacaatccat ggcattggac ggcatttgcg gggattttga 1080

gacaatcttt catgttctat ggagccggat accaagtaga aaaagtgatt tctcatccaa 1140

attatgactc caagaccaag aacaatgaca ttgcgctgat gaagctgcag aagcctctga 1200

ctttcaacga cctagtgaac ccagtgtgtc tgcccaacct aggcagtgat ctgcagccag 1260

aacagctctg ctggatttcc ggggtggggg ccaccgagga gaaaggggaa acctcagaag 1320

tgctgaacgc tgccaaggtg cttctcattg agacacagag atgcaacagc agatatgtct 1380

atgacaacct gatcacacca gccatgatct gtgccggctt cctgcagggg aacgtcgatt 1440

cttgccaggg tgacagtgga gggcctctgg tcacttcgaa gaacaatatc tgggtggctga 1500

taggggatac aagctggggt tctggctgtg ccaaagctta cagaccagga gtgtacggga	1560
atgtgatggg attcacggac tggatttatc gacaaatgag ggcagacggc taatccacat	1620
ggctcttcgtc cttgacgtcg ttttacaaga aaacaatggg gctgggtttg cttccccgtg	1680
catgatttac tcttagagat gattcagagg tcacttcatt tttattaaac agtgaacttg	1740
tctggctttg gcactctctg ccattctgtg caggctgcag tggctcccct gccagcctg	1800
ctctccctaa ccccttgctc gcaaggggtg atggccggct ggttggtggc actggcggtc	1860
aagtgtggag gagaggggtg gaggctgccc cattgagatc ttcctgctga gtcctttcca	1920
ggggccaatt ttggatgagc atggagctgt cacctctcag ctgctggatg acttgagatg	1980
aaaaaggaga gacatgaaa gggagacagc cagggtggcac ctgcagcggc tgccctctgg	2040
ggccacttgg tagtgtcccc agcctacctc tccacaaggg gattttgctg atgggttctt	2100
agagccttag cagccctgga tgggtggccag aaataaaggg accagccctt catgggtggt	2160
gacgtggtag tcacttgtaa ggggaacaga aacatttttg ttcttatggg gtgagaatat	2220
agacagtgcc cttggtgcga ggggaagcaat tgaaaaggaa cttgccctga gcactcctgg	2280
tgcaggtctc cacctgcaca ttgggtgggg ctccctgggag ggagactcag ccttcctcct	2340
catcctccct gaccctgctc ctagcaccct ggagagtgc catgccctt ggtcctggca	2400
gggcgccaaag tctggcacca tgttggcctc ttcaggcctg ctagtactg gaaattgagg	2460
tccatggggg aaatcaagga tgctcagttt aaggtacact gtttccatgt tatgtttcta	2520
cacattgcta cctcagtgt cctggaaact tagcttttga tgtctccaag tagtccacct	2580
tcatttaact ctttgaaact gtatcatctt tgccaagtaa gagtgggtggc ctatttcagc	2640
tgctttgaca aaatgactgg ctccctgactt aacgttctat aaatgaatgt gctgaagcaa	2700
agtgcccatg gtggcggcga agaagagaaa gatgtgtttt gttttggact ctctgtggtc	2760
ccttccaatg ctgtgggttt ccaaccaggg gaagggtccc ttttgcattg ccaagtgcc	2820
taaccatgag cactactcta ccatggttct gcctcctggc caagcaggct ggtttgcaag	2880
aatgaaatga atgattctac agctaggact taaccttgaa atggaaagtc atgcaatccc	2940
atgtgcagga tctgtctgtg cacatgcctc tgtagagagc agcattccca gggaccttgg	3000
aaacagttgg cactgtaagg tgcttgctcc ccaagacaca tcctaaaagg tgttgtaatg	3060
gtgaaaacgt cttccttctt tattgcccct tcttatttat gtgaacaact gtttgtcttt	3120
ttttgtatct tttttaaact gtaaagttca attgtgaaaa tgaatatcat gcaaataaat	3180
tatgcaattt ttttttcaaa gtaaaaaaaaa 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	75	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	155	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	235	240

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

<400> 5

10

```
gcagagtcta agaaatcgct gtgttttagcc ctgcacctgg gcaactgtcct cacgggagct      60
gctgtggctg ctgtcttgct ttggaagttc agtaagtgca gggagcctcg atcccaccat      120
gtgctcctgc agtccccagt gctctgagcc agacctgct ctctgggcta ttgagacctc      180
tggaggccct ccgtgagggt cctctcttac ataacgaggc tgtctctctt cccttctctt      240
gttttagctat gagattgaca catcatgggg aaagcattta gaatgtaccc agtgctttgg      300
ggtgcttggg gccaccacagc actgtgagca caggttcttc taccttgggg ccacaccacag      360
ttacctgtat ctactgcac agcagtggct gttggggacc aggcccaccc ctccatgtcc      420
cacctcctgc aactgcagcc tgagccttcc catcagcctg gggtggtgca gacctatgtg      480
ccattgtgga tccttcaagt tacctgtgtg gcagagagga cgtgtgagtg ccgtccaaac      540
ccaaacactg agagggtcct tcccattgcc cccacggaag taagggtgcc cagtgctaata      600
tccacttata cttgctggtg gcaaggacac ttctcctcct tattaaagtg ggggattggc      660
tgggtgaggt ggctcacgcc tgttatccca gcaatttaag aggccaaggc aggtggacca      720
cctgaggtca ggagtttgag accacaagcc tggccaacat gttgaaactc catctctact      780
aaaaatacaa aaattagtca ggcgtggtgg cgtgcacctg taatcccagc tacttaggag      840
gctggggcag gaggatcact tgaaccacag agttggaggt tgcaagtgagc caagattgtg      900
cccctgcact ccagcctggg tgacagaatg agacttcatc tcaaaaacaa acaaaaacaa      960
aacacagtgg ggccaggagt tggaggctgc agcgagctac agtaatgcca cgggtgttct      1020
cactccatga ggctcattgc gtttctcagc ctgaagggca cctctcttct gtttctctctg      1080
caagtgggca gcaagtgtc caactctggg atagagtgcg actcctcagg tacctgcac      1140
aaccctcta actggtgtga tggcgtgtca cactgccccg gcggggagga cgagaatcgg      1200
tgtggtgagt cagccttgac cttgggaagg gactcctctg ctcaccttgg agacagcagc      1260
cgggtccagg ggcctttggg tgactgggcc tggcgtgcgt ccagtacgct gacacatgat      1320
gtcattgaat ccctgctcca ggctgagccc tggggctcag agagggtgtg tttccggccc      1380
aacctcacc agcaggtggg agatgacagg gccaccgagg actgtgtcat tggaaccaca      1440
cgtgctctga actgccacag gaagtcagtt aagatgagca aactgtttat aaagttggag      1500
atgcaggcta ggaacggtgg ctcatgcctg taatcccagc actttgggag gccgaggcag      1560
atggatcacc tgagggtcagg agtttgagac cagcctgacc aatatggtga aaccttatct      1620
```


ccactaaaaa	tacaaaaaatt	agccaagcgc	ggtggcgggt	gcctgtaatt	ccagctattc	1680
aggaggctga	ggcaggagaa	tcacttgaac	ctgggaggcg	gaggttgcag	tgagctgaga	1740
tcacgccact	gcattccagc	ctgggagaca	gagctggctc	aaaaaataaa	ttaattaatt	1800
aaaaacaaaa	ttggagatgc	actatgttat	tttcaaaaca	agctgccttt	aaagatctat	1860
ctgttgtcac	agggtgggct	catctgtttc	attttatttt	ctgtggttta	tctattttatt	1920
cattttaatg	aactaggaag	cattgctcct	atttatggca	taccacatga	tgtttgata	1980
cgtgtatgcc	tgtggcatgg	ctaagtcaag	ctagaacatg	ggccttacct	catatacgtg	2040
tcttattaag	aacacataaa	acctactctt	gtagtgattt	tcaaatatgc	aacatatagt	2100
ttattaactg	cagtcactat	gatgtacaat	agattgctcg	aacttattcc	tctgtctaa	2160
ctaagatttt	gtgacctctg	accaacatct	ccccagtgtt	gtcaccccc	gccccagcc	2220
tctgatagct	gcctttctac	tctctgcttc	tgtgagtttg	atgtttatac	attccacatg	2280
taagtggcct	catgcagtgt	ttctgtctct	gtgtctggct	tgttcaacta	gcgtaatgtc	2340
ctccagcttc	atctatgttg	ttggaaatga	caggatttcc	ttctttcttg	tggtgaata	2400
gtattgcctt	gtgcatatac	accacatttt	ctttatccct	tcattcaactg	atggactcct	2460
aggttgatgt	catgtcttgg	ctgttgtgaa	aaatgccgca	gtgagcgtgg	gcgtgcaggt	2520
ccctcttcaa	cacacggatt	tcctttcctt	tggaataaaa	cccagcagtg	agattgctgg	2580
atcacatggc	agttctgttt	ctcacctttt	gaggaaactc	catactgttt	tccataatgg	2640
ctgtagcaac	ttccactccc	acccccacgg	tgcaaagtct	ccatttctct	tctacaacct	2700
caccaactcc	tgttattttc	catctttctg	atagtagcca	tttgaagagg	tatgagatga	2760
tacctcattg	tggttttcat	ttgcattttt	atgtgtattt	ttcatgaatt	tttgaggggtg	2820
atttcaaggg	tagttagtga	ctcgaacagg	gaaacgatcc	tgagtatgag	ggttgtgcta	2880
atcatcccc	tcctgccagc	tgcgtacgga	atggggctct	gcagatggca	gggagctggc	2940
tcgtttctct	ttaagagctg	ccttttactt	ttcttctctt	tcctttaaaa	cttatttctt	3000
ggccggacgc	agtggctcat	gcctgtaatc	ccagcacttt	gggaggccga	ggtgggcgga	3060
tcacgaggtc	aggaattcca	gaccagcctg	gccaacatgg	tgaaaccccg	tctctactaa	3120
aaatacaaaa	attagccaga	cgtgggtggtg	cgggcctata	gtcccagcta	ctcgggagggc	3180
tgaggcagga	gaatcacttg	aacctgggag	gaggggggtt	cagtgagccg	agattgcgcc	3240
actgcactcc	agcctgggcg	acagagccag	actccatctc	aaaaaaciaa	aaaaagttat	3300
ttcccaagca	cagccatgta	ttccaggctt	gtggatcagc	gttgggtggtg	gtgtgtgctc	3360
tcatatctta	gttccagcta	agcacactct	gacatgttta	cactagaacc	atltgttttt	3420
tctagaaata	gaaatttcag	aattgtagag	tcagaggact	taccagaaat	ctcttaggta	3480

gttctcctcc	cctccctcaa	gtgcagtcct	aacctcctgg	agttttctgt	agaaaccaca	3540
agcctcagag	ctggccgaga	attctagcca	aagatttttc	catgccaaag	taatccccc	3600
tctcctaagg	gccatccttg	gtggggactg	gtttcctggt	aagccctcgc	tgtcagtcct	3660
ggctgtggaa	tttcctggtg	aggagcactg	gcccgtggag	ctcggccctc	gtgccggcct	3720
tgagcaggcc	caagtgttcc	gtgttcttga	tacctttcct	ccagcacagt	cttgcttccc	3780
agaaaaaggt	ttgcacttga	aatgatgca	tttgctgatt	aaacatagtt	cttttgcttt	3840
atttggtttc	taaaataaag	tgggagtttt	tgagattgag	taacgtgagg	ttaagatagc	3900
acgtggaatg	gctttttctt	ttctttctat	tttttttttt	tttttctgg	agacagggtt	3960
tcactctggt	gcccaggctg	gagtgcagag	gcatgaccat	ggctcactgc	aacttcgatg	4020
tcctgggggt	aagcgatccc	ccagcctcag	cccccaagt	ggctgggact	acaggtgctc	4080
gccaccacac	ctggctaatt	tttgtatttt	ttgtagaaaa	tgggtttcat	caatgttgtc	4140
cagactggtc	tcgaactcct	gacctcaagc	aattctcctg	cctcagcctc	ccagactgct	4200
gggattacag	gcgtgaacta	ccacgcctgg	cctggaatgg	cttttgatgt	tctcctatgt	4260
gcacatgtgg	gtgaataaac	accaacaaag	tccttatggt	acctgaagag	ttgctctctt	4320
cttaatat	aagtcgtatt	tatttaata	ctttaatagt	tgtacactat	taaagtatta	4380
ttaggtcaaa	atcaaggaag	tacaaaagg	tatgctgtga	aaaatctctt	cttccttgct	4440
ctgcttactt	acctaccccg	catccccca	tacccccag	acacacacac	acacacacac	4500
acacacacac	acacacgcat	cactcccata	catgccacc	tgtttaccag	ccaatcacat	4560
ttcttggggc	aactcatctg	agttgcttct	ctttccagag	agtttttgca	taaagaagca	4620
caggatattc	tgcgttacca	tgacctatt	tcccagtggt	tcctagccag	ttgactctcc	4680
tgactggat	accatcctgg	acagcattcc	ttagggaaat	gagccccctg	ttttttccca	4740
ccatggcaca	gttgtcctt	tgcatggacg	caccattatt	gcccctgtct	cttcttggtg	4800
gaccttaagg	ttttctccat	ccttttgctg	taacacacac	tgctccaagt	gtgtgagcat	4860
atcagtagga	aacgcttcca	ggagtagaac	tgctaggtca	gagggcgtgt	ggatctgtaa	4920
cctgacagac	ctagaccggc	ttcagtttgg	ttttatccag	tttccatatt	gattattcat	4980
ataaaaggaa	acagacaaac	ataacgctgt	gcatgtattc	tctcttagac	cagaacaggc	5040
atagggtgca	cttttaattt	gtccatttcg	tagagtagaa	attgtttttg	ctgaaatgaa	5100
caccttagga	tgctgaagaa	tatgaccctg	cccatggaaa	acattcaaaa	atgtgtgtag	5160
cgctttcttc	ccaagggtgt	gtgtgcgcat	attttaacac	taattcactt	tctacttccg	5220
ttgctatcct	ttctgtgagt	ctttctcaga	atctcagaaa	agaaactaaa	ttgttctctc	5280
tagttatcaa	tgctgtactc	tatacctgga	atttgctaaa	agggcagatt	ttaagtattc	5340
tcaccacaga	aaagagaaaa	gaaaatggta	attatgtgac	gtgggtggaca	tgtaactag	5400

ctttattatg	gtgagcattt	cacagcggat	atccagtcac	cacgctgtac	acattaaaca	5460
tgtacaattg	ggtttttttg	agacaaggtc	tccttctgtc	acccagtcctg	gagtgacgtg	5520
gctcagtcac	ggctcattgc	agcctcgacc	tcctgggctc	aatccatcct	tccccctcag	5580
cctcctgaaa	agctggggcc	acaggcatgt	accatcatgc	caggctaata	catatatatt	5640
tatatatttt	ggcggagatg	gggttggtct	cgaactctgg	gctcaagtga	tcctcccgcc	5700
ttgcccttcc	aaagtgcctg	gattacaggc	atgaaccaca	gcaccaggcc	tacatgtaaa	5760
atttttattt	gtcaactata	ctttgacaaa	gctgagaaaa	aaaatcctaa	tatttaaaaa	5820
aaaaaaaaaa	aggactagct	tgagaccttt	tcagctcttc	tggttatca	gctgccgtct	5880
cttcggggtg	cagatagctg	gaagggaaag	aaaatcccta	aaattacca	caagccaaga	5940
atgaagtgtc	tccttttgag	ccacagtggc	agttttgttt	ttaatcatag	aagtgtattt	6000
tgagccgggt	gtgctggctc	acgcctgtaa	tccccgcact	ttgggaggcc	gaggtggggg	6060
gcggaggggg	tggggatcgc	ctgaggtcag	gagttcgaga	ccagcctgac	caacatggag	6120
aaaccccgtc	tctactaaaa	atacaaaatt	agccggcgtg	gtggtgcatg	cctgtaatcc	6180
cagctactca	tgaggctgag	tcaggagaat	ctcttgaacc	caggaggtgg	aggttgcggt	6240
gagctgagat	catgccattg	cactccagcc	tggaacaag	aaaaaaaaag	aagaagaaga	6300
agaagtgtat	tcatttcagt	tacttttaaa	aaagtgaaca	gactttatat	tttagagcgg	6360
ttttaggttt	acagaaaatg	aaacagacag	ggcagcgagc	tccttgtact	cctccccagc	6420
acacagttgc	cctgttatga	acatcccaca	tcagtgtgtg	gcgttcatta	acaccgatga	6480
acctgatgca	tacattatga	tgaactgaag	tcctggactt	caccctttct	cttgtacagt	6540
tctgtgggat	ttgacaaatg	cataatgctg	tacagccaca	atgatagtat	cgtccagagt	6600
agttctcctg	ccttaaaaacc	tcttttgctg	cacctgtttc	tctctcccca	ctcaccaccag	6660
ctatctgatc	ttcttagtgc	ctccgaagtt	ttggtctttt	caggatgttg	tagcgttgga	6720
atcatggagt	atgtagcctt	caccacatac	accttccttc	actttgttgg	cttcctttac	6780
ttagtaatat	gcattcaagt	ttcctccatg	ccttttcatg	gcttgatagc	tcatttcctt	6840
ttagcaccaa	ataatattcc	gttgtccaga	tgtagcacia	tgtttatcca	ttcatgtaac	6900
ctgtgaccga	ctcacagata	ggatgtggaa	tcactcacca	cagaggcatt	agacaataat	6960
cagacccaag	tcatttcatg	ggggaacaag	cccacaggta	ccagactgtc	cagtgaagtca	7020
gggccactcg	taggaagtaa	gaagagaggc	tagagcatag	ccaggtcctc	actttatact	7080
ttaagcccat	gtgtattttct	cccaaaccac	acagcattgt	ttccatgctt	tcagctttgc	7140
atgaataacg	tgatacttga	acgcatacatt	tatcacttgc	tctctttccc	acagcgctgt	7200
tttcaagctt	cttcctgttc	atgatgctct	gcttaaccct	taagctgcat	gggattctgt	7260

tctgtgaata	cgcccccccc	atgtattatc	ctgcccagca	aaaagtcccc	aaaactctgg	7320
atggtgggta	cctctaggga	gggagagaag	agattgggaa	tagggagcga	cttcaacggt	7380
gtttgtaatg	ttttgtttct	ttaaataaaa	gagctgagat	catttcagca	gaatgttgat	7440
ttagagtctc	ctggacaatt	tgttgctcaa	agtgtctctc	taaagagcac	tttaaaaaaa	7500
aaaacctttt	atcttattat	ttatttattt	atttattgag	acggagtttt	gctctgtcac	7560
ccaggctgga	gtggagtggg	gtgatctcag	ctcactgcaa	cctttacctc	ctgggttcaa	7620
gcaattcccc	tgcctcagcc	tccaagtag	gtgggattac	agatgcgtgc	caccacactt	7680
ggctaatttt	tgcatttttag	tagagatcgg	tttctccatg	ttggccaggc	tgatctcaaa	7740
cgctgacct	caggtgatct	gcccgccttg	gcctcccaaa	gtgctggtat	tacaggcgtg	7800
agctaccatg	cctggcttat	cttatatatt	tttaaaaaca	gcttattgag	atctaattta	7860
tgtaccataa	aattcaagta	tataattcag	tgcttttata	tataaaacat	atatatgaaa	7920
tagcttattg	agatataa	ttttatataa	aacagcttat	tgatatgtaa	tgtatgtacc	7980
ataaaattta	aatatataat	tactggctt	ttatatattc	acgaatatgt	gcaactatca	8040
ccacagtcaa	ttttagcata	ttttcatcag	ctcataaaga	aacccaagc	ccttgaacta	8100
tcaccccata	tccctcctcc	cagcccgtcc	ctcctactca	taagcaacca	ctaactact	8160
tagtgtctat	agatttccta	ctctaggcat	tccatgtgag	cgggatcatg	caatacgtgg	8220
gctcacacaa	tataagtggc	attccatgtg	agtcggctca	tgcatgatgt	ccggtcctt	8280
tactgagca	taaggtcttc	agcactcatc	caggttgacg	cctgtgtctg	aatttcattc	8340
cctcttctgg	ctgaatcgta	ttccattgtg	tatcttggac	atctctatt	ctgtcacccc	8400
agccgttggg	gggctgttgg	agtgttttcg	cctttcagct	gttttaagag	ggttgacgtg	8460
aacatttgta	caagtgttgg	acccaatgcc	tgttttcaat	tctcttgtgt	agagagcact	8520
ttttagcaga	aaaagaatag	atttgtggcc	tccctttgtg	tgcggtcagt	gccttgagaa	8580
gagtgaactg	tgctgccacc	tccggagccg	tggagagcgc	ggggcttggg	tagcagctag	8640
gacgatacaa	gttgggacaa	ggccaggtgc	aatggctcac	gcctgtaatt	ccaacacttt	8700
gggagaccga	ggcaggggga	tcacctgagg	tcaggagttc	aagaccagcc	tggccaacat	8760
ggtgaaaccc	catctcta	aaaacagaaa	aattaactgg	acgggggtgg	ggacgcctgt	8820
aatcccagct	actcgggagg	ctgaggcagg	agaatcactt	gaacctggga	ggcggaggct	8880
gcagtgagtg	gagatcagac	cactgcactt	cagcctaggt	gacagagcga	gactccgtct	8940
caaaaaaaag	aaaaaaaag	aaagaaactc	atggataatc	ctccctctcg	tgagttcgc	9000
ctctacggac	caaacttcat	ccttcagggtg	tactcatctc	agaggaagtc	ctggcaccct	9060
gtgtgccaag	acgactggaa	cgagaactac	gggcgggcgg	cctgcaggga	catgggctat	9120
aagtgagtat	ggggcagcac	ccgccaggtg	acagtaacag	acagcagaaa	cacgagaaga	9180

ccctctctct	gcctccctgt	gaaagcacccg	gcacatgagt	gctggggaca	attgtcacct	9240
tccaaaagct	gagccctata	accagcaggt	ggaatttgct	ctgctagggc	tgtgccccagc	9300
acacagacct	tggtcactg	ccaccttgcc	ctgcctcctc	cttggcctct	atagactcct	9360
ggttgctcgg	gagtgccag	tgctgtggtc	atctggctcag	aggggtaggc	tgagggcggt	9420
aggtgcctct	ttttccaagg	tgctctcag	ccagggtcca	ttcacctccc	tggttagagg	9480
ttggaccaga	acagctggcg	aggagggttg	ggctggggag	agcagcagag	acaaatcctg	9540
tgccagtttc	acttcattcg	ggagccatgg	aagccttttg	agctggggag	agaatcaatc	9600
aatcagactg	atacttaaaa	aatgtcattc	ctgctcgtag	ctctgaggga	aggtgggaag	9660
gcttaacagg	gtgtgtgtcg	cctgacagtg	attcctaacg	gggggtggggc	ggtggttacc	9720
atttaccagc	actgcctggg	gagatgcggc	agccctcagg	catcggggga	gaggggtggt	9780
ggatgctact	gccactttgt	tttccatggg	aggggtcccca	ggtgatttct	atgcaacttt	9840
agggatttca	atatgccagt	tttcagaatg	aattaccact	cggtgagaaa	gttggcatct	9900
tagctagtca	ctgtgacatc	cctaaacagc	aggggtgaat	tacacagcaa	agccccccca	9960
tcacagtcca	ggaacctggt	ggaattgata	actggggcca	tgtaacatc	tgtacctttt	10020
attagattaa	atgtgtgtat	gattatacaa	tcctatgtcc	ttctcatagt	ttcttgatcc	10080
taacctggat	aagaaacacg	accaatgaag	gaattttgct	tgacacttta	gggttattga	10140
atcgaaaaat	cgttacaata	ttctagcact	tggttagaac	gtgtgatttt	ttttcctaaa	10200
tgctaagggt	tttccctctt	attctgaatg	tcgtatgagc	ggtattatga	catagtatag	10260
gatttggtgt	tgcttatgcc	ttaaccatta	tcacaaataa	ggttttcttt	tttaggaata	10320
atctttactc	tagccaagga	atagtggatg	acagcggatc	caccagcttt	atgaaactga	10380
acacaagtgc	cggcaatgtc	gatatctata	aaaaactgta	ccacaggat	gcagcaattt	10440
cttcttgaaa	aattttggaa	tgaaatcaac	taggagacac	catggggaat	cgttgtcctg	10500
agtctgattt	ctctgagctg	caatactcgg	tctggatggg	ttttgcattg	ggaggagatt	10560
agagtctgac	caggcctggt	tactctaagc	agcccttggt	ttattcatag	gaagtggctg	10620
aggtttctct	gctatttcat	tttcagcctc	taccgtctgc	ccttggttgg	agcggctcac	10680
acttgcaaca	tcgacattca	actctattta	gttttctttc	ctcttcagac	atttagaggt	10740
gtacctattt	tgtcagggcg	tggttctagg	aatccaagat	aatgtctcag	tgtcccagcc	10800
agggtgaccg	gctcattcca	gtttgccagg	gacttcactg	gcttgagcaa	gggaagtcct	10860
gctccattcc	aggcagctgg	gctggctggt	cccgttagcc	ccaaccccg	gacagcagtg	10920
ccagaggggtg	ctctgtgagg	gatgggcagc	attctggcgg	cctgggaatg	agttgtgggtg	10980
tttccagggg	gtagaagtgg	gtacaagcca	caggtcacat	gatgagtggc	tgacctggct	11040

ES 2 886 958 T3

gggagggcag	aagaggggat	ggacttaggc	tcttcctttt	gctttgcaca	tatttaggat	11100
gtttgcagac	ttgctatgat	tgttgctgtt	atgtgttttc	tgatgtgaaa	gatacacagt	11160
gtcctttgcc	catgagctct	ccttgccctc	caggtcccca	gggcttatgc	ctggtgtcta	11220
ggcatcacct	ccctgcctgc	caggtgccag	gtgctgcatt	tcgggggagg	atgaactaat	11280
caccccgccg	cacctttcct	ctgagtggga	gcctggggca	ggtttgcatt	cctggaggcc	11340
gctggtggag	gggtctgggg	gcctgacttc	caactgcagc	tgctgtcctg	gggaatgtgg	11400
cagggcaagc	ccagtgggga	gggctgtgca	cggccaggtg	cacccatcaa	aacagcaggg	11460
ctgcggtttg	tccctgtgga	gaagctaaac	acagctgcct	gggcactttg	taaatgctga	11520
gtggttcttt	gtctttctgg	gttacacacg	gaatcagggg	gccaagtcca	gccgggcagg	11580
gacgggggga	ggggaggagg	tgctgccgtc	ccttgccaag	agccttgggg	actcacaagg	11640
aggctggagg	gcttggaaga	aagaagagaa	ggccattgtc	tggtaggctc	tattctatct	11700
cggtggtggt	ggtgggggga	ggcgcaactc	ttttcctctt	tctgtgcagc	agttgcctt	11760
tgatgcctga	gttcttggtt	tgttttctgt	cgggcttctg	tgaataacca	catgtgccct	11820
ggcgctgtga	ccacacaggg	ctatccctac	cgaccttagg	attcttagga	aatgtcttct	11880
cttaaagggg	acatgtcttc	acttggccgt	gtcagtgcc	cagagccaga	gtccacctgg	11940
aatgcacctg	tagtcaactga	gaacccgggg	ggtgtgcctt	agtaagaagg	tgtcaggaag	12000
gacctattat	tgtagggcct	gggctcctgc	aagggtggtt	gggggtggtt	ggaggaagca	12060
gagatttgct	ctggattgga	tgctgtcagg	aagcaggggt	aattctgtga	ggctgcttta	12120
ttatTTTTTT	tctaggagga	ggttggaatg	aggctaggct	aaagctgtga	ttggtaaaga	12180
aacgtccgtc	gctcaagtta	gccaggacag	gaggagacat	cagatcgtga	ttttgtggtt	12240
gtgagcacia	ggttcctgtt	ctgtctgttc	agacatcatt	tcggaggagg	ctccttgtgt	12300
cttgcccat	ctcaggcatg	gaggggccta	gtccgatatt	gacgctcagt	gaaataattc	12360
aggttccgca	gagcacacgg	cccagctatc	agggcggggc	agctctgcat	gccagggggc	12420
gcgtcttccc	ttctcagcat	agcctgggaa	attcactgca	ggacaaaatg	catcagttac	12480
ttcctcttca	tccataacct	gggatgtttg	actcccaa	gagtaactct	tacgtttctt	12540
ctaactctag	ggaaactatt	ggttatattg	ctttcaacac	tacaaattta	aagcagttat	12600
aggagcccag	aggtttccaa	atggcttcct	taaaaattag	aagatgattt	taaattccaa	12660
gaggaaaaac	aaaactagca	ttattgtata	cttaccctca	caaccgtcct	aggagctggt	12720
acaattttta	gagagggtta	gtaacttgcc	caaggtcaca	ctgtggggat	gtgagccgcg	12780
taccttggtt	cagtgtctgg	tctttgccac	tgtccctata	tggatttact	taccttattg	12840
gagttgtaac	tagcagaccc	ttctatgtct	cagaagacag	gagaggggaa	atcggaagaa	12900
atgactgatt	tctaagcatg	tgagaggcag	gtgactccgc	actatcgtga	ccagaatttc	12960

ccctgttctt	tttgcagtga	tgctgttct	tcaaaagcag	tggtttcttt	acgctgtata	13020
ggtaagttca	tctggagtcc	cccttttgat	acttctaact	aggaaaagct	ctctactttc	13080
agaacagtag	tccctgtgtc	tctgggggcg	tgggaggga	gaaggtgggg	tcacgggttg	13140
gaatgtgcc	agcggcgtct	cgctctttcc	aaggagctcc	tggtttagat	ttccatggcc	13200
tgtagacacc	ttcagccttg	ggtccaaggg	acacccctg	agatcaggca	cgctcaagaa	13260
gctgacaaag	ccctacactt	tatgccaccc	atgagctgga	ggcccggcag	gtctctttct	13320
ccagaaagca	aaggggggtg	gcgttagtga	gccctggcag	ccacctaacg	tggacttgga	13380
gcatctgcgg	ggctgtggtc	cagcaccacc	gtgtggccac	caggtgctca	tcagccagtg	13440
ggaccoggga	ggagggacaa	gaccagagaa	caacagtgtc	cttgctctt	ctctcctgaa	13500
ttttggacgg	tggttagac	ttgggtgtcc	ccatctctgt	gtttagagt	cttacagttt	13560
ccaaactgtt	tgcaaagtgt	gaagccaccg	tccctctcct	ctgggatggc	ccagtgtgt	13620
cgtggggccg	tggtcctgag	ctcagctttt	catttgaaga	ggtggaagga	gctgacaccg	13680
tcccatcccg	gcagggtgg	ctcaggtctt	ctttaggtcc	tgagtggggg	tccagcacag	13740
ccccaaaggt	gcgtggcacc	cgccctgcc	tctgcccatg	cactcatctc	ctggtggaga	13800
agacactcac	acacaggaag	cagggaaggc	agcagacctc	actcaccct	cacccctca	13860
ctcaccctct	actcaccctc	tcaacctctc	attcaccacc	cacccctcg	ccccctcact	13920
cacccctca	ctccctcaac	cctcactcac	ctcctcactc	cctcaaccct	cactcacctc	13980
ctcacctcct	cactctcccc	ctcatccctc	cctcacccca	ccccgtcacc	tcctcactca	14040
cctcctcacc	ccctcactca	cccttcaccc	cctcactcac	cacctcacct	cctcactcac	14100
cccctactca	acccctcatt	cacccctcac	cccctcactc	acccctgcac	cccctcactc	14160
accccttcat	ccactcaccc	acctgctcac	ctcctcactc	aacccctcac	cccctcacta	14220
atccctcact	ccctcacccc	ctcagccct	cactcacacc	ttcacctcct	cactcacccc	14280
ctcacccct	caaccctta	cttaccctct	cactcatccc	ttcacccctc	actcaccccc	14340
tctctcacc	attcaccccc	tactcatgc	cttcaccccc	tactcacct	cctcactcac	14400
accttcaccc	ctcagtcacc	ccctcactca	ccccttcacc	ccctcaatca	tgcttctact	14460
ccctcactca	ccccttcacc	ctctgaatta	ctccctcatc	ccctcactca	cccctcact	14520
caccccttca	ccccctcacc	caccacctca	cccacccctc	accaccccc	tcacctcctt	14580
acccctcacc	cccctcactc	acccctcacc	ccctcactca	ccacctcacc	cacccctcac	14640
ccacccctc	actcactccc	tcatccctc	actcaccccc	tcacccctc	actcaccccc	14700
tcacccaccc	ctcacccacc	ccctcacccc	ctcactcacc	ccttcacccc	ctcactcacc	14760
ccctcactca	ccccttcacc	ccctcactca	ccacctcacc	cacccctcac	ccacccctc	14820

actcactccc	tcacccccctc	actcaccccc	tcacccccctc	actcaccccc	tcacctcctc	14880
actcaccccc	tcacctcctc	actcaccgcg	tcacctcctc	actcaccccc	tcgccccctc	14940
actcaccctc	cacccccctca	ccccctcact	cacccctcac	cccctcgccc	cctcactcac	15000
cccctcgccc	cctcactcac	ccctcacccc	ctcacccccct	cactcatccc	ctcacctcct	15060
cactcacccc	ctcacctcct	cactcacccc	ctcacctcct	cactcacccc	ctcacctcct	15120
caccaccccc	ctcactcact	ccctcacccc	ctcacccccct	cactcacccc	ctcacctcct	15180
cactcacccc	ctcacctcct	caccaccccc	ctcactcact	ccctcacccc	ctcacccccct	15240
cactcacccc	ctcacctcct	cactcacccc	ctcacctcct	cactcacccc	ctcacctcct	15300
cactcatgcc	ctcacccccct	cactcacctc	ttcacctcct	tgctcatccc	ctcacttacc	15360
ccctcacttc	gtcaatcacc	cccccaacctc	gtcaatcacc	ccctcacctt	ttcactcacc	15420
ccctcactca	cccccttact	tcctcactta	cctcctcacc	cccactcac	cccctcaccc	15480
cccactcacc	ccctcacccc	acactcacc	cctcaccccc	cactcacccc	ctcacccccctc	15540
tcacctcctc	actcaccccc	tcacctcctc	acttatcccc	tcacccccctc	aattaccccc	15600
tcacccccctc	aattactccc	tcacctcttc	aattacccac	tcacccccctc	acctcctcac	15660
tcctcactca	ctccctcact	caccccttca	ccttctcact	cacctcctcg	tctcctcacc	15720
ccctcactca	cttcagccc	tgccccctccc	atcttctctt	tctttgtgtg	agaatctggg	15780
gtccctgagt	ggtgtcagtc	cctccaagac	tcaaggagtc	cccagggcct	tggtatccag	15840
aacacccccca	cctgggtccc	gggagacccc	atgggatcac	aggagtgttc	agggaagtgg	15900
tgcttctctg	gtctgggtgg	gctggagggg	catcctccct	tccccagag	gagacccccca	15960
ggagccccct	aagtccatcc	ccagcagtg	tgccccctgcc	ctgtccttgc	agcctgggag	16020
acccttggga	ggggcgggcg	ctgggtggct	gggcggcttc	tgctgggtctc	acccactgg	16080
cctcctgttt	gtcatcctca	gcctgcgggg	tcaacttgaa	ctcaagccgc	cagagcagga	16140
ttgtggggcg	cgagagcgcg	ctcccggggg	cctggccctg	gcaggtcagc	ctgcacgtcc	16200
agaacgtcca	cgtgtgcgga	ggctccatca	tcacccccga	gtggatcgtg	acagccgccc	16260
actgcgtgga	aaagtatgcc	aggggcggcg	cgggcgggt	gggggctcag	ggctggccta	16320
cagccaccct	gtgacctga	gcaggtctca	acccttgag	ccccggcatc	cttgtgttta	16380
aatggggaga	gtattgcacc	tgcttcttag	ggctgtgaga	catcaagtgc	gctcatgcca	16440
ggcagtgc	ggctgtatgc	actgagtgtc	ccctgcacgc	agggcacagg	gtgcaggtgg	16500
aacattctcc	acgatgtcgc	cgtgaccagc	gttccttcca	gccactgtcc	tctgagctct	16560
gtcctgccct	tgagcaaagc	ccctgcccc	tgaggtatcc	tgtctccggg	acgctagtcc	16620
caggagaggg	cacactcaga	caggcttcag	gctgccctgc	tggaagggtcc	ctgggggttaa	16680
gcgttcttgg	ccacagcatt	gctcatgcag	aggggttaggt	aggggtgagg	ctagccgtga	16740

cagtattagc	atztatggac	gctaccaccc	cctccccctt	tccttaaaca	catagtgtt	16800
ttgggtcacat	gctgtcttgg	aggaggcctc	acttggcgga	tgtatttttc	tgcccttagag	16860
agaggctgaa	ctgggtttga	ctgttggccc	agccctctct	tgctgcgtgc	ccttagacga	16920
ttcactcaac	gtctctgata	catggcatgt	acaactataa	gatgggcatg	cccttctcct	16980
ctcgggctgt	tatgaaggtc	aaggaagcaa	gggctgttac	ccaagggtgc	tccttctct	17040
ccccctcttc	acacccccag	gtgctctggg	ccctctagga	actgggtttc	tctcaagggc	17100
tgttacccaa	gggtgctccc	ttctctcccc	ctcttcacac	cactgggtgc	tctgggcccc	17160
ctaggagctg	ggattctctt	aagagggaaa	ctcttggata	aaggaaatgg	tttgattgat	17220
atcggacaag	tctgttcatt	agtatccatt	tattaagcac	ctaccatgtg	ccaggaaatg	17280
ctttggcgta	caaaggaaaa	taagggccag	tcctgctaga	aatggccttg	aaaccccagg	17340
gagggatgtc	ggcccatgtt	gggtgctgca	gattccttga	aggtgatgca	agagccagaa	17400
agaaggatga	tgtggggggc	tgaggcaggg	agtcgggggt	gggggagtgt	gggggagaag	17460
gggagaccga	gcacctcttc	cactatctcc	ctgtgtgggt	tttgggtgaac	catcctgcct	17520
ctgggtgtct	tgccctccagc	ttctgacgtt	ggaagtcat	ccactgagag	ctctgtgttt	17580
atggctctga	gatactgagt	ccttcttctc	tcccagacct	cttaacaatc	catggcattg	17640
gacggcattt	gcggggattt	tgagacaatc	tttcatgttc	tatggagccg	gataccaagt	17700
agaaaaagtg	atctctcatc	caaattatga	ctccaagacc	aagaacaatg	acattgcgct	17760
gatgaagctg	cagaagcctc	tgactttcaa	cggtacgtgt	ggctcaggct	tggaagcag	17820
gttggcagaa	tcttaaagag	atgttgattg	gaaatgacac	ttgtgctatg	ccaaatggaa	17880
gggaggcatt	tgctttgagc	gagggtagcg	tgacgcgggt	ggccaatggg	agaggctcac	17940
agaggctaag	agcacctgcc	gcatttttggg	ggaggcagca	gccaccacat	ctgttctgta	18000
ctgtactgag	tggtggtgat	tcaagccagg	catggaaaag	gctagaacag	ggctttccca	18060
ctgcagcacc	cttgacatct	gggtggttct	ctgttgtagg	gctctcttgt	gccttgtagg	18120
atgtttaaca	gcgtccccag	cctctaccca	ctggaggcca	gtagctacca	agctgtgaca	18180
accagtgttg	cctgctgaca	ttgccaacaa	tccgctttga	ggcaaagtca	cttccagttg	18240
agaactactg	gcctaaaaatg	tgtaaagatc	cttgattttt	aaagatacat	tctaaaacca	18300
agttgcttaa	ttcaggacaa	acatgctttc	tcttagcctc	ttattcggtc	ccactctggt	18360
ccatccaagg	gtctggaatg	ttctagcccc	atgtggatac	agaagaagca	aaacctcagc	18420
cctccctaca	gcatgtctgt	attcacattg	ggaaatgggt	cacatataga	agagcgaatg	18480
cctgagcaat	ggcgtgggtc	ctctggggcg	aaagctgact	ccattgactc	catcggtttt	18540
ttggctgttg	cctcctgtgt	gtctttcccg	tcttgatcac	ctggagatat	gtaattttgg	18600

ES 2 886 958 T3

aagcagagct	agcaaataat	tcctcttata	agcagagcta	gcaaataatt	ctacttataa	18660
gtagcataac	gtcttgctg	ccagaaggag	aggtctggca	gggggagaaa	gtgagaatgt	18720
gggacttggt	gggatgcagg	gtcctctggg	cagggtggcc	agggtgccag	gcccagcagc	18780
ctgcatgtgg	gaaggccagg	tggagacata	ggtgataccc	gcctggctca	ctgtgttttc	18840
tcttcttgaa	acagacctag	tgaaccagct	gtgtctgccc	aaccagggca	tgatgctgca	18900
gccagaacag	ctctgctgga	tttcgggtg	gggggccacc	gaggagaaag	gtgaggctgc	18960
tcctgggcac	acaggactgc	agggcccaca	gatggagcat	tgggttcgga	agtgggaggt	19020
ccaggtttta	atcccagttc	tactactcaa	tgactggatg	actttggttg	attccccag	19080
tccttggtgc	tcagtttctc	catctgctaa	gtgggagaaa	tcctgcccag	cctacctaat	19140
acactgtgtt	cttatcgtga	tcacacagag	cagcatgtgg	aatggctttt	gaagtatctg	19200
ggccatacga	gtttagaggt	gcaggatctc	ctgtgttgca	ctcattgtga	gtttagagct	19260
gccctggaga	tcccaccaag	gcctgctggg	ctgagtgaca	gggggcttgg	tgaggacggg	19320
catcctggac	ccatggtggc	cacatctaag	cctgtcctct	gccctgataa	ccacagagag	19380
aggctctctc	cacccacttc	ctttgcaatc	tgcatttctc	tctgacagtc	tttcaaata	19440
agggagcctg	gctgcttcat	ttttatggag	ggttggaagt	gcttagtggc	aggcacaaa	19500
gttcatttta	catattgttt	atatccttct	caaaagcgtc	taggccatac	agacaacaaa	19560
tcctttcaaa	caaggggaaa	agtacaaagg	ttgggtgatt	tctggggagc	gtcagggaa	19620
gtagtggggg	gcatcctggc	tcctcatcag	cagaaactta	ctacagtaga	gccacaggct	19680
gggcaaaaga	cctcatggaa	tccaagatga	agggaatctc	gacaaatatt	tgtgcgcacc	19740
tgcacctagt	acaggctggg	tgctactcag	gtgctgggaa	tgacagaagt	aacagagtaa	19800
gacaaatgtc	tctgctgtca	ggagctttac	ctctcttctg	gatgtcgggt	gtggggacgg	19860
ggcaggtgtg	gtcagacaga	tgggagacaa	acaactgagc	gaggtacttc	caaacatctg	19920
aggggtggga	tcacaaggct	ccggctatct	tgaaggggtg	gtcaggaaag	gcttctcgga	19980
agaggtggca	tttgagctga	gactcaaatg	gcaaaaatgt	gtacacatca	aaaaggctag	20040
tgcattgtat	ttcaggtgtg	gtcaaggggc	caaggaggtg	ggctggggcc	agattgcata	20100
ggctcctgtg	gattatgggt	aagacaccag	cttctcatct	gcttgaggtg	gggagatcgt	20160
gagccgggga	gtgccatgat	ctggcagctg	cgtggggagt	ggggatgaat	ggatggagac	20220
gaggatgatg	gtgacaagtc	cattgctgtg	gttccttgag	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	tgggggtggg	gaaggtgtga	cggagggttc	tggttttgca	cactaagccc	20520

acggtgcccc	gaagatgccc	gaggggaggc	agcaaagcga	gagtgggaaa	tgcagaggtg	20580
gcaagtgcag	gccgtgtctt	gagaagctct	aatgtgcagg	ggagccgaga	agcaggcggc	20640
ctagggaggg	tcacgtgtgc	tccagaagag	tgtgtgcatg	ccagagggga	aacaggcgcc	20700
tgtgtgtcct	gggtgggggtt	cagtgaggag	tgggaaattg	gttcagcaga	accaagccgt	20760
tgggtgaata	agagggggat	tccatggcac	tgatagagcc	ctatagtttc	agagctggga	20820
atttctttcc	ctgaagctga	actccagagc	tgcattcagc	acaggcaccg	ccagttgtaa	20880
ggagaatcca	ggtttcccag	gagaggggtt	ggtgctggga	tgagctgacc	ggggcagggc	20940
tggaaaatag	ggctgtgacc	atctgtgtag	tgcgtgtgga	ggtctcaggg	aggggaagtgt	21000
gctctccctg	cgagagctgc	aggcaacact	gggagctcaa	caagtctccc	tgtccttagg	21060
gaagacctca	gaagtgtctga	acgtgccaa	ggtgcttctc	attgagacac	agagatgcaa	21120
cagcagatat	gtctatgaca	acctgatcac	accagccatg	atctgtgccg	gcttcctgca	21180
ggggaacgtc	gattcttgcc	aggtaattca	acatttttat	tctacctttg	gtccttacca	21240
gctcctactg	aaccccccat	gagagagagg	gcattcttgg	ggtcagcaga	gcctcctcag	21300
tgacacggag	ccagctcggg	gcagtcatgg	gaagtgacgg	ccacaaacag	tgcgaacgct	21360
tctggtggca	gaaggaagta	cagtcaacaa	atcacacaca	ccctctgaaa	aaccgggtatt	21420
tggtaaaagt	gccagtggaa	cagaaacaag	tatttagact	attttaaatt	atgaacggca	21480
atttatttag	taacttttag	cttgaacaga	ttaaaattca	ggatgggggc	tatctctttg	21540
ggggttacat	ctctgttacc	atcacccctt	gatggtggag	attcgaagcc	cacacagtca	21600
ctcgtaaact	acactgcgac	ccccgcccc	caactcctct	aggcctggtc	agtgggtgtgc	21660
ggcagattgt	gacttgattt	tctgctctct	gtaccttgct	gtgtcccaca	gggtgacagt	21720
ggagggcctc	tggtcacttc	gaagaacaat	atctggtggc	tgatagggga	tacaagctgg	21780
ggttctggct	gtgccaaagc	ttacagacca	ggagtgtacg	ggaatgtgat	ggtattcacg	21840
gactggattt	atcgacaaat	gagggttaact	atcctgtcct	ccttctgact	gtgttctccg	21900
attcctcgag	ccaaagccag	acatctgtta	ggcgtgggtc	tgctgctgga	agctgactgg	21960
tgaccactgg	tcagcatgaa	gcaaactctg	cttcctccag	ccacagcccc	atccccccag	22020
tgtccaccca	ttgccattg	cctctcactg	gcttcacttg	catatttccc	ctggtgtttg	22080
gatgaaaagc	gctggggctc	agcttgtgtg	aaattccttg	gtgctctgcc	aaccacactt	22140
cgttctggct	cagctgactc	agctgttcca	cccaggccac	ctcacatcaa	actttttttt	22200
tttttttttg	agatggagtc	tcactgtgtc	gcccaggctg	gagtgcagtg	gcacaatctc	22260
gactcactgc	aacctttgcc	tcctgggttc	aagtgattct	cctgcctcag	cctcccaagt	22320
agctgggact	acaggcatgc	gccaccacgc	ccagctactt	tttgtatttt	tagtagagat	22380

ggggtttctc	catgttggcc	aggctggtct	cgaagccctg	acctcaggtg	attcaccac	22440
ctcagcctcc	cacagtgcctg	ggattacaag	tgtgaaccac	ggtgcccggc	ctcacatgaa	22500
acttttgatt	tatagagagc	agaggggaaga	gccggctgtg	cccatccttt	tctggggcca	22560
tcgagtggct	cctgggcagc	ccccaagggtt	aggaaggga	ggagcagcca	gggttctctg	22620
atgccccaga	ctcaagcacg	agggaaggtc	tcaggggttc	catgtgagcc	tcattgatgt	22680
ctctgcttag	cagagccctg	gctttgggca	ttgtccagat	aggggtgag	aaccagatct	22740
tctcatctcc	aggacctcag	acgtatagtt	ttctcagatt	tctgtgcttt	ctggggctgg	22800
gctactagtg	gaagaaagca	gtctattctg	tcttctccca	aatctcccag	atgccagtc	22860
tgttgaagga	ggagcagaac	cagggggcct	ttcccgtga	ggcccagcct	gtgtctcctt	22920
caaatgacac	gcgggactca	gggccttccc	atgaccatgg	ggcccagggg	gcgtcacctg	22980
gccagggcc	cagtgcctaga	aacagatgac	cccaggagga	ggaggcaggg	caggaggga	23040
gctggcaggg	ctgggatggt	cagccaggct	gaggggcgga	ctcgcaccag	gatggagcta	23100
ggaaatgac	caggtgtgtt	tggcggtgc	aggtgggtcc	gcattgctgt	gcaggagggg	23160
aagggtgcg	tggcaggaga	gcagccgggg	gagggccaga	ctctgctgaa	gagatgcctg	23220
ttgtgccggc	ctccacatcc	gctgccgct	ccttccggag	ctcctgcccc	gccatgctca	23280
gcctgactct	gaccaacacg	ttggagagaa	gaatgatccc	tttgtgctat	taagcttgct	23340
tatttggttt	ctaagtgcct	catgcgaacc	tagaggaaaa	aattattttc	cacctttgtt	23400
tgtcttaaga	aaataacaca	cttttttttt	tcctatttga	acaggcagac	ggctaatacca	23460
catggtcttc	gtccttgacg	tcgttttaca	agaaaacaat	ggggctggtt	ttgcttcccc	23520
gtgcatgatt	tactcttaga	gatgattcag	aggtcacttc	atttttatta	aacagtgaac	23580
ttgtctggct	ttggcactct	ctgccattct	gtgcaggctg	cagtggctcc	cctgccagc	23640
ctgctctccc	taacccttg	tccgcaaggg	gtgatggccg	gctggttgtg	ggcactggcg	23700
gtcaagtgtg	gaggagaggg	gtggaggctg	ccccattgag	atcttcctgc	tgagtccttt	23760
ccaggggcca	attttggtg	agcatggagc	tgtcacctct	cagctgctgg	atgacttgag	23820
atgaaaaagg	agagacatgg	aaaggagac	agccagggtg	cacctgcagc	ggctgccttc	23880
tggggccact	tggtagtgct	cccagcctac	ctctccacaa	ggggattttg	ctgatgggtt	23940
cttagagcct	tagcagccct	ggatggtggc	cagaaataaa	gggaccagcc	cttcattgggt	24000
ggtgacgtgg	tagtcacttg	taagggaac	agaaacattt	ttgttcttat	ggggtgagaa	24060
tatagacagt	gcccttggtg	cgagggaagc	aattgaaaag	gaacttgccc	tgagcactcc	24120
tgggtgcaggt	ctccacctgc	acattgggtg	gggctcctgg	gaggagact	cagccttcct	24180
cctcatcctc	cctgacctg	ctcctagcac	cctggagagt	gcacatgcc	cttggtcctg	24240
gcagggcgcc	aagtctggca	ccatgttggc	ctcttcaggc	ctgctagtca	ctggaaattg	24300

aggtccatgg	gggaaatcaa	ggatgctcag	tttaaggtac	actgtttcca	tggtatgttt	24360
ctacacattg	ctacctcagt	gctcctggaa	acttagcttt	tgatgtctcc	aagtagtcca	24420
ccttcattta	actctttgaa	actgtatcat	ctttgccaa	taagagtggg	ggcctatttc	24480
agctgctttg	acaaaatgac	tggtcctga	cttaacgttc	tataaatgaa	tggtctgaag	24540
caaagtgcc	atgggtggcg	cgaagaagag	aaagatgtgt	tttgttttg	actctctgtg	24600
gtcccttcca	atgctgtggg	tttccaacca	ggggaagggt	cccttttgca	ttgccaagt	24660
ccataacat	gagcactact	ctaccatggt	tctgcctcct	ggccaagcag	gctggtttgc	24720
aagaatgaaa	tgaatgattc	tacagctagg	acttaacctt	gaaatggaaa	gtcatgcaat	24780
cccatttgca	ggatctgtct	gtgcacatgc	ctctgtagag	agcagcattc	ccagggaacct	24840
tggaacagct	tggcactgta	aggtgcttgc	tccccaagac	acatcctaaa	aggtgttgta	24900
atggtgaaaa	cgtcttcctt	ctttattgcc	ccttcttatt	tatgtgaaca	actgtttgtc	24960
tttttttgta	tcttttttaa	actgtaaagt	tcaattgtga	aaatgaatat	catgcaaata	25020
aattatgcaa	tttttttttc	aaagtaacta	ctgcatcttt	gaagttctgc	ctggtgagta	25080
ggaccagcct	ccatttcctt	ataagggggg	gatgttgagg	ctgctggtca	gaggacccaa	25140
ggtgaggcaa	ggccagactt	ggtgctcctg	tggttctcga	gataacttcg	tataatgtat	25200
gctataogaa	gttatatgca	tggtcctcgc	gccgggtttt	ggcgcctccc	gcgggcgccc	25260
ccctcctcac	ggcgagcgct	gccacgtcag	acgaagggcg	cagcgagcgt	cctgatcctt	25320
ccgcccggac	gctcaggaca	gcggcccgc	gctcataaga	ctcggcctta	gaacccccagt	25380
atcagcagaa	ggacatttta	ggacgggact	tggttgactc	tagggcactg	gttttctttc	25440
cagagagcgg	aacaggcgag	gaaaagtagt	cccttctcgg	cgattctgcg	gagggatctc	25500
cgtggggcgg	tgaacgccga	tgattatata	aggacgcgcc	gggtgtggca	cagctagttc	25560
cgtcgcagcc	gggatttggg	tcgcggttct	tgtttggtga	tcgctgtgat	cgtcacttgg	25620
tgagtagcgg	gctgctgggc	tggtcggggc	tttcgtggcc	gccgggccgc	tcggtgggac	25680
ggaagcgtgt	ggagagaccg	ccaagggtg	tagtctgggt	ccgcgagcaa	ggttgccctg	25740
aactgggggt	tggggggagc	gcagcaaaat	ggcggctgtt	cccagtcctt	gaatggaaga	25800
cgcttgtag	gcgggctgtg	aggtcggtga	aacaagggtg	ggggcatggt	gggcggcaag	25860
aacccaaggt	cttgaggcct	tcgctaatac	gggaaagctc	ttattcgggt	gagatgggct	25920
ggggcaccat	ctggggaccc	tgacgtgaag	tttgctactg	actggagaac	tcggtttgtc	25980
gtctgttgcg	ggggcggcag	ttatggcggt	gccgttgggc	agtgcacccg	tacctttggg	26040
agcgcgcgcc	ctcgtcgtgt	cgtgacgtca	cccgttctgt	tggtctataa	tgacgggtgg	26100
ggccacctgc	cggtaggtgt	gcggtaggct	tttctccgtc	gcaggacgca	gggttcgggc	26160

ctagggtagg ctctcctgaa tcgacaggcg ccggacctct ggtgagggga gggataagtg	26220
aggcgtcagt ttctttgggtc ggttttatgt acctatcttc ttaagtagct gaagctccgg	26280
ttttgaacta tgcgctcggg gttggcgagt gtgttttgtg aagtttttta ggcacctttt	26340
gaaatgtaat catttggttc aatatgtaat tttcagtgtt agactagtaa attgtccgct	26400
aaattctggc cgtttttggc ttttttgtta gacgtgttga caattaatca tcggcatagt	26460
atatcggcat agtataatac gacaagggtga ggaactaaac catgggatcg gccattgaac	26520
aagatggatt gcacgcagggt tctccggccg cttgggtgga gaggctattc ggctatgact	26580
gggcacaaca gacaatcggc tgctctgatg ccgccgtgtt ccggctgtca gcgcaggggc	26640
gcccggttct ttttgtcaag accgacctgt ccggtgccct gaatgaactg caggacgagg	26700
cagcgcggt atcgtggctg gccacgacgg gcgttccttg cgcagctgtg ctcgacgttg	26760
tcactgaagc ggggaaggac tggctgctat tgggcgaagt gccggggcag gatctcctgt	26820
catctcacct tgctcctgcc gagaaagtat ccatcatggc tgatgcaatg cggcggctgc	26880
atacgcttga tccggctacc tgcccattcg accaccaagc gaaacatcgc atcgagcgag	26940
cacgtactcg gatggaagcc ggtcttgtcg atcaggatga tctggacgaa gagcatcagg	27000
ggctcgcgcc agccgaactg ttcgccaggc tcaaggcgcg catgcccgcac ggcgatgac	27060
tcgtcgtgac ccatggcgat gcctgcttgc cgaatatcat ggtggaaaat ggccgctttt	27120
ctggattcat cgactgtggc cggctgggtg tggcggaccg ctatcaggac atagcgttgg	27180
ctaccgctga tattgctgaa gagcttggcg gcgaatgggc tgaccgcttc ctcgtgcttt	27240
acggtatcgc cgctcccgat tcgcagcgca tcgccttcta tcgccttctt gacgagttct	27300
tctgagggga tccgctgtaa gtctgcagaa attgatgac tattaacaa taaagatgtc	27360
cactaaaatg gaagtttttc ctgtcatact ttgttaagaa gggtgagaac agagtaccta	27420
cattttgaat ggaaggattg gagctacggg ggtgggggtg ggggtgggatt agataaatgc	27480
ctgctcttta ctgaaggctc ttactattg ctttatgata atgtttcata gttggatata	27540
ataattttaa caagcaaac caaattaagg gccagctcat tcctcccact catgatctat	27600
agatctatag atctctcgtg ggatcattgt ttttctcttg attcccactt tgtggttcta	27660
agtactgtgg ttccaaatg tgtcagtttc atagcctgaa gaacgagatc agcagcctct	27720
gttccacata cacttcattc tcagtattgt ttgccaagt tctaattcca tcagacctcg	27780
acctgcagcc cctagataac ttcgtataat gtatgctata cgaagttatg ctagtaacta	27840
taacggtcct aaggtagcga gctagctcca cgtggctttg tcccagactt cctttgtctt	27900
caacaacctt ctgcaagaaa accaagggcc tgaattttaa ctctctg	27947

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

5 <220>
 <223> Polinucleótido recombinante

 <400> 6

gcagagtcta agaaatcgct gtgttttagcc ctcgccctgg gcaactgtcct cacgggagct	60
gctgtggctg ctgtcttgct ttggaagttc agtaagtgc gggagcctcg atcccaccat	120
gtgtcctgc agtccccagt gctctgagcc agaccctgct ctctgggcta ttgagacctc	180
tggaggccct ccgtgagggt cctctcttac ataacgagggc tgtctctctt cccttctctt	240
gtttagctat gagattgaca catcatgggg aaagcattta gaatgtacct agtgctttgg	300
gggtgcttgg gccaccacag actgtgagca caggttcttc taccttgggg ccacaccacg	360
ttacctgtat ctcaactgcac agcagtggct gttggggacc agggccaccc ctccatgtcc	420
cacctcctgc aactgcagcc tgagcccttc catcagcctg ggggtggtgca gacctatgtg	480
ccattgtgga tccttcaagt tacctgtgtg gcagagagga cgtgtgagtg ccgtccaaac	540
ccaaacactg agaggggtcct tcccattgcc ccacggaag taagggtgcc cagtgtctaat	600
tccacttata ctgtgtgggt gcaaggacac ttctcctcct tattaaagtg ggggattggc	660
tgggtgaggt ggctcacgcc tgttatccca gcactttaag aggccaaagg aggtggacca	720
cctgaggtca ggagtttgag accacaagcc tggccaacat gttgaaactc catctctact	780
aaaaatacaa aaattagtca ggcgtgggtg cgtgcacctg taatccagc tacttaggag	840
gctggggcag gaggatcact tgaaccacag agttggagggt tgcagtgagc caagattgtg	900
cccctgcact ccagcctggg tgacagaatg agacttcac tcaaaaacaa aacaaaacaa	960
aacacagtgg ggccaggagt tggaggctgc agcgagctac agtaatgcca cgggtgtcct	1020
cactccatga ggctcattgc gtttctcagc ctgaagggca cctctcttct gtttctctg	1080
caagtgggca gcaagtgtc caactctggg atagagtgcg actcctcagg tacctgcac	1140
aacccctcta actggtgtga tggcgtgtca cactgccccg gcggggagga cgagaatcgg	1200
tgtggtgagt cagccttgac cttgggaagg gactcctctg ctcaccttgg agacagcagc	1260
cgggtccagg ggcccttggg tgactgggccc tggcgtgcgt ccagtaagct gacacatgat	1320
gtcattgaat ccctgctcca ggctgagccc tggggctcag agaggttgtg tttccggccc	1380
aacctacccc agcagggtgg agatgacagg gccaccgagg actgtgtcat tggaaaccaca	1440
cgtgctctga actgccacag gaagtcaagt aagatgagca aactgtttat aaagttggag	1500
atgcaggcta ggaacgggtg ctcatgcctg taatccagc actttgggag 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 gaggttgacag tgagctgaga	1740
tcacgccact gcattccagc ctgggagaca gagctggctc aaaaaataaa ttaattaatt	1800
aaaaacaaaa ttggagatgc actatgttat tttcaaaaca agctgccttt aaagatctat	1860
ctgttgtcac aggggtgggct catctgtttc attttatttt ctgtggttta tctattttatt	1920
cattttaatg aactaggaag cattgtcctt atttatggca taccacatga tgtttggata	1980
cgtgtatgcc tgtggcatgg ctaagtcaag ctagaacatg ggccttacct catatacgtg	2040
tcttattaag aacacataaa acctactctt gtagtgattt tcaaatatgc aacatatagt	2100
ttattaactg cagtcactat gatgtacaat agattgctcg aacttattcc tcctgtctaa	2160
ctaagatttt gtgacctctg accaacatct cccagtggt gtcaccccc gcccccagcc	2220
tctgatagct gcctttctac tctctgcttc tgtgagtttg atgtttatac attccacatg	2280
taagtggcct catgcagtgt ttctgtctct gtgtctggct tgttcactta gcgtaatgtc	2340
ctccagcttc atctatgttg ttggaaatga caggatttcc ttctttcttg tggctgaata	2400
gtattgcctt gtgcatatac accacatttt ctttatccct tcattcactg atggactctt	2460
aggttgatgt catgtcttgg ctgttgtgaa aaatgccgca gtgagcgtgg gcgtgcaggt	2520
ccctcttcaa cacacggatt tcctttcctt tggatataaa cccagcagtg agattgctgg	2580
atcacatggc agttctgttt ctcacctttt gaggaaactc catactgttt tccataatgg	2640
ctgtagcaac ttccactccc acccccacgg tgcaaagtct ccatttctct tctacaacct	2700
caccaactcc tgttattttc catctttctg atagtagcca tttgaagagg tatgagatga	2760
tacctcattg tggttttcat ttgcattttt atttgtattt ttcatgaatt tttgagggtg	2820
atttcaaggg tagttagtga ctgcaacagg gaaacgatcc tgagtatgag ggttgtgcta	2880
atcatcccc tcctgccagc tgcgtacgga atggggctct gcagatggca gggagctggc	2940
tcgtttctct ttaagagctg cttttactt ttcttctct tcctttaaaa cttatttctt	3000
ggccggacgc agtggctcat gcctgtaatc ccagcacttt gggaggccga ggtgggaggga	3060
tcacgaggtc aggaattcca gaccagcctg gccaacatgg tgaaaccccg tctctactaa	3120
aaatacaaaa attagccaga cgtgggtgtg cgggcctata gtcccagcta ctcgggaggc	3180
tgaggcagga gaatcacttg aacctgggag gaggggggtg cagtgagccg agattgcgcc	3240
actgcactcc agcctgggcg acagagccag actccatctc aaaaaacaaa aaaaagttat	3300
ttcccaagca cagccatgta ttccaggctt gtggatcagc gttggtggtg gtgtgtgctc	3360
tcatatctta gttccagcta agcacactct gacatgttta cactagaacc atttgttttt	3420
tctagaaata gaaatttcag aattgtagag tcagaggact taccagaaat ctcttaggta	3480
gttctctctc cctccctcaa gtgcagtcct aacctcctgg agttttctgt agaaaccaca	3540
agcctcagag ctggccgaga attctagcca aagatttttc catgccaaag taatcccccc	3600

tctcctaagg gccatccttg gtggggactg gtttcctggt aagccctcgc tgtcagtcct	3660
ggctgtggaa tttcctggtg aggagcactg gcccgaggag ctcgccctc gtgccggcct	3720
tgagcaggcc caagtgttcc gtgttcttga tacctttcct ccagcacagt cttgcttccc	3780
agaaaaaggt ttgcacttga aaatgatgca tttgctgatt aaacatagtt cttttgcttt	3840
atttggtttc taaaataaag tgggagtttt tgagattgag taacgtgagg ttaagatagc	3900
acgtggaatg gctttttctt ttctttctat tttttttttt tttttcctgg agacagggtt	3960
tcactctggt gcccaggctg gagtgcagag gcatgaccat ggctcactgc aacttcgatg	4020
tcctgggggt aagcgatccc ccagcctcag cccccaagt ggctgggact acaggtgctc	4080
gccaccacac ctggctaatt tttgtatttt ttgtagaaaa tgggtttcat caatgttgtc	4140
cagactggtc tcgaactcct gacctcaagc aattctcctg cctcagcctc ccagactgct	4200
gggattacag gcgtgaacta ccacgcctgg cctggaatgg cttttgatgt tctcctatgt	4260
gcacatgtgg gtgaataaac accaacaag tccttatggt acctgaagag ttgctctctt	4320
cttaatatTTT aagtcgtatt tatttaaata ctttaatagt tgtacactat taaagtatta	4380
ttaggtcaaa atcaaggaag tacaaaaggg tatgctgtga aaaatctctt cttccttgct	4440
ctgcttactt acctaccccg catccccca tacacccag acacacacac acacacacac	4500
acacacacac acacacgcat cactcccata catgccacc tgtttaccag ccaatcacat	4560
ttcttggggc aactcatctg agttgcttct ctttcagag agtttttgca taaagaagca	4620
caggatattc tgcgttacca tgacctatt tcccagtggt tcctagccag ttgactctcc	4680
tgcactggat accatcctgg acagcattcc ttagggaaat gagccccctg ttttttccca	4740
ccatggcaca gttggtcctt tgcattggacg caccattatt gccctgtct cttcttggtg	4800
gaccttaagg ttttctccat ctttttgctg taacacacac tgctccaagt gtgtgagcat	4860
atcagtagga aacgcttcca ggagtagaac tgctaggtca gagggcgtgt ggatctgtaa	4920
cctgacagac ctagaccggc ttcagtttgg ttttatccag tttccatatt gattattcat	4980
ataaaaggaa acagacaaac ataacgctgt gcatgtattc tctcttagac cagaacaggc	5040
atagggtgca cttttaattt gtccatttcg tagagtagaa attgtttttg ctgaaatgaa	5100
caccttagga tgctgaagaa tatgaccctg cccatggaaa acattcaaaa atgtgtgtag	5160
cgttttcttc ccaaggggtg gtgtgcat attttaacac taattcactt tctacttccg	5220
ttgctatcct ttctgtgagt ctttctcaga atctcagaaa agaaactaaa ttgttctctc	5280
tagttatcaa tgctgtactc tatacctgga atttgctaaa agggcagatt ttaagtattc	5340
tcaccacaga aaagagaaaa gaaaatggta attatgtgac gtgggtggaca tgtaaactag	5400
ctttattatg gtgagcattt cacagcggat atccagtcac cacgctgtac acattaaaca	5460

tgtacaattg ggtttttttt agacaaggtc tccttctgtc acccagtctg gagtgcagtg	5520
gctcagtcac ggctcattgc agcctcgacc tcctgggctc aatccatcct tccccctcag	5580
cctcctgaaa agctggggcc acaggcatgt accatcatgc caggctaata catatatatt	5640
tatatattttt ggtggagatg gggttggtct cgaactctgg gctcaagtga tcctcccgcc	5700
ttgcccttcc aaagtgctga gattacaggc atgaaccaca gcaccaggcc tacatgtaaa	5760
atttttattt gtcaactata ctttgacaaa gctgagaaaa aaaatcctaa tatttaaaaa	5820
aaaaaaaaaa aggactagct tgagaccttt tcagctctc tggcttatca gctgccgtct	5880
cttcgggtg cagatagctg gaagggaag aaaatcccta aaattaccca caagccaaga	5940
atgaagtgtc tccctttgag ccacagtggc agttttgttt ttaatcatag aagtgtattt	6000
tgagccgggt gtgctggctc acgcctgtaa tccccgact ttgggaggcc gaggtggggg	6060
gcggaggggg tggggatcgc ctgaggtcag gagttcgaga ccagcctgac caacatggag	6120
aaaccccgct tctactaaaa atacaaaatt agccggcgctg gtggtgcatg cctgtaatcc	6180
cagctactca tgaggctgag tcaggagaat ctcttgaacc caggaggtgg aggttgcggt	6240
gagctgagat catgccattg cactccagcc tgggaacaag aaaaaaaaag aagaagaaga	6300
agaagtgtat tcatttcagt tactttttaa aaagtgaaca gactttatat ttagagcgg	6360
ttttaggttt acagaaaatg aaacagacag ggcagcgagc tccttgtaact cctccccagc	6420
acacagttgc cctgttatga acatcccaca tcagtgtgtg gcgttcatta acaccgatga	6480
acctgatgca tacattatga tgaactgaag tcctggactt caccctttct cttgtacagt	6540
tctgtgggat ttgacaaatg cataatgctg tacagccaca atgatatgat cgtccagagt	6600
agttctcctg ccttaaaacc tcttttgctg cacctgtttc tctctcccca ctcacccag	6660
ctatctgata ttcttagtgc ctccgaagt ttggtctttt caggatgttg tagcgttgga	6720
atcatggagt atgtagcctt caccacatac accttccttc actttgttgg cttcctttac	6780
ttagtaatat gcattcaagt ttccctccatg ccttttcatg gcttgatagc tcatttcttt	6840
ttagcaccaa ataattattc gttgtccaga tgtagcacia tgtttatcca ttcagttaac	6900
ctgtgaccga ctacagata ggatgtggaa tcaactacca cagaggcatt agacaataat	6960
cagacccaag tcatttcatg ggggaacaag cccacaggta ccagactgtc cagtgaagtca	7020
gggccaactc taggaagtaa gaagagaggc tagagcatag ccaggtcctc actttatact	7080
ttaagcccat gtgtatttct cccaaaccac acagcattgt ttccatgctt tcagctttgc	7140
atgaataacg tgatacttga acgcattcatt tatcacttgc tctctttccc acagcgctgt	7200
tttcaagctt cttcctgttc atgatgctct gcttaaccct taagctgcat gggattctgt	7260
tctgtgaata cgcccacccc atgtattatc ctgccagca aaaagtcccc aaaactctgg	7320
atggtggtta cctctagga gggagagaag agattgggaa tagggagcga cttcaacggt	7380

gtttgtaatg	ttttgtttct	ttaaataaaa	gagctgagat	catttcagca	gaatgttgat	7440
ttagagtctc	ctggacaatt	tgttgctcaa	agtgctctct	taaagagcac	tttaaaaaaa	7500
aaaacctttt	atcttattat	ttatttat	atttattgag	acggagtttt	gctctgtcac	7560
ccaggctgga	gtggagtgg	gtgatctcag	ctcactgcaa	cctttacctc	ctgggttcaa	7620
gcaattcccc	tgcctcagcc	tcccaagtag	gtgggattac	agatgcgtgc	caccacactt	7680
ggctaatttt	tgcatttttag	tagagatcgg	tttctccatg	ttggccaggc	tgatctcaaa	7740
cgctgacct	caggtgatct	gcccgccttg	gcctcccaaa	gtgctggtat	tacaggcgtg	7800
agctaccatg	cctggcttat	cttatatat	tttaaaaaca	gcttattgag	atctaattta	7860
tgtaccataa	aattcaagta	tataattcag	tgctttttata	tataaaacat	atatatgaaa	7920
tagcttattg	agatataatt	ttttatataa	aacagcttat	tgatatgtaa	tgtatgtacc	7980
ataaaaattta	aatatataat	tactggctt	ttatatattc	acgaatatgt	gcaactatca	8040
ccacagtcaa	tttttagcata	ttttcatcag	ctcataaaga	aacccaagc	ccttgaacta	8100
tcaccccata	tccctcctcc	cagcccgtcc	ctcctactca	taagcaacca	ctaacttact	8160
tagtgtctat	agatttccta	ctctaggcat	tccatgtgag	cgggatcatg	caatacgtgg	8220
gctcacacaa	tataagtggc	attccatgtg	agtcggctca	tgcagtatgt	ccggctcctt	8280
tactgagca	taaggctctc	agcactcatc	caggttgtag	cctgtgtctg	aatttcattc	8340
cctcttctgg	ctgaatcgta	ttccattgtg	tatcttggac	atatactatt	ctgctcacc	8400
agccgttgg	gggcgtttgg	agtgtttctg	cctttcagct	gttttaagag	ggttgtagtg	8460
aacatttgta	caagttttgg	acccaatgcc	tgttttcaat	tctcttgtgt	agagagcact	8520
tttttagcaga	aaaagaatag	atttgtggcc	tccctttgtg	tgcggtcagt	gccttgagaa	8580
gagtgaactg	tgtgccacc	tccggagccg	tggagagcgc	ggggcttggg	tagcagctag	8640
gacgatacaa	gttgggacaa	ggccagggtgc	aatggctcac	gcctgtaatt	ccaacacttt	8700
gggagaccga	ggcaggggga	tcacctgagg	tcaggagttc	aagaccagcc	tggccaacat	8760
ggtgaaaccc	catctcta	aaaacagaaa	aattaactgg	acgggggtgg	ggacgcctgt	8820
aatcccagct	actcgggagg	ctgaggcagg	agaatcactt	gaacctggga	ggcggaggct	8880
gcagtgagtg	gagatcagac	cactgcactt	cagcctagg	gacagagcga	gactccgtct	8940
caaaaaaag	aaaaaaaag	aaagaaactc	atggataatc	ctccctctcg	tgcagttcgc	9000
ctctacggac	caaacttcat	ccttcagggtg	tactcatctc	agaggaagtc	ctggcaccct	9060
gtgtgccaag	acgactggaa	cgagaactac	gggcgggcgg	cctgcaggga	catgggctat	9120
aagtgagtat	ggggcagcac	ccgccgagtg	acagtaacag	acagcagaaa	cacgagaaga	9180
ccctctctct	gcctccctgt	gaaagcaccg	gcacatgagt	gctggggaca	attgtcacct	9240

tccaaaagct	gagccctata	accagcaggt	ggaatttgct	ctgctagggc	tgtgccagc	9300
acacagacct	tggctcactg	ccaccttgcc	ctgcctcctc	cttggcctct	atagactcct	9360
ggttgctcgg	gagtgccag	tgtgtggtc	atctggctcag	aggggtaggc	tgagggcggt	9420
aggtgcctct	ttttccaagg	tgcctctcag	ccaggggtcca	ttcacctccc	tgggtagagg	9480
ttggaccaga	acagctggcg	aggaggggtg	ggctggggag	agcagcagag	acaaatcctg	9540
tgccagtttc	acttcattcg	ggagccatgg	aagccttttg	agctggggag	agaatcaatc	9600
aatcagactg	atacttaaaa	aatgtcattc	ctgctcgtag	ctctgagggg	aggtgggaag	9660
gcttaacagg	gtgtgtgtcg	cctgacagtg	attcctaacg	ggggtggggc	ggtgggtacc	9720
atttaccagc	actgcctggg	gagatgcggc	agccctcagg	catcggggga	gaggggtggt	9780
ggatgctact	gccactttgt	tttccatggg	aggggtcccca	ggtgatttct	atgcaacttt	9840
aggggtattca	atatgccagt	tttcagaatg	aattaccact	cggtgagaaa	gttggcatct	9900
tagctagtca	ctgtgacatc	cctaaacagc	aggggtgaat	tacacagcaa	agcccccca	9960
tcacagtcca	ggaacctggg	ggaattgata	actggggcca	tgtaaacatc	tgtacctttt	10020
attagattaa	atgtgtgtat	gattatacaa	tcctatgtcc	ttctcatagt	ttcttgatcc	10080
taacctggat	aagaaacacg	accaatgaag	gaattttgtc	tgacacttta	gggttattga	10140
atcgaaaaat	cgttacaata	ttctagcact	tggttagaac	gtgtgatttt	ttttcctaaa	10200
tgctaagggt	tttccctctt	attctgaatg	tcgtatgagc	ggtattatga	catagtatag	10260
gatttggtgt	tgcttatgcc	ttaaccatta	tcacaaataa	ggttttcttt	tttaggaata	10320
atttttactc	tagccaagga	atagtggatg	acagcggatc	caccagcttt	atgaaactga	10380
acacaagtgc	cggcaatgtc	gatatctata	aaaaactgta	ccacaggtat	gcagcaattt	10440
cttcttgaaa	aattttggaa	tgaaatcaac	taggagacac	catggggaat	cgttgtcctg	10500
agtctgattt	ctctgagctg	caatactcgg	tctggatggg	ttttgcattg	ggaggagatt	10560
agagtctgac	caggcctggg	tactctaagc	agcccttggg	ttattcatag	gaagtggctg	10620
aggtttctct	gctatttcat	tttcagcctc	taccgtctgc	ccttggtggg	agcggctcac	10680
acttgcaaca	tcgacattca	actctattta	gttttctttc	ctcttcagac	atttagaggt	10740
gtacctattt	tgtcagggcg	tggttctagg	aatccaagat	aatgtctcag	tgtcccagcc	10800
aggggtgaccg	gctcattcca	gtttgccagg	gacttcaactg	gcttgagcaa	gggaagtcct	10860
gctccattcc	aggcagctgg	gctggctggg	cccgttagcc	ccaaccccg	gacagcagtg	10920
ccagaggggtg	ctctgtgagg	gatgggcagc	attctggcgg	cctgggaatg	agttgtgggtg	10980
tttccagggg	gtagaagtgg	gtacaagcca	caggtcacat	gatgagtggc	tgacctggct	11040
gggagggcag	aagaggggat	ggacttaggc	tcttcctttt	gctttgcaca	tatttaggat	11100
gtttgcagac	ttgctatgat	tgttgctggt	atgtgttttc	tgatgtgaaa	gatacacagt	11160

gtcctttgcc	catgagctct	ccttgccctcc	caggtcccca	gggcttatgc	ctggtgtcta	11220
ggcatcacct	ccctgcctgc	caggtgccag	gtgctgcatt	tcgggggagg	atgaactaat	11280
caccccgcg	cacctttcct	ctgagtggga	gcctggggca	ggtttgcatt	cctggaggcc	11340
gctggtggag	gggtctgggg	gcctgacttc	cactgcagcc	tgctgtcctg	gggaatgtgg	11400
cagggcaagc	ccagtgggga	gggctgtgca	cggccaggtg	cacccatcaa	aacagcaggg	11460
ctgcggtttg	tccctgtgga	gaagctaaac	acagctgcct	gggcactttg	taaatgctga	11520
gtggttcttt	gtctttctgg	gttacacacg	gaatcaggga	gccaagtcca	gccgggcagg	11580
gacgggggga	ggggaggagg	tgctgccgtc	ccttggcaag	agccttggga	actcacaagg	11640
aggctggagg	gcttgggaaga	aagaagagaa	ggccattgtc	tggtaggctc	tattctatct	11700
cgggtggtggt	ggtgggggga	ggcgcaacttc	ttttcctctt	tctgtgcagc	agttgccctt	11760
tgatgcctga	gttcttggtc	tgttttctgt	cgggcttctg	tgaataacca	catgtgccct	11820
ggcgctgtga	ccacacaggg	ctatccctac	cgaccttagg	attccttagga	aatgtcttct	11880
cttaaagggg	acatgtcttc	acttggccgt	gtcagtgcc	cagagccaga	gtccacctgg	11940
aatgcacctg	tagtcaactga	gaacccgggg	ggtgtgcctt	agtaagaagg	tgtcaggaag	12000
gacctattat	tgtagggcct	gggctcctgc	aagggtggtt	gggggtggtt	ggaggaagca	12060
gagatttgct	ctggattgga	tgctgtcagg	aagcaggggt	aattctgtga	ggctgcttta	12120
ttattttttt	tctaggagga	ggttggaatg	aggctaggct	aaagctgtga	ttggtaaaga	12180
aacgtccgtc	gctcaagtta	gccaggacag	gaggagacat	cagatcgtga	ttttgtggtt	12240
gtgagcacia	ggttcctggt	ctgtctgttc	agacatcatt	tcggaggagg	ctccttgtgt	12300
cttgcccat	ctcaggcatg	gaggggccta	gtccgatatt	gacgctcagt	gaaataattc	12360
aggttccgca	gagcacacgg	cccagctatc	agggcggggc	agctctgcat	gccagggggc	12420
gcgtcttccc	ttctcagcat	agcctgggaa	attcactgca	ggacaaaatg	catcagttac	12480
ttcctcttca	tccataacct	gggatgtttg	actcccaa	gagtaactct	tacgtttctt	12540
ctaatacctag	ggaaactatt	ggttatattg	ctttcaacac	tacaaattta	aagcagttat	12600
aggagcccag	aggtttccaa	atggcttcct	taaaaattag	aagatgattt	taaattccaa	12660
gaggaaaaac	aaaactagca	ttattgtata	cttaccctca	caaccgtcct	aggagctggt	12720
acaattttta	gagagggtta	gtaacttgcc	caaggtcaca	ctgtggggat	gtgagccgcg	12780
taccttggtc	cagtgtctgg	tctttgccac	tgtccctata	tggtttact	taccttattg	12840
gagttgtaac	tagcagaccc	ttctatgtct	cagaagacag	gagagggaac	atcggaagaa	12900
atgactgatt	tctaagcatg	tgagaggcag	gtgactccgc	actatcgtga	ccagaatttc	12960
ccctgttctt	tttgacgtga	tgctgttct	tcaaaagcag	tggtttcttt	acgctgtata	13020

ggtaagttca tctggagtc cccttttgat acttctaact aggaaaagct ctctactttc	13080
agaacagtac tccctgtgtc tctgggggcg tgggagggaa gaaggtgggg tcacgggttg	13140
gaatgtgccc agcggcgtct cgctctttcc aaggagctcc tggtttagat ttccatggcc	13200
tgtagacacc ttcagccttg ggtccaaggg acaccccctg agatcaggca cgctcaagaa	13260
gctgacaaag ccctacactt tatgccaccc atgagctgga ggcccggcag gtctctttct	13320
ccagaaagca aaggggggtg gcgttagtga gccctggcag ccacctaacg tggacttgga	13380
gcatctgctg ggctgtggtc cagcaccacc gtgtggccac caggtgctca tcagccagt	13440
ggacccggga ggagggacaa gaccagagaa caacagtgt cttgcctctt ctctcctgaa	13500
ttttggacgg tggcttagac ttgggtgtcc ccatctctgt gtttagagt cttacagttt	13560
ccaaactgtt tgcaaagtgt gaagccaccg tccctctcct ctgggatggc ccagtgtgt	13620
cgtggggccg tggctcctgag ctacgctttt catttgaaga ggtggaagga gctgacaccg	13680
tcccatcccg gcagggtctg ctacggtctt ctttaggtcc tgagtggggg tccagcacag	13740
ccccaaaggt gcgtggcacc cgccctgccc tctgcccctg cactcatctc ctggtggaga	13800
agacactcac acacaggaag cagggaaggc agcagacctc actcaccctt caccctctca	13860
ctcaccctt actcaccctt tcaacctctc attcaccacc caccctctcg cccctcact	13920
caccctctca ctccctcaac cctcactcac ctccctcact cctcaaccct cactcacctc	13980
ctcactcct cactctcccc ctcatccctc cctcaccctc ccccgctacc tctcactca	14040
cctcctcacc cctcactca cccttcacc cctcactcac cacctcacct cctcactcac	14100
cccctactca accctcatt caccctcac cccctcactc accctgcac cccctcactc	14160
acccttcat ccaactaccc acctgctcac ctccctcactc aaccctcac cccctcacta	14220
atccctcact ccctcaccct ctacgacct cactcacacc ttacctcct cactcaccct	14280
ctcaccctt caaccctta cttaccctt cactcatccc ttaccctc actcaccctt	14340
tctctcacc attcaccctt tcatcatgc cttcaccctt tactcacct cctcactcac	14400
accttcacc ctacgtcacc ccctcactca ccccttcacc ccctcaatca tgccctcact	14460
ccctcactca ccccttcacc ctctgaatta ctccctcact ccctcactca cccctcact	14520
cacccttca cccctcacc caccactca cccaccctc accaccctt taccctctt	14580
accctcacc cccctcactc accctcacc ccctcactca ccactcacc caccctcac	14640
ccaccctc actcactccc tcatccctc actcaccctt tcaccctc actcaccctt	14700
tcaccaccc ctacccacc ccctcaccct ctactcacc ccttcaccct ctactcacc	14760
ccctcactca ccccttcacc ccctcactca ccactcacc caccctcac ccaccctc	14820
actcactccc tcaccctc actcaccctt tcaccctc actcaccctt tcatctctc	14880
actcaccctt tcactctc actcaccctt tcactctc actcaccctt tcgccccctc	14940

actcaccct	cacccctca	ccccctact	cacccctcac	ccctcgccc	cctcactcac	15000
ccctcgccc	cctcactcac	ccctcacccc	ctcacccct	cactcatccc	ctcacctcct	15060
cactcacccc	ctcacctcct	cactcacccc	ctcacctcct	cactcacccc	ctcacctcct	15120
cacccacccc	ctcactcact	ccctcacccc	ctcacccct	cactcacccc	ctcacctcct	15180
cactcacccc	ctcacctcct	cacccacccc	ctcactcact	ccctcacccc	ctcacccct	15240
cactcacccc	ctcacctcct	cactcacccc	ctcacctcct	cactcacccc	ctcacctcct	15300
cactcatgcc	ctcacccct	cactcacct	ttcacctcct	tgctcatccc	ctcacttacc	15360
ccctcacttc	gtcaatcacc	ccccacctc	gtcaatcacc	ccctcacctt	ttcactcacc	15420
ccctcactca	cccccttact	tcctcactta	cctcctcacc	cccactcac	cccctcaccc	15480
cccactcacc	ccctcacccc	aactcacccc	cctcaccccc	cactcacccc	ctcacccctc	15540
tcacctcctc	actcaccccc	tcacctcctc	acttatcccc	tcacccctc	aattaccccc	15600
tcacccctc	aattactccc	tcactctttc	aattaccac	tcacccctc	acctcctcac	15660
tcctcactca	ctccctcact	caccccttca	ccttctcact	cacctcctcg	tctcctcacc	15720
ccctcactca	cttccagccc	tgccctccc	atcttctttt	tctttgtgtg	agaatctggg	15780
gtccctgagt	ggtgtcagtc	cctccaagac	tcaaggagtc	cccagggcct	tgttatccag	15840
aacaccccca	cctgggtccc	gggagacccc	atgggatcac	aggagtgttc	aggggaagtgg	15900
tgcttcctgg	gtctgggtgg	gctggagggg	catcctccct	tccccaagag	gagaccccca	15960
ggagccccct	aagtccatcc	ccagcagtgg	tgccctgcc	ctgtccttgc	agcctgggag	16020
acccttgggg	ggggcgggcg	ctgggtggct	gggcggcttc	tgctggctc	acccactgg	16080
cctcctgttt	gtcatcctca	gcctgcgggg	tcaacttgaa	ctcaagccgc	cagagcagga	16140
ttgtgggcgg	cgagagcgcg	ctcccgggg	cctggccctg	gcaggtcagc	ctgcacgtcc	16200
agaacgtcca	cgtgtgcgga	ggctccatca	tcacccccga	gtggatcgtg	acagccgccc	16260
actgcgtgga	aaagtatgcc	aggggcggcg	cgggcggggt	gggggctcag	ggctggccta	16320
cagccaccct	gtgaccttga	gcaggtctca	acccttgag	ccccggcatc	cttgtgttta	16380
aatggggaga	gtattgcacc	tgcttcctag	ggctgtgaga	catcaagtgc	gctcatgcca	16440
ggcagtgcac	ggctgtatgc	actgagtgtc	ccctgcacgc	agggcacagg	gtgcaggtgg	16500
aacattctcc	acgatgtcgc	cgtgaccagc	gttccttcca	gccactgtcc	tctgagctct	16560
gtcctgccct	tgagcaaagc	ccctgcccc	tgaggtatcc	tgtctccggg	acgctagtcc	16620
caggagaggg	cacactcaga	caggcttcag	gctgccctgc	tggaagggtcc	ctgggggttaa	16680
gcgttcttgg	ccacagcatt	gctcatgcag	aggggttaggt	aggggtgagg	ctagccgtga	16740
cagtattagc	atttatggac	gctaccaccc	cctccccctt	tccttaaaca	catagtgcct	16800

ES 2 886 958 T3

ttggtcacat	gctgcttttg	aggaggcctc	acttggcgga	tgtatttttc	tgccttagag	16860
agaggctgaa	ctgggtttga	ctgttgccc	agccctctct	tgctgcgtgc	ccttagacga	16920
ttcactcaac	gtctctgac	catggcatgt	acaactataa	gatgggcatg	cccttctcct	16980
ctcgggctgt	tatgaaggtc	aaggaagcaa	gggctgttac	ccaagggtgc	tcccttctct	17040
ccccctcttc	acacccccag	gtgctctggg	ccctctagga	actgggtttc	tctcaagggc	17100
tgttacccaa	gggtgctccc	ttctctcccc	ctcttcacac	cactgggtgc	tctgggccca	17160
ctaggagctg	ggattctctt	aagagggaaa	ctcttgata	aaggaaatgg	tttgattgat	17220
atcggacaag	tctgttcatt	agtatccatt	tattaagcac	ctaccatgtg	ccaggaaatg	17280
ctttggcgta	caaaggaaaa	taagggccag	tcctgctaga	aatggccttg	aaaccccagg	17340
gagggatgtc	ggccattgt	gggtgctgca	gattccttga	aggtgatgca	agagccagaa	17400
agaaggatga	tgtggggggc	tgaggcagg	agtcgggggt	gggggagtgt	gggggagaa	17460
gggagaccga	gcacctcttc	cactatctcc	ctgtgtgggt	tttgggtgac	catcctgcct	17520
ctgggtgtct	tgcctccagc	ttctgacgtt	ggaagttcat	ccactgagag	ctctgtgttt	17580
atggctctga	gatactgagt	ccttcttctc	tcccagacct	cttaacaatc	catggcattg	17640
gacggcattt	gcggggattt	tgagacaatc	tttcatgttc	tatggagccg	gataccaagt	17700
agaaaaagtg	atttctcatc	caaattatga	ctccaagacc	aagaacaatg	acattgogct	17760
gatgaagctg	cagaagcctc	tgactttcaa	cggtacgtgt	ggctcaggct	tggcaagcag	17820
gttggcagaa	tcttaaagag	atgttgattg	gaaatgacac	ttgtgctatg	ccaaatggaa	17880
gggaggcatt	tgcgttgagc	gagggtagcg	tgcagcgggt	ggccaatggg	agaggctcac	17940
agaggctaag	agcacctgcc	gcattttggg	ggaggcagca	gccaccacat	ctgttctgta	18000
ctgtactgag	tggtggtgat	tcaagccagg	catggaaaag	gctagaacag	ggctttccca	18060
ctgcagcacc	cttgacatct	gggtggttct	ctgttgtagg	gctctcttgt	gccttgtagg	18120
atgtttaaca	gcgtccccag	cctctaccca	ctggaggcca	gtagctacca	agctgtgaca	18180
accagtgttg	cctgctgaca	ttgccaaaaca	tccgctttga	ggcaaagtca	cttcagttg	18240
agaactactg	gcctaaaatg	tgtaaagatc	cttgattttt	aaagatacat	tctaaaacca	18300
agttgcttaa	ttcaggacaa	acatgctttc	tottagcctc	ttattcggtc	ccactctggg	18360
ccatccaagg	gtctggaatg	ttctagcccc	atgtggatac	agaagaagca	aaacctcagc	18420
cctccctaca	gcatgtctgt	attcacattg	ggaaatgggt	cacatataga	agagcgaatg	18480
cctgagcaat	ggcgtgggtgc	ctctggggcg	aaagctgact	ccattgactc	catcggtttt	18540
ttggctgttg	cctcctgtgt	gtctttcccg	tcttgatcac	ctggagatat	gtaattttgg	18600
aagcagagct	agcaaataat	tcctcttata	agcagagcta	gcaaataatt	ctacttataa	18660
gtagcataac	gtcttgcttg	ccagaaggag	aggtctggca	gggggagaaa	gtgagaatgt	18720

gggacttggt	gggatgcagg	gtcctctggg	cagggtggcc	agggtgccag	gccagcagc	18780
ctgcatgtgg	gaaggccagg	tggagacata	ggtgataccc	gcctggctca	ctgtgttttc	18840
tcttcttgaa	acagacctag	tgaaccaggt	gtgtctgccc	aaccagggca	tgatgctgca	18900
gccagaacag	ctctgctgga	tttccgggtg	gggggccacc	gaggagaaag	gtgaggctgc	18960
tcctgggcac	acaggactgc	agggccaca	gatggagcat	tgggttcgga	agtgggaggt	19020
ccaggtttta	atcccagttc	tactactcaa	tgactggatg	actttggttg	attccccag	19080
tccttggtgc	tcagtttctc	catctgctaa	gtgggagaaa	tcctgccag	cctacctaat	19140
acactgtggt	cttatcgtga	tcacacagag	cagcatgtgg	aatggctttt	gaagtatctg	19200
ggccatacga	gtttagaggt	gcaggatctc	ctgtgttgca	ctcattgtga	gtttagagct	19260
gccctggaga	tcccaccaag	gcctgcgtgg	ctgagtgaca	gggggcttgg	tgaggacggg	19320
catcctggac	ccatggtggc	cacatctaag	cctgtcctct	gccctgataa	ccacagagag	19380
aggctctctc	caccacttcc	ctttgcaatc	tgcatttctc	tctgacagtc	tttcaaatga	19440
agggagcctg	gctgcttcat	ttttatggag	ggttggaagt	gcttagtggc	aggcaciaaag	19500
gttcatttta	catattgttt	atatccttct	caaaagcgtc	taggccatac	agacaacaaa	19560
tcctttcaaa	caaggggaaa	agtacaaagg	ttgggtgatt	tctggggagc	gtcaggggaag	19620
gtagtggggg	gcatcctggc	tcctcatcag	cagaaactta	ctacagtaga	gccacaggct	19680
gggcaaaaga	cctcatggaa	tccaagatga	agggaatatc	gacaaatatt	tgtgcgcacc	19740
tgcacctagt	acaggctggg	tgctactcag	gtgctgggaa	tgcagaagtg	aacagagtaa	19800
gacaaatgtc	tctgctgtca	ggagctttac	ctctcttctg	gatgtcgggtg	gtggggacgg	19860
ggcaggtgtg	gtcagacaga	tgggagacaa	acaactgagc	gaggtacttc	caaacatctg	19920
aggggtggga	tcacaaggtc	ccggctattt	tgaaggggtg	gtcaggaaag	gcttctcgga	19980
agaggtggca	tttgagctga	gactcaaatg	gcaaaaatgt	gtacacatca	aaaaggctag	20040
tgcatgtatc	ttcaggtgtg	gtcaaggggc	caaggaggtg	ggctggggcc	agattgcata	20100
ggtccttggtg	gattatggtg	aagacaccag	cttctcatct	gcttgaggtg	gggagatcgt	20160
gagccgggga	gtgccatgat	ctggcagctg	cgtggggagt	gggatgaat	ggatggagac	20220
gaggatgatg	gtgacaagtc	cattgctgtg	gttccttgag	acaggaagcc	agctcatagc	20280
agagtgcggg	cgtggatgtg	aagagatgag	ggtacactag	ggctagagcc	accagactta	20340
ctgatgggtt	gcatgtctgt	gggagagaga	gtgagaagtc	agggacgatg	gctttccact	20400
ctgtggctga	agccccaggg	tggcgggtgg	tgccattttt	caagccagga	aatattgggt	20460
ggtgagaatt	tgggggtggga	gaaggtgtga	cggaggggtc	tggttttgca	cactaagccc	20520
acgggtgcca	gaagatgcc	gaggggaggc	agcaaagcga	gagtgggaaa	tgcagaggtg	20580

gcaagtgcag gccgtgtctt gagaagctct aatgtgcagg ggagccgaga agcaggcggc	20640
ctagggaggg tcacgtgtgc tccagaagag tgtgtgcatg ccagagggga aacaggcgcc	20700
tgtgtgtcct ggggtggggtt cagtgaggag tgggaaattg gttcagcaga accaagccgt	20760
tgggtgaata agagggggat tccatggcac tgatagagcc ctatagtttc agagctggga	20820
atttctttcc ctgaagctga actccagagc tgcattcagc acaggcaccg ccagttgtaa	20880
ggagaatcca ggtttcccag gagaggggtt ggtgctggga tgagctgacc ggggcagggc	20940
tggaaaatag ggctgtgacc atctgtgtag tgcgtgtgga ggtctcaggg agggaaagtgt	21000
gctctccctg cgagagctgc aggcaacact gggagctcaa caagtctccc tgtccttagg	21060
gaagacctca gaagtgtga acgctgccaa ggtgcttctc attgagacac agagatgcaa	21120
cagcagatat gtctatgaca acctgatcac accagccatg atctgtgccg gcttcctgca	21180
ggggaacgtc gattcttgcc aggtaattca acatttttat tctacctttg gtccttacca	21240
gctcctactg aaccccccat gagagagagg gcattcttgg ggtcagcaga gcctcctcag	21300
tgacacggag ccagctcggg gcagtcattg gaagtgcagg ccacaaacag tgcgaacgct	21360
tctgggtggca gaaggaagta cagtcaacaa atcacacaca ccctctgaaa aaccggtatt	21420
tggtaaaagt gccagtggaa cagaaacaag tatttagact attttaaatt atgaacggca	21480
atttatttag taacttttag cttgaacaga ttaaaattca ggatgggggc tatctctttg	21540
ggggttacat ctctgttacc atcaccctt gatggtggag attcgaagcc cacacagtca	21600
ctcgtaactc aactgcgac ccccgcccc caactcctct aggcctggtc agtgggtgtgc	21660
ggcagattgt gacttgattt tctgtctct gtaccttgt gtgtcccaca gggtgacagt	21720
ggagggcctc tggtcacttc gaagaacaat atctggtggc tgatagggga tacaagctgg	21780
ggttctggct gtgccaaagc ttacagacca ggagtgtacg ggaatgtgat ggtattcacg	21840
gactggattt atcgacaaat gagggtaact atcctgtcct ccttctgact gtgttctccg	21900
attcctcgag ccaaagccag acatctgtta ggcgtggttc tgctgctgga agctgactgg	21960
tgaccactgg tcagcatgaa gcaaactctg cttcctccag ccacagcccc atccccag	22020
tgtccacca ttgccattg cctctcactg gcttcacttg catatttccc ctgggtgttg	22080
gatgaaaagc gctggggctc agcttgtgtg aaattccttg gtgctctgcc aaccacactt	22140
cgttctggct cagctgactc agctgttcca ccaggccac ctcacatcaa actttttttt	22200
tttttttttg agatggagtc tcaactgtgtc gccaggctg gagtgacgtg gcacaatctc	22260
gactcactgc aacctttgcc tcctgggttc aagtgattct cctgcctcag cctcccaagt	22320
agctgggact acaggcatgc gccaccacgc ccagctactt tttgtatttt tagtagagat	22380
ggggtttctc catgttggtc aggtgtgtc cgaagccctg acctcaggtg attcaccac	22440
ctcagcctcc cacagtgtgt ggattacaag tgtgaaccac ggtgcccggc ctcacatgaa	22500

acttttgatt	tatagagagc	agagggaaga	gccggctgtg	cccatccttt	tctggggcca	22560
tcgagtggct	cctgggcagc	ccccaaaggtt	aggaagggca	ggagcagcca	gggttctctg	22620
atgccccaga	ctcaagcacg	agggaaggtc	tcaggggttc	catgtgagcc	tcatggatgt	22680
ctctgcttag	cagagccctg	gctttgggca	ttgtccagat	agggggtgag	aaccagatct	22740
tctcatctcc	aggacctcag	acgtatatgtt	ttctcagatt	tctgtgcttt	ctggggctgg	22800
gctactagt	gaagaaagca	gtctattctg	tcttctccca	aatctcccag	atgccagtc	22860
tgttgaagga	ggagcagaac	cagggggcct	ttcccgtga	ggcccagcct	gtgtctcctt	22920
caaatgacac	gcgggactca	gggccttccc	atgaccatgg	ggcccagggg	gcgtcacctg	22980
gcccagggcc	cagtgcctaga	aacagatgac	cccaggagga	ggaggcaggg	caggagggaa	23040
gctggcaggg	ctgggatgg	cagccaggct	gaggggcgga	ctcgaccag	gatggagcta	23100
ggaaatgatc	caggtgtgtt	tggcggctgc	aggtgggtcc	gcatggctgt	gcagggaggg	23160
aagggtctgc	tggcaggaga	gcagccgggg	gagggcccaga	ctctgctgaa	gagatgcctg	23220
ttgtgccggc	ctccacatcc	gctgcccgct	ccttccggag	ctcctgcccc	gccatgctca	23280
gcctgactct	gaccaacacg	ttggagagaa	gaatgatccc	tttgtgctat	taagcttgct	23340
tatttggttt	ctaagtgctt	catgcgaacc	tagaggaaaa	aattattttc	cacctttgtt	23400
tgtcttaaga	aaataacaca	cttttttttt	tcctatttga	acaggcagac	ggctaatacca	23460
catggtcttc	gtccttgacg	tcgttttaca	agaaaacaat	ggggctgggt	ttgcttcccc	23520
gtgcatgatt	tactcttaga	gatgattcag	aggtcacttc	atttttatta	aacagtgaac	23580
ttgtctggct	ttggcactct	ctgccattct	gtgcaggctg	cagtggctcc	cctgcccagc	23640
ctgctctccc	taacccttg	tccgcaagg	gtgatggccg	gctggttgtg	ggcactggcg	23700
gtcaagtgtg	gaggagaggg	gtggaggctg	ccccattgag	atcttcctgc	tgagtccttt	23760
ccagggggcca	attttggatg	agcatggagc	tgtcacctct	cagctgctgg	atgacttgag	23820
atgaaaaagg	agagacatgg	aaaggagagc	agccagggtg	cacctgcagc	ggctgccctc	23880
tggggccact	tggtagtgtc	cccagcctac	ctctccacaa	ggggattttg	ctgatgggtt	23940
cttagagcct	tagcagccct	ggatggtggc	cagaaataaa	gggaccagcc	cttcatgggt	24000
ggtgacgtgg	tagtcacttg	taagggggaa	agaaacattt	ttgttcttat	ggggtgagaa	24060
tatagacagt	gcccttggtg	cgagggaagc	aattgaaaag	gaacttgccc	tgagcactcc	24120
tggtgcaggt	ctccacctgc	acattgggtg	gggtcctctg	gagggagact	cagccttcct	24180
cctcatcctc	cctgaccctg	ctcctagcac	cctggagagt	gcacatgccc	cttggctctg	24240
gcagggcgcc	aagtctggca	ccatgttggc	ctcttcaggc	ctgctagtca	ctggaaattg	24300
aggtccatgg	gggaaatcaa	ggatgctcag	tttaagggtac	actgtttcca	tgttatgttt	24360

ES 2 886 958 T3

```

ctacacattg ctacctcagt gtccttgga acttagcttt tgatgtctcc aagtagtcca 24420
ccttcattta actctttgaa actgtatcat ctttgccaag taagagtggg ggcctatttc 24480
agctgctttg acaaaatgac tggctcctga cttaacgttc tataaatgaa tgtgctgaag 24540
caaagtgcc atggtggcgg cgaagaagag aaagatgtgt tttgttttgg actctctgtg 24600
gtcccttcca atgctgtggg tttccaacca ggggaagggg cccttttgca ttgccaagtg 24660
ccataaccat gagcactact ctaccatggg tctgcctcct ggccaagcag gctggtttgc 24720
aagaatgaaa tgaatgattc tacagctagg acttaacctt gaaatggaaa gtcattgcaat 24780
cccatattgca ggatctgtct gtgcacatgc ctctgtagag agcagcattc ccagggacct 24840
tggaacacagt tggcactgta aggtgcttgc tccccaaagac acatcctaaa aggtgttgta 24900
atggtgaaaa cgtcttcctt ctttattgcc ccttcttatt tatgtgaaca actgtttgtc 24960
tttttttgta tcttttttaa actgtaaagt tcaattgtga aaatgaatat catgcaaata 25020
aattatgcaa tttttttttc aaagtaacta ctgcatcttt gaagttctgc ctggtgagta 25080
ggaccagcct ccatttcctt ataagggggg gatgttgagg ctgctggtca gaggaccaa 25140
ggtgaggcaa ggccagactt ggtgctcctg tggttctcga gataacttcg tataatgtat 25200
gctatacgaa gttatgctag taactataac ggtcctaagg tagcgagcta gctccacgtg 25260
gctttgtccc agacttcctt tgtcttcaac aaccttctgc aagaaaacca agggcctgaa 25320
ttttaacttc ctg 25333

```

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

<220>
 <223> Proteína recombinante
 <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

ccggttggtg	tataggactt	gaccagcccc	aatagtcctc	aagtcactcc	tagatacagt	60
ggcaggtggt	agctggcttg	cggaaggaag	aggaagaaga	gaatgtgggc	catcaaggag	120
caaggccagc	cttgcaacttg	ggccccctct	gctcagtgc	gaccagggct	ttctgagccg	180
cttcctaata	aggctcattt	gaagaccccc	ccccaccccc	ctcctgctgt	cttgggtggc	240
agagctagct	ccaggctgta	agaaaattag	gaggattacc	aaagcagtat	ggagtcagac	300
agtggccaac	ccctcaacaa	ccgtgatatt	gttccctttc	gcaaaccctg	aaggccccag	360
gagaccttca	aaaaggtggg	gatccccatc	attgcagtgc	tgctgagcct	gatagccctc	420

gtgattgtgg	cccttctcat	caaggtgatt	ctggataaat	actacttcat	ctgcggcagt	480
cccctgacct	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
agccagcccc	ctttcagagc	agtggagatc	cgtccagatc	agaacctccc	tgttgctcaa	780
gtcacaggaa	acagccagga	acttcaggtg	cagaatggaa	gcagatcctg	cctctcaggc	840
tccctggttt	ccttgcgctg	ccttgactgt	ggaaagagcc	tgaagactcc	tcgtgtggtg	900
ggtgggggtg	aggcccctgt	ggattcttgg	ccgtggcagg	tcagcatcca	gtacaacaag	960
cagcatgtct	gtggtgggag	catcctggat	ccccactgga	tcctcacagc	agccccactg	1020
ttcaggaagt	atcttgatgt	gtcaagctgg	aaggtcaggg	caggctcaaa	catactgggt	1080
aactctccat	ccttgccctgt	ggccaagatc	ttcatcgctg	aaccaatcc	tctgtacccc	1140
aaagagaagg	acattgccct	tgtaagctg	cagatgccac	tcacattctc	aggctcagtc	1200
aggcccatct	gcctgccctt	ctctgatgag	gtgcttgtcc	cagccacacc	agtctgggtc	1260
attggatggg	gctttacaga	agaaaacgga	ggaaagatgt	ctgacatgct	actgcaggca	1320
tcagtccagg	tcattgacag	cacacggtgc	aatgcagagg	atgcctacga	aggggaagtg	1380
accgctgaga	tgctgtgtgc	aggtaccca	caggggtggca	aggacacctg	ccagggtgac	1440
agtgggtggc	ctttgatgta	ccattctgac	aagtggcagg	tagtaggcat	cgtgagctgg	1500
ggccatggat	gcggcggccc	aagtactcct	ggagtgtata	ccaaggtcac	tgcctatctc	1560
aactggatct	acaatgttcg	gaagtctgag	atgtaacgct	gccgtcccc	acatccagaa	1620
gctgcttccc	ttcagaccta	cctacggcat	gaccctcaa	agtcagatat	gggacaagag	1680
cctccttgaa	caaactctgg	tatccctgca	gcaagcaagg	atacattgca	gaggtgcccc	1740
gagtggagtc	agatgggcta	gctcagccac	ccctgcatct	cccaaaccct	gggagacatg	1800
tggcccatgg	gagtaaattcc	aggacattga	ctcaactctc	agaagtgtta	ttcagtcaag	1860
gaggctctcc	cttccactga	aggaaggaaa	gtcagctctc	tcctgaaagg	ccagatcact	1920
ggctgagtag	atgagacaag	ggtatgaaag	gcctttgcca	tcttctttgc	ccagtcctga	1980
aagcactgac	gtaagagacc	agtcagttct	aatgtaagg	gtatatttta	gtgtcaggg	2040
attgcaattg	tcacctctgt	ggtcaatatc	attaaacagg	tatgagaatt	cgctggcata	2100
gacttctctg	tctgcttaat	aagaatccaa	ctaaggatgt	cacatgacag	tttcccagaa	2160
aatgtgaaca	agtgtccatc	tgacacacgg	caccaatgac	aaaccaaaga	agttattctg	2220
cctgagtctc	agttgctgaa	ctaataaatt	agctgcgggt	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

atcattccag ttggaact tcaattgtag ggctgttta atcaagctgc ccaaagtccc

60

ES 2 886 958 T3

ccaatcactc	ctggaataca	cagagagagg	cagcagcttg	ctcagcggac	aaggatgctg	120
ggcgtgaggg	accaaggcct	gccctgcact	cgggcctcct	ccagccagtg	ctgaccaggg	180
acttctgacc	tgctggccag	ccaggacctg	tgtggggagg	ccctcctgct	gccttgggggt	240
gacaatctca	gctccaggct	acagggagac	cgggaggatc	acagagccag	catggatcct	300
gacagtgatc	aacctctgaa	cagcctcgat	gtcaaaccct	tgcgcaaacc	ccgtatcccc	360
atggagacct	tcagaaaggt	ggggatcccc	atcatcatag	cactactgag	cctggcgagt	420
atcatcattg	tggttgtcct	catcaaggtg	attctggata	aatactactt	cctctgcggg	480
cagcctctcc	acttcatccc	gaggaagcag	ctgtgtgacg	gagagctgga	ctgtcccttg	540
ggggaggacg	aggagcactg	tgtcaagagc	ttccccgaag	ggcctgcagt	ggcagtccgc	600
ctctccaagg	accgatccac	actgcagggtg	ctggactcgg	ccacagggaa	ctggttctct	660
gcctgtttcg	acaacttcac	agaagctctc	gctgagacag	cctgtaggca	gatgggctac	720
agcagcaaac	ccactttcag	agctgtggag	attggcccag	accaggatct	ggatgttggt	780
gaaatcacag	aaaacagcca	ggagcttcgc	atgcggaact	caagtgggcc	ctgtctctca	840
ggctccctgg	tctccctgca	ctgtcttgcc	tgtgggaaga	gcctgaagac	cccccggtg	900
gtgggtgggg	aggaggcctc	tgtggattct	tggccttggc	aggtcagcat	ccagtacgac	960
aaacagcacg	tctgtggagg	gagcatcctg	gacccccact	gggtcctcac	ggcagccac	1020
tgcttcagga	aacataccga	tgtgttcaac	tggaaggtgc	gggcaggctc	agacaaactg	1080
ggcagcttcc	catccctggc	tgtggccaag	atcatcatca	ttgaattcaa	ccccatgtac	1140
cccaaagaca	atgacatcgc	cctcatgaag	ctgcagttcc	cactcacttt	ctcaggcaca	1200
gtcaggccca	tctgtctgcc	cttctttgat	gaggagctca	ctccagccac	cccactctgg	1260
atcattggat	ggggctttac	gaagcagaat	ggagggaaga	tgtctgacat	actgctgcag	1320
gcgtcagtcc	aggtcattga	cagcacacgg	tgcaatgcag	acgatgcgta	ccagggggaa	1380
gtcaccgaga	agatgatgtg	tgcaggcatc	ccggaagggg	gtgtggacac	ctgccaggggt	1440
gacagtgggtg	ggcccctgat	gtaccaatct	gaccagtggc	atgtggtggg	catcgtagt	1500
tggggctatg	gctgcggggg	cccagcacc	ccaggagtat	acaccaaggt	ctcagcctat	1560
ctcaactgga	tctacaatgt	ctggaaggct	gagctgtaat	gctgctgcc	ctttgcagtg	1620
ctgggagccg	cttccttcct	gccctgccca	cctggggatc	ccccaaagtc	agacacagag	1680
caagagtccc	cttgggtaca	cccctctgcc	cacagcctca	gcatttcttg	gagcagcaaa	1740
gggcctcaat	tcctataaga	gaccctcgca	gccagaggc	gccagagga	agtcagcagc	1800
cctagctcgg	ccacacttgg	tgtctccagc	atcccaggga	gagacacagc	ccactgaaca	1860
aggtctcagg	ggtattgcta	agccaagaag	gaactttccc	acactactga	atggaagcag	1920

gctgtcttgt aaaagcccag atcactgtgg gctggagagg agaaggaaag ggtctgcgcc 1980
agccctgtcc gtcttcaccc atccccaagc ctactagagc aagaaaccag ttgtaataata 2040
aaatgcactg ccctactggt ggtatgacta ccgttaccta ctgttgtcat tgttattaca 2100
gctatggcca ctattattaa agagctgtgt aacatctctg gcataggcta gctggaatgc 2160
ttgataagaa ctgagctggg atgattgaac tttcattctt tggcttgggg agaaaagaag 2220
tcctggggaa gcaattgagt ctcaaagtag aggcagggga aaaaagagtt agggagacca 2280
gatctgctga gtggcagcaa gagtgaagctg cagattacag aaaccagggt gagcaagttt 2340
gagtcccaca cagggccttc tccctttgcc tctttccctc cctccctgcc tgtgataatc 2400
agccaggagc cagggataac ctatgacttg ggaaagagat gagttaggca gtcaagggtg 2460
acattcaatc agggatccac aagtggctgg aaagaaatgc tggtcctgtg tcctaacttt 2520
ttccgcctgg agagccctca gtgtggcttc ttacatttaa aaaacaaaaa ggatcagctg 2580
ccagggtgtga ggcagtcacc aagctgagtt gtgaggatgt aagcatgaat aagtcctgc 2640
actcaaaatg gtcaaagaat taaaccccat ggactttttt ggcatctgta tgaaagcttg 2700
ggttttctga ggactgtctt gctatagtta agtcagatcc tagatgaaat atacttgttc 2760
atactgtact aggttcttag gaaacaacag aattcctcaa atgccaaaaa caaagaaaat 2820
agaaaccag aaaacaaaac aaaataaaac aaaaccatca gaactgtgag tggaaactaa 2880
ggtgatgatc tgggagcaat aactaaaaat cttgggtcga gacctatatg aaggctggca 2940
gtggagctaa acctggacac actgaagaca agggagctga accagggctc ctacatgaag 3000
cagggataac tgatggcagt aaatgtggtc tcaaattgca gatggctctg aggaaaattt 3060
cccaaattta gagcctcagg attcccaaag atcctccaaa tatgagctca caatcaaaga 3120
tcagagacgt tgaaaaataa aaaacacctt aagtgggcag cataaaaaac agctaattta 3180
gaaccccaaa ggcttcagat gtcagaatat tagagactta tgataataag caatatttgc 3240
agagtatttg tatgtgccag aactattgt aagtgcctca tcatgtactg attcatttaa 3300
tactcacaga aatctgtgag atgggtatta ttcttatcct cactctatgg attaaaaaaa 3360
ctaaggcaca aagtggttaa gtccttgcc tgagattata gactgtaagt tgaacgtgag 3420
cacttggaat acagagttca tgctgtaaac taccacacta tagggcctcc aatatgataa 3480
tttataaaat atttgaataa aaaatgaata ctagttccac attttaaaaa 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 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 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 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 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 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

<400> 12

ccacccgcac aactacagt cgagataact tcgtataatg tatgctatac gaagttatat 60
gcattggcctc cgcgccgggt ttggcgccct cccgcggggc ccccccctct caccggcgagc 120

gctgccacgt	cagacgaagg	gcgcagcgag	cgctcctgatc	cttccgcccc	gacgctcagg	180
acagcggccc	gctgctcata	agactcggcc	ttagaacccc	agtatcagca	gaaggacatt	240
ttaggacggg	acttgggtga	ctctagggca	ctggttttct	ttccagagag	cggaacaggc	300
gaggaaaagt	agtcccttct	cggcgattct	gcggagggat	ctccgtgggg	cggtgaacgc	360
cgatgattat	ataaggacgc	gccgggtgtg	gcacagctag	ttccgtcgca	gccgggattt	420
gggtcgcggt	tcttgtttgt	ggatcgctgt	gatcgctact	tggtgagtag	cgggctgctg	480
ggctggccgg	ggctttcgtg	gccgccgggc	cgctcggtag	gacggaagcg	tgtggagaga	540
ccgccaaggg	ctgtagtctg	ggtccgcgag	caaggttgcc	ctgaactggg	ggttgggggg	600
agcgcagcaa	aatggcggct	gttcccagat	cttgaatgga	agacgcttgt	gaggcgggct	660
gtgaggtcgt	tgaacaagg	tggggggcat	ggtgggcggc	aagaacccaa	ggtcttgagg	720
ccttcgctaa	tgcgggaaag	ctcttattcg	ggtgagatgg	gctggggcac	catctgggga	780
ccctgacgtg	aagtttgtca	ctgactggag	aactcggttt	gtcgtctgtt	gcgggggagg	840
cagttatggc	ggtgccgttg	ggcagtgac	ccgtaccttt	gggagcggcg	gccctcgtcg	900
tgtcgtgacg	tcaccggttc	tgttggctta	taatgcaggg	tggggccacc	tgccggtagg	960
tgtgcggtag	gcttttctcc	gtcgcaggac	gcagggttgc	ggcctagggg	aggctctcct	1020
gaatcgacag	gcgccggacc	tctggtgagg	ggagggataa	gtgaggcgct	agtttctttg	1080
gtcggtttta	tgtacctatc	ttcttaagta	gctgaagctc	cggttttgaa	ctatgcgctc	1140
ggggttggcg	agtgtgtttt	gtgaagtttt	ttaggcacct	tttgaaatgt	aatcatttgg	1200
gtcaatatgt	aattttcagt	gttagactag	taaattgtcc	gctaaattct	ggccgttttt	1260
ggcttttttg	ttagacgtgt	tgacaattaa	tcacggcat	agtatatcgg	catagtataa	1320
tacgacaagg	tgaggaacta	aaccatggga	tcggccattg	aacaagatgg	attgcacgca	1380
ggttctccgg	ccgcttgggt	ggagaggcta	ttcggctatg	actgggcaca	acagacaatc	1440
ggctgctctg	atgccgccgt	gttccggctg	tcagcgcagg	ggcgcccggg	tctttttgtc	1500
aagaccgacc	tgtccggtgc	cctgaatgaa	ctgcaggacg	aggcagcgcg	gctatcgtgg	1560
ctggccacga	cgggcgttcc	ttgcgcagct	gtgctcgacg	ttgtcactga	agcgggaagg	1620
gactggctgc	tattgggcga	agtgccgggg	caggatctcc	tgtcatctca	ccttgctcct	1680
gccgagaaag	tatccatcat	ggctgatgca	atgcggcggc	tgcatacgct	tgatccggct	1740
acctgcccac	tcgaccacca	agcgaaacat	cgcacgcagc	gagcacgtac	tcggatggaa	1800
gccggtcttg	tcgatcagga	tgatctggac	gaagagcatc	aggggctcgc	gccagccgaa	1860
ctgttcgcca	ggctcaaggc	gcgcacgccc	gacggcgatg	atctcgtcgt	gacccatggc	1920
gatgcctgct	tgccgaatat	catggtggaa	aatggccgct	tttctggatt	catcgactgt	1980
ggccggctgg	gtgtggcgga	ccgctatcag	gacatagcgt	tggtacctcg	tgatattgct	2040

gaagagcttg gcggcgaatg ggctgaccgc ttctctgtgc tttacggtat cgccgctccc	2100
gattcgcagc gcatcgctt ctatcgctt cttgacgagt tcttctgagg ggatccgctg	2160
taagtctgca gaaattgatg atctattaaa caataaagat gtccactaaa atggaagttt	2220
ttctgtcat actttgttaa gaagggtag aacagagtac ctacattttg aatggaagga	2280
ttggagctac ggggggtggg gtgggggtgg attagataaa tgcctgctct ttactgaagg	2340
ctctttacta ttgctttatg ataatgtttc atagttggat atcataattt aaacaagcaa	2400
aaccaaatta agggccagct cattcctccc actcatgac tatagatcta tagatctctc	2460
gtgggatcat tgtttttctc ttgattccca ctttgtggtt ctaagtactg tggtttccaa	2520
atgtgtcagt ttcatagcct gaagaacgag atcagcagcc tctgttccac atacacttca	2580
ttctcagtat tgttttgcca agttctaatt ccatcagacc tcgacctgca gccctagcc	2640
cgggcgccag tagcagcacc cacgtccacc ttctgtctag taatgtccaa cacctccctc	2700
agtccaaaca ctgctctgca tccatgtggc tcccatttat acctgaagca cttgatggg	2760
cctcaatgtt ttactagagc ccacccccct gcaactctga gaccctctgg atttgtctgt	2820
cagtgcctca ctggggcggt ggataatttc ttaaaaggctc aagttccctc agcagcattc	2880
tctgagcagt ctgaagatgt gtgcttttca cagttcaa atccatgtggct gtttcaccca	2940
cctgcctggc cttgggttat ctatcaggac ctagcctaga agcaggtgtg tggcacttaa	3000
cacctaaagt gagtgactaa ctgaacactc aagtggatgc catctttgtc acttcttgac	3060
tgtgacacaa gcaactcctg atgccaaagc cctgccacc cctctcatgc ccatatttg	3120
acatggtaca ggtcctcact ggccatggc tgtgaggtcc tggtcctctt tgacttcata	3180
attcctaggg gccactagta tctataagag gaagaggggtg ctggctccca ggccacagcc	3240
cacaaaattc cacctgctca caggttggct ggctcgacc aggtgggtgc ccctgctctg	3300
agccagctcc cgccaagcc agcaccatgg gtaccccaa gaagaagagg aaggtgcgta	3360
ccgatttaaa ttccaattta ctgaccgtac accaaaattt gcctgcatta ccggtcgatg	3420
caacgagtga tgaggttcgc aagaacctga tggacatgtt cagggatcgc caggcgttt	3480
ctgagcatac ctggaaaatg cttctgtccg tttgccggtc gtgggcggca tgggtgcaagt	3540
tgaataaccg gaaatggtt cccgcagaac ctgaagatgt tcgcgattat cttctatata	3600
ttcaggcgcg cggctctggc gtaaaaacta tccagcaaca tttgggccag ctaaactgc	3660
ttcatcgtcg gtccgggctg ccacgaccaa gtgacagcaa tgctgtttca ctggttatgc	3720
ggcggtaccg aaaagaaaac gttgatgccg gtgaacgtgc aaaacaggct ctagcgttcg	3780
aacgcactga tttcgaccag gttcgttcac tcatggaaaa tagtgatcgc tgccaggata	3840
tacgtaatct ggcatttctg gggattgctt ataaccacct gttacgtata gccgaaattg	3900

ccaggatcag	ggttaaagat	atctcacgta	ctgacggtgg	gagaatgtta	atccatattg	3960
gcagaacgaa	aacgctggtt	agcaccgcag	gtgtagagaa	ggcacttagc	ctgggggtaa	4020
ctaaactggt	cgagcgatgg	atttcogtct	ctggtgtagc	tgatgatccg	aataactacc	4080
tgttttgccg	ggtcagaaaa	aatggtgttg	ccgcgccatc	tgccaccagc	cagctatcaa	4140
ctcgcgccct	ggaagggatt	tttgaagcaa	ctcatcgatt	gatttacggc	gctaaggtaa	4200
atataaaatt	tttaagtgtg	taatgtgtta	aactactgat	tctaattggt	tgtgtatttt	4260
aggatgactc	tggtcagaga	tacctggcct	ggtctggaca	cagtgcccg	gtcggagccg	4320
cgcgagatat	ggcccgcgct	ggagtttcaa	taccggagat	catgcaagct	ggtggctgga	4380
ccaatgtaaa	tattgtcatg	aactatatcc	gtaacctgga	tagtgaaaca	ggggcaatgg	4440
tgcgcctgct	ggaagatggc	gattgatcta	gataagtaat	gatcataatc	agccatatca	4500
catctgtaga	ggttttactt	gctttaaaaa	acctcccaca	cctccccctg	aacctgaaac	4560
ataaaatgaa	tgcaattggt	gttggttaa	ctgccctagt	tgcggccaat	tccagctgag	4620
cgtgcctccg	caccattacc	agttggtctg	gtgtcaaaaa	taataataac	cgggcagggg	4680
ggatctaagc	tctagataag	taatgatcat	aatcagccat	atcacatctg	tagaggtttt	4740
acttgcttta	aaaaacctcc	cacacctccc	cctgaacctg	aaacataaaa	tgaatgcaat	4800
tgttggtgtt	aacttgttta	ttgcagctta	taatggttac	aaataaagca	atagcatcac	4860
aaatttcaca	aataaagcat	ttttttcact	gcattctagt	tgtggtttgt	ccaaactcat	4920
caatgtatct	tatcatgtct	ggaataactt	cgtataatgt	atgctatacg	aagttatgct	4980
agtaactata	acggtcctaa	ggtagcgagc	tagccaagtc	tgtgtgctac	caagtagcaa	5040
aactgagcct	ggaactcaca	catgcgtgtc	tgagagccca	gcactatcgc	caggaaaacc	5100
cagcgtctcc	ctgctcaagc	ctgaccctca	gccctctctg	cctctccctg	cacttgccct	5160
ccagtcaagg	tgattctgga	taaatactac	ttcctctgcy	ggcagcctct	ccacttcac	5220
ccgaggaagc	agctgtgtga	cggagagctg	gactgtccct	tgggggagga	cgaggagcac	5280
tgtgtcaaga	gcttccccga	agggcctgca	gtggcaggtg	agtgcagggg	ctgaggcaca	5340
agagaagtgg	gccagcagg	aggtctgctc	aggcccccac	ggcccactgc	atagtatctg	5400
ccccctactt	gtcacttttc	atccttggtg	tataagggtc	tttggttggt	tgttggtgtg	5460
tgttttgagg	cagagtgtct	tgtggcccaa	gatggagtgc	agtgtcttgg	tctcggctca	5520
ctgcaacctc	tgctcccag	tttcaagtga	ttcttctgcc	tcagcctcat	gagtagctgg	5580
gattacaggt	gccagccacc	acgcctggct	aatttttata	tttttagtag	agacgggggt	5640
ttgccacatt	ggtcaggctg	atcttgaact	cctgacctca	ggtgatctgc	ccgcctcagc	5700
ctcccaaagt	gctgggatta	caggcgtgag	ccaccgtgcc	cagctgtgta	agtttcttga	5760
gagcaggacc	ctgtcttgtc	tacctttaa	tcctagtact	taacacacag	caaacagtaa	5820

ctatttgatg	accaaagtgt	agccagaaaag	gacaggaaat	tgtaaactgag	gctgccccat	5880
gcgtgctgcg	cctggtggat	ttcaggcaga	gggctagact	gggtgacctt	ggggcattcc	5940
tcctttctat	gaaatttggt	atttcaagga	gactagaaaa	gagacttctc	agccacttcg	6000
ccagctattg	gtccttctat	tcattagtgt	ttgctgagac	atgctatgtg	acaggactga	6060
gccaggtcct	ttcaatggat	aggagatggt	ttgagcataa	aatccacggt	ctctcttggg	6120
ctgggctctt	ctaccttctt	ccccctggtg	cttgggctct	gaagaaaaaa	agataggtag	6180
gagatgagtg	atggggcttc	tgagggcagg	gctgagtgac	tttctgtgta	tttgcctttt	6240
ctttatcaga	agtcaaatgc	ccacaggcac	ctgtcatcct	actgccagta	ggacttctca	6300
ctcaaccttc	ccctctgacc	ttacttggag	aaggacttag	gtccctctct	cagacatttc	6360
cccaggctgg	gcaagttgtg	tggaccatgg	atgggtatgt	ggtccataca	atttaaacia	6420
gctgtatatg	gtcgctgggt	agagtgacca	cataattgat	catcaaaact	gatacctgta	6480
agagcaaaaag	ggggcactat	taaccattgg	gtcagggcaa	caggtcaaaa	tggagaccta	6540
ccctgggact	tctggtcaca	ctagctactg	tcaaaatggg	gcccacaaatag	acaaagccaa	6600
atggaagaaa	ttcccttgac	attgaaagtg	ttggggctct	gtggcacccc	cagttctagg	6660
ttgggggagc	ttgggctggt	ctcatgatga	gttctgaggg	ggatggggcca	gttggggccc	6720
ccgttccatc	taactcaggt	tcctttcctc	ccagtccgcc	tctccaagga	ccgatccaca	6780
ctgcagggtg	tggactcggc	cacagggaac	tggttctctg	cctgtttcga	caacttcaca	6840
gaagctctcg	ctgagacagc	ctgtaggcag	atgggctaca	gcaggtaacc	aacctggggc	6900
tctctccttt	ttccctcctt	cctccttctt	cctcttctct	ctttccttcc	tccttctctc	6960
tctctttcct	aaaaattacg	ggcattggag	ccaggcagaa	tggcttttga	atcccagcat	7020
ttcacttata	agcaacatga	agttaaattt	cctaagcctc	aggttcctca	ggagtttaatt	7080
gggggaacta	atgccaacct	cataggatag	ttttgcaatg	ccagtgagag	aatgtgtgct	7140
gccctccaac	acacacacac	acacttctag	cgtctatgca	gtcctctcct	ttcctttact	7200
cctcaacctt	cactcctttg	tgctggcttt	gcaagaaact	gttcctgccc	agtaatacaa	7260
aagctaagtt	aacttattca	aagtttcgtt	agttaagatt	tagcttaagt	gagcctagtt	7320
tcagtggggc	cccatcttca	gcaatcccag	ctctctctgc	aaatttcaaa	agcagttcca	7380
aatctggagt	ggatgaaaag	gtgtaagatg	atagtaagag	taatttgcat	tctatatatt	7440
tatattcact	tgattttggc	agaaaaccaa	aaagatagtt	attatatctt	atatatagat	7500
atatattata	tctatttcat	aaataggctc	aaacaaagta	agtaacttgc	tagggtaacta	7560
gctgggaggt	agagggctag	aatttgagcc	caagaccctt	aattcttgcg	cattaggagt	7620
tcccacattg	tttctgtttc	tagactgagt	aattctttat	tctcatgtag	gacatcatct	7680

ctaaggggaag	gggctaata	gatgggtgat	cactcagaga	gtttagctgg	agaggatgga	7740
aaagaaccca	tacattcagt	tgacagattga	gatagcctat	ctctggcagg	cctcagattt	7800
cttcaggatt	ctaacagact	ggacccagag	actaggccaa	acaaacaaac	aaacaaaaac	7860
tctactaggc	agacatcacc	aaccaatcac	agaactctct	cccatggatc	cctaatacag	7920
cctcaaagtc	cttttcagta	aatgctccag	gcagccatta	caaatacaat	agaattattt	7980
gcctttctct	tctctgctca	acgggcttct	gctgctctct	actttccata	gggggcaact	8040
tccattaccc	tctagaaagc	acaccccacc	accttcattt	caaggagagt	gaggaactca	8100
tgcccagcac	ctgctattct	cccctcttcc	tgacagccag	gagcccagcc	tcgctgcagc	8160
cagccctgcc	tccccactgt	agtccagtca	actgctgcat	cagccgttcc	tggcacagca	8220
ggctgagcct	tgattatgaa	acctgggtgt	ctccaggggt	tcttaagatg	ataggctcct	8280
ggaatttctg	tccttttggg	gctcagtaag	gcaccaaacc	acctgagtct	tgtgcttcac	8340
aaaatcaaag	ttcatcagaa	tcattcattg	ggatggaatt	ggtgaacaga	agttaacttt	8400
cctgggaatg	tccatttcca	ccatattccg	tccttctagg	tctcagactt	ctctactttc	8460
tttcctctct	ctagatcgga	ggcccttctt	gtcctagaac	cataggcatt	tcaagatgtg	8520
ggagacccta	gggatcatct	agtccacgca	tctttttttt	ttttttttga	cagagtctca	8580
ctctgtcacc	caggctggag	tgcaatggca	ccatctctgc	ttactgcaac	ctccacctcc	8640
caggttcaag	tgattctttc	gcctcagcct	cccaagtagc	tgggattaca	ggcacgcacc	8700
atcatgcccc	gctaattttt	atatttttgt	agagaccgag	tttcaccatg	ttggccaggc	8760
tggtcttgaa	ctcctgacct	caggtgatcc	acccacctcg	gcctcccaa	gtgctgggat	8820
tacaggcgtg	agccactgca	cccagccccg	tgcatctttt	tatagagggg	gaaactgagg	8880
cttgagagga	cccagaaaaa	gaatatgacc	tgcccaaggc	cacacatcaa	actagtcca	8940
gagccaggga	cagaacctag	atcatgagga	ctcttaaaat	gcactctagt	cctcccagg	9000
ctgagacttg	ggtccttcca	ggaagtgcc	gcattcctgc	ctgagaatgt	gccaatccac	9060
cagtattgcc	aatgactcag	ccctccatgg	agagcttcta	ctaacattac	tagcatagtt	9120
agggatggaa	ggaaaagatt	tagaagaggc	agattcagta	aaggaacaat	cagagagatg	9180
gaattaatca	aggaaggctt	cctggaggag	gaaaaacttc	aaccaagggt	ttgaaagtag	9240
caagcatgga	ttagcaggga	gaaagaggga	gagtgggtcca	gttgagagaa	acgtttgtct	9300
ggattcatat	gaagacagat	ctagtcctgt	tctattaaat	atctctaagg	gggccaaaaa	9360
catacccccg	ctatcaaagt	cagaccagat	gctttgtttg	gagaacgaaa	tatccacatt	9420
ccaactccct	cccagggtgag	aaggagagcta	acctgagccc	ctatgcctct	ttgtttccct	9480
gctgtgaacc	agaagacatt	gctgggatat	ttgaaatagg	gacagagctg	ggaatatgga	9540
aaggagaccc	ctaacatttc	tccagggctc	tgggttctgg	atttgatttc	cccacccaag	9600

aaagcaagtt	acatcagcaa	tgcactgagg	gttgagtcct	gggatgccaa	gggtcggttc	9660
tttattgtat	agcaaagcag	gccccatctt	cactgactaa	gaccatctcc	actccctggc	9720
cactccccac	caagcattct	ctgccactct	ttctcctgaa	agtggggggc	aactctacca	9780
tcttgttcta	accccctgcc	ccagctcaca	actctctctc	cctcttgatg	tgagcagcaa	9840
acccactttc	agagctgtgg	agattggccc	agaccaggat	ctggatgttg	ttgaaatcac	9900
agaaaacagc	caggagcttc	gcatgcggaa	ctcaagtggg	taagtgaggg	gacaccttct	9960
ggcctacaga	aggccccac	atggacgctg	ctcttcaggt	tgcaaccagc	tcacctggaa	10020
ccccaaagcag	ccagggggaat	gtaagcagac	atcaggaaga	actcctagcc	agatggatca	10080
ttcaatgcca	agagctatag	actcacattt	tggagagggt	ttctgtgttg	acttggtttt	10140
aatacaatgg	acagctggac	aaagtgtgtt	gtcctactca	gagccagagg	gatggataat	10200
gtgacctttc	catcaatctg	gatagtaa	agtttttgct	actgctgtag	gttttcta	10260
aaattgcccc	ataggcaaga	ttccaaagtc	actttgtcct	tccctaccac	ttaccagcc	10320
agagctcccc	accttcttga	tgtctcaggg	aagaggctcc	atggcccttg	tgggtggcct	10380
gttcctgagc	ctcgccaccc	tgtgttagag	cagagcatcc	agatgaaatc	tgtcacactg	10440
tggcaaagtg	gctcagagag	gaggctggct	tcctagcatt	cagggacgtt	gctgagggcc	10500
gcttattcac	cgaaaataaa	tcttgaaaag	gacagggctg	gtagcagaat	gatcctttac	10560
ctaaaattct	atcaaaatcc	cattcttcca	tttggaaagc	ccacagtgtc	acagactctg	10620
ttccgggctc	tgtcctcttc	cctcttgggt	cccaggagcc	caggctgggc	tttgaagcag	10680
gcagggccca	gcacacagta	ggtactcagc	agtgggggtg	ttgaatccaa	tcaaacggaa	10740
gtgtcaatgc	aggaaatgca	atggatgtca	atgcagtctc	caaagtgtcc	ccactgtgca	10800
gcttccacat	tcccgaggta	ttgggagggg	acttgaatta	acagcttcgg	gaggcctgag	10860
tccctgcctc	ccagctgagg	aagaagctta	aatcacaggg	cgctgtgtct	gtcttccagg	10920
ccctgtctct	caggctccct	ggtctccctg	cactgtcttg	gtgagtacct	ccaatctctg	10980
agggtttggg	gcctgggcca	gcaatgagca	gggaggaaga	ccttcatctt	cactcctaaa	11040
tttctgggac	tccaagtttc	attctgcott	ggtctacagc	ccttgggctt	gtcgggtcaat	11100
gccccctcga	gttggttggtg	gccttgggca	ggtcacattc	tttttctggg	tctttccaag	11160
ccccagtttc	cccttcttac	catctgtgca	tggctccatg	acctaaagtg	agacctggga	11220
gagagtgtta	ggaagaccga	aaagggcagg	acggggcctc	cactgcctcc	catccctggt	11280
ccgggccccac	atagccttct	ttgtcacaat	cagctcaggt	atccaagatc	agattaccca	11340
cattcattat	ttgagcaact	attcattgaa	cagttagaat	atgtctcact	ctgtcagttg	11400
ctggctagaa	gtagaaagta	ccagatgagt	gaaataattg	gccactatcc	ttggtagctg	11460

atgactaagt	aagagagaga	tgcaagacaa	catgtggaaa	atgccaaact	gagtagcagt	11520
cacagttgac	atgctgcaga	gagagctggc	cgggggtcag	aagacctggg	caccagtcct	11580
gttcattttcc	agtgtggcct	cgagtcattc	acctgacctc	cctgaagttc	attttcccaa	11640
gaagttgttt	agtccaactg	cccatcaagg	atcttttaggg	acccttctag	ctctaacaga	11700
ggagatcaga	aaagaaaaca	agcaatgtgg	ctcagctcat	cctacaagct	tcatagagaa	11760
ctgagactgg	cctggaagca	tagccagaaa	ttagaacgcc	taagggaaga	aggtcacaac	11820
gctgcctctg	caatttagga	gtgtatatgc	tttcctgcag	gatgttgaga	gtttcattca	11880
ttatcgtatg	ccccctaccc	cggccccaca	atacctagtg	cgtgggatct	gacacgtggt	11940
ggctggtcaa	tgaatgaatg	aatgaatggt	cacaccatct	gaggttctgc	actgagtagc	12000
cctgaaggct	tgaagcagca	taagtgcag	gtcctccctt	gaggggcctc	tgttttacca	12060
ataagccaag	acctaagctc	aacaacactg	aaagggtggc	caatacccag	gacagcctgt	12120
gggaattcca	gagaaaggga	gattcccagg	gactgggggc	ccaggctaaa	cactgaaaaa	12180
tgcactctgta	ggctcaagga	ggaaaagccc	atgtctgtct	gtcttgccca	ccactctctc	12240
ccagcaccca	gcactgcccc	aggacagaga	gcacttgaca	caagttgggt	agattaatga	12300
atgatttaga	gttcagtggg	ccccaacctt	tttggcacia	gagactgggt	gcatggaaga	12360
caatttttcc	gcaaaccaag	agggggatag	agagcattag	attctctctt	tttttttttt	12420
ttgagaccaa	gtctggctct	tgtcactcag	cctggagtaa	agtgttgcca	tctcggtcca	12480
ctgcaacctc	cgctcctgg	attcaagcga	ttctcctgcc	tcagccccct	aaatagctgg	12540
gattacaggc	acctgtcacc	agcccagctg	ggactatagg	catgtgccac	catgcccggc	12600
taatttttgt	atttttagta	gagacggcgt	ttcaccatgt	tggccaggct	agtctcgaac	12660
tcctgacctc	aggtgatctg	ccgcctgag	cctcccaaag	tgctgggatt	acaggcatga	12720
gctgcctcac	ccagcctaaa	gtctcataag	gaacgtacag	catagatccc	tcacatgtgc	12780
agttcacaat	aaggttgtgc	tcctacaaga	atctaagccc	acctctgatc	tgacaggagg	12840
tgaagctcag	gtggatcatg	tcgcttgtcc	ctgccactca	cttcctaattg	tacagccagg	12900
ttcctaacag	gccacgaacc	agtgggaagg	gcactctttt	ggatcaaaaa	cagaattact	12960
ttttagagaa	ctacaagcag	atcaatttgg	ctagacagag	actttatatg	aaacagcagg	13020
aggctgctag	gaggagtggg	aactctactt	tgccctcaag	ggagatcccg	aagggtttg	13080
caggagcggg	caagggtggc	tgaagaaagc	agtgtttgaa	atcaggtggg	atttgaaaag	13140
cccagccctt	ccccttagaa	tggcccttct	accatctgtg	catggctcca	caaccgtggt	13200
ggtggctgcc	agaagaattg	gaaaggcaga	gcatgggtgg	agagggggga	cctgagggct	13260
ttacaggagt	tccgggggtg	gtgagggtgt	gaaagccagg	tcagtcagta	ggaagacagg	13320
atgtcagatt	gagagactcc	cctggccggg	gaaacagact	tggagaaggg	ggagttttgg	13380

atgagacagt	ccacttccga	gtcacaaaat	agcttgtggg	tgtctgttta	ctgttactca	13440
gtgggagtg	ctggggacac	gccacctggg	cagggctttc	gtaattctgc	atcacttgtg	13500
aaggtcacag	attcccagca	caacggacac	acccatgttc	atagtctgaa	ctcctaaaca	13560
catcttaaac	caaaataaaa	aaaaaagaaa	gaaagaaaga	aaaaggagag	ggagggtttga	13620
ggaaagccta	tggctctggga	caactcaatac	ctcccatgaa	tatctcatat	tgggctggtc	13680
ctctctccac	tctggcccca	gccataaggg	ccctgcttag	agcagatttt	gggtgctgag	13740
tggaggcagc	ctcatcccca	acagcctgac	ttcctgcctc	ctccctgcct	ctgcctgtgt	13800
ccagcctgtg	ggaagagcct	gaagaccccc	cgtgtggtgg	gtgtggagga	ggcctctgtg	13860
gattcttggc	cttggcaggt	cagcatccag	tacgacaaac	agcacgtctg	tggagggagc	13920
atcctggacc	cccactgggt	cctcacggca	gccactgct	tcaggtaaga	ccccagctgt	13980
aaggaggtct	ctggggacca	aggccagtca	gggaccagag	agcttggggg	cctgtctcct	14040
ggcaccgtcc	ttctcttcac	tctcccacta	gagacgtttt	ccaggttgtg	gtggcccca	14100
tgagacaatg	gccatgatgc	cctttgttag	gcttttgggt	gtctgagcag	agggtgctgg	14160
tcaccaagca	tggcctcttc	ctggtgggac	accagcagat	accagagtc	ctcacccac	14220
ccccatatcg	ttcaagctac	aaaagctctt	cccacctgcc	tcaacttcca	agaactcact	14280
ctctttttgc	ttgtttccag	gaagttgttc	cagggctctag	agtcatagcc	acgtcctcat	14340
tatgtctgga	aacttttaaaa	aaattaaaga	gcataggttc	ctttcagtcc	acagagaagc	14400
ctggccttac	ctcagggag	ggctactccc	agaccccctt	cacttttttt	tttttttttt	14460
tttttttttt	ttttgagaca	gagtcttgct	ctgttgctta	ggctggagcg	cagcagcatg	14520
atcttggtc	actgcaacct	ccgcctcctg	agttcaagca	attctcctgc	ctcagcttcc	14580
caagtagctg	ggactatagg	catgggccac	catgcccggc	taatttttgt	atttttggt	14640
gagacaggg	ttcaccatgt	tggccaggct	gatctctaac	tcctgacctc	aagtgatctg	14700
cccacctcag	cctcccaaac	tgctgggatt	acaggcatga	gccagggcat	ccggctttta	14760
tttattcatt	cattcaatat	ctaatagaca	cctaccagg	accaaacc	agatgatgcg	14820
cccaagttca	ttagacccca	ccgctgtctt	caaggcactc	atgatctagg	ccagcgtttt	14880
ttaaccactt	tttttttttt	tttttttgag	attctgggtga	gagctataaa	ttctttcctg	14940
gaaaaacatc	tctgcacact	aagctgtgcc	tggcattggg	aaaaagaaag	cacgtaatgt	15000
aactgacagc	atgagtaaca	cagtgaagaa	ggttgaggga	gagagcgcca	ggacctcaga	15060
actcaggcat	tagaggagcc	ccttccccag	ccctccttga	ggtttcgttg	ggcagggttc	15120
actgaggaaa	aagggtcaaa	tccctttttc	gaatttgact	tcttgtaagt	gccagaagac	15180
tgccccttct	ccaccatccc	tgcctcacca	tcacttttcc	tcccaaggca	gtgacatcca	15240

ES 2 886 958 T3

gcaccccgat	ccctagggcc	ctggggaccc	agcctttggc	aaagtctcct	caggcttggg	15300
tcaggcctga	accagctgt	ctctaccccc	aggaaacata	ccgatgtgtt	caactggaag	15360
gtgcgggcag	gctcagacaa	actgggcagc	ttcccatccc	tggctgtggc	caagatcatc	15420
atcattgaat	tcaaccccat	gtaccccaaa	gacaatgaca	tgcacctcat	gaagctgcag	15480
ttcccactca	ctttctcagg	tgagaagcag	ggcccaaggc	cactcaagcc	tcttacatca	15540
gttttcacgc	ccactctgct	attagctcac	tgaccgccct	tggcacataa	tgtctcctct	15600
caagtcctca	gcttgcccat	ttgtctctaa	tacgtcagcc	taacatcact	gatgccatga	15660
ggcctcctca	agctgtcagc	taacacctcc	actccattcc	ctgccagaga	ttcttccaag	15720
gcctgtcttc	cctatgtgga	gcccctcgag	tgagaactgg	agtttcatcc	aatcttggag	15780
ttttaggaga	ccttttaaaa	agattatcga	gctaattccc	caccactgac	caacacgcaa	15840
gagcctgctc	agtatccctg	ccaaggagtc	attgtgcccc	tgtttgctct	cctccagggg	15900
cagggaaacc	attacctgtg	aggcagccca	cagagtcttt	gaacagctct	gttggatgcc	15960
ttgtgcttat	actgaaatgt	atthagatca	ggattcccaa	ctgtggggtc	cacaagacac	16020
tggccccttg	gagaagagag	gattccattg	tcaaataagt	ttggggaaca	ttttcatact	16080
acagctccct	tcttgggaaca	cattagttta	ttaaaggtag	gagaagtttt	taaaataatc	16140
tgttttattg	cgtttaacct	acatttttta	aattttattg	accacagaat	ccttttttca	16200
tgctacttct	attagcatcc	catagaacaa	gtgttctaga	gaccctgggtg	tgaccctttt	16260
cagagagctt	aactgccagg	ctctcctgag	ccctgggtgtg	tgtttcaaga	tttgtgcctg	16320
ggaattgttt	taatcaggta	tggcaagggtg	acagatacag	acacagctat	ctttgaaaga	16380
agagtttatt	atttataatt	cctgagagaa	agggacatac	cccacccccc	aacacaggga	16440
cacccgggga	agcagctggg	tccaccagga	ggcaggagtg	aggggaaggc	atggcccaga	16500
gccacctgtg	gcttccatgg	gcaggctctg	ccaaggtagg	gtaggcaaga	ttgagcatgc	16560
tcaggattgg	atagtgtgga	caattctcta	ggctatagat	gtcagcctct	ggttgtctag	16620
tatctgtccc	tgggggtgatt	tagggcaggg	aaaatattgg	cttgggtgtct	gagagtcaga	16680
taaaggaagt	ggttggggat	atgggcctttg	ggttggctgg	tttgcctatt	aaaggcgtgc	16740
ccaaagccaa	gttgtttact	atctgcagga	attagctaac	ccagtctctc	ccagaccagc	16800
aagatcccca	taatcataaa	gcatcataat	ttacagaaaa	ttaacactta	tgatgaataa	16860
aagatctcct	tcttcctctg	tgctcctggc	aggcacagtc	aggcccatct	gtctgccctt	16920
ctttgatgag	gagctcactc	cagccacccc	actctggatc	attggatggg	gctttacgaa	16980
gcagaatgga	ggtaagtcct	gggtgcagga	ccacagggca	ggagatgccc	ttgtatgagg	17040
gagcagcttc	cagaagtaat	gggaaggagg	accacccctc	agagaaaccc	atcctggagg	17100
accaagcacc	aaggcgccag	gcagaaagca	aagtggtttg	gcaatccagg	gctgggggat	17160

agaaggcaag	gatgggaatg	tgagtgtttt	taccctccca	gggaagatgt	ctgacatact	17220
gctgcaggcg	tcagtccagg	tcattgacag	cacacggtgc	aatgcagacg	atgcgtacca	17280
gggggaagtc	accgagaaga	tgatgtgtgc	aggcatcccc	gaagggggtg	tggacacctg	17340
ccaggtgggg	cctccaagaa	tcattggggag	ttctaagaat	agggtttagg	tcctagagag	17400
atgagaaaac	ccagaggctg	catgccctac	aggaagcctt	gcataatcatg	ggcactcaat	17460
gtgtgatgat	gggaggaaga	gagggaggga	aggaaaggat	agtcagataa	aagtgtacca	17520
atagatgagt	gggtggatgg	atggatgcag	acaagcagag	agatttcaaa	tgtctctttc	17580
acattcgaag	atgatgttac	tggcctggca	tggtggtctca	cgcttgtaat	cccagcactt	17640
tgggaggctg	aggcgggcag	gtgatttgag	gtcaggaatt	caagaccagc	ctggccaaca	17700
tggtgaaatc	ccatctctac	taaaaagaat	acaaaaatta	gctgggcgtg	gtggcacgtg	17760
cctgtaatcc	cagctacttg	ggaggctgag	gcaggagaat	tgcttgaacc	caggaggcag	17820
aggttgcagt	aagctgagat	tgcgccactg	cactccagcc	tgggtgaccc	agcaagactc	17880
catctgaaaa	caacaacaac	aacaaagatg	acattactca	tccacccccc	ccacccttct	17940
cactagctac	agaatgatta	gccccttgag	gtcaggaatc	ccaggtctat	tttctctgtg	18000
actctcccca	agctgctgaa	ctacactagg	aaagaattac	cgctgcaga	atgctggaag	18060
cacatctgtg	tgtgccctca	ccccggcctc	attggccatc	aggactgctt	agcaatccct	18120
gtagaccttc	ttcctccccc	atacttcag	aggatcttct	gaactatttt	ctttttttat	18180
tttttctttt	atgtttttta	acagagacag	ggtcactatg	ttgccagtc	tggctctcaa	18240
ctcctgggtt	caagggattc	tcccacctca	gctttccaaa	atgctgggat	tacaggcatg	18300
agccatcgtg	cttggcctga	accattttca	ttaaaacccc	taccctactc	tcacctccat	18360
ttccagtcat	taaattcctt	catttaagag	gcattctctta	gtcatcgcat	gtgtgccatg	18420
aacatggtag	tctttggaga	cccctcaggg	agctcacagt	ggttggggga	aaggggggca	18480
ttaaacagac	atttaagcta	tagttttggg	ttcagaggga	ggaagcccca	ggggctaaaa	18540
cagctgataa	ggactcccag	ataagtgcac	ttttcactat	ctggcatttt	cttgttttgt	18600
tatttgcttg	ttcactgtct	ctcaccat	ttgatcctaa	gctttctgag	ggcagggatc	18660
tttgtttttt	ttcatcagtt	ggatcccaat	tgcttagaac	actacctggc	acaaaatagg	18720
cactctataa	gtgattacac	aaattttgga	acgactaggt	taaacaatga	taaccaggct	18780
tttttttttt	tttttgagac	tgagtctcac	tctgttgccc	aggctagagt	gaagtggttt	18840
gatctcggct	cactgcagcc	tccgcctctg	ggttcgaatg	attctccacc	tcagcctcct	18900
gagtagctgg	gattacaggt	gcctgccact	atgccagct	aatttttgta	tttgtagtag	18960
agacgggttt	caccatgttg	gccaggctgg	tcttgaactc	ctgacctcaa	gtgattcacc	19020

ES 2 886 958 T3

cgccctcagcc tcccaaggtg ctgggattac aggtgtgagc caccgctcct ggccaacaac 19080
caggcttttt taagacatca ctgagagcct ttaatttgct aatgtgagtt gtgaatctct 19140
gagagaaggg taacggcatg cttgcaactt acttgtccac agacaagcct ttctgccccca 19200
gaagagaaga ccattctagg gtgctaata gcaaagaggg tgaggggtgga atatcggaga 19260
gcagcagggg gtgcagggga acagataggc cagttcaggg agcagagaag gagaagcccc 19320
ccacactcac ctgccctccc cagcagtcctc tgtttctgggc tctcacaggg tgacagtggg 19380
gggcccctga tgtaccaatc tgaccagtgg catgtgggtg gcacgttag ttggggctat 19440
ggctgcgggg gcccgagcac ccaggagta tacaccaagg tctcagccta tctcaactgg 19500
atctacaatg tctggaaggt aaggtacctt tgccctaccc actgtgcctt ccctccagtc 19560
ctctacctgg ggggtgcca tccatcctca ggtttgattt aaatggttct gacaactctt 19620
tacatcccaa ataactttcc ctccaagcaa gggacagcct gagattgcac tattaaggct 19680
gaaattcctt aggtcagaga tttctgataa atgcaaatac cttagggaat agaacacacc 19740
aagcctttct ttctcttttc tgacagaatg agactatcag atcctttcta gagagaagat 19800
tctgataagg aagagagtgg aaaggctcat gagacctcct ggccctctgc agggtaggga 19860
gagaagcaaa gtgtttcaga aaaggaagac tcacgttaca catgtcacca ctttgtccag 19920
tttcagataa tctgactttc tcttcacgg tctctcttat tctaggctga gctgtaacgc 19980
tgccgtcccc cacatccaga agctgcttcc cttcagacct acctacggca tgaccctca 20040
aagtcagata tgggacaaga gcctccttga acaaactc 20078

<210> 13
<211> 15159
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido recombinante

<400> 13

ccaccgcac aactacagt cgagataact tcgtataatg tatgctatac gaagttatgc 60
tagtaactat aacggctcta aggtagcgag ctagccaagt ctgtgtgcta ccaagtagca 120
aaactgagcc tggaactcac acatgcgtgt ctgagagccc agcactatcg ccaggaaaac 180
ccagcgtctc cctgctcaag cctgaccctc agccctctct gcctctccct gcacttgcct 240
tccagtcaag gtgattctgg ataaatacta ctctctctgc gggcagcctc tccacttcat 300
cccgaggaag cagctgtgtg acggagagct ggactgtccc ttgggggagg acgaggagca 360
ctgtgtcaag agcttccccg aagggcctgc agtggcaggt gagtgcaggg tctgaggcac 420
aagagaagtg ggccagcag gaggtctgct caggccccca cggccactg catagtatct 480
gccccctact tgtcactttt catccttgtt gtataagggt ctttgtttgt ttgtttgttg 540

ttgttttgag	gcagagtgct	ctgtggccca	agatggagtg	cagtgtcttg	gtctcggctc	600
actgcaacct	ctgcctccca	gtttcaagtg	attcttctgc	ctcagcctca	tgagtagctg	660
ggattacagg	tgccagccac	cacgcctggc	taatttttat	attttttagta	gagacggggt	720
tttgccacat	tggtcaggct	gatcttgaac	tcctgacctc	aggtgatctg	cccgccctcag	780
cctcccaaag	tgctgggatt	acaggcgtga	gccaccgtgc	ccagctgtgt	aagtttcttg	840
agagcaggac	cctgtcttgt	ctacctttaa	atcctagtag	ttaacacaca	gcaaacagta	900
actatttgat	gaccaaagt	gagccagaaa	ggacaggaaa	ttgtaactga	ggctgccccca	960
tgctgtctgc	gcctgggtga	tttcaggcag	agggctagac	tgggtgacct	tggggcattc	1020
ctcctttcta	tgaaatttgt	tatttcaagg	agactagaaa	agagacttct	cagccacttc	1080
gccagctatt	ggctcttcta	ttcattagtg	tttgctgaga	catgctatgt	gacaggactg	1140
agccagggtcc	tttcaatgga	taggagatgt	tttgagcata	aatccacgt	tctctcttgg	1200
gctgggctct	tctaccttct	tccccctgg	gcttgggctc	tgaagaaaaa	aagataggta	1260
ggagatgagt	gatggggctt	ctgagggcag	ggctgagtga	ctttctgtgt	atttgctctt	1320
tctttatcag	aagtcaaagt	cccacaggca	cctgtcatcc	tactgccagt	aggacttctc	1380
actcaacctt	cccctctgac	cttacttga	gaaggactta	ggccctctc	tcagacattt	1440
ccccaggctg	ggcaagttgt	gtggaccatg	gatgggtatg	tgggtccatac	aatttaaaaa	1500
agctgtatat	ggctgctggg	tagagtgacc	acataattga	tcataaaaa	tgataacctgt	1560
aagagcaaaa	gggggcacta	ttaaccattg	ggtcagggca	acagggtcaaa	atggagacct	1620
accctgggac	ttctgggtcac	actagctact	gtcaaaatgg	ggcccaaata	gacaaagcca	1680
aatggaagaa	attcccttga	cattgaaagt	gttggggctc	tgtggcaccc	ccagttctag	1740
gttgggggag	cttgggctgg	tctcatgatg	agttctgagg	gggatgggcc	agttggggcc	1800
cccgttccat	ctaactcagg	ttcctttcct	cccagtcgc	ctctccaagg	accgatccac	1860
actgcagggtg	ctggactcgg	ccacaggga	ctggttctct	gcctgtttcg	acaacttcac	1920
agaagctctc	gctgagacag	cctgtaggca	gatgggctac	agcaggtaac	caacctgggc	1980
ctctctcctt	tttccctcct	tcctccttcc	tcctcttctc	cctttccttc	ctcccttctt	2040
ctctctttcc	taaaaattac	gggcattgga	gccaggcaga	atggcttttg	aatcccagca	2100
tttcacttat	aagcaacatg	aagttaaatt	tcctaagcct	caggttcctc	aggagttaat	2160
tgggggaact	aatgccaacc	tcataggata	gttttgcaat	gccagtgaga	gaatgtgtgc	2220
tgccctccaa	cacacacaca	cacacttcta	gcgtctatgc	agtcctctcc	tttcttttac	2280
tcctcaacct	tcactccttt	gtgctggctt	tgcaagaaac	tgctcctgcc	cagtaataca	2340
aaagctaagt	taacttattc	aaagtttcgt	tagttaagat	ttagcttaag	tgagcctagt	2400

ttcagtgggg	ccccatcttc	agcaatccca	gctctctctg	caaatttcaa	aagcagttcc	2460
aaatctggag	tggatgaaaa	ggtgtaagat	gatagtaaga	gtaatttgca	ttctatatat	2520
ttatatccac	ttgatttttg	cagaaaacca	aaaagatagt	tattatatct	tatatataga	2580
tatatattat	atctatttca	taaataggct	caaacaaagt	aagtaacttg	ctaggggtact	2640
agctgggagg	tagagggcta	gaatttgagc	ccaagacccc	taattcttgc	gcattaggag	2700
ttcccacatt	gtttctgttt	ctagactgag	taattcttta	ttctcatgta	ggacatcatc	2760
tctaagggaa	ggggctaata	agatgggtga	tactcagag	agtttagctg	gagaggatgg	2820
aaaagaacct	atacattcag	ttgcagattg	agatagccta	tctctggcag	gcctcagatt	2880
tcttcaggat	tctaacagac	tggaccacga	gactaggcca	aacaaacaaa	caaacaaaaa	2940
ctctactagg	cagacatcac	caaccaatca	cagaactctc	tcccatggat	ccctaataca	3000
gcctcaaagt	ccttttcagt	aatgctcca	ggcagccatt	acaaatcaat	cagaattatt	3060
tgctttctc	ttctctgctc	aacgggcttc	tgctgctctc	tactttccat	agggggcaac	3120
ttccattacc	ctctagaaag	cacacccac	caccttcatt	tcaaggagag	tgaggaaactc	3180
atgccagca	cctgctattc	tcccctcttc	ctgcagccac	ggagcccagc	ctcgtgcag	3240
ccagccctgc	ctcccactg	tagtccagtc	aactgctgca	tcagccgttc	ctggcacagc	3300
aggctgagcc	ttgattatga	aacctgggtg	tctccagggg	ttcttaagat	gataggctcc	3360
tggaatttct	gtccttttg	agctcagtaa	ggcaccaaac	cacctgagtc	ttgtgcttca	3420
caaaatcaaa	gttcatcaga	atcattcatt	gggatggaat	tggatgaacag	aagttaactt	3480
tcttggaat	gtccatttcc	accatattcc	gtccttctag	gtctcagact	tctctacttt	3540
ctttcctctc	tctagatcgg	aggcccttct	tgtcctagaa	ccataggcat	ttcaagatgt	3600
gggagaccct	agggatcatc	tagtccacgc	atcttttttt	tttttttttg	acagagtctc	3660
actctgtcac	ccaggctgga	gtgcaatggc	accatctctg	cttactgcaa	cctccacctc	3720
ccaggttcaa	gtgattcttt	cgcctcagcc	tcccaagtag	ctgggattac	aggcacgcac	3780
catcatgccc	agctaatttt	tatatatttg	tagagaccga	gtttcaccat	gttggccagg	3840
ctggtcttga	actcctgacc	tcaggtgatc	caccacctc	ggcctcccaa	agtgtggga	3900
ttacaggcgt	gagccactgc	accagcccc	gtgcatcttt	ttatagaggg	ggaaactgag	3960
gcttgagag	accagaaaa	agaatatgac	ctgccaagg	ccacacatca	aactagtgcc	4020
agagccaggg	acagaacct	gatcatgagg	actcttaaaa	tgactctag	tcctcccagg	4080
tctgagactt	gggtccttcc	aggaagtgcc	agcattcctg	cctgagaatg	tgccaatcca	4140
ccagtattgc	caatgactca	gccctccatg	gagagcttct	actaacatta	ctagcatagt	4200
tagggatgga	aggaaaagat	ttagaagagg	cagattcagt	aaaggaacaa	tcagagagat	4260
ggaattaatc	aaggaaggct	tcctggagga	ggaaaaactt	caaccaagg	tttgaaagta	4320

gcaagcatgg attagcaggg agaaagaggg agagtgggcc agttgagaga aacgtttgtc	4380
tggattcata tgaagacaga tctagtcctg ttctattaaa tatctctaag ggggccaaaa	4440
acataccccc gctatcaaag tcagaccaga tgctttgttt ggagaacgaa atatccacat	4500
tccaactccc tcccaggtga gaagggagct aacctgagcc cctatgcctc tttgtttccc	4560
tgctgtgaac cagaagacat tgctgggata tttgaaatag ggacagagct gggaaatatg	4620
aaaggagacc cctaacatct ctccagggct ctgggttctg gatttggtt cccaccccaa	4680
gaaagcaagt tacatcagca atgcaactgag ggttgagtc tgggatgcca agggtcgggt	4740
ctttattgta tagcaaagca ggccccatct tcaactgacta agaccatctc cactccctgg	4800
ccactcccca ccaagcattc tctgccactc tttctcctga aagtgggggc caactctacc	4860
atcttgttct aacccccctgc ccagctcac aactctctct ccctcttgat gtgagcagca	4920
aaccacttt cagagctgtg gagattggcc cagaccagga tctggatgtt gttgaaatca	4980
cagaaaacag ccaggagctt cgcattcgga actcaagtgg gtaagtgagg ggacaccttc	5040
tggcctacag aaggcccca catggagct gctcttcagg ttgcaaccag ctacacctga	5100
acccaagca gccaggggaa tgtaagcaga catcaggaag aactcctagc cagatggatc	5160
attcaatgcc aagagctata gactcacatt ttggagaggt tttctgtgtt gacttgtttt	5220
taatacaatg gacagctgga caaagtgtgt tgtcctactc agagccagag ggatggataa	5280
tgtgaccttt ccatcaatct ggatagtaaa tagtttttgc tactgctgta ggttttctaa	5340
taaattgccc aataggcaag attccaaagt cactttgtcc ttccctacca cttaccacgc	5400
cagagctccc caccttcttg atgctccagg gaagaggctc catggccctt gtgggtggcc	5460
tgctcctgag cctcgccacc ctgtgttaga gcagagcatc cagatgaaat ctgtcacact	5520
gtggcaaagt ggctcagaga ggaggctggc ttccctagcat tcagggacgt tgctgagggc	5580
cgcttattca ccgaaaataa atcttgaaaa ggacagggct ggtagcagaa tgatccttta	5640
cctaaaattc tatcaaaatc ccattcttcc atttggaag cccacagtgt cacagactct	5700
gttcggggct ctgtcctctt ccctcttggg tcccaggagc ccaggctggg ctttgaagca	5760
ggcagggccc agcacacagt aggtactcag cagtgggggt gttgaatcca atcaaacgga	5820
agtgtcaatg caggaaatgc aatggatgac aatgcagtct ccaaagtgtc cccactgtgc	5880
agcttcacac ttcccaggt attgggaggg gacttgaatt aacagcttcg ggaggcctga	5940
gtccctgcct cccagctgag gaagaagctt aaatcacagg gcgctgtgtc tgtcttcag	6000
gccctgtctc tcaggctccc tggctccct gcactgtctt ggtgagtacc cccaatctct	6060
gagggtttgg ggcctgggccc agcaatgagc agggaggaag accttcatct tcaactcctaa	6120
atttctggga ctccaagttt cattctgcct tggctctacag cccttgggct tgtcgggtcaa	6180

ES 2 886 958 T3

tgccccctcg	agttgttgg	ggccttgggc	aggtcacatt	ctttttctgg	gtctttccaa	6240
gccccagttt	cccccttcta	ccatctgtgc	atggctccat	gacctaaagt	gagacctggg	6300
agagagtgtt	aggaagaccg	aaaagggcag	gacggggcct	ccactgcctc	ccatccctgg	6360
tccgggcca	catagccttc	tttgtcaca	tcagctcagg	tatccaagat	cagattaccc	6420
acattcatta	tttgagcaac	tattcattga	acagttagaa	tatgtctcac	tctgtcagtt	6480
gctggctaga	agtagaaagt	accagatgag	tgaataaatt	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	gcccatacag	gatctttagg	gaccttcta	gctctaacag	6780
aggagatcag	aaaagaaaac	aagcaatgtg	gctcagctca	tcctacaagc	ttcatagaga	6840
actgagactg	gcctggaagc	atagccagaa	attagaacgc	ctaagggaag	aaggtcaca	6900
cgctgcctct	gcaatttagg	agtgtatatg	ctttcctgca	ggatgttgag	agtttcattc	6960
attatcgtat	gccccctacc	ccggccccac	aatacctagt	gcgtgggatc	tgacacgtgg	7020
tggctggtca	atgaatgaat	gaatgaatgg	tcacaccatc	tgaggttctg	cactgagtag	7080
ccctgaaggc	ttgaagcagc	ataagtgaca	ggtcctccct	tgaggggcct	ctgttttacc	7140
aataagccaa	gacctaaagt	caacaacact	gaaaggggtg	ccaataccca	ggacagcctg	7200
tgggaattcc	agagaaaggg	agattcccag	ggactggggg	cccaggctaa	acactgaaaa	7260
atgcatctgt	aggctcaagg	aggaaaagcc	catgtctgtc	tgtcttgccc	accactctct	7320
cccagcacc	agcactgccc	caggacagag	agcacttgac	acaagttggt	tagattaatg	7380
aatgatttag	agttcagtgg	tccccaacct	ttttggcaca	agagactggt	tgcatggaag	7440
acaatttttc	cgcaaaccac	gagggggata	gagagcatta	gattctctct	tttttttttt	7500
tttgagacca	agtctggctc	ttgtcactca	gcctggagta	aagtgttgcg	atctcggtc	7560
actgcaacct	ccgcctcctg	gattcaagcg	attctcctgc	ctcagcccc	taaatagctg	7620
ggattacagg	caccgcgtcac	cagcccagct	gggactatag	gcatgtgcca	ccatgcccgg	7680
ctaatttttg	tatttttagt	agagacggcg	tttcacccatg	ttggccaggc	tagtctcgaa	7740
ctcctgacct	caggtgatct	gcccgcctga	gcctcccaaa	gtgctgggat	tacaggcatg	7800
agctgcctca	cccagcctaa	agtctcataa	ggaacgtaca	gcatagatcc	ctcacatgtg	7860
cagttcaca	taaggttgtg	ctcctacaag	aatctaacgc	cacctctgat	ctgacaggag	7920
gtgaagctca	ggtggtcatg	ctcgcttgct	cctgccactc	acttccta	gtacagccag	7980
gttcctaaca	ggccacgaac	cagtgggaag	ggcatctttt	tggatcaaaa	acagaattac	8040
tttttagaga	actacaagca	gatcaatttg	gctagacaga	gactttatat	gaaacagcag	8100

gaggctgcta ggaggagtgg aaactctact ttgccctcaa gggagatccc gaagggtttt	8160
gcaggagcgg gcaaggtggc atgaagaaag cagtgtttga aatcaggtgg tatttgaaaa	8220
gcccagccct tccccttaga atggcccttc taccatctgt gcatggctcc acaaccgtgg	8280
tggtggctgc cagaagaatt ggaaaggcag agcatgggtg gagagggggg acctgagggc	8340
tttacaggag ttccgggggt ggtgaggggtg tgaaagccag gtcagtcagt aggaagacag	8400
gatgtcagat tgagagactc ccctggccgg ggaaacagac ttggagaagg gggagttttg	8460
gatgagacag tccacttccg agtcacaaaa tagcttgtgg gtgtctgttt actgttactc	8520
agtgggagtg gctggggaca cgccacctgg gcagggtttt cgtaattctg catcacttgt	8580
gaaggtcaca gattcccagc acaacggaca caccatgtt catagtctga actcctaaac	8640
acatcttaaa ccaaaataaa aaaaaagaa agaaagaaag aaaaaggaga gggaggtttg	8700
aggaaagcct atggtctggg aactcaata cctcccatga atatctcata ttgggctggt	8760
cctctctcca ctctggcccc agccataagg gccctgctta gagcagattt tgggtgctga	8820
gtggaggcag cctcatcccc aacagcctga ctctctgctt cctccctgcc tctgcctgtg	8880
tccagcctgt gggaagagcc tgaagacccc ccgtgtggtg ggtgtggagg aggcctctgt	8940
ggattcttgg ccttggcagg tcagcatcca gtacgacaaa cagcacgtct gtggaggag	9000
catcctggac cccactggg tccctacggc agcccactgc ttcaggtaag accccagctg	9060
taaggaggtc tctggggacc aaggccagtc agggaccaga gagcttgggg tctgtctcc	9120
tggcacctgc ctctcttcca ctctccact agagacgttt tccaggttgt ggtggcccca	9180
atgagacaat ggccatgatg ccctttgtta ggcttttggg tgtctgagca gaggggtgctg	9240
gtcaccaagc atggcctctt cctggtggga caccagcaga taccagagt cctcacccca	9300
cccccatatc gttcaagcta caaaagctct tcccacctgc ctcaacttcc aagaactcac	9360
tctctttttg ctgttttcca ggaagttgtt ccagggtcta gagtcatagc cacgtcctca	9420
ttatgtctgg aaactttaaa aaaattaaag agcatagggt cctttcagtc cacagagaag	9480
cctggcctta cctcaggga gggctactcc cagacccct tcaacttttt tttttttttt	9540
tttttttttt tttttgagac agagtcttgc tctgttgctt aggctggagc gcagcagcat	9600
gatcttggct cactgcaacc tccgcctcct gagttcaagc aattctcctg cctcagcttc	9660
ccaagtagct gggactatag gcatgggcca ccatgccgg ctaatttttg tatttttgggt	9720
agagacaggg ttccaccatg ttggccaggc tgatctctaa ctctgacct caagtgatct	9780
gccacctca gcctcccaa ctgctgggat tacaggcatg agccaggga tccggctttt	9840
atttattcat tcattcaata tctaattgagc acctaccagg taccaaacac cagatgatgc	9900
gccaagttc attagacccc accgctgtct tcaaggcact catgatctag gccagcgttt	9960

tttaaccact	tttttttttt	ttttttttga	gattctggtg	agagctataa	attctttcct	10020
ggaaaaacat	ctctgcacac	taagctgtgc	ctggcattgg	gaaaaagaaa	gcacgtaatg	10080
taactgacag	catgagtaac	acagtgagaa	aggttggagg	agagagcgcc	aggacctcag	10140
aactcaggca	ttagaggagc	cccttcccca	gccctccttg	aggtttcgtt	gggcagggtt	10200
cactgaggaa	aaagggtaaa	atcccttttt	cgaatttgac	ttcttgtaag	tgccagaaga	10260
ctgccccctc	tccaccatcc	ctgcctcacc	atcatctttc	ctcccaaggc	agtgcacatcc	10320
agcaccocga	tccctagggc	cctggggacc	cagccttttg	caaagtctcc	tcaggcttgg	10380
atcaggcctg	aaccagctg	tctctacccc	caggaaacat	accgatgtgt	tcaactggaa	10440
ggcgccggca	ggctcagaca	aactgggcag	cttcccatcc	ctggctgtgg	ccaagatcat	10500
catcattgaa	ttcaacccca	tgtaccccaa	agacaatgac	atcgccctca	tgaagctgca	10560
gttcccactc	actttctcag	gtgagaagca	gggcccgaag	ccactcaagc	ctcttacatc	10620
agttttcacg	cccactctgc	tattagctca	ctgaccgccc	ttggcacata	atgtctcctc	10680
tcaagtcctc	agcttgccca	tttgtctcta	atacgtcagc	ctaacatcac	tgatgccatg	10740
aggcctcctc	aagctgtcag	ctaacacctc	cactccattc	cctgccagag	attctttcaa	10800
ggcctgtctt	ccctatgtgg	agcccctcga	gtgagaactg	gagtttcatc	caatcttgga	10860
gttttaggag	accttttaaa	aagattatcg	agctaattcc	ccaccactga	ccaacacgca	10920
agagcctgct	cagtatccct	gccaaaggag	cattgtgccc	ctgtttgctc	tcctccaggg	10980
gcagggaacc	cattacctgt	gaggcagccc	acagagtctt	tgaacagctc	tgttggtatgc	11040
cttgtgttta	tactgaaatg	tatttagatc	aggattccca	actgtggggg	ccacaagaca	11100
ctggccccct	ggagaagaga	ggattccatt	gtcaaataag	tttgggggaa	attttcatac	11160
tacagctccc	ttcttggaac	acattagttt	attaaaggta	ggagaagttt	ttaaaataat	11220
ctgttttatt	gcgtttaacc	tacatttttt	aaattttatt	gaccacagaa	tccttttttc	11280
atgctacttc	tattagcatc	ccatagaaca	agtgttctag	agaccctggt	gtgacccttc	11340
tcagagagct	taactgccag	gctctcctga	gccctggtgt	gtgtttcaag	atgtgtgcct	11400
gggaattgtt	ttaatcaggt	atggcaaggt	gacagataca	gacacagcta	tccttgaaag	11460
aagagtttat	tattttataat	tcctgagaga	aagggaacata	ccccaccccc	caacacaggg	11520
acaccggggg	aagcagctgg	gtccaccagg	aggcaggagt	gaggggaagg	catggcccag	11580
agccacctgt	ggcttccatg	ggcaggtctg	gccaaaggtag	ggtaggcaag	attgagcatg	11640
ctcaggattg	gatagtgtgg	acaattctct	aggctataga	tgtcagcctc	tggttgtcta	11700
gtatctgtcc	ctgggggtgat	ttagggcagg	gaaaatattg	gcttggtgtc	tgagagtcag	11760
ataaaggaag	tgggtgggga	tatgggcttt	gggttggtctg	gtttgcctat	taaaggcgtg	11820
cccaaagcca	agttgtttac	tatctgcagg	aattagctaa	cccagtctct	cccagaccag	11880

caagatcccc	ataatcataa	agcatcataa	tttacagaaa	attaacactt	atgatgaata	11940
aaagatctcc	ttcttcctct	gtgctcctgg	caggcacagt	caggcccatc	tgtctgccct	12000
tctttgatga	ggagctcact	ccagccaccc	cactctggat	cattggatgg	ggctttacga	12060
agcagaatgg	aggtaagtcc	tgggtgcagg	accacagggc	aggagatgcc	cttgtatgag	12120
ggagcagctt	ccagaagtaa	tgggaaggag	gaccaccctt	cagagaaacc	catcctggag	12180
gaccaagcac	caaggcgcca	ggcagaaagc	aaagtggttt	ggcaatccag	ggctggggga	12240
tagaaggcaa	ggatgggaat	gtgagtgttt	ttaccctccc	aggaagatg	tctgacatac	12300
tgctgcaggc	gtcagtccag	gtcattgaca	gcacacggtg	caatgcagac	gatgcgtacc	12360
agggggaagt	caccgagaag	atgatgtgtg	caggcatccc	ggaagggggt	gtggacacct	12420
gccaggtggg	gcctccaaga	atcatgggga	gttctaagaa	tagggtttag	gtcctagaga	12480
gatgagaaaa	cccagaggct	gcatgcccta	caggaagcct	tgcatatcat	gggcactcaa	12540
tgtgtgatga	tgggaggaag	agagggaggg	aaggaaagga	tagtcagata	aaagtgtacc	12600
aatagatgag	tgggtggatg	gatggatgca	gacaagcaga	gagatttcaa	atgtctcttt	12660
cacattcgaa	gatgatgtta	ctggcctggc	atggtggctc	acgcttgtaa	tcccagcact	12720
ttgggaggct	gaggcgggca	ggtgatttga	ggtcaggaat	tcaagaccag	cctggccaac	12780
atggtgaaat	cccattctta	ctaaaaagaa	tacaaaaatt	agctgggctg	ggtggcacgt	12840
gcctgtaatc	ccagctactt	gggaggctga	ggcaggagaa	ttgcttgaa	ccaggaggca	12900
gaggttgtag	taagctgaga	ttgcgccact	gcactccagc	ctgggtgacc	cagcaagact	12960
ccatctgaaa	acaacaacaa	caacaaagat	gacattactc	atccacccca	cccacccttc	13020
tactagcta	cagaatgatt	agccccttga	ggtcaggaat	cccaggtcta	ttttctctgt	13080
gactctcccc	aagctgctga	actacactag	gaaagaatta	ccgcctgcag	aatgctggaa	13140
gcacatctgt	gtgtgccctc	accccgccct	cattggccat	caggactgct	tagcaatccc	13200
tgtagacctt	cttcctcccc	catacttcca	gaggatcttc	tgaactatct	tcttttttta	13260
ttttttcttt	tatgtttttt	aacagagaca	gggtcactat	gttgcccagt	ctggtctcaa	13320
actcctgggt	tcaagggtat	ctcccacctc	agcttttcaa	aatgctggga	ttacaggcat	13380
gagccatcgt	gcttggcctg	aaccattttc	attaaaaccc	ctaccctact	ctcacctcca	13440
tttccagtca	ttaaattcct	tcatttaaga	ggcatctctt	agtcacgca	tgtgtgccat	13500
gaacatggta	gtctttggag	acccctcagg	gagctcacag	tgggtggggg	aaaggggggc	13560
attaaacaga	catttaagct	atagttttgg	gttcagaggg	aggaagcccc	aggggctaaa	13620
acagctgata	aggactccca	gataagtgca	cttttcaacta	tctggcattt	tcttgttttg	13680
ttatttgctt	gttcactgtc	tctcacccca	tttgatccta	agctttctga	gggcagggat	13740

ES 2 886 958 T3

```

ctttgttttt tttcatcagt tggatcccaa ttgcttagaa cactacctgg cacaaaatag 13800
gcactctata agtgattaca caaatTTTtgg aacgactagg ttaaacaatg ataaccaggc 13860
TTTTTTTTTT TTTTTtgaga ctgagtctca ctctgttgcc caggctagag tgaagtgggt 13920
tgatctcggc tcaactgcagc ctccgcctct ggggtcgaat gattctccac ctcagcctcc 13980
tgagtagctg ggattacagg tgcctgccac tatgccagc taatttttgt atttgtagta 14040
gagacggggt tcaccatgtt ggccaggctg gtcttgaact cctgacctca agtgattcac 14100
ccgcctcagc ctccaaggt gctgggatta cagggtgtgag ccaccgctcc tggccaacaa 14160
ccaggctttt ttaagacatc actcagagcc tttaatTTtgc taatgtgagt tgtgaatctc 14220
tgagagaagg ctaacggcat gcttgcaact tacttgtcca cagacaagcc tttctgcccc 14280
agaagagaag accattctag ggtgctaag agcaaagagg gtgaggggtg aatatcggag 14340
agcagcaggg agtgcagggg aacagatagg ccagttcagg gagcagagaa ggagaagccc 14400
ccccacctca cctgccctcc ccagcagtct ctgttctggt ctctcacagg gtgacagtgg 14460
tgggccccctg atgtaccaat ctgaccagtg gcatgtggtg ggcatcgta gttggggcta 14520
tggtctcggg ggcccgagca cccaggagt atacaccaag gtctcagcct atctcaactg 14580
gatctacaat gtctggaagg taaggtaact ttgccctacc cactgtgcct tccctccagt 14640
cctctacctg gggggtgcca atccatcctc aggtttgatt taaatggttc tgacaactct 14700
ttacatccca aataactttc cctccaagca agggacagcc tgagattgca ctattaaggc 14760
tgaaattcct taggtcagag atttctgata aatgcaaata ccttagggaa tagaacacac 14820
caagcctttc tttctctttt ctgacagaat gagactatca gatcctttct agagagaaga 14880
ttctgataag gaagagagt gaaaggctca tgagacctcc tggccctctg cagggtaggg 14940
agagaagcaa agtgtttcag aaaaggaaga ctcacgttac acatgtcacc actttgtcca 15000
gtttcagata atctgacttt ctcttcatcg gtctctctta ttctaggctg agctgtaacg 15060
ctgcogtccc ccacatccag aagctgcttc ccttcagacc tacctacggc atgacccctc 15120
aaagtcagat atgggacaag agcctccttg aacaaactc 15159

```

<210> 14
 <211> 435
 <212> PRT
 <213> Secuencia artificial

<220>
 <223> Proteína recombinante

<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

Lys Ile 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*

<400> 15

ES 2 886 958 T3

cagaaacaag gacctcttca ttattcaaga gtaaaatgta taggccaaga ccaatgctat	60
caccgtcaag attcttcact ccctttgcag tagctttcgt tgtcataata acggtagggc	120
tcctggccat gatggcaggt ctacttattc acttttttagc ttttgacaag aaagcttact	180
tttatcatag cagctttcaa atcctaaacg ttgaatacac tgaggcttta aactcaccag	240
ctacacacga atacagaacc ttgagtgaag gaattgaggc tatgattact gatgaatttc	300
gaggatcaag tctaaaaagt gagtttatca ggacacatgt tgtcaaacta agaaaagaag	360

ES 2 886 958 T3

```

ggactggtgt ggttgcggat gttgtcatga aatttcgatc tagtaaactg aacaacagaa 420
aggtaatgaa aaccagaatt caatctgtgc tacgaagact cagcagctct ggaaacttgg 480
aaatagcccc ttcgatgag ataacatcac tcaactgacca ggatacagaa aatgttttga 540
ctcaagaatg tggagcacgt ccagacctta taacactgtc agaagagaga atcattggag 600
gcatgcaagc tgagcccggg gactggccct ggcaagtcag tctacagctc aataatgtcc 660
accactgtgg aggtgccctg atcagtaaca tgtgggtcct gacagcagct cattgcttca 720
aaagctatcc taatcctcaa tattggacag ccacctttgg ggtttctaca atgagcccta 780
ggctgagagt gagagtaagg gctattttag cccacgacgg gtacagctcc gtaactcgtg 840
acaatgacat cgcagttgta caacttgaca gatctgtcgc cttttccaga aatatccata 900
gggtatgtct cccagcagca acccaaaata tcatccctgg ttctgtcgca tatgttacag 960
gatggggatc tctcacatat ggaggcaacg cagtcacaaa tctacggcaa ggagagggtca 1020
gaataataag ttcagaggaa tgcaatacgc cagctgggtta cagtgggaagt gtcttgccag 1080
gaatgctgtg tgctggaatg cgttcagggg ccgtggatgc atgccagggt gattcagggtg 1140
gcccgtagt acaagaagac tcaaggcggc tttggtttgt tgtgggcatt gtgagctggg 1200
gatatcagtg tggcctccca aataagccag gcgtgtatac tcgagtgaca gcctaccgca 1260
actggatcag acagcagacg ggaatctagt gcaaccgagg aaaaaacgtg ccatgaggtc 1320
tctgtatcca agtgtgactg actcggatgc catggcttca catttcaact gcaaaggaga 1380
ctggaaatgc cccttctgaa cgtcccatta cataaatatg gtttaactgt ttagtatttc 1440
tttgtcggta cagattttta ctttcttgag gaaaaaaaaa acatgaacat ggctaagtaa 1500
gaattatgtt aggctagtaa caggaagaca tttattacat ggggtggtcag gtgtagtagt 1560
gagaagtcag gtaagttaag tcaataatth acagaaaata atgtcaggta gtcctaactg 1620
taaatatgtg aggccacaga acaaatagtg ttagaactga agccatccca agtatttaac 1680
atthgttttc aagtgaact aagaaacaga cttacatata gttttaatgg tgaattttca 1740
ttttaaatat tttatctaca tagaaaagac atatctcctt catgaagaag ctgaggtgat 1800
gaatcaacac agcctcttca gctatgtttg caaccacaag atthgtggga aagaaatccc 1860
tactaccaac ttcctactgt tggcattatt ttttagagta acacgacgca caatagcaaa 1920
atthtaagtaa caaattaaaa gttaatgatg aagaagaagt aaagagtttg tttgcaaaga 1980
caaaaattaa acagattaat atcaataaat 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

ES 2 886 958 T3

atttgagtgg gaatctcaa gcagttgagt aggcagaaaa aagaacctct tcattaagga	60
ttaaaatgta taggccagca cgtgtaactt cgacttcaag atttctgaat ccatatgtag	120
tatgtttcat tgcgtcgca ggggtagtga tcctggcagt caccatagct ctacttggtt	180
acttttttagc ttttgatcaa aaatcttact tttataggag cagttttcaa ctcctaaatg	240
ttgaatataa tagtcagtta aattcaccag ctacacagga atacaggact ttgagtggaa	300

gaattgaatc tctgattact aaaacattca aagaatcaaa ttttaagaaat cagttcatca	360
gagctcatgt tgccaaactg aggcaagatg gtagtgggtg gagagcggat gttgtcatga	420
aatttcaatt cactagaaat aacaatggag catcaatgaa aagcagaatt gagtctgttt	480
tacgacaaat gctgaataac tctggaaacc tggaaataaa cccttcaact gagataacat	540
cacttactga ccaggctgca gcaaattggc ttattaatga atgtggggcc ggtccagacc	600
taataacatt gtctgagcag agaatccttg gaggcactga ggctgaggag ggaagctggc	660
cgtggcaagt cagtctgcgg ctcaataatg cccaccactg tggaggcagc ctgatcaata	720
acatgtggat cctgacagca gctcactgct tcagaagcaa ctctaatacct cgtgactgga	780
ttgccacgtc tgggtatttc acaacatttc ctaaactaag aatgagagta agaaatattt	840
taattcataa caattataaa tctgcaactc atgaaaatga cattgcactt gtgagacttg	900
agaacagtgt cacctttacc aaagatatcc atagtgtgtg tctcccagct gctaccaga	960
atattccacc tggctctact gcttatgtaa caggatgggg cgctcaagaa tatgctggcc	1020
acacagttcc agagctaagg caaggacagg tcagaataat aagtaatgat gtatgtaatg	1080
caccacatag ttataatgga gccatcttgt ctggaatgct gtgtgctgga gtacctcaag	1140
gtggagtgga cgcagtgcag ggtgactctg gtggccact agtacaagaa gactcacggc	1200
ggctttggtt tattgtgggg atagtaagct ggggagatca gtgtggcctg ccggataagc	1260
caggagtgtg tactcgagtg acagcctacc ttgactggat taggcaacaa actgggatct	1320
agtgaacaa gtgcatccct gttgcaaagt ctgtatgcag gtgtgcctgt cttaaattcc	1380
aaagctttac atttcaactg aaaaagaaac tagaaatgtc ctaatttaac atcttgttac	1440
ataaatatgg ttttaacaaac actgtttaac ctttctttat tattaaaggt tttctatttt	1500
ctccagagaa ctatatgaat gttgcatagt actgtggctg tgtaacagaa gaaacacact	1560
aaactaatta caaagttaac aatttcatta cagttgtgct aaatgccgt agtgagaaga	1620
acaggaacct tgagcatgta tagtagagga acctgcacag gtctgatggg tcagaggggt	1680
cttctctggg tttcactgag gatgagaagt aagcaactg tggaaacatg caaaggaaaa	1740
agtgatagaa taatattcaa gacaaaaaga acagtatgag gcaagagaaa taatatgtat	1800
ttaaaatttt tggttactca atatcttata cttagtatga gtcctaaaat taaaaatgtg	1860
aaactgttgt actatacgta taacctaacc ttaattattc tgtaagaaca tgcttcata	1920
ggaaatagtg gataattttc agctatttaa ggcaaaagct aaaatagttc actcctcaac	1980
tgagacccaa agaattatag atatttttca tgatgaccca tgaaaaatat cactcatcta	2040
cataaaggag agactatata 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 aaagggctgt caaacgctaa atctaagtaa cgaattgaaa gtttaaagag 2280
ggggaagagt tggtttgcaa aggaaaagtt taaatagctt aatatcaata gaatgacct 2340
gaagacagaa aaaactttgt cactcttcct ctctcatttt ctttctctct ctctcccctt 2400
ctcatacaca tgctccccc accaaagaat ataatgtaa tttaatccac taaaatgtaa 2460
tggcatgaaa atctctgtag tctgaatcac taatattcct gagtttttat gagctcctag 2520
tacagctaaa gtttgcctat gcatgatcat ctatgcgtca gagcttcctc cttctacaag 2580
ctaactccct gcatctgggc atcaggactg ctccatacat ttgctgaaaa cttcttgat 2640
ttcctgatgt aaaattgtgc aaacacctac aataaagcca tctactttta gggaaagga 2700
gttgaaaatg caaccaactc ttggcgaact gtacaaacaa atctttgcta tactttat 2760
caaataaatt ctttttaaaa taaaaaaaaa 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																			
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					
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
<212> ADN
<213> Secuencia artificial

<220>
<223> Polinucleótido recombinante

<400> 19

```
gagggagggt ggtgctttgc taatggtgaa ttactaactc ctcaataaag aatattatatt 60
gaaataatatt ttgaaatttc ataattactt tgggttcctt cttaatgata aataaataat 120
agtatattac aaacatacat taatatttcc tgaatgaata caccacaaat ctcccttaaa 180
atatagcaag aataaaaatt atactatttc tgacaatttt taatttctca aataataata 240
ccactctgat ttttaaacat ctacaccact ctggctttgc caatcttttt aaaaattgaa 300
aagataataa ttttatcata attacactga agcatagaac tttttctttc aaggaaagca 360
aatttttgaa attctataat ataacctccc ataatcctga ataaattaaa ggttcaacaa 420
cttagtaaag taagactgac cttccctttt atttcttttt cagatcaaaa atcttacttt 480
tataggagca gttttcaact cctaaatggt gaatataata gtcagttaaa ttcaccagct 540
acacaggaat acaggacttt gagtggaaga attgaatctc tggtaagtta atatttgtct 600
ttgctcttta ttccattata aatgaatat gataataaac ctaatgtttt gtaatatatt 660
ttcagttgct aagtgtctca catattttcc ttccttgaat ggtgaaacat gtgtttctct 720
ctgcttttat ccagttagtt tactcatata ctggttctta ttcacatctt tgtcatgagt 780
aaaaagtgtt agaaaggcca cgagtaaata tgcattttat ttgtttatga attcaaatac 840
taaaagtgtt ttatttggtt aattaagcat tgacattgtc tttttaaatt cttttcattt 900
taccttcttc cctcttcctt atccaactaa agacgcaaag caggaggtgt taaaaaacag 960
gtttaccata tcagcagtaa catagtttgg acaacattac actttggttc aatgatagac 1020
atagaagttt gaacagaaat atgcaaagca agtttgagct ctaacttgaa gagagcctct 1080
gggtgcctgc caggaaacct cacgagtgga cccttaacat tcatgtgtca ccacaaacta 1140
ggggctgccc tttagttttg accagtctca gtgtcactca cttaccctta ctttttcaaa 1200
aaaaagtcct aagaatataa agtaattcaa tgggtctaca attttagcat gtaactgagt 1260
```


cacctggcag ggttgctttg gtgagctcaa gataaaattht tatcagcatt tctacatttht	1320
ctggaatatt ccttaatcca ggctthttaat cccttggtgc thttctgaac cactgcaatg	1380
agcttctaac tgttctcact gtgtgcaggc tctthttcctt ctaatctaath ttacacactt	1440
ctgaacacaa atctctcaca gcctgthttcc thcatgttac ctccagctca agactthttg	1500
cctacaaaat aaaattcaaa cttgttagct aagcaccttc tcatgtctat gctthggctc	1560
atatttcagc catcgtgtgc ccactthatt cttatagcca acctgaaaag ccatctthta	1620
taagaaacta cctctgctct ccatgattgg atataattaa tctctcttcc acatcacctc	1680
gccacaaaat tgtatctgtg ttgatctcat gccacatacc tgtatgtatt ttatattata	1740
aatatttgca gacttgthta atthgccatg ttagactaag thccatgaag acagctccat	1800
atccattcca thtttatata tccacaacat thggctcgggt tgatgcttaa taaatgthta	1860
ttgaaggaac aggagtctcc cacttctgac ataatgaact taththcccc agtgthtaacc	1920
ctacatctgg thcctgtcca agagtctctt cccaaatcat tctgattcaa ctgttcattc	1980
tgatctcatt aaacattthaa atgatatact taacttcgct tgctthattc tatgctcatc	2040
ctgcagtctc ctcataactt ggtthcaatg atgcttgctt ctagagaaaa aaatgtatta	2100
aataagctta tgattcagtc ctccagctgt gatggthctc actgaacatt agctcagtg	2160
thttcgaagt atggctctca gcataaccta gaaacttgth agaaatgcaa atthctgggc	2220
tcaccaagac atactaaatc aaaaattctg acattggggc ctagaaatct gtgththaac	2280
aagcctgcca gtgcagcctg gtccctthtc thctcggagc ccactcaaa gctthcagtg	2340
ctcatctccc accaatgaca gggctctcta tggaaaccgg caggacggt thccaactca	2400
actacgthtt agagthtgct thctagggt atccaggcac caagtatcac aggttagtht	2460
cccaggaag cagactctga gacttgcatg caggagtggt ctctggggtg ctctcaacca	2520
acaccttcag gaagagaagg aagcagcatt gggcagaggc atagtcaaac tacagtgtg	2580
thggcacaga agactgaagg gagtcaagc cagggggtag aggtgggccc ttagcatcca	2640
thcttcacca ttaggtgtga gthgccccac thccttgatg gtgtaacctc agthccaagg	2700
tgggtgggag tgcagcagag cagcccctac aagggccaaa ccagagatac accaggcgcc	2760
agaagtgtg ccaggaata gagaggaaag gatgggctta aggtaggatc cacagaactt	2820
ggcaatggat tagaagacag gatgagaagt gacaggthaa cactaacaca gaaatgtcta	2880
acttcggtag ataatggtgc cattggctag aagaggaaac cgaaatgaaa gcaggttgth	2940
caggagaca aaagthcact gtggacatct cagcagatg atthcagtg gaaaggaatg	3000
gatgccaga ccacctcaga ggaagatcta agctggagcc agcaataaag atacaagatg	3060
aaaatccct aacgaactgc thctcagcca tgthccccag acacgtgtg thcagatttat	3120
agthcgggtg aggttaggag gtgcgcctcc thcagtgag gacagcaaag caccagtggc	3180

tccagggagt taaaatcttt tgataatfff tgttctagca tctgtctgca gagctgtctc	3240
tcagccattg cctgccttta cacaggagtg cagtccgaaa ttgggagatg agtgaaatff	3300
attatgccta gagatctgga tccccagttg tttgggagta tattttctga accacttggt	3360
ggtttaagta atgcagatff attgatgcca cttctcttga atctgtgact ctggaccac	3420
catctaagtg aatgtgcaga gggaacggaa tggctgcaat agatctccat taaaaccagt	3480
gcatcctccc agacacatac agtagtaggg aggtgagtca atgtcaggac agcaccagct	3540
cccgttcgg tacatttcca aagttctcag tctgtgtaca aaggtttgct ctggggcagc	3600
agaaatagcc ctgggcaggt agtcaaaggc ctggtttgat ttctccact tccaggcaag	3660
tcactcgaag gctcacaggc tttttcctca cctgccacat gggctcagt agatctactg	3720
agctgtaaat aatgaaatga gtgtgtgtgc agtcatctat aagttgtaaa gtactagaaa	3780
atggtgaaac tttgggattt gggctattta aggctgaatg ctaaaaatgt caggcattgt	3840
ggagaaagga atttaaatat aagattgatt gactgggatt taaagacaaa tgaaggcaca	3900
cacgcaagtg cacaccaca ctgacactgc acagctcccg ttggaggcat atcctgacca	3960
tgcagacctg gggctctgcc tgtccaagtg cactccttta ctacataaac cctccttctc	4020
ttttggggct gtcacccac cagagctggc accgagccct tgctgctgcg cttccctggg	4080
gtgtcagctt ttgacagggg gtttcctccc tctgcaggag ccttaacatc ccttggaact	4140
ccttcccccc acccaccccc agcagtttta tctcttccta actcgggacc ctttttttcc	4200
cacacaaagt ttattgtcag ttgctggttt catctgtttg agcggctgca acaaaatacc	4260
atagactggg tggcatatgc acgacaaaaa tttatttctc acaggagaag tcaaagatta	4320
atgcaccagc agatctggtg tctgaggggc caccttctg tttgtagatg atgctttcta	4380
gttaaaacac ctatttaaca cactattaaa cactaagtgt gttaaatagt gcagttgatg	4440
tatttgtcat gtcaccttta tcatacacta aatccttctt tgtctttttt tctgtactct	4500
aatctctttc tgtaagtaat ctttgcttgc agcagtagga tatttagagt actgtggctt	4560
gacaatatat ttagtatttc aagatttcca tgaaattctt ctgatgtatg agttccctag	4620
ttaatcttac atatgtatcc ctttgtaaaa acactttgaa catttaaaat gatacatgaa	4680
tagtactcta atacaatgcc ataaaaatta taaatcattt gtatagactg gtaagtaaag	4740
attgtgagat taagaaacgc atcaaaggcc attgagctgg aaagtggat aatgagaatt	4800
caaaccaggg tctcttgact caaaatctaa ggatcatacc atttctcatg ataatatgag	4860
tattattggt atctctatcc catagacaaa gtgttaacac tgaatgagca gtgaaatagt	4920
ctcagaatff tttattttat ttagcaattc acttgtcatt tctggtcctc agttttattca	4980
cgagtaaaat aaaatagttg gactagataa tttctatagt acattcttac acaaaaaatc	5040

tatgattttg	ttatttttaa	tgtgatatac	tcatggcact	cattcacctc	attttcccag	5100
cctgcctcac	tggtcattac	ttctctgtgt	tctttacagg	ctccccctcc	tctacactgc	5160
cattaaatat	tgaaacacct	caaagcttta	cttatgtcca	cctctcctct	gacactatca	5220
ttctgtctag	atgatcccat	acatacatgc	ccattacttc	aacctgtatt	tatacgccaa	5280
tgattcacta	tatttccagc	ctagacattc	ttttgtactc	tagttaccag	cttgatatcc	5340
ttacatggct	gtttcaaaac	aactcaaata	tattatctct	caaaatcaaa	ctcatgatgt	5400
ccccacacca	tcctagcttt	ccaccaacaa	tacctatccc	tattaatagc	aataccattt	5460
attcagttat	ccaaatcaaa	aacctagaat	tcatccttaa	aattctacta	tcattccaaa	5520
tatcctatcc	atcagcagcc	actgtattct	taatcccctg	tatttccttc	aaatccattc	5580
acctctctcc	atatccattg	ctgcatgact	atccaagcca	tgcctcttac	cctagggtac	5640
caaaatagca	acaaacctaa	tctgttcatt	tgcatatttt	tttctccaaa	actgattatc	5700
tatatgtagc	aagacagatt	gttctcaaat	tgcaaatccc	actatattat	cctcttgctt	5760
caaacacttc	catggtttcc	cattgtttat	gataaaacca	aatgcttcaa	gttcgaagac	5820
cggcatgatt	gggaatttcc	tgtcaccccta	gcctacttgc	tctccatggg	acagttgcac	5880
tggctttctt	tcattcctta	agtacaacct	gtttcctccc	acctcaggac	tgtgcatgtg	5940
ccattcattc	tgctgaggag	cctttttcct	tccacttcaa	tcagctaagt	ctgattcttc	6000
ctgacaatct	cagctcaata	agcatttcct	ctaagaaatg	tctctaatat	cattaattgg	6060
ctcaggtccc	tctactgtat	tgctgcactt	ttcacagtta	taattttact	taattatgaa	6120
tgattatttg	attaggtcta	tttccatcca	ttagacataa	gcttcatgat	ggccagatta	6180
ctgttttcta	tccatcgttg	tattccaata	cctgacagaa	ggagggcggg	aggtgggtggc	6240
acacaagaga	tgctcaaaaa	caattgttga	ataagtaa	gaatgaggcc	atttagaaat	6300
aacgaaagta	cctgtttaca	aagtacatgt	atcaaaacta	tgaatgcatt	ctacttacat	6360
ggttttctcc	aaataaaaca	aaagacttca	atcaggatta	atacctggga	taaactgagt	6420
cattaaatct	ctcctttgcc	atcaggagtg	acattgaaac	aaatgtctgc	aaacaacaaa	6480
tacttttttc	ccaaaatata	ttgaatggca	tttccataaa	caaactagaa	catgggagga	6540
gaaagaaagc	aatattaatt	taaaattaat	cttatcacat	aacttatacc	atcagggatt	6600
tcgggtaaaa	ttcctttcag	gcacatccat	ttaacaagaa	ttgattgtta	ctgaaagcct	6660
agaagagaat	ttggcacata	cttggtgttc	aaatatattg	tgactgagtg	aataaatgat	6720
gcaagtgtct	aagaaacaca	aaataaggac	atgattacag	tcacgggtga	gttcacagtc	6780
atctccaaaa	tgaggatatg	catcccaggg	aggaccaaca	attcattgga	gtgctgaaat	6840
aaaatactca	aaggtcattt	tacatgtatt	ttttctctaa	attacttttc	ttaagacaca	6900
gaaaacaaaa	aaagaaactt	agctttgtta	ctttctaaca	aatagttaaa	tcattaaaca	6960

ggattgacac tagcatcctt gtttgggtctt atgccttagg ggaacatgaa atgtgtgaag	7020
acattctgag atctgaggggagggttagaca gtaatacagt gggactgacc aggcttcagc	7080
acacctttac ctctctcag cagatttcag tgatgagcag ttacaaacta gattgaaaga	7140
ttatattatc tagttctaaa agaaaaactaa gcctcccaaa agcaacaagg gaactgagag	7200
gaatcctgca aaacaaaaac aaatttttaa acttgcaactt tgtaataacc ctaatatgta	7260
atcacagtaa tgaacagtaa gataatgaca gaactgacat atttccttat ctattaaagc	7320
catattaaca ggtaaagcaa tgccagtcag tggtagactt cttagaagat atttaataca	7380
tactagacac atacacacac acaacatttt ccttcaagggt gtatgtatca gaaaatcact	7440
ttttaaggcc ggatgacgtg gctcaggcct gtaatcccag cactttggga ggccgacgtg	7500
ggcggatcat ctgaggtcag gagttcaaga ccagcctgcc caacatggcg aaaccccatc	7560
tctacaaaaa tacaaaaatt agccagggat gatgggtggat gcttgtagtc ccagctactc	7620
aagaggcaga ggcaggagaa tcacttgaac ctgggaggca gaggttgcag tgagccaaga	7680
tcacccattg cactccagcc tgggcaacag agtgagactc tgtctcaaaa aaaaaaaat	7740
cacttttttag ataaaaattca tgctatagag agaagactat gaaaatatgt ttagcaatgt	7800
gtccatcatt aggtgattga gtttcctttt gttttgtttt actgaaaatc atataaagta	7860
tgttatctgt aaaagtcttc tgacatgcac acataaaaaat ttgggagaaa agattaacta	7920
taatgtttaa tagattttgt acacatttct ttaaaaatat ataaaacaca acacctttca	7980
attggtttgc aagaataacc aattgacatc atggaaaatg gaaattcact 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 ccacttatat tataatatta aaaatatcat tacaaaaaaa	8280
acctgtgttt attttataac tcagcctttt taatttctaa tttcataaat atattataat	8340
ggatattgtt agtaatgtag tattattaca tgtatataat ttataagtaa atatacatgt	8400
tttggctact catgcataaa atgtttcacc cataggagca cataatcaga aatgtctgga	8460
gaccattata gtaatagata gatcatattg ccacatattt tatctcctcc ttgacaactg	8520
agctttccag atcttctggt gaaacgaaag agaaagttgt aacagaagag tgattaaaat	8580
gacaaaagca ttacttctat tacttctatt ctaataatat gagcaaagct ataactatca	8640
agtaataatg cactaaagaa ggtgattaat ctgatatatt cacaggcaac taataagacc	8700
tttctattgc agcatgaaa aatatgtgac aattatagat atcctgtgtg cagtgtttca	8760
acctttatgt gacctgttct actaacagat ttagtgatgt tcactttgtt agaattttct	8820

tacacatgcc	ataacttgct	tcagtctttt	gattatgaat	attatggata	ttaaggattc	8880
tagactattc	tagatttaaa	aaataatatt	gtcacctcaa	tcagaaggga	aatattaaat	8940
agttctcatt	ttttcaatgt	ttactcagtt	tttgtccaat	gtaatgaaag	tgtcagcagt	9000
acaggttaca	aaataaaatg	tgtattaaag	taaactcatt	tgaacagggt	aataattgta	9060
gagggagggg	aaaggctaaa	agattgaatg	taaaacttat	gaaaagtaga	tacatcgtct	9120
ctatgatttg	cagtagtcaa	ctgcatacag	atgaatcatt	ttaatacacg	ttaactactt	9180
tcctttttaca	gatggagaaa	ctgagaggaa	gaaagtttat	atgggttcatt	aaactttgtg	9240
atgcaagcta	aactaacctg	tctctgtatt	ttccatctac	tgcccttatc	actatctcat	9300
tagaatactc	ttcaagcatc	tccttactga	ttttcttacc	aagcatttgt	taagttctaa	9360
tgagagttgg	tagtaacatt	ttcaccctact	ctgtgaaata	tgaaatctta	ttcataggcc	9420
tcttcttttta	ttcttgtatt	tgcataatcaa	ccaattaatc	aacttgcttt	ctttatgttg	9480
cttattatct	tagtccttac	taaattgcct	cttaatgttg	tccacataac	agaaatgtta	9540
aggtggatac	ttaacathtt	agtccagtct	agccgggtgcc	agtgcaatgc	caaatcatga	9600
attaaaatat	aattacaaga	accacttatc	aaattttaac	aattccttca	gctttgtgac	9660
agttttttct	acttcgatta	aagtcaagta	aaattaaagt	taaatathtt	tattaaaata	9720
tctcctttta	cattccatat	taataaacat	attaaagctc	atgcttctaa	gtagattact	9780
agaagttact	ttatcgaatt	acagcaatgg	ttaattctag	atcatagaat	ttagaatgac	9840
tttttgccct	cttctttttt	ttcctttttt	ttaaacagag	tcttgctctg	ttgtccaggc	9900
tggagtgtac	tggcgcgatc	ttgactcaat	gcgacctctg	ccctgcagggt	tcaagtgttt	9960
ctcctgcccc	agcctcttaa	gtagttggga	ttacaggtgc	ctgccaccac	acctggctaa	10020
tttttttttt	gtatttttag	gagagacagg	gtttcaccat	gttggccaga	ctgggtctga	10080
actcctgacc	tcaagtgtat	cacttgccct	agcctcccaa	agtgtgggga	ttacaggtgt	10140
gagccactgt	gcctggcctg	actttttgtt	ttcttcttaa	tacttactag	tattttctga	10200
attttttaaa	aagaaacata	aagtactttg	ataaaaccaa	cagtctcatt	gttcttaaaa	10260
ttgttcaaag	gttctctgga	aaaaaaaaag	aaaattatca	tttggttaag	aatcatgttg	10320
gtctgacatc	aatcatccta	taggagtga	tattgaaaaa	gtaagatata	ttgtgggtata	10380
atcgagattg	cataaathtt	accatttttg	agaagaatct	gctccaaatc	ctggcttaat	10440
gtaatatcca	gcatgctact	taattttctt	gtcttcacct	tttcatatcc	acatccacct	10500
aggtgccacc	tcacagtata	agccagcata	atccattctt	ctcaatgaaa	ccacaataca	10560
tctgaccctg	catctcagga	gaactgtatc	agccacagca	cttccagttg	actatgaatc	10620
tgaatgttat	gcctcaggag	aaacatcctt	gctgggactg	agtagtgatt	caaggagata	10680
gttatgattc	agtcaagaaa	ttaataatta	gtgttathtt	tattattgag	acagagtctc	10740

gttctgtagc	ccaggctgga	gtacagtggc	atgatctcgg	ctcactgcaa	cctctacctc	10800
cccggttcaa	gtgattctcc	tgcctcagcc	tcccaaataa	ctgggacagc	aggcacttgc	10860
caccacgcct	agctaatttt	ttgtattttt	agtagagacg	gagtttcacc	gtgttagcca	10920
ggatgggtctc	gatctcctga	cctcaaggtc	cacctgcctc	agcctcccaa	agtgctggga	10980
ttacaggcgt	gagccactgc	gcccggccat	aaattattaa	ctgagccagg	cacagtggta	11040
cacacttata	gtcccagata	ctcaggagac	tgaggttgga	gtatcctttt	ttatgttatt	11100
ttatttttaa	ttattatggg	tacataatag	gtgtacatac	ccatggagta	caagtcatgt	11160
tctgatacag	acacataatg	tttaataatc	acatcagggt	aattgggata	tccatcacct	11220
caagcattta	tctttctttg	tgtaggaac	attccacctc	cactcttgga	ataggcaccc	11280
tgttggtgcta	ttaaatacga	ggtcttattc	atttcatcta	actatatttt	tctacccatt	11340
aaccatcacc	tcttttcccc	tcttccccac	tacctttcct	gtgaggctgc	aggattctta	11400
agcacaacag	ttagaggcca	gcctggacaa	catagtgaga	ctcaatttct	aaaaaataaa	11460
aaagaaatta	ccaactaatg	ctaaaaaaat	agtctctgat	gcttaggtat	gaattagaaa	11520
tgaccaaaaa	aaaaaaaaaa	aaaaagactg	ccctttgctt	ccttctcccc	ttctcttcaa	11580
gttttccatt	gctactcatt	ttagtctggt	ttaatcaggt	ttcatccatt	aaaagcaatt	11640
gttgggatca	cacattttga	gttgtgtcag	tggacttccc	tcatgctggc	atgattcctg	11700
ccccaaagccc	ttagtaaaag	ccaccaagcc	atataacata	atctctcatt	gagtaaaaca	11760
tctgatgtgt	ttagaatgac	ttctagcaaa	aaaccagcct	gtccagcatc	atctctgtat	11820
aacagataaa	ggaataggta	ctgcatcaaa	aggttataga	acctgcccac	atcaatccca	11880
tgtgttttgc	aatggaatta	ggttgaacta	aagtgaaaat	tcagttttct	actcctcatt	11940
aacatgtctc	atgttgcaag	gttgagagga	aggagaagaa	gaactgtatt	tacagagaga	12000
ttccccctct	ctttctttct	acagattact	aaaacattca	aagaatcaaa	tttaagaaat	12060
cagttcatca	gagctcatgt	tgccaaactg	aggtgagtgg	aactgtagaa	aaaatattta	12120
agtatagata	caatgtggca	tacttgactt	tttgtcacag	aatgaatagt	aaatgacatg	12180
ttcagataag	ttgttgtaat	attatgaaaa	tagtatttta	gtcagcttaa	aaaccaatgc	12240
caaaaaagcc	aaacatatga	tctatttagc	tactaatgta	aataaccata	ttatatctat	12300
tcttattggg	aagaggaaga	aggggtggag	agagagttgg	ggtgaaggta	cagtaacaag	12360
gccatcctat	tgtaaaactc	cagtggatat	cattcacagt	gcagcctatg	taaacagtcc	12420
ctcctggagt	tgtacaatgc	tgtggtttgg	gtgtatccat	ccaagatcaa	gacactatga	12480
ccaacatcaa	aagtggcttt	ttggttttat	ctgcctgatg	tgctataata	aaaggggtatt	12540
atggccaaat	ccaaggcatg	tctatcatga	attaataata	ggaggagtag	cagcatgcat	12600

gctagttatt	tgccattcct	gccttagtta	aatatgatgt	gataaaacca	gcctttccaa	12660
ctgaaatagt	cacctttact	gactctcccg	caaagtctc	aatgaccac	attgctctag	12720
tctttaaata	atatgcaata	gttctttggt	agaagaggaa	ttataactaat	tctttctcaa	12780
atactagcat	cacaagaaaa	ttaattcttg	ttctctggag	agtcacctag	taagtatctg	12840
gagcacagat	gtctggtcag	gtaagttttg	atgaggagtt	aaagggataa	gaagagtcca	12900
tgagaagggt	attttccaaa	acacctttcg	gtcaattcag	tgacacattca	cttagtactt	12960
tcttgtcagt	atctgtatca	gccactaatg	ttcaaaagtg	agtaagccct	gaaaacctgt	13020
aggactacat	gagccttctg	ccttttctct	ccttttgttc	acttcccact	tatcactcaa	13080
tcctctgcaa	cctggcttca	ataccaccat	aaaatatcaa	ctgctcttgc	cgattcaaca	13140
atgacatcca	gataacaaaa	tccaaagaaa	ccacatcagt	cctattcttg	gacctttcaa	13200
cagtatttgg	tcctgttggc	ctgtcactcc	ttgaaatagg	actatccctt	ggtttgcatg	13260
gccttgtata	ccctgatttt	ccccttacct	ccctagctat	tccttcttag	tttcctttac	13320
taggtcttac	ttctttgtat	attccttaaa	tggtgctgaa	catcaggctg	tgctctaggc	13380
ctctcatctt	ctcaggtcac	actctctcct	ttccttggcc	ttcactgcca	cccatatgct	13440
gagtgcctct	aaagttgtat	ctctaggcca	gtcctctttt	gcctccaaac	atgaatatat	13500
gcagccatct	acttggtacc	atcacatgga	taattctcat	gatctcttcc	agtatgactg	13560
cttctttatt	tttttctggg	ctctttttta	gcattgcttt	acatggaact	ttatcatgtc	13620
tctcaacctc	tattttatct	tttatctatg	tatgtagagt	ctgtgtaatt	tcttcatctc	13680
ttttagataa	ctaatatctc	ttcagctttg	acttgatttc	tgtgtaaccc	atttattgcg	13740
ttttcaattt	caatgagtat	gttttcctat	ctgcaagttc	tatttgtttc	ttttgagaat	13800
cttcttggtc	ttttaaacac	atttcttatt	ttaatttttg	ggggtaccta	gtagttgtat	13860
gtatttttgg	agtacatgag	atgttttgat	acaagcaaac	aatgcataat	aatcacattg	13920
tgtaaaatgg	ggtatccatc	ccctcaagca	tttatccttt	gtgttacaaa	caatccaatt	13980
atattctttt	agttattttt	aaatgtacaa	ttaaattatt	attgaccata	gtgactctgt	14040
tgtgctatca	gatactaggt	gatcttttaa	aaataatggt	ttctacttaa	tctcattttt	14100
atgattccct	cttttacgtc	atttgtcatt	tcaaatacag	tcacttgtct	gttgattcta	14160
ttatgtgaag	tttttgagga	taatcttttt	gttactttga	ttccaccttg	gtatggtttg	14220
gctgtgcccc	cactaaaatc	tcactctgaa	ctctggttcc	cataataccc	acatgttgtg	14280
ggagggacct	tgtgggaggt	gattagatta	tagggacgtt	ttcccccttt	gctctgttct	14340
ttttctgccc	accatgtaag	aaagatgtgt	ttgcttcccc	ttctgccatg	attgtaaatt	14400
tcctgaggcc	tcgcagcca	tgaggacct	cttttctttg	taaattaccc	agtctccggc	14460
ggttctttat	agctccgtga	gaaaaaacta	atacacacct	catgatgtat	tgtttaccac	14520

tgaattgta tgcttaaatt taatctcact tgggaccctg tacaacctag acttaacata	14580
tctacctcca gagcagttac atctgtcaga cattctagag gaatcagcag cacatggact	14640
ttgttgttgt taatttggtg tcggggggagg ggggagggat agcattagga gatacaccta	14700
atgctaaatg acgagttaat ggggtgcagca caccaacatg gcacatgtat acatatgtaa	14760
caaacctgca cgttgtgcac atatacccta aaacttaaaag tataataata ataaaattaa	14820
aaaaaaaaag gttctgggag tattcaggta gtattaatga agattcagac atcgtgcagc	14880
caggcccatg cttatgaatt ttcagggtgat acttcttttt cttttttctt aatttaaagc	14940
tggatctcgg aaacagataa atttattttt ttatgacatg acgagcattt ttttcattct	15000
agttcatgct gttattgggt gtttagttct ttgagactcc tggccttttt ctaaaacctc	15060
aagttcaact tcctattttg cactggccca aggtcccatc tccagtctct atgtaaatgc	15120
taaacataag cctgtggaat attctagtct caccacatac tattcacatt cttctttggt	15180
tttgggtcttc caggattttc cttacttttc tatgaaccca gtcttgcat tgaatggaa	15240
tttattatat attatctatc ctttctattt gttttatgca gaaagtgttt tctaaaatta	15300
tttaggcttc catattgcta gacatggaag ttgtaattat ttgttcagtg cctgtttcta	15360
catctaaact gcaagaccca tatggcaact gtgaatctta gtcccagcta atttctgaag	15420
cttagaatag tgcctagcac aagaagttgt ttatctaaca tttttaaaaa taaatattaa	15480
attcatatct ggaatgaata ttaagttaga gctggtcatt gaggtgagag gaggaagcca	15540
agagagaata tgagagcctc aaagccaaat atctttaatg tactttttca gaaaagaaga	15600
cagccaatgt cagggtggagg aactggttta tgaggtaact ttcttggaag aaaatagaaa	15660
ttactgaggt tttagataat ccaaataatt aatcaagtca ccaaggttta ttgtggggaa	15720
tctttattat taattaaaat gagtgatgaa atcttaatat acgacaaaag ttaaaatttg	15780
cttttgagg cagatgaatg gtctaggtat caaaaaatta agttgagtct ctaactcaca	15840
caaatttaca accctatcac tttatgaatt tgtttaggag attattttta ataactgg	15900
tgaagtctaa gaatagctaa aatttatagt acacttattg tgtgctattg actcttcttt	15960
gaagttttgc atatagtgat tcatctaate ttcataaccc attttacatg tgaagaaact	16020
tagatataga aagattaaga aacttacata acttatccaa agttacacag taaaactctg	16080
gcattataac ttcaaaatca gctatcctac agtgagtaca gtgttctgtg cattgaaatc	16140
aaataagtga gatagcatcg tgatatagta ttacgtatgc aaacactggt acagagatct	16200
gtctaaagtt aaattccaca aatgaattct ttaaaagggt ttaatcaaga agaatatata	16260
aacaggatgg tgaaaaattg tcatattatt tgttttttta aatatcttta tgatttacag	16320
gcaagatggt agtgggtgta gagcggatgt tgtcatgaaa tttcaattca ctagaaataa	16380

caatggagca tcaatgaaaa gcagaattga gtctgtttta cgacaaatgc tgaataactc	16440
tggaacactg gaaataaacc cttcaactga gataacatgt aagtataatt ttccataaac	16500
aatttttattt caatatatcc ctcaagttta ccaattcaaa ttcataatttt aattgagagg	16560
ctgacttttc tttctttgaa actaaactgt gaaaacaatc cattaaaaag ctaaataatc	16620
catatagctc cctaacgtaa atcattctaa gacttaaaga atcatttggc atttatatag	16680
taaattttat ttgctaaaaa ttctcattaa ttatccctgc aacattcctt atgagtgatg	16740
ttactgtcag atgtcattag tggataggcc ataggagggg tacatagatg ctcaaggtca	16800
gagaactatt taattaatga tccacctcag aggcttcttc atttttcttt gtaacattta	16860
tcacaattga aattacaaag ttatctgtgt aaattttgta ttgtttggct tcatcctaca	16920
ctgtaatcat cctaaaagaa agaaccagtc aaccttcttc atcctactac cctcctacca	16980
cccagtctcc atcatataac acatattcaa taaataattc ttgcatgact gaaagaaaag	17040
aaataatata tgcatagaat ttaaggacat tctccaagt tggttacatt ctgctagttt	17100
aataagccat tatttcttct cgatgagctc aagattaaaa ggattttgat gattoccata	17160
ctagactggg aggtaccagt tacagatgta ctaactgtta aatattgaaa tgctttccta	17220
tttgttggtt aacaattact gcatcaggcc cacaagttg tcttccgaga tgtttcaa	17280
ccactgcccc tgctgctaaa gagttatgct tagcaaagca aagcactcta agacactgct	17340
ccaactccat ggccctgattg catcttttat gactggccaa tgctcacgca ctgcagtttg	17400
ttaggtagtt gaatattacc tctgcttcca cacattaagg aatgctcccg aacgcacttc	17460
ccaagtgttt atttatttat cattatacta gacaatatgg tgatacgatg gtcacagaat	17520
agcggtttcc acctccagag ccataatct agttgaaggg aaagatattc caacacaaga	17580
gtgttgacaa tcaagataga atatgatcaa gggcccagtg tgaggcccag gcaatgatca	17640
ctgcaggaat ctggggaaga aagagaccag cgtgcttggg atatctagca aaagtttcat	17700
gaaggagaat ggactttgac tttgaaatat gggtaggatt tacatatttt gagatgagaa	17760
aaagaaagtt ccagagaag gaaagcatga aaaggcaaac agtctgtact gaacgcgatg	17820
ctttgacaga ataatgaaga aagggacctg ctggaatgat tgatcagtgt tcatcattca	17880
caccatcatc atcaaacac ttatttaatg agaacttact gttttttagg catggcttta	17940
atgccctata tgaatttttt tcttgattaa tcttacaac aaacatatcc catagatagt	18000
tttattgtcc cccttagaaa agataaattg cctaggctga cacagtcagt atatgaggca	18060
gtcaggattc aaactaagtc tgtttgttca aaaaattaag aatggccagc tttttaaaat	18120
tttctgtctc cagaagtatg atttggctcc actgaagttt gcaaaacaaa tgtgataccc	18180
aaacctgtg aaacttttag tgggaaataa ctttgcataa gtcggtttga gagagcgtgg	18240
aaacctgtct tgaaaagttt taatttaact tgcaggaaat aaaaatgatg ggtttctcaa	18300

ttaaaaattt	caatcaagga	aggatatgag	ctaacataac	atTTTTTTTaa	aaagatcagt	18360
ctggtaaggt	agagggtgcat	aaactgaaaa	ggagcaaaaag	tgggtggaatt	cagttagaaa	18420
attattgtaa	ctgtactgat	gtcaaagtat	gaaacccatga	actaaagtag	tacccaaaagg	18480
agtgaggagg	atggaataat	tcaaaaagata	gaggacagat	gtgcagaacc	tggagattat	18540
aagatgtgaa	aggaggagtt	tgagaaaatt	tcagatttttg	gaagtgggtgt	catttttacta	18600
aaaggatata	ataagtagca	aatttttggat	aaagttgggt	cccactgagt	ttgagatggc	18660
tgttggacat	gcagagaaaa	ctgtcttgta	tgctgttctt	aaattgaaat	agacagacct	18720
ttaccctctg	atactgacat	atTTTtcttt	ccaggctcac	cctccatttc	cctaaacaca	18780
acacatgcac	tagctctcct	tactttattg	ctccacaaaac	atcttacacc	tccaagcatt	18840
tgtgccact	gtaccttcta	tctggaatct	cttttgtcct	cttgtgtgcc	tgaaaaattc	18900
ctttcagatc	ttcaaaaatac	agtgcagatg	ctattttcttc	tagctcaaat	attatctcct	18960
ccatataatt	taattactct	ctTTTTtctt	ttctctactt	tgactttaca	tttatttgaa	19020
tgattgcttg	attaatttct	acctgtaaat	tatgtgaggg	caggtcctct	atattttgct	19080
cgcagttaaa	tctgcagcac	ttattataga	gtggatcat	tagagtaata	tacatatatt	19140
tgaggacatg	ataaattaac	ttcccctata	gtatttatca	cattgcatct	caatgacttg	19200
cttatgtttc	tgTTTTccca	tataaattga	gtaacttgaa	aaaagagata	tctattaagt	19260
atttaatgag	aaattaaagt	acaaacttta	gtatgcataa	caacaaattg	ggaaaagggt	19320
gtaaacaaag	agatttgtag	ggcccatgag	ttagagatcg	tttcagcagg	tctgaaagga	19380
agcctaggaa	tctgcatttt	agaggaccac	ctcccaaccc	caacaagtaa	ttctgcttct	19440
tgttgtctgg	gtactgtact	ttaagaaatt	atggtgaaat	gatatcagcc	tttattgtat	19500
ttatcttatt	ctcatttttt	aataactagca	cttactgacc	aggctgcagc	aaattggctt	19560
attaatggta	agttttaata	ttattttgta	actgtaattt	gccaaatcat	aaagagtaaa	19620
agtgcaagtc	ttttgtgtac	ttttggccaa	ggcagtatct	atcaagttga	tgtctttgtt	19680
cttagttcgc	tcaggtggtg	ttgaaacaag	acagtgctga	tcccaagtgt	cccatggagt	19740
ggacttttagg	tttccccttt	ccttttagaa	aaaggaagaa	gttgtagtgg	aggactaccc	19800
actctgcact	caaaattgcc	ctcatgaaaa	tttctttggc	agctttgaga	accttttact	19860
gccctggttc	taagggtggca	tttctgtaga	cttacaaatt	atgtttgatg	acaccgttta	19920
tgtagcttct	cctaaccacc	agagtagctt	gctttgttgt	gaattcaggt	taatcacaaa	19980
gtataataaa	aaagaattgt	cagaagtctt	cccagctttg	ggtctataac	ctgaaggaaa	20040
agtcactact	cttcaacatc	atcctatgta	ctctcaggct	aggatagcag	aaatgcaatc	20100
cctagaaaac	agcaacttac	ttctctgacc	aaaaaaatgc	agttaaaaat	tagttcaatg	20160

tacctggtag	ctggcctatc	ttaggtactt	cagtgat	ttt acaaagtgat	ggtagtccta	20220
tggtgtgttt	tcagcttcac	tacgtattta	attcatgctt	attgttaatg	aaactgtgat	20280
aagcaattta	ctaggggtatt	tggttgggag	atgccacaaa	ggaacacatg	tatctcttaa	20340
tggaagcctg	gtcctccttt	atccaggaaa	tttgctagga	aaaaaaagcc	tttaggtggt	20400
tgtgctatta	aaccagggca	ctacttaaaa	gccagcccag	caatagttgt	gtgatttacc	20460
attaatttct	tagtaataga	ccacacaaaa	gaagaaaatt	atgggaatgc	gagttgagag	20520
gaattgggtg	atcagcctac	cccagcccgt	ttcagctctg	gccagtagac	tattcacgag	20580
ctctttgaaa	acattttaat	aaaccttatt	tagatactag	aaacctctg	tcacctcaa	20640
gaatattctg	tggtatagcg	actcctttat	gagggcatgt	ttggtaatag	agcatcagtc	20700
ttggaggtgg	actggattct	acaaggtgaa	ctgcagtcac	taaggagtct	tttggtgag	20760
accagttttc	ctccaacttc	aatgtgtgca	tgaacctcac	atcaaatgt	agcttttagat	20820
ttgtcccatg	atgtggttcc	aagaatcagc	acttctaata	agtttccagg	ggatgcccac	20880
gctgcaggcc	cacaaaccac	actgagcata	gcaagactat	tgagaaaaag	gaaatttccc	20940
aggagtctgt	ggcctgagct	ggcacatcca	ataatgacct	atcttaacct	caactcatga	21000
ggaattccag	ggaactctga	agctgctcaa	aatttgaagc	ctatatgcca	actaaattca	21060
gaaatgttct	ccaaaatgct	atctataagc	aacagtagtc	acaaatgcat	tgtagaaata	21120
tatcgatcat	gctttttgga	aaatccagca	tgtcctgagg	aagaatgtat	aagacataaa	21180
agtcataaat	tatggaaaga	ctcttcagct	tcttccaaat	gtaaaggaat	catgatcttc	21240
ccagcacatt	aatgcccttt	ctcattagaa	tgtggggccg	gtccagacct	aataacattg	21300
tctgagcaga	gaatccttgg	aggcactgag	gctgaggagg	gaagctggcc	gtggcaagtc	21360
agtctgcggc	tcaataatgc	ccaccactgt	ggaggcagcc	tgatcaataa	catgtggatc	21420
ctgacagcag	ctcactgctt	cagaaggtga	ggccaccact	acctacctat	ctgggaacaa	21480
ttagaataga	caggtcatga	agactgcacc	ctctacccta	ggattgaatt	gagccagaaa	21540
taattcaatg	caaaaaaatc	agtaagaatt	ttcttcctat	tcatgaaagg	aaaaggattt	21600
ttccccttta	gcatgcta	ttagtgtat	ttctctgttt	caggtaataa	tatattagca	21660
cagtaaagaa	caaagattta	tatgtcagaa	tggttttttaa	atcctagcta	taaaagctta	21720
agaaatttac	taaatctcca	taagctttat	tttttttcca	aattaaggga	caacactgtt	21780
atctgtgact	tagtgttact	ggtagcattg	agtacactaa	tgtaaacata	cgttaaatgt	21840
tagcgaaacg	aattgctgtg	gaagatttgc	acattatata	atgggagctg	atggctaacc	21900
tagagactgc	cccatgccat	taattttatt	attcataaag	attattgagt	atctagtatg	21960
agcacagtgt	tatatattgt	agaagctact	agtataaaca	aagtattgcc	tctgccttca	22020
aagagcttac	actcgaatgt	tggaatcaga	atgcacaaaa	ataatgatca	attacaatga	22080

gtagcataaa	taaaattaat	gtaggcaact	tacaagaatt	cttaattgag	gtgactaaac	22140
tattgccaac	actaggggtga	tatgctacca	gtggcgagta	ggttgcataa	acttacctta	22200
ttggtaaaaa	gaaaagttca	cattgctcat	aaaagaagga	ttttagattt	cagcataact	22260
aaaatctggt	tcaaacctgc	cttggtactg	gggcatcgca	gaccacaaca	gttggtggga	22320
acttaactca	aaaagttcac	ccagaaaaat	aatggagatt	tgaactcgtg	tgcccctgac	22380
catatcaatt	ttcttctcag	actcttactc	taaactggac	ctccttatca	cacacacaaa	22440
gccttccata	ggcagatcaa	tccagtctta	tttctcaaag	catgtacctt	gagcttcaga	22500
taaacagcat	tgttctcttc	ccctggactc	ttcctacatt	tccctaccta	tgagtatctg	22560
atcaatctgc	ttatccttga	aatgttaata	tatttaccac	atctctattt	gaattttatg	22620
aaatttttga	taatttctaa	gtagtttttt	cagatttata	ggcactactt	catggtacag	22680
tgactgttac	aaacgtattt	gttaaattta	gaaggaataa	agatttaaaa	gactagggta	22740
gttactgaac	taaagtttta	ggaaatccca	aattatttca	aatttttctt	atggtaattt	22800
tatgacttaa	tatttttata	tgcagtgaac	aaatttgaaa	ctttaaaaga	tactcccaga	22860
attatcagtt	ttctgatgta	gattggcaaa	tttattacta	tatcccaaat	aaccaagag	22920
acaaaattca	caaaaacatt	tcaattttca	ttgccacttg	aaaggccaaa	aagcagaaat	22980
ggcacgcatt	gatttcaatc	gtactcttga	gtgtgggaac	caggaattaa	aatacctgga	23040
cttatcaggc	acttagcata	accaagaacg	gaatagaaac	ctccctggat	tctaagccct	23100
attcagtccc	aatcaccaaa	aaccaagtaa	acgatatcac	tataatgaaa	gccacagtta	23160
taaatatcga	caacgattac	caaaggaatc	catggaactt	tgaattttgc	caccccacat	23220
ccttctattc	attaccatga	ttgatccact	aaagctaaca	gactctgtga	accttgtatt	23280
ggacccctcc	ctaaagacct	gattgtcact	gagaaccatc	agtgaggatt	tgtttggggc	23340
atgaccagcc	ttacatcaaa	gtacatagaa	gtgatgaggt	cttatcaaag	aggattattg	23400
aattatcacc	tcttctatgt	agctttccct	gatactctct	ttcctctcca	ttgagttcca	23460
cagaaatttt	tttatctgcc	tttaacagtt	gtcctcatga	tttgtgatat	ttgacttacc	23520
tcttgtcagt	ttccttcact	agtgtagagt	tcctcaaaga	aagagaccat	aattacttat	23580
atttttattc	ctggagactc	atactattcc	ttatacaaag	tagacactta	acaatggctt	23640
gttgaactat	aattaatgaa	aataatagct	accttcatga	aagttcactt	tgtgccaaac	23700
actatagttg	acataataca	tttgtctcat	taatacttaa	caattgtgtg	agaaggtatc	23760
accaatcaca	ttttatatgt	aaataaaccc	cagagctatt	aattaacttg	tcataaataa	23820
cacttttcat	atgtggcata	gccaagattt	aaatataaat	gttactgggt	ccaaaatgat	23880
gctctaattc	acttgctgga	aagaaggaaa	ggaagaaaat	aaacgagtgg	aaggaagaga	23940

gggaggggaag agagaaaagg aaggaaaagaa aaaagagtct cttcagaacc ttcactgtaa	24000
agactccgag caaaagaagt tgaatataaa aacaacatag gtttgtttgt tttctaatat	24060
tttttcttca aaatttttaa ctcaggttca ctcttacaca aactactgtg tcttataaaa	24120
gtatttccgg tcatagaatt tttattttct gtattaactc cactatctaa tctccataaa	24180
actcctaaat tgggtattatc ggtaacattt tgtttttact caacccttag gaacaatggt	24240
aagttaatca gccctccaca tcacagatcc ttattttcat cagtctgtac aaggcatttc	24300
tctcatttta attttttttc ctctgtcat ccctggattt cactttcact gccctccttc	24360
cacccatatg cctcatacta atatattcga aatatacatg tcttaaaggt acatgcacgc	24420
acctacaaaa cctatagtgt ttttttgtat gtatatgtct ttaatttaaa taagtagcat	24480
tgtgtaaaag tctaataattg tttcttactg ttttactca attcttgga ttttcatctg	24540
atgcactgct gcatagcacc ccatggtagt cagccaccat atttccttca tccaattagg	24600
ttgcatgacc taccttccca ttgccacaaa gagtacacac aaaatatttg tacttatctt	24660
tctgtaaacc ttcaggaatt tcagaagcac acatgcaggc tgctaaatat accagaatac	24720
tttccagcca cttaaactct taccagtatt gcaaaagagg ccccatcttc ctccacatca	24780
acatttagta ttattctttt gttaagttt tatcaatctt ttaaattgtac acaagatgct	24840
catttttata attttaattt ctcagattac tagtttgagt atcttttcat atatctaaga	24900
gctgttttga tctcccctac catgaactgc cactaatatt ctttgcctat tttacaatgg	24960
tttttctgct tattttattac tggtttacag acttttaaaa tatattctac aaaaatttta	25020
gacattaaac attaccaata ttttcccatg gtctctcatc catctggtaa acttgtctat	25080
ggtatatcta attttgattt aatagaattc attctatttt taccttttag tttgtgtttt	25140
tgttgttttag ccaaaaagtc cccattccta ggtcataaag gtaatgtcct tttttttttt	25200
ttaacgctac tgttctctct ctgtctcccc ctatgtatat aggtgcacat atacttgtac	25260
acacatacat atacctatat atgaggggag ttcgataagt ttatggaaaa taaaattaaa	25320
agataaaata aaaaattata aactttattt ctcaacataa gctccttcaa gttcaagaca	25380
cttttgtaag caataatacc agccatatcg tccatcccta aagaactgag ggtcctgaga	25440
atttaactat gtcaatgcag tcttttttac attacttttt tacagtactt attgatgaaa	25500
aatgggtgcc ttttaaagat tgttttaaga ttagggaaca aaaataagtc agaggaagtc	25560
aaatcaggac tgaaaggtgg atgcctagtg atttattgct gaaactttca taaaactaac	25620
cttatttgat gagaggaatg agcatgagca tggttgtgat ggagaagaac tctggtggag	25680
ctttcctgga cactttttct actaaagctt tggctaactt tcttactctc ataagaagaa	25740
gatgttattt ttcactgacc ctttagaagg tcaacaagca aaatgccttc agcatcccaa	25800
atgtctgttg tcatgacttt tgttcttgac tagtctgggt ttgctttgac tggaccactt	25860

ctacctcttt atagccattg ctttgatggt gctttgtctt caagattgta ttagtaaagc	25920
catatttcat cttctgttac aattcttcaa agaaatactt cagaatcttg atctgacatg	25980
tttaaaattt ctattggaag ctctgacctt ggggtgcagct gatctgggcg aaacagtttt	26040
ggcatccatc aagtagaaag tttgctcaac tttagttttt cagtcagaat tgtataagct	26100
gaaccagttg agatgtctat ggtgttgtct attgtttctc acagttaatt gttggtcctc	26160
tttgagacat gaacaagatg aaatttttcc tagcaaactg atgtggatga tctgttgctg	26220
cgggcttcac cctcaacaac atctctttct ttcttgaaac aaattatcca ttagtaaact	26280
gatgattggg ggagatgctg tccccataaa ctttttgtaa ggcataaata atttcacat	26340
tcttccagtt tcaccataaa tttgacgttt ttttgcttca attttagcag cattcatggt	26400
gctttgataa gagctctttt caaattcatg tcttattcct cttagtgcct caaactagat	26460
cttgttcagt atgacaagtt agtatgagtt tatctgcatg caaaaatctt tgaaatccat	26520
gcatagtttg tttataatat acattttcaa tgaacttttg aagaccccat acatacatat	26580
gtatatatat gcacacacac acacacacac acacaaaaat cttcaaccat tatcagactt	26640
agtgcagaaa aattattcat ccattaacaa gataagaatg ccccttatca tcaactactat	26700
ttaaatggag ctcttggtta aaggaaaaga cagggattga aaaaaattag ttaaatctaa	26760
aatgtttatt atttcaggtt tcttagttgc ttaaatggga agggaggtat ggacaaaaga	26820
gaaatcaaag atatttgtgt tatgctactt atcattaaag tatcagaata acttcattgg	26880
aatagaaaaa caccaagatc accccacgat atgttttcta aaatcttctc catttcttta	26940
gacaagtgac catgtattcg gccagtgaag aattaaactc acttgccagc ttataatgca	27000
ggaaaatata gcaaagagat gtggatccaa tagtttctag atagtgttac aggatggcta	27060
agatgaattt atatatctga aatgttcaca aattccctac tcatatagca tgttttcata	27120
atgttttagc aactctaact ctctgactg gattgccacg tctggtattt ccacaacatt	27180
tcctaaacta agaatgagag taagaaatat tttaattcat aacaattata aatctgcaac	27240
tcatgaaaat gacattgcac ttgtgagact tgagaacagt gtcaccttta ccaaagatat	27300
ccatagtgtg tgtctcccag ctgctaccca gaatattcca cctggctcta ctgcttatgt	27360
aacaggatgg ggcgctcaag aatatgctgg taagtgtctc ggaaaaaaa attacaata	27420
gaaatgtctt atatttgcta ttaggtaatt ttttaaatta ggaaacatct ggaataggtg	27480
tttctattct tctacagaca gaaccattct atattctgct cagcccaagc tctggctacc	27540
cctgagtctc cttagcaaag caaagcaatg ctccagaaac tatgggaatt ctcaaata	27600
gtaataggaa aatgtaaaag aaagttatga agacacgagt tctttaataa tccagagatt	27660
ctataagatt caaatagctt ccctataaac aataaaaaag attttgtttg tttgtttgtt	27720

tgcttggttt	ttagagacaa	agactttctc	agactggagt	gcagtgggtgc	aatcatggct	27780
tactgcagcc	tcaaactctg	gtcttaagaa	atcctcttgc	ttcagcctcc	caagtagcta	27840
gaattataaa	taagtgtgta	ccaccatacc	cagctttttt	tttttttttc	tacagacagg	27900
ttcttgctct	gttgcccagg	ctggtctgga	attcctgccc	tcaagccatc	ctcctgcctt	27960
gttggcctcc	caaagcaatg	ggaggattta	gattagacat	tgtatgaggg	cttaataatc	28020
cttaaggtat	taactgccct	ttaaagtatt	ctgggatatg	gcaaaaactc	gatgtgtata	28080
taaacattgg	tcatatttgt	ttattgaatg	aataaaatgg	aaactaaaat	gaggacaatg	28140
cacaagagct	actagaacca	gtaagagtat	cagcgaagga	gtggaagggt	agcattgaca	28200
atttccctgg	gcttttacct	atgttgtaga	ttgtctctcc	aaggaataat	acaaagcctt	28260
aatagtccta	gaacacattc	tattgtgttc	ttatggccca	aagtaaattg	gtgtagtaga	28320
taacatttgc	accagtcatg	aaaaactatt	ggtgtcattc	tgagagtaca	tcaatataaa	28380
atagactagt	tcttttagcct	tgaaactaga	ctggtttctc	ttttgctgct	aggttaaagg	28440
ttattcaata	tgtaatcttc	caatccaaaa	tctgtcagtg	gataatttaa	aagcttttag	28500
tcaattttta	gatatttggt	ttcttaaaat	tttaaggggc	actgtgtcac	aaagctaaag	28560
aaaaaaaaaga	aaaaaaaaact	gatctgtgaa	aggggttatc	ctcatctact	tggggaattt	28620
tggctgcgaa	gaaactccaa	agtaaactct	tagaagcctt	cattgttaaa	tatgaaataa	28680
tgtttggagt	acatttatct	cttctcaaat	ttattatagg	gtcaataatg	tacacatctt	28740
gaagtccatt	tttttctctg	ttttataaca	aacaggccac	acagttccag	agctaaggca	28800
aggacaggtc	agaataataa	gtaatgatgt	atgtaatgca	ccacatagtt	ataatggagc	28860
catcttgtct	ggaatgctgt	gtgctggagt	acctcaaggt	ggagtggacg	catgtcaggt	28920
aagctcaaga	caatctcatc	catgtcatca	tccaagaagt	gtataagcac	ttcctagtat	28980
gtgataatgt	gatagacata	agtgtaacag	ttacaataca	cagccctgtt	cctctaaaat	29040
ttataatcta	gatttttagaa	ataaattttt	ttatgaatga	agtttatcta	tcatgaaagc	29100
attaactctg	agaggccaaa	ttacagagta	gttaaccatc	caaagctcaa	gaatcagaaa	29160
gacctcgatt	tgaattcctt	aacctctatt	accaagtctc	taactaaaag	ctggggataa	29220
tcataatagc	acctaaactt	ttgggtacta	agaaaagtta	aatgaagact	aaatatatca	29280
ggcacatggg	aaacaacaaa	gaaatctcat	ctatttcact	attattaatg	tagaccatgg	29340
tcactcgtgt	taataacttt	aacctcaacc	ttttaactgc	tgtgaaggat	taaataaaaa	29400
attaatcact	atattataaa	aattaattga	tatataataa	atgaatttta	agagatacgt	29460
aataattcat	ggactccttg	aagatagaaa	atttatacaa	aatcctagta	atttgagtca	29520
caaaagctcc	tacaataatg	aaacagtatg	aatgaaaaag	aaaagaaata	actattatat	29580
ttggatctag	cccataatct	ttaaccaaat	gcacaaaaac	aaacaacaaa	tatgaaattc	29640

tcaactgtaaa	gtgattaaaa	tcaaatttga	attctaaaaat	tttaaattaa	attatctaaa	29700
cataattgat	gcagttatat	gttttaatat	gttttggtca	catatctgaa	atccaactcc	29760
acacagtagc	aggaacagct	ggtgtcagaa	attaaatatt	cttttagtct	ggagttttaa	29820
aaaatcaatc	tgtttacttg	agtaatttgt	tgctgttttc	atgggtgaat	tgtatacaga	29880
aggataagaa	ttattcttcg	catcaaaagg	tcaactgactt	tcatatttag	tgctcatggt	29940
ctttaaaaag	tggataaaaa	gtagttctca	catttcatgg	aaagcccca	atccatgagc	30000
acatttccca	aatgaaaca	tttttatcaa	ctgcaagttg	tgtgtaggtg	gagatttggt	30060
tttcaattgt	caagatactg	ttaattaccc	agtcccttat	ctccttttgg	tggagatgtc	30120
tctgtgctag	gaaacccttc	ttgctctcct	tccgtgttct	cttttactac	tggccctgaa	30180
acaacaaatt	ctcaagtttc	atgacagctt	tccaaagaat	ccatcaatca	aataagcaac	30240
acaactcgac	actgacaatt	ccagacctac	taagagcatt	aattaagact	taaaaataaa	30300
catgagtttt	aaaaggggtg	tattcattat	tttcccat	ataacgtccc	ttaccttctg	30360
tccttcagt	catacaaatt	attatcttcc	ttgaagccca	gttcaagccg	tacctcacca	30420
tgataccttc	catgtatatt	ccactctagg	cctcactgat	ttttaactga	aatactataa	30480
tgcatagttc	acacttaaaa	aaaaaaaaaa	aacacagcac	tttacataag	agcttacagg	30540
atcctatttg	ttttatccat	tcttttggtc	atttttacaa	tcattaattc	aaaggaatta	30600
tattaattac	tttctatgca	cccgacgttg	tggttaacaca	acaatactat	ccctgcattc	30660
agcaagtcta	tggtctacaa	gagaggacac	aaattcaaat	gtctgtagtc	aagcagtga	30720
gctggctaga	tatggaaaaa	ttacaagtcc	ctcttgcttt	aacatttgct	tgcccacatt	30780
tggtcagaca	tcatgcaaaa	taatttctca	ctatagaaaa	aaaaaacta	caaaaacaat	30840
aatataaaga	actgagaact	ggttaactga	agcatgcata	tgtcatctaa	aagaagcagg	30900
tgacgaccag	cttcatgaag	tacttgccat	gcatattggc	acttcacaca	ctgacccttc	30960
tccccaccta	gaccagtaat	taaacaggta	tggtatgagct	agctactaag	agcagccaac	31020
tgaatagctg	actaacttag	aagcacactt	ggtaataata	gctgactttt	attagtactg	31080
actatactat	atgctaagct	gtactcaaag	tgctttgagt	tttaaactga	tacaaacatt	31140
atatgaggaa	acagaggtag	agagagctat	tcaccagctt	accaaaggct	acatagctgg	31200
taagtggagg	acttaaacc	agactatcta	gtttcagaac	ccacagactt	aatccatcgt	31260
gcagaacata	agacatactc	catctgtctc	cccaactagg	ttattatgtg	cacaaatatt	31320
tattggttgg	ttggttcatt	attatgactg	ggtggtaagt	atgtcattag	gagtgttttg	31380
cttatgacta	tataaatttc	ttcaccaaaa	gaagactttc	tgatgatata	ctatgcatca	31440
gacaccacgc	agggtgctaa	ggttaggaag	ataagtgaga	cttctagaaa	ctcattcatt	31500

caacaaatat	ctcctaaggg	ctagaagctt	aggtttcagc	agtgaacaga	ataggtatgt	31560
tctctttcgt	gttggacctt	atagtatatc	tgggaaaaca	gacattgaat	aaatatcaca	31620
aatgcaagtg	agtgtttcag	agacatgcag	ctgctacatc	aaaacaaaac	agaacaaaac	31680
aaacaaacaa	aaactgacca	gtgggattaa	gtgtaaatag	gcacacaaat	gcacaaatat	31740
gcttttataa	aatagtgaag	cagtgcacaga	gacacacaca	agatataaag	acacaatgaa	31800
gaacaattga	gccccaaagct	ggaaaggggtg	agagtgtgaa	ggaaaaaggt	tgatcagaga	31860
agttttcccg	aaggagagaa	agcctggatg	attaggaggc	aaccactcgg	tgactgaggg	31920
aaatctgaaa	aatgtatttg	tcatcttctc	agacttgctg	aaggaatgac	ttgggtactt	31980
tgaggatttc	agtaattttt	ccatgacttg	gtataatatt	tcaaaaggaa	ataggctgac	32040
tttatttgta	taatgaatgt	gactccttcc	tcgactgcca	tagaaataaa	ctccttaata	32100
ttttgggttt	gtctttgcac	ttaagtaatc	agtcattctg	tttttttaca	gggtgactct	32160
gggtggccac	tagtacaaga	agactcacgg	cggctttggg	ttattgtggg	gatagtaagc	32220
tggggagatc	agtgtggcct	gccggataag	ccaggagtgt	atactcgagt	gacagcctac	32280
cttgactgga	ttaggcaaca	aactgggatc	tagtgcaaca	agtgcacccc	tgttgcaaag	32340
tctgtatgca	gggtgtgcctg	tcttaaattc	caaagcttta	catttcaact	gaaaaagaaa	32400
ctagaaatgt	cctaatttta	catcttgta	cataaatatg	gtttaacaaa	cactgtttta	32460
cctttcttta	ttattaaagg	ttttctatct	tctccagaga	actatatgaa	tgttgcatag	32520
tactgtggct	gtgtaacaga	agaaacacac	taaactaatt	acaaagttaa	caatttcatt	32580
acagttgtgc	taaatgcccg	tagtgagaag	aacaggaacc	ttgagcatgt	atagtagagg	32640
aacctgcaca	ggctctgatg	gtcagagggg	tcttctctgg	gtttcactga	ggatgagaag	32700
taagcaaact	gtggaaacat	gcaaaggaaa	aagtgataga	ataatattca	agacaaaaag	32760
aacagtatga	ggcaagagaa	ataatatgta	tttaaaattt	ttggttactc	aatatcttat	32820
acttagtatg	agtcctaaaa	ttaaaaatgt	gaaactgttg	tactatacgt	ataacctaac	32880
cttaattatt	ctgtaagaac	atgcttccat	aggaaatagt	ggataatttt	cagctattta	32940
aggcaaaagc	taaaatagtt	cactcctcaa	ctgagaccca	aagaattata	gatatttttc	33000
atgatgacc	atgaaaaata	tcactcatct	acataaagga	gagactatat	ctattttata	33060
gagaagctaa	gaaatatacc	tacacaaact	tgctcaggtgc	tttacaacta	catagtactt	33120
tttaacaaca	aaataataat	tttaagaatg	aaaaatttaa	tcatcgggaa	gaacgtccca	33180
ctacagactt	cctatcactg	gcagttatat	ttttgagcgt	aaaagggctg	tcaaacgcta	33240
aatctaagta	acgaattgaa	agtttaaaga	gggggaagag	ttggtttgca	aaggaaaagt	33300
ttaaatagct	taatatcaat	agaatgatcc	tgaagacaga	aaaaactttg	tcactcttcc	33360
tctctcatct	tctttctctc	tctctcccct	tctcatacac	atgcctcccc	caccaaagaa	33420

tataatgtaa attaaatcca ctaaaatgta atggcatgaa aatctctgta gtctgaatca	33480
ctaataattcc tgagtttttta tgagctccta gtacagctaa agtttgccta tgcgatgatca	33540
tctatgcgtc agagcttcct ccttctacaa gctaactccc tgcattctggg catcaggact	33600
gctccataca tttgctgaaa acttcttgta tttcctgatg taaaattgtg caaacaccta	33660
caataaagcc atctactttt agggaaaggg agttgaaaat gcaaccaact cttggcgaac	33720
tgtacaaaca aatctttgct atactttatt tcaaataaat tctttttaaa ataatttccc	33780
tgctaatta tttatggaag ttatgacttt tgaaggacaa ttcaaaacca tttatttaat	33840
tggttctgca atgaaagaac tgcccatat actctactaa aggcttggca ctttctgctg	33900
ccttttaatc cagcgctata attgaggcaa gcgtccagct tgacacctcg agataacttc	33960
gtataatgta tgctatacga agttatatgc atggcctccg cgccggggtt tggcgcctcc	34020
cgccgggcgc cccctcctca cggcgagcgc tgccacgtca gacgaagggc gcagcgagcg	34080
tcctgatcct tccgcccgga cgctcaggac agcgggccgc tgctcataag actcggcctt	34140
agaaccccag tatcagcaga aggacatttt aggacgggac ttgggtgact ctagggcact	34200
ggttttcttt ccagagagcg gaacaggcga ggaaaagtag tcccttctcg gcgattctgc	34260
ggagggatct ccgtggggcg gtgaacgccg atgattatat aaggacgcgc cgggtgtggc	34320
acagctagtt ccgtcgcagc cgggatttgg gtgcggttc ttgtttgtgg atcgtgtga	34380
tcgtcacttg gtgagtagcg ggctgctggg ctggccgggg ctttcgtggc cgccgggccc	34440
ctcggtggga cggaagcgtg tggagagacc gccaaaggct gtagtctggg tccgcgagca	34500
aggttgccct gaactggggg ttggggggag cgcagcaaaa tggcggctgt tcccaggtct	34560
tgaatggaag acgcttgtga ggcgggctgt gaggtcgttg aaacaagggt gggggcatgg	34620
tgggcggcaa gaaccaagg tcttgaggcc ttcgctaatt cgggaaagct cttattcggg	34680
tgagatgggc tggggcacca tctggggacc ctgacgtgaa gtttgtcact gactggagaa	34740
ctcggtttgt cgtctgttgc gggggcggca gttatggcgg tgccgttggg cagtgcaccc	34800
gtacctttgg gagcgcgcgc cctcgtcgtg tcgtgacgtc acccgttctg ttggcttata	34860
atgcagggtg gggccacctg ccggtaggtg tgccgtaggc ttttctccgt cgcaggacgc	34920
agggttcggg cctagggtag gctctcctga atcgacaggc gccggacctc tggtagggg	34980
agggataagt gaggcgtcag tttctttggg cggttttatg tacctatctt cttaagtagc	35040
tgaagctccg gttttgaact atgcgctcgg ggttggcgag tgtgttttgt gaagtttttt	35100
aggcaccttt tgaaatgtaa tcatttgggt caatatgtaa ttttcagtgt tagactagta	35160
aattgtccgc taaattctgg ccgtttttgg cttttttgtt agacgtgttg acaattaatc	35220
atcggcatag tatatcgga tagtataata cgacaagggt aggaactaaa ccatgggatc	35280

ggccattgaa	caagatggat	tgacgcagc	ttctccggcc	gcttgggtgg	agaggctatt	35340
cggctatgac	tgggcacaac	agacaatcgg	ctgctctgat	gccgccgtgt	tccggctgtc	35400
agcgcagggg	cgcccggttc	tttttgtcaa	gaccgacctg	tccggtgccc	tgaatgaact	35460
gcaggacgag	gcagcgcggc	tatcgtggct	ggccacgacg	ggcgttcctt	gcgcagctgt	35520
gctcgacgtt	gtcactgaag	cgggaaggga	ctggctgcta	ttgggcgaag	tgccggggca	35580
ggatctcctg	tcatctcacc	ttgctcctgc	cgagaaagta	tccatcatgg	ctgatgcaat	35640
gcggcggctg	catacgcttg	atccggctac	ctgcccattc	gaccaccaag	cgaaacatcg	35700
catcgagcga	gcacgtactc	ggatggaagc	cggctctgtc	gatcaggatg	atctggacga	35760
agagcatcag	gggctcgcgc	cagccgaact	gttcgccagg	ctcaaggcgc	gcatgcccga	35820
cggcgatgat	ctcgtcgtga	cccattggcg	tgcttgcctt	ccgaatatca	tggtggaaaa	35880
tggccgcttt	tctggattca	tcgactgtgg	ccggctgggt	gtggcggacc	gctatcagga	35940
catagcgttg	gctaccctg	atattgctga	agagcttggc	ggcgaatggg	ctgaccgctt	36000
cctcgtgctt	tacgggtatc	ccgctcccg	ttcgcagcgc	atcgcttctt	atcgcttctt	36060
tgacgagttc	ttctgagggg	atccgctgta	agtctgcaga	aattgatgat	ctattaaaca	36120
ataaagatgt	ccactaaaat	ggaagttttt	cctgtcatac	tttgtaaga	aggtgagaa	36180
cagagtacct	acattttgaa	tggaaggatt	ggagctacgg	gggtgggggt	gggtgggat	36240
tagataaatg	cctgctcttt	actgaaggct	ctttactatt	gctttatgat	aatgtttcat	36300
agttggatat	cataatttaa	acaagcaaaa	caaattaag	ggccagctca	ttcctccac	36360
tcatgatcta	tagatctata	gatctctcgt	gggatcattg	tttttctctt	gattcccact	36420
ttgtggttct	aagtactgtg	gtttccaaat	gtgtcagttt	catagcctga	agaacgagat	36480
cagcagcctc	tgttccacat	acacttcatt	ctcagtattg	ttttgccaag	ttctaattcc	36540
atcagacctc	gacctgcagc	ccctagccc	ggcgccagta	gcagcaccca	cgtccacctt	36600
ctgtctagta	atgtccaaca	cctccctcag	tccaaacact	gctctgcac	catgtggctc	36660
ccatttatac	ctgaagcact	tgatggggcc	tcaatgtttt	actagagccc	accccctgc	36720
aactctgaga	ccctctggat	ttgtctgtca	gtgcctcact	ggggcgttgg	ataatttctt	36780
aaaagggtcaa	gttccctcag	cagcattctc	tgagcagttc	gaagatgtgt	gcttttcaca	36840
gttcaaatcc	atgtggctgt	ttcaccacc	tgcttggcct	tgggttatct	atcaggacct	36900
agcctagaag	caggtgtgtg	gcacttaaca	cctaagctga	gtgactaact	gaacactcaa	36960
gtggatgcca	tctttgtcac	ttcttgactg	tgacacaagc	aactcctgat	gccaaagccc	37020
tgcccacccc	tctcatgccc	atatttgac	atggtacagg	tcctcactgg	ccatggtctg	37080
tgaggtcctg	gtcctctttg	acttcataat	tcctaggggc	cactagtatc	tataagagga	37140
agagggtgct	ggctcccagg	ccacagccca	caaaattcca	cctgctcaca	ggttggtggtg	37200

ctcgaccag	gtggtgtccc	ctgctctgag	ccagctcccg	gccagccag	caccatgggt	37260
acccccaa	agaagaggaa	ggtgcgtacc	gatttaaatt	ccaatttact	gaccgtacac	37320
caaaatttgc	ctgcattacc	ggtcgatgca	acgagtgatg	aggttcgcaa	gaacctgatg	37380
gacatgttca	gggatcgcca	ggcgttttct	gagcatacct	ggaaaatgct	tctgtccgtt	37440
tgccggtcgt	gggcggtcatg	gtgcaagttg	aataaccgga	aatggtttcc	cgcagaacct	37500
gaagatgttc	gcgattatct	tctatatctt	caggcgcgcg	gtctggcagt	aaaaactatc	37560
cagcaacatt	tgggccagct	aaacatgctt	catcgctcgt	ccgggctgcc	acgaccaagt	37620
gacagcaatg	ctgttttact	ggttatgctg	cggatccgaa	aagaaaacgt	tgatgccggt	37680
gaacgtgcaa	aacaggctct	agcgttogaa	cgcactgatt	tcgaccaggt	tcgttcactc	37740
atggaaaata	gtgatcgctg	ccaggatata	cgtaatctgg	catttctggg	gattgcttat	37800
aacaccctgt	tacgtatagc	cgaaattgcc	aggatcaggg	ttaaagatat	ctcacgtact	37860
gacggtggga	gaatgttaat	ccatattggc	agaacgaaaa	cgctggttag	caccgcaggt	37920
gtagagaagg	cacttagcct	gggggtaact	aaactggctg	agcgatggat	ttccgtctct	37980
ggtgtagctg	atgatccgaa	taactacctg	ttttgccggg	tcagaaaaaa	tggtgttgcc	38040
gcgccatctg	ccaccagcca	gctatcaact	cgcgccctgg	aagggtttt	tgaagcaact	38100
catcgattga	tttacggcgc	taaggtaa	ataaaattt	taagtgtata	atgtgttaaa	38160
ctactgattc	taattgtttg	tgtatttttag	gatgactctg	gtcagagata	cctggcctgg	38220
tctggacaca	gtgcccggtg	cggagccgcg	cgagatatgg	cccgcgctgg	agtttcaata	38280
ccggagatca	tgcaagctgg	tggctggacc	aatgtaaata	ttgtcatgaa	ctatatccgt	38340
aacctggata	gtgaaacagg	ggcaatggtg	cgcctgctgg	aagatggcga	ttgatctaga	38400
taagtaatga	tcataatcag	ccatatcaca	tctgtagagg	ttttacttgc	tttaaaaaac	38460
ctcccacacc	tccccctgaa	cctgaaacat	aaaatgaatg	caattgttgt	tggttaaacct	38520
gccctagttg	cggccaattc	cagctgagcg	tgctccgca	ccattaccag	ttggtctggt	38580
gtcaaaaata	ataataaccg	ggcagggggg	atctaagctc	tagataagta	atgatcataa	38640
tcagccatat	cacatctgta	gaggttttac	ttgctttaaa	aaacctccca	cacctcccc	38700
tgaacctgaa	acataaaatg	aatgcaattg	ttgttggtta	cttgtttatt	gcagcttata	38760
atggttacaa	ataaagcaat	agcatcacia	atttcacaaa	taaagcattt	ttttcactgc	38820
attctagttg	tggtttgtcc	aaactcatca	atgtatctta	tcatgtctgg	aataacttcg	38880
tataatgtat	gctatacgaa	gttatgctag	taactataac	ggctctaagg	tagcgagcta	38940
gctgcaaccg	aggaaaaaac	gtgccatgag	gtctctgtat	ccaagtgtga	ct	38992

<210> 20
 <211> 34073
 <212> ADN
 <213> Secuencia artificial
 <220>

<223> Polinucleótido recombinante

<400> 20

gagggagggt ggtgctttgc taatggtgaa ttactaactc ctcaataaag aatattat	60
gaaataat	120
agatatattac aaacatacat taatatttcc tgaatgaata caccacaaat ctcccttaaa	180
atatagcaag aataaaaatt atactatttc tgacaatttt taatttctca aataataata	240
ccactctgat ttttaaacat ctacaccact ctggctttgc caatcttttt aaaaattgaa	300
aagataataa ttttatcata attacactga agcatagaac tttttctttc aaggaaagca	360
aatttttgaa attctataat ataacctccc ataactctga ataaattaaa gggtcaacaa	420
cttagtaaag taagactgac cttccctttt atttcttttt cagatcaaaa atcttacttt	480
tataggagca gttttcaact cctaaatggt gaatataata gtcagttaaa ttcaccagct	540
acacaggaat acaggacttt gagtggaaga attgaatctc tggtaagtta atatttgtct	600
ttgctcttta ttccattata aaatgaatat gataataaac ctaatgtttt gtaatatatt	660
ttcagttgct aagtgtctca catattttcc ttccttgaat ggtgaaacat gtgtttctct	720
ctgcttttat ccagttagtt tactcatata ctggttctta ttcacatctt tgtcatgagt	780
aaaaagtgtt agaaaggcca cgagtaaata tgcattttat ttgtttatga attcaaatac	840
taaaagtgtt ttatttggtt aattaagcat tgacattgtc tttttaaatt cttttcattt	900
taccttcttc cctcttcctt atccaactaa agacgcaaag caggagggtgt taaaaaacag	960
gtttaccata tcagcagtaa catagtttgg acaacattac actttgggtc aatgatagac	1020
atagaagttt gaacagaaat atgcaaagca agtttgagct ctaacttgaa gagagcctct	1080
gggtgcctgc caggaaacct cacgagtgga cccttaacat tcatgtgtca ccacaaacta	1140
ggggctgccc tttagttttg accagtctca gtgtcactca cttaccctta ctttttcaaa	1200
aaaaagtcct aagaatataa agtaattcaa tggttctaca attttagcat gtaactgagt	1260
cacctggcag ggttgctttg gtgagctcaa gataaaattt tatcagcatt tctacatttt	1320
ctggaatatt ccttaatcca ggcttttaat cccttgggtgc ttttctgaac cactgcaatg	1380
agcttctaac tgttctcact gtgtgcaggc tcttttcctt ctaactaat ttacacactt	1440
ctgaacacaa atctctcaca gcctgtttcc ttcattgttac ctccagctca agactttttg	1500
cctacaaaat aaaattcaaa cttgttagct aagcaccttc tcatgtctat gctttggctc	1560
atatttcagc catcgtgtgc ccacttatt cttatagcca acctgaaaag ccatctttta	1620
taagaaacta cctctgctct ccatgattgg atataattaa tcctccttcc acatcacctc	1680

gccacaaaat	tgtatctgtg	ttgatctcat	gccacatacc	tgtatgtatt	ttatatattata	1740
aatatttgca	gacttggtta	atttgccatg	ttagactaag	ttccatgaag	acagctccat	1800
atccattcca	tttttatata	tccacaacat	ttggtcgggt	tgatgcttaa	taaatgttta	1860
ttgaaggaac	aggagtctcc	cacttctgac	ataatgaact	tatttcccc	agtgttaacc	1920
ctacatctgg	ttcctgtcca	agagtctctt	cccaaactcat	tctgattcaa	ctgttcattc	1980
tgatctcatt	aaacatttaa	atgatataat	taacttcgct	tgctttattc	tatgctcatc	2040
ctgcagtctc	ctcataactt	ggtttcaatg	atgcttgctt	ctagagaaaa	aaatgtatta	2100
aataagctta	tgattcagtc	ctccagctgt	gatggttctc	actgaacatt	agctcagtgg	2160
ttttcgaagt	atggtctcta	gcataaccta	gaaacttggt	agaaatgcaa	attcttgggc	2220
tcaccaagac	atactaaatc	aaaaattctg	acattggggc	ctagaaatct	gtgttttaac	2280
aagcctgcca	gtgcagcctg	gtcccttttc	ttctcggagc	cccactcaaa	gctttcagtg	2340
ctcatctccc	accaatgaca	gggtcctcta	tggaaaccgg	caggacggtt	tccaactcta	2400
actacgtttt	agagtttgct	tcctagggct	atccaggcac	caagtatcac	aggtagttt	2460
cccaggaag	cagactctga	gacttgcatg	cagggagtgt	ctctgggggtg	ctctcaacca	2520
acaccttcag	gaagagaagg	aagcagcatt	gggcagaggc	atagtcaaac	tacagtgctg	2580
ttggcacaga	agactgaagg	gagtcagagc	cagggggtag	aggtgggccc	ttagcatcca	2640
tccttcacca	ttaggtgtga	gttgccccac	ctccttgatg	gtgtaacctc	agtcccaagg	2700
tgggtgggag	tgagcagag	cagcccctac	aaggggcaaa	ccagagatac	accaggcgcc	2760
agaagtgctg	ccaggaata	gagaggaaag	gatgggctta	aggtaggatc	cacagaactt	2820
ggcaatggat	tagaagacag	gatgagaagt	gacagggtta	cactaacaca	gaaatgtcta	2880
acttcggtag	ataatgggtc	cattggctag	aagaggaaac	cgaaatgaaa	gcaggttggt	2940
cagggagaca	aaagttcact	gtggacatct	cagcagagtg	attcagtggg	gaaaggaatg	3000
gatgcccaga	ccacctcaga	ggaagatcta	agctggagcc	agcaataaag	atacaagatg	3060
aacaatccct	aacgaactgc	tcctcagcca	tgctccccag	acacgctgct	tcagatttat	3120
agtccgggtg	aggctaggag	gtgcgcctcc	ctcagtggag	gacagcaaag	caccagtggc	3180
tccagggagt	taaaatcttt	tgataatttt	tgttctagca	tctgtctgca	gagctgtctc	3240
tcagccattg	cctgccttta	cacaggagtg	cagtccgaaa	ttgggagatg	agtgaaattt	3300
attatgccta	gagatctgga	tccccagttg	tttgggagta	tattttctga	accacttggt	3360
ggtttaagta	atgcagattt	attgatgcc	cttctcttga	atctgtgact	ctggaccac	3420
catctaagt	aatgtgcaga	gggaacggaa	tggctgcaat	agatctccat	taaaaccagt	3480
gcacccctcc	agacacatac	agtagtaggg	aggtgagtca	atgtcaggac	agcaccagct	3540


```

cccgttctcg tacatttcca aagtctctcag tctgtgtaca aaggtttgct ctggggcagc 3600
agaaatagcc ctgggcaggt agtcaaaggc ctggtttgat ttcctccact tccaggcaag 3660
tactcogaag gctcacaggc tttttcctca cctgccacat ggggccagtg agatctactg 3720
agctgtaaat aatgaaatga gtgtgtgtgc agtcatctat aagttgtaaa gtactagaaa 3780
atgggtgaaac tttgggattt gggctattta aggctgaatg ctaaaaatgt caggcattgt 3840
ggagaaagga atttaaatat aagattgatt gactgggatt taaagacaaa tgaaggcaca 3900
cacgcaagtg cacaccaca ctgacactgc acagctcccg ttggaggcat atcctgacca 3960
tgcagacctg gggctctgcc tgtccaagtg cactccttta ctacataaac cctccttctc 4020
ttttggggct gtcacccccc cagagctggc accgagccct tgctgctgcg cttccctggg 4080
gtgtcagctt ttgacagggt gtttcctccc tctgcaggag ccttaacatc ccttggactt 4140
ccttcccccc acccaccccc agcagtttta tctcttcta actcgggacc ctttttttcc 4200
cacacaaagt ttattgtcag ttgctggtt catctgtttg agcggctgca acaaaatacc 4260
atagactggg tggcatatgc acgacaaaaa tttatttctc acaggagaag tcaaagatta 4320
atgcaccagc agatctggtg tctgaggggc caccttctgg tttgtagatg atgctttcta 4380
gttaaaacac ctatttaaca cactattaaa cactaagtgt gttaaatagt gcagttgatg 4440
tatttgtcat gtcaccttta tcatacacta aatccttctt tgtctttttt tctgtactct 4500
aatctctttc tgtaagtaat ctttgcttgc agcagtagga tatttagagt actgtggctt 4560
gacaatatat ttagtatttc aagatttcca tgaaattctt ctgatgtatg agttccctag 4620
ttaatcttac atatgtatcc ctttgtaaaa acactttgaa catttaaaat gatacatgaa 4680
tagtactcta atacaatgcc ataaaaatta taaatcattt gtatagactg gtaagtaaag 4740
attgtgagat taagaaacgc atcaaaggcc attgagctgg aaagtggat aatgagaatt 4800
caaaccaggg tctcttgact caaaatctaa ggatcatacc atttctcatg ataatatgag 4860
tattattggt atctctatcc catagacaaa gtgttaacac tgaatgagca gtgaaatagt 4920
ctcagaattt tttattttat ttagcaattc acttgtcatt tctggtcctc agtttattca 4980
cgagtaaaat aaaatagttg gactagataa tttctatagt acattcttac acaaaaaatc 5040
tatgattttg ttatttttaa tgtgatatac tcatggcact cattcacctc attttcccag 5100
cctgcctcac tggtcattac ttctctgtgt tctttacagg ctccccctcc tctacactgc 5160
cattaaatat tgaaacacct caaagcttta cttatgtcca cctctcctct gacactatca 5220
ttctgtctag atgatcccat acatacatgc ccattacttc aacctgtatt tatacgccaa 5280
tgattcacta tatttccagc ctagacattc ttttgtactc tagttaccag cttgatatcc 5340
ttacatggct gtttcaaaac aactcaaata tattatctct caaaatcaaa ctcatgatgt 5400
ccccacacca tcctagcttt ccaccaacaa tacctatccc tattaatagc aataccattt 5460

```

attcagttat ccaaatacaa aacctagaat tcatccttaa aattctacta tcattccaaa	5520
tatcctatcc atcagcagcc actgtattct taatcccctg tatttccttc aaatccattc	5580
acctctctcc atatccattg ctgcatgact atccaagcca tcgcctctac cctagggtac	5640
caaaatagca acaaacctaa tctgttcatt tgcattatct tttctccaaa actgattatc	5700
tatatgtagc aagacagatt gttctcaaat tgcaaatccc actatattat cctcttgctt	5760
caaacacttc catgggttcc cattgtttat gataaaacca aatgcttcaa gttcgaagac	5820
cggcatgatt gggaatttcc tgtcacccta gcctacttgc tctccatggc acagttgcac	5880
tggctttctt tcattcctta agtacaacct gtttcctccc acctcaggac tgtgcatgtg	5940
ccattcattc tgctgaggag cctttttcct tccacttcaa tcagctaagt ctgattcttc	6000
ctgacaatct cagctcaata agcatttcct ctaagaaatg tctctaatat cattaattgg	6060
ctcaggtccc tctactgtat tgctgcactt ttcacagtta taattttact taattatgaa	6120
tgattatttg attaggtcta tttccatcca ttagacataa gcttcatgat ggccagatta	6180
ctgttttcta tccatcgttg tattccaata cctgacagaa ggaggggcggg aggtgggtggc	6240
acacaagaga tgctcaaaaa caattgttga ataagtaaat gaatgaggcc atttagaaat	6300
aacgaaagta cctgtttaca aagtacatgt atcaaaaacta tgaatgcatt ctacttacat	6360
ggttttctcc aaataaaaaca aaagacttca atcaggatta atacctggga taaactgagt	6420
cattaaatct ctcttttgcc atcaggagtg acattgaaac aaatgtctgc aaacaacaaa	6480
tacttttttc caaaaatata ttgaatggca tttccataaa caaactagaa catgggagga	6540
gaaagaaagc aatattaatt taaaattaat cttatcacat aacttatacc atcagggtatt	6600
tcgggtaaaa ttcttttcag gcacatccat ttaacaagaa ttgattgtta ctgaaagcct	6660
agaagagaat ttggcacata cttgggtgttc aaatatattg tgactgagtg aataaatgat	6720
gcaagtgtct aagaaacaca aaataaggac atgattacag tcacgggtgga gttcacagtc	6780
atctccaaaa tgaggatatg catcccaggg aggaccaaca attcattgga gtgctgaaat	6840
aaaataactca aaggtcattt tacatgtatt ttttctctaa attacttttc ttaagacaca	6900
gaaaacaaaa aaagaaactt agctttgtta ctttctaaca aatagttaaa tcattaaaca	6960
ggattgacac tagcatcctt gtttggctct atgccttagg ggaacatgaa atgtgtgaag	7020
acattctgag atctgaggga agggtagaca gtaatacagt gggactgacc aggcttcagc	7080
acacctttac ctctctcag cagatttcag tgatgagcag tttacaacta gattgaaaga	7140
ttatattatc tagttctaaa agaaaactaa gcctcccaaa agcaacaagg gaactgagag	7200
gaatcctgca aaacaaaaac aaattttaaa acttgcaact tgaataaacc ctaatatgta	7260
atcacagtaa tgaacagtaa gataatgaca gaactgacat atttccttat ctattaaagc	7320

catattaaca	ggtaaagcaa	tgccagtcag	tggtacactt	cttagaagat	atttaataca	7380
tactagacac	atacacacac	acaacatttt	ccttcaaggt	gtatgtatca	gaaatcact	7440
ttttaaggcc	ggatgcagtg	gctcaggcct	gtaatcccag	cactttggga	ggccgacgtg	7500
ggcggatcat	ctgaggtcag	gagttcaaga	ccagcctgcc	caacatggcg	aaaccccatc	7560
tctacaaaaa	tacaaaaatt	agccagggat	gatggtggat	gcttgtagtc	ccagctactc	7620
aagaggcaga	ggcaggagaa	tcacttgaac	ctgggaggca	gaggttgag	tgagccaaga	7680
tcacccattg	cactccagcc	tgggcaacag	agtgagactc	tgtctcaaaa	aaaaaaaaat	7740
cactttttag	ataaaattca	tgctatagag	agaagactat	gaaaatatgt	ttagcaatgt	7800
gtccatcatt	aggtgattga	gtttcctttt	gttttgtttt	actgaaaatc	atataaagta	7860
tgttatctgt	aaaagttctc	tgacatgcac	acataaaaat	ttgggagaaa	agattaacta	7920
taatgtttaa	tagattttgt	acacatttct	ttaaaaatat	ataaaacaca	acacctttca	7980
attggtttgc	aagaataacc	aattgacatc	atggaaaatg	gaaattcact	tgctgaattt	8040
taacaaaaat	ttgcatgatg	agtgagactg	acaacttagt	gtcatgattt	aatgaattat	8100
gccaatggta	aacttcatgc	acatggggcc	aggtaattat	gtggaaactt	tttcaatgct	8160
taaagccaag	tattgaaatt	aaacttagaa	tcagaccttt	gaaccatttt	atgacaatgt	8220
tcaaaaatta	taaattctat	ccacttatat	tataatatta	aaaatatcat	tacaaaaaaa	8280
acctgtgttt	attttataac	tcagcctttt	taattttctaa	tttcataaat	atattataat	8340
ggatattggt	agtaatgtag	tattattaca	tgtatataat	ttataagtaa	atatacatgt	8400
tttggtact	catgcataaa	atgtttcacc	cataggagca	cataatcaga	aatgtctgga	8460
gaccattata	gtaatagata	gatcatattg	ccacatattt	tatctcctcc	ttgacaactg	8520
agctttccag	atcttctggt	gaaacgaaag	agaaagttgt	aacagaagag	tgattaaaat	8580
gacaaaagca	ttacttctat	tacttctatt	ctaataatat	gagcaaagct	ataactatca	8640
agtaataatg	cactaaagaa	ggtgattaat	ctgatataat	cacaggcaac	taataagacc	8700
tttctattgc	agccatgaaa	aatatgtgac	aattatagat	atcctgtgtg	cagtgtttca	8760
acctttatgt	gacctgttct	actaacagat	ttagtgatgt	tcactttggt	agaattttct	8820
tacacatgcc	ataacttgct	tcagtctttt	gattatgaat	attatggata	ttaaggattc	8880
tagactattc	tagattttaa	aaataatatt	gtcacctcaa	tcagaaggga	aatattaaat	8940
agttctcatt	ttttcaatgt	ttactcagtt	tttgtccaat	gtaatgaaag	gtgcagcagt	9000
acaggttaca	aaataaaatg	tgtattaaag	taaactcatt	tgaacagggt	aataattgta	9060
gaggaggga	aaaggctaaa	agattgaatg	taaaacttat	gaaaagtaga	tacatcgctc	9120
ctatgatttg	cagtagtcaa	ctgcatacag	atgaatcatt	ttaatacacg	ttaactactt	9180
tccttttaca	gatggagaaa	ctgagaggaa	gaaagtttat	atggttcatt	aaactttgtg	9240

atgcaagcta aactaacctg tctctgtatt ttccatctac tgcccttata actatctcat	9300
tagaatactc ttcaagcatc tccttactga ttttcttacc aagcatttgt taagttctaa	9360
tgagagttgg tagtaacatt ttcacccact ctgtgaaata tgaaatctta ttcataaggcc	9420
tcttctttta ttcttgtatt tgcatatcaa ccaattaatc aacttgcttt ctttatgttg	9480
cttattatct tagtccttac taaattgcct cttaatgttg tccacataac agaaatgtta	9540
aggtggatac ttaacatttt agtccagtct agccggtgcc agtgcaatgc caaatcatga	9600
attaaaatat aattacaaga accacttata aaattttaac aattccttca gctttgtgac	9660
agttttttct acttcgatta aagtcaagta aaattaaagt taaatatattt tattaataa	9720
tctcctttaa cattccatat taataaacat attaaagctc atgcttctaa gtagattact	9780
agaagttact ttatcgaatt acagcaatgg ttaattctag atcatagaat ttagaatgac	9840
tttttgctt cttctttttt ttcctttttt ttaaacagag tcttgctctg ttgtccaggc	9900
tgagagtgtac tggcgcgatc ttgactcact gcgacctctg ccctgcagggt tcaagtgatt	9960
ctcctgcccc agcctcttaa gtagttggga ttacagggtgc ctgccaccac acctggctaa	10020
tttttttttt gtatttttag gagagacagg gtttcaccat gttggccaga ctgggtctga	10080
actcctgacc tcaagtgatc cacttgctc agcctcccaa agtgctggga ttacagggtgt	10140
gagccactgt gcctggcctg actttttgct ttcttcttaa tacttactag tatttcttga	10200
atttttaaaa aagaaacata aagtactttg ataaaaccaa cagtctcatt gttcttaaaa	10260
ttgttcaaag gttctctgga aaaaaaaaaag aaaattatca tttgggttaag aatcatgttg	10320
gtctgacatc aatcatccta taggagtga tattgaaaaa gtaagatata ttgtggtata	10380
atcgagattg cataaatttt accatttttg agaagaatct gctccaaatc ctggcttaat	10440
gtaatatcca gcatgctact taattttctt gtcttcacct tttcatatcc acatccacct	10500
aggtgccacc tcacagtata agccagcata atccattctt ctcaatgaaa ccacaataca	10560
tctgaccctg catctcagga gaactgtatc agccacagca cttccagttg actatgaatc	10620
tgaatgttat gcctcaggag aaacatcctt gctgggactg agtagtgatt caaggagata	10680
gttatgattc agtcaagaaa ttaataatta gtgttatttt tattattgag acagagtctc	10740
gttctgtagc ccaggctgga gtacagtggc atgatctcgg ctcaactgcaa cctctacctc	10800
cccggttcaa gtgattctcc tgcctcagcc tcccaaataa ctgggacagc aggcacttgc	10860
caccacgcct agctaatttt ttgtattttt agtagagacg gagtttcacc gtgttagcca	10920
ggatggctct gatctcctga cctcaaggtc cacctgcctc agcctcccaa agtgctggga	10980
ttacaggcgt gagccactgc gcccgccat aaattattaa ctgagccagg cacagtggta	11040
cacacttata gtcccagata ctcaggagac tgagggttga gtatcctttt ttatgttatt	11100

ttatTTTTaa	ttattatggg	tacataatag	gtgtacatac	ccatggagta	caagtcattgt	11160
tctgatacag	acacataatg	tttaataatc	acatcagggg	aattgggata	tccatcacct	11220
caagcattta	tctttctttg	tgttaggaac	attccacctc	cactcttgga	ataggcacc	11280
tgttgtgcta	ttaaatacga	ggctttattc	atttcattct	actatatttt	tctaccatt	11340
aaccatcacc	tcttttcccc	tcttccccac	tacctttcct	gtgaggctgc	aggattctta	11400
agcacaacag	ttagaggcca	gcctggacaa	catagtgaga	ctcaatttct	aaaaaataaa	11460
aaagaaatta	ccaactaatg	ctaaaaaaat	agtctctgat	gcttaggtat	gaattagaaa	11520
tgaccaaaaa	aaaaaaaaaa	aaaaagactg	ccctttgctt	ccttctcccc	ttctcttcaa	11580
gttttccatt	gctactcatt	ttagtctggg	ttaatcaggt	ttcatccatt	aaaagcaatt	11640
gttgggatca	cacattttga	gttgtgtcag	tggacttccc	tcatgctggc	atgattcctg	11700
ccccaaagcc	ttagtaaaag	ccaccaagcc	atataacata	atctctcatt	gagtaaaaca	11760
tctgatgtgt	ttagaatgac	ttctagcaaa	aaaccagcct	gtccagcatc	atctctgtat	11820
aacagataaa	ggaataggta	ctgcatcaaa	aggttataga	acctgcccc	atcaatccca	11880
tgtgttttgc	aattggaatta	ggttgaacta	aagtgaaaat	tcagttttct	actcctcatt	11940
aacatgtctc	atgttgcaag	gttgagagga	aggagaagaa	gaactgtatt	tacagagaga	12000
ttccccctct	ctttctttct	acagattact	aaaacattca	aagaatcaaa	tttaagaaat	12060
cagttcatca	gagctcatgt	tgccaaaactg	aggtagtggt	aactgtagaa	aaaatattta	12120
agtatagata	caatgtggca	tacttgactt	tttgtcacag	aatgaatagt	aatgacatg	12180
ttcagataag	ttgttgtaat	attatgaaaa	tagtatttta	gtcagcttaa	aaaccaatgc	12240
caaaaaagcc	aaacatatga	tctatttagc	tactaatgta	aataaccata	ttatatctat	12300
tcttattggg	aagaggaaga	aggggtggag	agagagttgg	ggtgaaggta	cagtaacaag	12360
gccatcctat	tgtaaaactc	cagtggatat	cattcacagt	gcagcctatg	taaacagtcc	12420
ctcctggagt	tgtacaatgc	tgtggtttgg	gtgtatccat	ccaagatcaa	gacactatga	12480
ccaacatcaa	aagtggcttt	ttggttttat	ctgcctgatg	tgctataata	aaagggatt	12540
atggccaaat	ccaaggcatg	tctatcatga	attaataata	ggaggagtag	cagcatgcat	12600
gctagtatt	tgccattcct	gccttagtta	aatatgatgt	gataaaacca	gcctttccaa	12660
ctgaaatagt	cacctttact	gactctcccg	caaagtgtct	aatgaccac	attgctctag	12720
tctttaaata	atatgcaata	gttctttggg	agaagaggaa	ttataactaat	tctttctcaa	12780
atactagcat	cacaagaaaa	ttaattcttg	ttctctggag	agtcacctag	taagtatctg	12840
gagcacagat	gtctggtcag	gtaagttttg	atgaggagtt	aaagggataa	gaagagtcca	12900
tgagaagggg	attttccaaa	acacctttcg	gtcaattcag	tgacatttca	cttagtactt	12960
tcttgtcagt	atctgtatca	gccactaatg	ttcaaaagtg	agtaagccct	gaaaacctgt	13020

aggactacat	gagccttctg	ccttttctct	ccttttgttc	acttcccact	tatcactcaa	13080
tcctctgcaa	cctggcttca	ataccaccat	aaaatatcaa	ctgctcttgc	cgattcaaca	13140
atgacatcca	gataacaaaa	tccaaagaaa	ccacatcagt	cctattcttg	gacctttcaa	13200
cagtatttgg	tcctgttggc	ctgtcactcc	ttgaaatagg	actatccctt	ggtttgcatg	13260
gccttgata	ccctgatttt	ccccttacct	ccctagctat	tccttcttag	tttcctttac	13320
taggtcttac	ttctttgtat	attccttaaa	tggtgctgaa	catcaggctg	tgctctaggc	13380
ctctcatctt	ctcaggtcac	actctctcct	ttccttggcc	ttcactgcca	cccataatgct	13440
gagtgtcttc	aaagttgtat	ctctaggcca	gtcctctttt	gcctccaaac	atgaatatat	13500
gcagccatct	acttgggtacc	atcacatgga	taattctcat	gatctcttcc	agtatgactg	13560
cttctttatt	tttttctggg	ctctttttta	gcattgcttt	acatggaact	ttatcatgtc	13620
tctcaacctc	tattttatct	tttatctatg	tatgtagagt	ctgtgtaatt	tcttcatctc	13680
ttttagataa	ctaatatctc	ttcagctttg	acttgtattc	tgtgtaaccc	atttattgcg	13740
ttttcaattt	caatgagtat	gttttcctat	ctgcaagttc	tatttgtttc	ttttgagaat	13800
cttctgtgtc	ttttaaacac	atttcttatt	ttaatttttg	ggggtacct	gtagttgtat	13860
gtatttttgg	agtacatgag	atgttttgat	acaagcaaac	aatgcataat	aatcacattg	13920
tgtaaaatgg	ggtatccatc	ccctcaagca	tttatccttt	gtgttacaaa	caatccaatt	13980
atattctttt	agttattttt	aaatgtacaa	ttaaattatt	attgaccata	gtgactctgt	14040
tgtgctatca	gatactaggt	gatcttttaa	aaataatgtt	ttctacttaa	tctcattttt	14100
atgattccct	cttttacgtc	atttgtcatt	tcaaatacag	tcacttgtct	gttgattcta	14160
ttatgtgaag	tttttgagga	taatcttttt	gttactttga	ttccaccttg	gtatggtttg	14220
gctgtgcccc	cactaaaatc	tcactctgaa	ctctggttcc	cataataccc	acatgttgtg	14280
ggagggacct	tgtgggaggt	gattagatta	tagggacgtt	tcccccttt	gctctgttct	14340
ttttcctgcc	accatgtaag	aaagatgtgt	ttgcttcccc	ttctgccatg	attgtaaatt	14400
tcctgaggcc	tccgcagcca	tgcaggacct	cttttctttg	taaattaccc	agtctccggc	14460
ggttctttat	agctccgtga	gaaaaaacta	atacacacct	catgatgtat	tgtttaccac	14520
tgaaattgta	tgcttaaatt	taatctcact	tgggaccctg	tacaacctag	acttaacata	14580
tctacctcca	gagcagttac	atctgtcaga	cattctagag	gaatcagcag	cacatggact	14640
ttgttgttgt	taatttgttg	tcgggggagg	ggggagggat	agcattagga	gatacaccta	14700
atgctaaatg	acgagttaat	gggtgcagca	caccaacatg	gcacatgtat	acatatgtaa	14760
caaacctgca	cgttgtgcac	atatacccta	aaacttaaag	tataataata	ataaaaattaa	14820
aaaaaaaaag	gttctgggag	tattcaggta	gtattaatga	agattcagac	atcgtgcagc	14880

caggcccatg	cttatgaatt	ttcaggtgat	acttcttttt	cttttttctt	aatttaaagc	14940
tgatctcgg	aaacagataa	atttatTTTT	ttatgacatg	acgagcattt	ttttcattct	15000
agttcatgct	gttattgggt	gtttagttct	ttgagactcc	tggccttttt	ctaaaacctc	15060
aagttcaact	tcctattttg	cactggccca	aggtcccatc	tccagtctct	atgtaaagtc	15120
taaacataag	cctgtggaat	attctagtct	caccacatac	tattcacatt	cttctttggt	15180
tttggctctc	caggattttc	cttacttttc	tatgaaccca	gtcttgcatc	tgaaatggaa	15240
tttattatat	attatctatc	ctttctatct	gttttatgca	gaaagtgttt	tctaaaatta	15300
tttaggcttc	catattgcta	gacatggaag	ttgtaattat	ttgttcagtg	cctgtttcta	15360
catctaaact	gcaagaccca	tatggcaact	gtgaatctta	gtcccagcta	atttctgaag	15420
cttagaatag	tgcttagcac	aagaagttgt	ttatctaaca	tttttaaaaa	taaatattaa	15480
attcatatct	ggaatgaata	ttaagttaga	gctgggtcatt	gaggtgagag	gaggaagcca	15540
agagagaata	tgagagcctc	aaagccaaat	atctttaatg	tactttttca	gaaaagaaga	15600
cagccaatgt	caggtggagg	aactggttta	tgaggtaact	ttcctggaag	aaaatagaaa	15660
ttactgaggt	tttagataat	ccaaatattt	aatcaagtca	ccaaggttta	ttgtggggaa	15720
tctttattat	taattaaat	gagtgatgaa	atcttaatat	acgacaaaag	ttaaaatttg	15780
cttttgcagg	cagatgaatg	gtctaggtat	caaaaaatta	agttgagtct	ctaactcaca	15840
caaatttaca	accctatcac	tttatgaatt	tgtttaggag	attattttta	ataaactggg	15900
tgaagtctaa	gaatagctaa	aatttatagt	acacttattg	tgtgctattg	actcttcttt	15960
gaagttttgc	atatagtgat	tcatactaac	ttcataaccc	attttacatg	tgaagaaact	16020
tagatataga	aagattaaga	aacttacata	acttatccaa	agttacacag	taaaactctg	16080
gcattataac	ttcaaaatca	gctatcctac	agtgagtaca	gtgttctgtg	cattgaaatc	16140
aaataagtga	gatagcatcg	tgatatagta	ttacgtatgc	aaacactggt	acagagatct	16200
gtctaaagtt	aaattccaca	aatgaattct	ttaaaagggg	ttaatcaaga	agaatatata	16260
aacaggatgg	tgaaaaattg	tcatattatt	tgttttttta	aatatcttta	tgattttacag	16320
gcaagatggg	agtgggtgtg	gagcggatgt	tgtcatgaaa	tttcaattca	ctagaaataa	16380
caatggagca	tcaatgaaaa	gcagaattga	gtctgtttta	cgacaaatgc	tgaataactc	16440
tggaacacctg	gaaataaacc	cttcaactga	gataacatgt	aagtataatt	tttcataaac	16500
aattttatatt	caatatatcc	ctcaagttta	ccaattcaaa	ttcatatttt	aattgagagg	16560
ctgacttttc	tttctttgaa	actaaactgt	gaaaacaatc	cattaaaaag	ctaaatatac	16620
catatagctc	cctaacgtaa	atcattctaa	gacttaaaga	atcatttggc	atttatatag	16680
taaattttat	ttgctaaaaa	ttctcattaa	ttatccctgc	aacattcctt	atgagtgatg	16740
ttactgtcag	atgtcattag	tggaataggcc	ataggagggg	tacatagatg	ctcaaggtca	16800

gagaactatt	taattaatga	tccacctcag	aggcttcttc	atTTTTcttt	gtaacattta	16860
tcacaattga	aattacaaag	ttatctgtgt	aaattttgta	ttgtttggct	tcacccctaca	16920
ctgtaatcat	cctaaaagaa	agaaccagtc	aaccttcttc	atcctactac	cctcctacca	16980
cccagtctcc	atcatataac	acataattcaa	taaataattc	ttgcatgact	gaaagaaaaag	17040
aaataatata	tgcatagaat	ttaaggacat	tcctccaagt	tggttacatt	ctgctagttt	17100
aataagccat	tatttcttct	cgatgagctc	aagattaaaa	ggattttgat	gattcccata	17160
ctagactggg	aggtaccagt	tacagatgta	ctaactgtta	aatattgaaa	tgctttccta	17220
tttgttggtg	aacaattact	gcatcaggcc	cacaaagttg	tcttccgaga	tgtttcaa	17280
ccactgcccc	tgctgctaaa	gagttatgct	tagcaaagca	aagcactcta	agacactgct	17340
ccaactccat	ggcctgattg	catcttttat	gactggccaa	tgctcacgca	ctgcagtttg	17400
ttaggtagtt	gaatattacc	tctgcttcca	cacattaagg	aatgctcccg	aacgcacttc	17460
ccaagtgttt	atttatttat	cattatacta	gacaatatgg	tgatacgatg	gtcacagaat	17520
agcggtttcc	acctccagag	cccataatct	agttgaaggg	aaagatattc	caacacaaga	17580
gtgttgacaa	tcaagataga	atatgatcaa	gggcccagtg	tgaggcccag	gcaatgatca	17640
ctgcaggaat	ctggggaaga	aagagaccag	cgtagcttggg	atatctagca	aaagtttcat	17700
gaaggagaat	ggactttgac	tttgaaatat	gggtaggatt	tacatatattt	gagatgagaa	17760
aaagaaagtt	cccagagaag	gaaagcatga	aaaggcaaac	agtctgtact	gaacgcgatg	17820
ctttgacaga	ataatgaaga	aagggacctg	ctggaatgat	tgatcagtgt	tcattcattca	17880
caccatcatc	atcaaaacac	ttatttaatg	agaacttact	gttttttagg	catggcttta	17940
atgccctata	tgaatttttt	tcttgattaa	tccttacaac	aaacatatcc	catagatagt	18000
tttattgtcc	cccttagaaa	agataaattg	cctaggctga	cacagtcagt	atatgaggca	18060
gtcaggattc	aaactaagtc	tgtttgttca	aaaaattaag	aatggccagc	tttttaaaat	18120
tttctgtctc	cagaagtatg	atttggctcc	actgaagttt	gcaaaacaaa	tgtgataccc	18180
aaaccttgtg	aaacttttag	tgggaaataa	ctttgcataa	gtcggtttga	gagagcgtgg	18240
aaacctgtct	tgaaaagttt	taatttaact	tgaggaaat	aaaaatgatg	ggtttctcaa	18300
ttaaaaattt	caatcaagga	aggatatgag	ctaacataac	atTTTTttta	aaagatcagt	18360
ctggtaaggt	agaggtgcat	aaactgaaaa	ggagcaaaag	tggtggaatt	cagttagaaa	18420
attattgtaa	ctgtactgat	gtcaaatgat	gaaaccatga	actaaagtag	tacccaaaagg	18480
agtgaggagg	atggaataat	tcaaaagata	gaggacagat	gtgcagaacc	tggagattat	18540
aagatgtgaa	aggaggagtt	tgagaaaatt	tcagattttg	gaagtgggtg	cattttacta	18600
aaaggatata	ataagtagca	aattttggat	aaagttgggt	cccactgagt	ttgagatggc	18660

tggttgacat	gcagagaaaa	ctgtcttgta	tgctgttctt	aaattgaaat	agacagacct	18720
ttaccctctg	atactgacat	atcttccttt	ccaggctcac	cctccatttc	cctaaacaca	18780
acacatgcac	tagctctcct	tactttattg	ctccacaaac	atcttacacc	tccaagcatt	18840
tgtgcccact	gtaccttcta	tctggaatct	cttttgtcct	cttgtgtgcc	tgaaaaattc	18900
ctttcagatc	ttcaaaatac	agtgcagatg	ctatttcttc	tagctcaaat	attatctcct	18960
ccatataatt	taattactct	cttttttctt	ttctctactt	tgcaattaca	tttatttgaa	19020
tgattgcttg	attaatttct	acctgtaaat	tatgtgaggg	caggctcctct	atattttgct	19080
cgcagttaaa	tctgcagcac	ttattataga	gtggtatcat	tagagtaata	tacatatatt	19140
tgaggacatg	ataaattaac	ttcccctata	gtatttatca	cattgcatct	caatgacttg	19200
cttatgtttc	tgttttccca	tataaattga	gtaacttgaa	aaaagagata	tctattaagt	19260
atttaatgag	aaattaaagt	acaaacttta	gtatgcataa	caacaaattg	ggaaaagggt	19320
gtaaacaaag	agatttgtag	ggcccatgag	ttagagatcg	tttcagcagg	tctgaaagga	19380
agcctaggaa	tctgcatttt	agaggaccac	ctcccaaccc	caacaagtaa	ttctgcttct	19440
tgttgtctgg	gtactgtact	ttaagaaatt	atggtgaaat	gatatcagcc	tttattgtat	19500
ttatcttatt	ctcatTTTTT	aatactagca	cttactgacc	aggctgcagc	aaattggctt	19560
attaatggta	agttttaata	ttattttgta	actgtaattt	gccaaatcat	aaagagtaaa	19620
agtgcaagtc	ttttgtgtac	ttttggccaa	ggcagtatct	atcaagttga	tgtctttgtt	19680
cttagttcgc	tcaggtggtg	ttgaaacaag	acagtgctga	tcccaagtgt	cccatggagt	19740
ggacttttagg	tttccccttt	ccttttagaa	aaaggaagaa	gttgtagtgg	aggactaccc	19800
actctgcact	caaaattgcc	ctcatgaaaa	tttctttggc	agctttgaga	accttttact	19860
gccctggttc	taaggtggca	tttctgtaga	cttacaaatt	atgtttgatg	acaccgttta	19920
tgtagcttct	cctaaccacc	agagtagctt	gctttgttgt	gaattcaggt	taatcacaaa	19980
gtataataaaa	aaagaattgt	cagaagtctt	cccagctttg	ggtctataac	ctgaaggaaa	20040
agtcactact	cttcaacatc	atcctatgta	ctctcaggct	aggatagcag	aaatgcaatc	20100
cctagaaaaac	agcaacttac	ttctctgacc	aaaaaaaaatgc	agttaaaaat	tagttcaatg	20160
tacctggtag	ctggcctatc	ttaggtactt	cagtgatTTT	acaaagtgat	ggtagtcta	20220
tgggtgtttt	tcagcttcac	tacgtattta	attcatgctt	attgttaatg	aaactgtgat	20280
aagcaattta	ctaggttatt	tgtttgggag	atgccacaaa	ggaacacatg	tatctcttaa	20340
tggaagcctg	gtcctccttt	atccaggaaa	tttgctagga	aaaaaaagcc	tttaggtggt	20400
tgtgctatta	aaccagggca	ctacttaaaa	gccagcccag	caatagttgt	gtgatttacc	20460
attaatttct	tagtaataga	ccacacaaaa	gaagaaaatt	atgggaatgc	gagttgagag	20520
gaattgggtg	atcagcctac	cccagccgt	ttcagctctg	gccagtagac	tattcacgag	20580

ctctttgaaa	acattttaa	aaaccttatt	tagatactag	aaacctctg	tcacctcaa	20640
gaatattctg	tggtatagcg	actcctttat	gagggcatgt	ttggtaatac	agcatcagtc	20700
ttggaggtgg	actggattct	acaaggtgaa	ctgcagtcac	taaggagtct	tttggatgag	20760
accagttttc	ctccaacttc	aatgtgtgca	tgaacctcac	atcaaaatgt	agctttagat	20820
ttgtcccatg	atgtggttcc	aagaatcagc	acttctaata	agtttccagg	ggatgcccac	20880
gctgcaggcc	cacaaaccac	actgagcata	gcaagactat	tgagaaaaag	gaaatttccc	20940
aggagtctgt	ggcctgagct	ggcacatcca	ataatgacct	atcttaacct	caactcatga	21000
ggaattccag	ggaactctga	agctgctcaa	aatttgaagc	ctatatgcc	actaaattca	21060
gaaatgttct	ccaaaatgct	atctataagc	aacagtagtc	acaaatgcat	tgtagaaata	21120
tatcgatcat	gctttttgga	aaatccagca	tgtcctgagg	aagaatgtat	aagacataaa	21180
agtcataaat	tatggaaaga	ctcttcagct	tcttccaaat	gtaaaggaat	catgatcttc	21240
ccagcacatt	aatgcccttt	ctcattagaa	tgtggggccg	gtccagacct	aataacattg	21300
tctgagcaga	gaatccttgg	aggcactgag	gctgaggagg	gaagctggcc	gtggcaagtc	21360
agtctgcggc	tcaataatgc	ccaccactgt	ggaggcagcc	tgatcaataa	catgtggatc	21420
ctgacagcag	ctcactgctt	cagaaggtga	ggccaccact	acctacctat	ctgggaacaa	21480
ttagaataga	caggtcatga	agactgcacc	ctctacccta	ggattgaatt	gagccagaaa	21540
taattcaatg	caaaaaaatc	agtaagaatt	ttcttcttat	tcatgaaagg	aaaaggattt	21600
ttccctttta	gcatgcta	ttagtgttat	ttctctgttt	caggtaataa	tatattagca	21660
cagtaaagaa	caaagattta	tatgtcagaa	tgttttttaa	atcctagcta	taaaagctta	21720
agaaatttac	taaatctcca	taagctttat	tttttttcca	aattaaggga	caacactggt	21780
atctgtgact	tagtgttact	ggtagcattg	agtacactaa	tgtaaacata	cgttaaatgt	21840
tagcgaaacg	aattgctgtg	gaagatttgc	acattatata	atgggagctg	atggctaacc	21900
tagagactgc	cccatgccat	taattttatt	attcataaag	attattgagt	atctagtatg	21960
agcacagtgt	tatatattgt	agaagctact	agtataaaca	aagtattgcc	tctgccttca	22020
aagagcttac	actcgaatgt	tggaatcaga	atgcacaaaa	ataatgatca	attacaatga	22080
gtagcataaa	taaaattaat	gtaggcaact	tacaagaatt	cttaattgag	gtgactaaac	22140
tattgccaac	actaggggtga	tatgctacca	gtggcgagta	ggttgcataa	acttacctta	22200
ttggtaaaaa	gaaaagttca	cattgctcat	aaaagaagga	tttttagattt	cagcataact	22260
aaaatctggt	tcaaacctgc	cttggttactg	gggcatcgca	gaccacaaca	gttggtggga	22320
acttaactca	aaaagttcac	ccagaaaaat	aatggagatt	tgaactcgtg	tgcccctgac	22380
catatcaatt	ttcttctcag	actcttactc	taaactggac	ctccttatca	cacacacaaa	22440

gccttccata	ggcagatcaa	tccagtctta	tttctcaaag	catgtacctt	gagcttcaga	22500
taaacagcat	tgttctcttc	ccctggactc	ttcctacatt	tccctaccta	tgagtatctg	22560
atcaatctgc	ttatccttga	aatgttaata	tatttaccac	atctctattt	gaattttatg	22620
aaatttttga	taattttctaa	gtagtttttt	cagattttata	ggcactactt	catggtacag	22680
tgactgttac	aaacgtattt	gttaaattta	gaaggaataa	agatttaaaa	gactagggtta	22740
gttactgaac	taaagtttta	ggaaatccca	aattattttca	aatttttctt	atggtaattt	22800
tatgacttaa	tatttttata	tgacgtgaac	aaatttgaaa	ctttaaaaga	tactcccaga	22860
attatcagtt	ttctgatgta	gattggcaaa	tttattacta	tatcccaaata	aacccaagag	22920
acaaaattca	caaaaacatt	tcaattttca	ttgccacttg	aaaggccaaa	aagcagaaat	22980
ggcacgcatt	gattttcaatc	gtactcttga	gtgtgggaac	caggaattaa	aatacctgga	23040
cttatcaggg	acttagcata	accaagaacg	gaatagaaac	ctccctggat	tctaagccct	23100
attcagtcct	aatcaccaaa	aaccaagtaa	acgatatcac	tataatgaaa	gccacagtta	23160
taaatatcga	caacgattac	caaaggaatc	catggaactt	tgaattttgc	caccccacat	23220
ccttctattc	attaccatga	ttgatccact	aaagctaaca	gactctgtga	accttgtatt	23280
ggacccctcc	ctaaagacct	gattgtcact	gagaaccatc	agtgaggatt	tgtttggggc	23340
atgaccagcc	ttacatcaaa	gtacatagaa	gtgatgaggt	cttatcaaag	aggattattg	23400
aattatcacc	tcttctatgt	agctttccct	gatactctct	ttcctctcca	ttgagttcca	23460
cagaaatfff	tttatctgcc	tttaacagtt	gtcctcatga	tttgtgatat	ttgacttacc	23520
tcttgtcagt	ttccttcact	agtgtagagt	tcctcaaaga	aagagaccat	aattacttat	23580
atfttttattc	ctggagactc	atactattcc	ttatacaaag	tagacactta	acaatggctt	23640
gttgaactat	aattaatgaa	aataatagct	accttcatga	aagttcactt	tgtgccaaac	23700
actatagttg	acataataca	tttgtctcat	taatacttaa	caattgtgtg	agaaggatc	23760
accaatcaca	ttttatatgt	aaataaaacc	cagagctatt	aattaacttg	tcataaataa	23820
cacttttcat	atgtggcata	gccaaagattt	aaatataaat	gttactgggt	ccaaaatgat	23880
gctctaattc	acttgctgga	aagaaggaaa	ggaagaaaat	aaacgagtgg	aaggaagaga	23940
gggagggag	agagaaaagg	aaggaaagaa	aaaagagtct	cttcagaacc	ttcactgtaa	24000
agactccgag	caaaagaagt	tgaatataaa	aacaacatag	gtttgtttgt	tttctaatat	24060
tttttcttca	aaatfttttaa	ctcaggttca	ctcttacaca	aactactgtg	tcttataaaa	24120
gtatttccgg	tcatagaatt	tttatftttct	gtattaactc	cactatctaa	tctccataaa	24180
actcctaaat	tggtattatc	ggtaacattt	tgtfttttact	caacccttag	gaacaatgft	24240
aagttaatca	gccctccaca	tcacagatcc	ttatftttcat	cagtctgtac	aaggcatttc	24300
tctcatttta	atftttftttc	ctcctgtcat	ccctggattt	cacttttact	gccctccttc	24360

cacccatatg	cctcatacta	atatattcga	aatatacatg	tcttaaaggt	acatgcacgc	24420
acctacaaaa	cctatagtgt	ttttttgtat	gtatatgtct	ttaatttaaa	taagtagcat	24480
tgtgtaaaag	tctaataattg	tttcttactg	ttttcactca	attccttgaa	ttttcatctg	24540
atgcactgct	gcatagcacc	ccatgggatg	cagccaccat	atttccttca	tccaattagg	24600
ttgcatgacc	taccttccca	ttgccacaaa	gagtacacac	aaaatatttg	tacttatctt	24660
tctgtaaacc	ttcaggaatt	tcagaagcac	acatgcaggc	tgctaaatat	accagaatac	24720
tttccagcca	cttaaatctt	taccagtatt	gcaaaagagg	ccccatttcc	ctccacatca	24780
acatttagta	ttattctttt	gtttaagttt	tatcaatctt	ttaaattgtac	acaagatgct	24840
catttttata	attttaattt	ctcagattac	tagtttgagt	atcttttcat	atatctaaga	24900
gctgttttga	tctcccctac	catgaactgc	cactaatatt	ctttgcctat	tttacaatgg	24960
tttttctgct	tatttattac	tggtttacag	acttttaaaa	tatattctac	aaaaatttta	25020
gacattaaac	attaccaata	ttttcccatg	gttcctcatc	catctggtaa	acttgtctat	25080
ggtatatcta	attttgattt	aatagaattc	attctatttt	taccttttag	tttgtgtttt	25140
tgttgtttag	ccaaaaagtc	cccattccta	ggtcataaag	gtaatgtcct	tttttttttt	25200
ttaacgctac	tgttctctct	ctgtctcccc	ctatgtatat	aggtgcacat	atacttgtag	25260
acacatacat	atacctatat	atgaggggag	ttcgataagt	ttatggaaaa	taaaattaaa	25320
agataaaaata	aaaaattata	aactttattt	ctcaacataa	gctccttcaa	gttcaagaca	25380
cttttgtaag	caataatacc	agccatatcg	tccatcccta	aagaactgag	ggtcctgaga	25440
atttaactat	gtcaatgcag	tctttttttac	attacttttt	tacagtactt	attgatgaaa	25500
aatgggtgcc	ttttaaagat	tgttttaaga	ttagggaaca	aaaataagtc	agaggaagtc	25560
aaatcaggac	tgaaaggtgg	atgcctagtg	atttattgct	gaaactttca	taaaactaac	25620
cttatttgat	gagaggaatg	agcatgagca	tggttgatg	ggagaagaac	tctggtggag	25680
ctttcctgga	cactttttct	actaaagctt	tggttaactt	tcttactctc	ataagaagaa	25740
gatgttattt	ttcactgacc	ctttagaagg	tcaacaagca	aaatgccttc	agcatcccaa	25800
atgtctgttg	tcatgacttt	tgttcttgac	tagtctgggt	ttgctttgac	tggtaccactt	25860
ctacctcttt	atagccattg	ctttgatggg	gctttgtctt	caagattgta	ttagtaaagc	25920
catatttcat	cttctgttac	aattcttcaa	agaaatactt	cagaatcttg	atctgacatg	25980
tttaaaattt	ctattggaag	ctctgacctt	gggtgcagct	gatctgggcg	aaacagtttt	26040
ggcatccatc	aagtagaaaag	tttgctcaac	tttagttttt	cagtcagaat	tgtataagct	26100
gaaccagttg	agatgtctat	ggtgttgtct	attgtttctc	acagttaatt	gttggtcctc	26160
tttgagacat	gaacaagatg	aaatttttcc	tagcaaactg	atgtggatga	tctgttgctg	26220

cgggcttcac	cctcaacaac	atctctttct	ttcttgaaac	aaattatcca	ttagtaaact	26280
gatgattggg	ggagatgctg	tccccataaa	ctttttgtaa	ggcataaata	atttcaccat	26340
tcttcagtt	tcaccataaa	tttgacgttt	ttttgcttca	attttagcag	cattcatggt	26400
gctttgataa	gagctctttt	caaattcatg	tcttattcct	cttagtgcct	caaactagat	26460
cttgttcagt	atgacaagtt	agtatgagtt	tatctgcatg	caaaaatctt	tgaaatccat	26520
gcatagtttg	tttataatat	acattttcaa	tgaacttttg	aagaccccat	acatacatat	26580
gtatatatat	gcacacacac	acacacacac	acaccaaact	cttcaacat	tatcagactt	26640
agtgacagaaa	aattattcat	ccattaacaa	gataagaatg	ccccttatca	tcactactat	26700
ttaaatggag	ctcctggcta	aaggaaaaga	cagggattga	aaaaaattag	ttaaatctaa	26760
aatgtttatt	atttcagggt	tcttagttgc	ttaaatggga	agggaggtat	ggacaaaaga	26820
gaaatcaaag	atatattgtg	tatgctactt	atcattaaag	tatcagaata	acttcattgg	26880
aatagaaaaa	caccaagatc	acccacgat	atgttttcta	aaatcttctc	catttcttta	26940
gacaagtgac	catgtattcg	gccagtgaag	aattaaactc	acttgccagc	ttataatgca	27000
ggaaaatata	gcaaagagat	gtggatccaa	tagtttctag	atagtggtag	aggatggcta	27060
agatgaattt	atatatctga	aatgttcaca	aattccctac	tcatatagca	tgttttcata	27120
atgttttagc	aactctaata	ctcgtgactg	gattgccacg	tctgggtatt	ccacaacatt	27180
tcctaaacta	agaatgagag	taagaaatat	tttaattcat	aacaattata	aatctgcaac	27240
tcatgaaaat	gacattgcac	ttgtgagact	tgagaacagt	gtcaccttta	ccaaagatat	27300
ccatagtgtg	tgtctcccag	ctgctaccca	gaatattcca	cctggctcta	ctgcttatgt	27360
aacaggatgg	ggcgtcaag	aatatgctgg	taagtgtctc	ggaaaaaaaa	attaacaata	27420
gaaatgtctt	atatttgcta	ttaggtaatt	ttttaaat	ggaaacatct	ggaatagggtg	27480
tttctattct	tctacagaca	gaaccattct	atattctgct	cagcccaagc	tctggctacc	27540
cctgagtctc	cttagcaaag	caaagcaatg	ctccagaaac	tatgggaatt	ctcaaata	27600
gtaataggaa	aatgtaaaag	aaagttatga	agacacgagt	tctttaataa	tccagagatt	27660
ctataagatt	caaatagctt	ccctataaac	aataaaaaag	attttgtttg	tttgtttggt	27720
tgcttgtttt	ttagagacaa	agactttctc	agactggagt	gcagtggtag	aatcatggct	27780
tactgcagcc	tcaaactctg	gtcttaagaa	atcctcttgc	ttcagcctcc	caagtagcta	27840
gaattataaa	taagtgtgta	ccaccatacc	cagctttttt	tttttttttc	tacagacagg	27900
ttcttgctct	gttgcccagg	ctggctgga	attcctgccc	tcaagccatc	ctcctgcctt	27960
gttgccctcc	caaagcaatg	ggaggattta	gattagacat	tgtatgaggg	cttaataatc	28020
cttaagggtat	taactgccct	ttaaagtatt	ctgggatatg	gcaaaaactc	gatgtgtata	28080
taaacattgg	tcatatttgt	ttattgaatg	aataaaatgg	aaactaaaat	gaggacaatg	28140

cacaagagct	actagaacca	gtaagagtat	cagcgaagga	gtggaagggg	agcattgaca	28200
atttccctgg	gcttttacc	atgttgtaga	ttgtctctcc	aaggaataat	acaaagcctt	28260
aatagtccta	gaacacattc	tattgtgttc	ttatggccca	aagtaaattg	gtgtagtaga	28320
taacatttgc	accagtcatg	aaaaactatt	ggtgtcattc	tgagagtaca	tcaatataaa	28380
atagactagt	tctttagcct	tgaaactaga	ctggtttctc	ttttgctgct	aggttaaagg	28440
ttattcaata	tgtaatcttc	caatccaaaa	tctgtcagtg	gataatttaa	aagcttttag	28500
tcaattttaa	gatatttgtt	ttcttaaaat	tttaaggggc	actgtgtcac	aaagctaaag	28560
aaaaaaaaa	aaaaaaaact	gatctgtgaa	aggggttatc	ctcatctact	tggggaattt	28620
tggctgcgaa	gaaactccaa	agtaaactct	tagaagcctt	cattgttaaa	tatgaaataa	28680
tgtttggagt	acatttat	cttctcaa	ttattatagg	gtcaataatg	tacacatctt	28740
gaagtccatt	tttttcctgc	ttttataaca	aacaggccac	acagttccag	agctaaggca	28800
aggacaggtc	agaataataa	gtaatgatgt	atgtaatgca	ccacatagtt	ataatggagc	28860
catcttgtct	ggaatgctgt	gtgctggagt	acctcaaggt	ggagtggacg	catgtcaggt	28920
aagctcaaga	caatctcatc	catgtcatca	tccaagaagt	gtataagcac	ttcctagtat	28980
gtgataatgt	gatagacata	agtgtaacag	ttacaataca	cagccctggt	cctctaaaa	29040
ttataatcta	gatttttaga	ataaattttt	ttatgaatga	agtttatcta	tcatgaaagc	29100
attaactctg	agaggccaaa	ttacagagta	gttaaccatc	caaagctcaa	gaatcagaaa	29160
gacctcgatt	tgaattcctt	aacctctatt	accaagtctc	taactaaaag	ctggggataa	29220
tcataatagc	acctaacttt	ttgggtacta	agaaaagtta	aatgaagact	aaatatatca	29280
ggcacatggt	aaacaacaaa	gaaatctcat	ctatttcact	attattaatg	tagaccatgg	29340
tcactcgtgt	taataacttt	aacctcaacc	ttttaactgc	tgtgaaggat	taaataaaaa	29400
attaatcact	atattataaa	aattaattga	tatataataa	atgaatttta	agagatacgt	29460
aataattcat	ggactccttg	aagatagaaa	atttatacaa	aatcctagta	atttgagtca	29520
caaaagctcc	tacaataatg	aaacagtatg	aatgaaaaag	aaaagaaata	actattatat	29580
ttggatctag	cccataattt	ttaaccaa	gcacaaaaac	aaacaacaaa	tatgaaattc	29640
tcactgtaaa	gtgattaaaa	tcaaatttga	attctaaaa	tttaaattaa	attatctaaa	29700
cataattgat	gcagttatat	gttttaatag	gttttgttca	catatctgaa	atccaactcc	29760
acacagtagc	aggaacagct	ggtgtcagaa	attaaatatt	cttttagtct	ggagttttaa	29820
aaaatcaatc	tgtttacttg	agtaatttgt	tgctgttttc	atgggtgaat	tgtatacaga	29880
aggataagaa	ttattcttcg	catcaaaagg	tactgactt	tcatatttag	tgctcatggt	29940
ctttaaaaag	tggataaaaa	gtagttctca	catttcatgg	aaagcccca	atccatgagc	30000

acatttccca	aatgaaaca	tttttatcaa	ctgcaagttg	tgtgtaggtg	gagatttggt	30060
tttcaattgt	caagatactg	ttaattaccc	agtcctttat	ctccttttgg	tgagatgtc	30120
tctgtgctag	gaaacccttc	ttgctctcct	tctgttttct	cttttactac	tgccctgaa	30180
acaacaaatt	ctcaagtttc	atgacagctt	tccaaagaat	ccatcaatca	aataagcaac	30240
acaactcgac	actgacaatt	ccagacctac	taagagcatt	aattaagact	taaaaataaa	30300
catgagtttt	aaaaggggtg	tattcattat	tttcccatth	ataacgtccc	ttaccttctg	30360
tccttcagtg	catacaaatt	attatcttcc	ttgaagccca	gttcaagccg	tacctacca	30420
tgataccttc	catgtatatt	ccactctagg	cctcactgat	ttttaactga	aatactataa	30480
tgcatagttc	acacttaaaa	aaaaaaaaaa	aacacagcac	tttacataag	agcttacagg	30540
atcctatthg	ttttatccat	tcttttgthc	atthtttaca	tcattaatth	aaaggaatta	30600
tattaattac	tttctatgca	cccgacgttg	tgthaacaca	acaatactat	ccctgcattc	30660
agcaagtcta	tggtctacaa	gagaggacac	aaattcaaat	gtctgtagtc	aagcagtga	30720
gctggctaga	tatggaaaaa	ttacaagthc	ctcttgctth	aacatttgct	tgccacatt	30780
tggtcagaca	tcatgcaaaa	taattthctc	ctatagaaaa	aaaaacacta	caaaaacaat	30840
aatataaaga	actgagaact	ggttaactga	agcatgcata	tgthcatctaa	aagaagcagg	30900
tgacgaccag	cttcatgaag	tacttgccat	gcataatggc	acttcacaca	ctgacccttc	30960
tccccaccta	gaccagtaat	taaacaggta	tggtatgagct	agctactaag	agcagccaac	31020
tgaatagctg	actaacttag	aagcacacth	ggtaataata	gctgacttht	attagtactg	31080
actatactat	atgctaagct	gtactcaaag	tgctthgagth	thtaaaactga	tacaaacatt	31140
atatgaggaa	acagaggthc	agagagctat	tcaccagcth	accaaaggth	acatagctgg	31200
taagtggagg	acttaaaccc	agactatcta	gtthcagaac	ccacagacth	aatccatcgt	31260
gcagaacata	agacatactc	catctgtctc	cccaactagg	ttattatgtg	cacaaatatt	31320
tattggthtg	ttggthcatt	attatgactg	ggtggtaagth	atgtcattag	gagthththg	31380
cttatgacta	tataaatthc	thcacaaaaa	gaagactthc	tgatgatata	ctatgcatca	31440
gacaccacgc	aggtgctaa	ggttaggaag	ataagtgaga	cttctagaaa	ctcattcatt	31500
caacaaatat	ctcctaaggg	ctagaagctt	aggtthcagc	agtgaacaga	ataggtatgt	31560
tctctthcgt	gttgacctth	atagtatatc	tggaacaa	gacattgaat	aaatatcaca	31620
aatgcaagtg	agththtcag	agacatgcag	ctgctacatc	aaaacaaaac	agaacaaaac	31680
aaacaaacaa	aaactgacca	gtgggattaa	gtgtaaatag	gcacacaaat	gcacaaatat	31740
gctthtataa	aatagtgaag	cagtgcagaa	gacacacaca	agatataaag	acacaatgaa	31800
gaacaattga	gccccaaagct	ggaaagggth	agagthtgaa	ggaaaaggth	tgatcagaga	31860
agththcccg	aaggagagaa	agcctggatg	attaggaggc	aaccactcgg	tgactgaggg	31920

aaatctgaaa	aatgtatttg	tcatctttctc	agacttgctg	aaggaatgac	ttgggtactt	31980
tgaggatttc	agtaattttt	ccatgacttg	gtataatatt	tcaaaaggaa	ataggctgac	32040
tttatttgta	taatgaatgt	gactccttcc	tcgactgcca	tagaaataaa	ctccttaata	32100
ttttgggttt	gtctttgcac	ttaagtaatc	agtcattctg	tttttttaca	gggtgactct	32160
ggtggcccac	tagtacaaga	agactcacgg	cggctttggt	ttattgtggg	gatagtaagc	32220
tggggagatc	agtgtggcct	gccggataag	ccaggagtgt	atactcgagt	gacagcctac	32280
cttgactgga	ttaggcaaca	aactgggatc	tagtgcaaca	agtgcacccc	tgttgcaaag	32340
tctgtatgca	ggtgtgcctg	tcttaaattc	caaagcttta	catttcaact	gaaaaagaaa	32400
ctagaaatgt	cctaatttta	catcttggtt	cataaatatg	gtttaacaaa	cactgtttta	32460
cctttcttta	ttattaaagg	ttttctattt	tctccagaga	actatatgaa	tgttgcatag	32520
tactgtggct	gtgtaacaga	agaaacacac	taaactaatt	acaaagttaa	caatttcatt	32580
acagttgtgc	taaatgcccg	tagtgagaag	aacaggaacc	ttgagcatgt	atagtagagg	32640
aacctgcaca	ggtctgatgg	gtcagagggg	tcttctctgg	gtttcactga	ggatgagaag	32700
taagcaaaact	gtggaaacat	gcaaaggaaa	aagtgataga	ataatattca	agacaaaaag	32760
aacagtatga	ggcaagagaa	ataatatgta	tttaaaattt	ttggttactc	aatatcttat	32820
acttagtatg	agtcctaaaa	ttaaaaatgt	gaaactgttg	tactatacgt	ataacctaac	32880
cttaattatt	ctgtaagaac	atgcttccat	aggaaatagt	ggataatttt	cagctattta	32940
aggcaaaaagc	taaaatagtt	cactcctcaa	ctgagaccca	aagaattata	gatatttttc	33000
atgatgaccc	atgaaaaata	tcactcatct	acataaagga	gagactatat	ctattttata	33060
gagaagctaa	gaaatatacc	tacacaaact	tgtcaggtgc	tttacaacta	catagtactt	33120
tttaacaaca	aaataataat	tttaagaatg	aaaaatttta	tcatcgggaa	gaacgtccca	33180
ctacagactt	cctatcactg	gcagttatat	ttttgagcgt	aaaagggctg	tcaaacgcta	33240
aatctaagta	acgaattgaa	agtttaaaga	gggggaagag	ttggtttgca	aaggaaaagt	33300
ttaaatagct	taatatcaat	agaatgatcc	tgaagacaga	aaaaactttg	tcactcttcc	33360
tctctcattt	tctttctctc	tctctcccct	tctcatcac	atgcctcccc	caccaaagaa	33420
tataatgtaa	attaaatcca	ctaaaatgta	atggcatgaa	aatctctgta	gtctgaatca	33480
ctaattattcc	tgagttttta	tgagctccta	gtacagctaa	agtttgcccta	tgcatgatca	33540
tctatgcgtc	agagcttcct	ccttctacaa	gctaactccc	tgcatctggg	catcaggact	33600
gctccataca	tttgctgaaa	acttcttgta	tttctctgatg	taaaattgtg	caaacaccta	33660
caataaagcc	atctactttt	agggaagggg	agttgaaaat	gcaaccaact	cttggcgaac	33720
tgtacaaaca	aatctttgct	atactttatt	tcaataaat	tctttttaaa	ataatttccc	33780

ES 2 886 958 T3

```

tgcctaatta tttatggaag ttatgacttt tgaaggacaa ttcaaaacca tttattttaat 33840
tggttctgca atgaaagaac tgcccatat actctactaa aggcttggca ctttctgctg 33900
ccttttaatc cagcgctata attgaggcaa gcgtccagct tgacacctcg agataacttc 33960
gtataatgta tgctatacga agttatgcta gtaactataa cggtcctaag gtagcgagct 34020
agctgcaacc gaggaaaaaa cgtgccatga ggtctctgta tccaagtgtg act 34073

```

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

<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

5	<210> 22		
	<211> 257		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
10	<223> oligonucleótido sintético		
	<400> 22		
	agcaccctc tcttcgcag agtctaagaa atcgctgtgt ttagccctcg cctggggcac	60	
	tgtcctcacg ggagctgctg tggtgctgt cttgctttgg aagttcagta agtgcagggg	120	
	gcctcgatcc caccatgtgc tcctgcagtc cccagtgctc tgagccagac cctgctctct	180	
15	gggctattga gacctctgga ggccctcgt gaggttcctc tcttacataa cgaggctgtc	240	
	tctcttcct tctcttg	257	
	<210> 23		
	<211> 190		
	<212> ADN		
20	<213> Secuencia artificial		
	<220>		
	<223> oligonucleótido sintético		
	<400> 23		
	ggtcagagga ccaaaggtga ggcaaggcca gacttggtgc tcctgtggtt ctcgagataa	60	
25	cttcgtataa tgtatgctat acgaagttat atgcatggcc tccgcgccgg gttttggcgc	120	
	ctcccgcggg cgccccctc ctcacggcga gcgctgccac gtcagacgaa gggcgagcg	180	
	agcgtcctga	190	
	<210> 24		
	<211> 171		
30	<212> ADN		
	<213> Secuencia artificial		
	<220>		
	<223> oligonucleótido sintético		
35	<400> 24		
	attgttttgc caagttctaa ttccatcaga cctcgacctg cagcccctag ataacttogt	60	
	ataatgtatg ctatacgaag ttatgctagt aactataacg gtcctaaggt agcgagctag	120	
	ctccacgtgg ctttgtccca gacttccttt gtcttcaaca accttctgca a	171	
40	<210> 25		
	<211> 177		
	<212> ADN		
	<213> Secuencia artificial		
	<220>		
45	<223> oligonucleótido sintético		
	<400> 25		

ES 2 886 958 T3

```

ggtcagagga ccaaaggtga ggcaaggcca gacttggtgc tcctgtggtt ctcgagataa      60
cttcgtataa tgtatgctat acgaagttat gctagtaact ataacggtcc taaggtagcg      120
agctagctcc acgtggcttt gtcccagact tcctttgtct tcaacaacct tctgcaa      177

```

5	<210> 26 <211> 20 <212> ADN <213> Secuencia artificial	
10	<220> <223> oligonucleótido sintético	
	<400> 26 gccgtgactg tgacctctc	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 gccttgcct caatggaac	20
35	<210> 29 <211> 21 <212> ADN <213> Secuencia artificial	
40	<220> <223> oligonucleótido sintético	
	<400> 29 ggttgacag 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	
	<220>	
	<223> oligonucleótido sintético	
5	<400> 31 tggaatggaa ggagctggag	20
	<210> 32	
10	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
15	<223> oligonucleótido sintético	
	<400> 32 gtccacctc ctgcaactg	19
20	<210> 33	
	<211> 22	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> oligonucleótido sintético	
	<400> 33 tgagccttc catcagcctg gg	22
30	<210> 34	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
35	<220>	
	<223> oligonucleótido sintético	
	<400> 34 ccacaatggc acatgggtct g	21
	<210> 35	
	<211> 18	
	<212> ADN	
45	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido sintético	
50	<400> 35 ggtgcttgct cccaaga	18
	<210> 36	
	<211> 20	
55	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido sintético	
60	<400> 36 cctaaaaggt gttgtaatgg	20
	<210> 37	
65	<211> 24	
	<212> ADN	

	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido sintético	
5	<400> 37	
	ggcaataaag aaggaagacg tttt	24
10	<210> 38	
	<211> 120	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> oligonucleótido sintético	
	<400> 38	
	ccagtcaggg acacacatgc tcacacgccc gccacccgc acacactaca gtcgagataa	60
20	cttcgtataa tgtatgctat acgaagttat atgcatggcc tccgcgccgg gttttggcgc	120
	<210> 39	
	<211> 198	
	<212> ADN	
	<213> Secuencia artificial	
25	<220>	
	<223> oligonucleótido sintético	
	<400> 39	
30	attctagttg tggtttgtcc aaactcatca atgtatctta tcatgtctgg aataacttcg	60
	tataatgtat gctatacgaa gttatgctag taactataac ggtcctaagg tagcgagcta	120
	gccaaagtctg tgtgctacca agtagcaaaa ctgagcctgg aactcacaca tgcgtgtctg	180
	agagcccagc actatcgc	198
35	<210> 40	
	<211> 100	
	<212> ADN	
	<213> Secuencia artificial	
40	<220>	
	<223> oligonucleótido sintético	
	<400> 40	
	taatctgact ttctcttcat cggctctctt tattctaggc tgagctgtaa cgctgccgtc	60
	ccccacatcc agaagctgct tcccttcaga cctacctacg	100
45	<210> 41	
	<211> 177	
	<212> ADN	
	<213> Secuencia artificial	
50	<220>	
	<223> oligonucleótido sintético	

ES 2 886 958 T3

<400> 41

ccagtcaggg acacacatgc tcacacgccc gccacccgc acacactaca gtcgagataa	60
cttcgtataa tgtatgctat acgaagttat gctagtaact ataacgggcc taaggtagcg	120
agctagccaa gtctgtgtgc taccaagtag caaaactgag cctggaactc acacatg	177

5 <210> 42
 <211> 19
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> oligonucleótido sintético

<400> 42
 gagcagggcc atgacacat 19

15 <210> 43
 <211> 24
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <223> oligonucleótido sintético

<400> 43
 accattagat cccagcactg gaca 24

25 <210> 44
 <211> 20
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <223> oligonucleótido sintético

35 <400> 44
 aaaccctcc cgagagagaa 20

40 <210> 45
 <211> 23
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <223> oligonucleótido sintético

<400> 45
 gaggaacact gtgtcaagga ctt 23

50 <210> 46
 <211> 22
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <223> oligonucleótido sintético

<400> 46
 cctgaaaagc ccggagtggc ag 22

60 <210> 47

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

	<211> 19	
	<212> ADN	
	<213> Secuencia artificial	
5	<220>	
	<223> oligonucleótido sintético	
	<400> 53	
10	gctgaggggc aggcttgag	19
	<210> 54	
	<211> 21	
	<212> ADN	
	<213> Secuencia artificial	
15	<220>	
	<223> oligonucleótido sintético	
	<400> 54	
20	tctgcagggt agggagagaa g	21
	<210> 55	
	<211> 29	
	<212> ADN	
25	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido sintético	
30	<400> 55	
	tgtttcagaa aaggaagact cacgttaca	29
	<210> 56	
	<211> 24	
35	<212> ADN	
	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido sintético	
40	<400> 56	
	gagaccgatg aagagaaagt caga	24
	<210> 57	
45	<211> 100	
	<212> ADN	
	<213> Secuencia artificial	
	<220>	
50	<223> oligonucleótido sintético	
	<400> 57	
	gaccatttta aggttttgct tgggtgtttt ggagggaggg tggcgctttg ctaatggtga	60
55	attactaact cctcaataaa gaatattatt tgaaataatt	100
	<210> 58	
	<211> 190	
	<212> ADN	
60	<213> Secuencia artificial	
	<220>	
	<223> oligonucleótido sintético	
65	<400> 58	

ES 2 886 958 T3

		gctgcctttt aatccagcgc tataattgag gcaagcgtcc agcttgacac ctcgagataa	60
		cttcgtataa tgtatgctat acgaagttat atgcatggcc tccgcgccgg gttttggcgc	120
		ctccccgggg cgcacctc ctcacggcga gcgctgccac gtcagacgaa gggcgagcg	180
		agcgtcctga	190
5		<210> 59 <211> 171 <212> ADN <213> Secuencia artificial	
10		<220> <223> oligonucleótido sintético	
		<400> 59	
		attgttttgc caagttctaa ttccatcaga cctcgacctg cagcccctag ataacttcgt	60
		ataatgtatg ctatacgaag ttatgctagt aactataacg gtcctaaggt agcgagctag	120
		ctgcaaccga ggaaaaaacg tgccatgagg tctctgtatc caagtgtgac t	171
15		<210> 60 <211> 177 <212> ADN <213> Secuencia artificial	
20		<220> <223> oligonucleótido sintético	
		<400> 60	
		ccagtcaggg acacacatgc tcacacgcc gccacccgc acacactaca ctcgagataa	60
		cttcgtataa tgtatgctat acgaagttat gctagtaact ataacggtcc taaggtagcg	120
25		agctagctgc aaccgaggaa aaaacgtgcc atgaggtctc tgtatccaag tgtgact	177
30		<210> 61 <211> 21 <212> ADN <213> Secuencia artificial	
		<220> <223> oligonucleótido sintético	
35		<400> 61 tcctctccag acaagaaagc t	21
40		<210> 62 <211> 30 <212> ADN <213> Secuencia artificial	
45		<220> <223> oligonucleótido sintético	
		<400> 62	

	tcatagcagc ttcaaatcc taaacgttga	30
5	<210> 63 <211> 20 <212> ADN <213> Secuencia artificial	
10	<220> <223> oligonucleótido sintético	
	<400> 63 tcgtgtgtag ctggtgagtt	20
15	<210> 64 <211> 22 <212> ADN <213> Secuencia artificial	
20	<220> <223> oligonucleótido sintético	
	<400> 64 catgcatca caggaggaga tc	22
25	<210> 65 <211> 22 <212> ADN <213> Secuencia artificial	
30	<220> <223> oligonucleótido sintético	
35	<400> 65 aattgggccc gaagccagat gc	22
40	<210> 66 <211> 20 <212> ADN <213> Secuencia artificial	
	<220> <223> oligonucleótido sintético	
45	<400> 66 cggaaggctt ctgtgacttc	20
50	<210> 67 <211> 25 <212> ADN <213> Secuencia artificial	
	<220> <223> oligonucleótido sintético	
55	<400> 67 gtctcccact tctgacataa tgaac	25
60	<210> 68 <211> 27 <212> ADN <213> Secuencia artificial	
65	<220> <223> oligonucleótido sintético	
	<400> 68	

	cccagtggtta accctacatc tggttcc	27
5	<210> 69 <211> 20 <212> ADN <213> Secuencia artificial	
10	<220> <223> oligonucleótido sintético	
	<400> 69 tgggaagaga ctctggaca	20
15	<210> 70 <211> 25 <212> ADN <213> Secuencia artificial	
20	<220> <223> oligonucleótido sintético	
	<400> 70 atgagctcct agtacagcta aagtt	25
25	<210> 71 <211> 26 <212> ADN <213> Secuencia artificial	
30	<220> <223> oligonucleótido sintético	
35	<400> 71 atgcatgatc atctatgcgt cagagc	26
40	<210> 72 <211> 21 <212> ADN <213> Secuencia artificial	
	<220> <223> oligonucleótido sintético	
45	<400> 72 tgcccagatg caggagtgta g	21

REIVINDICACIONES

1. Un roedor cuyo genoma comprende un gen *Tmprss* humanizado, en donde el gen *Tmprss* humanizado

- 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.

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.

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

- (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

en donde el gen *Tmprss2* 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.

5. El roedor de la reivindicación 1, en donde el gen *Tmprss* humanizado es un gen *Tmprss4* humanizado, el gen *Tmprss* de roedor endógeno es un gen *Tmprss4* de roedor endógeno y el gen *TMPRSS* humano afín es un gen *TMPRSS4* humano.

6. El roedor de la reivindicación 5, en donde la secuencia de nucleótidos del 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.

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.

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.

11. El roedor de acuerdo con una cualquiera de las reivindicaciones anteriores, en donde el roedor es homocigoto para el gen *Tmprss* humanizado.

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,
- 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 *TPRSS* humana codificada por el gen *TPRSS* humano afín y
 5 (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,
 15 (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 *TPRSS2* humano y (iii) el exón 4 codificante hasta el exón 13 codificante del gen *TPRSS2* 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 *TPRSS4*
 20 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 *TPRSS11D* humano.
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.
- 25 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:
- 30 proporcionar un roedor de acuerdo con una cualquiera de las reivindicaciones 1-11 o 15;
 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 *TPRSS* humana; y
 monitorizar la presencia y la gravedad de la infección por el virus de la gripe A en el roedor para determinar la
 35 eficacia terapéutica del compuesto candidato.

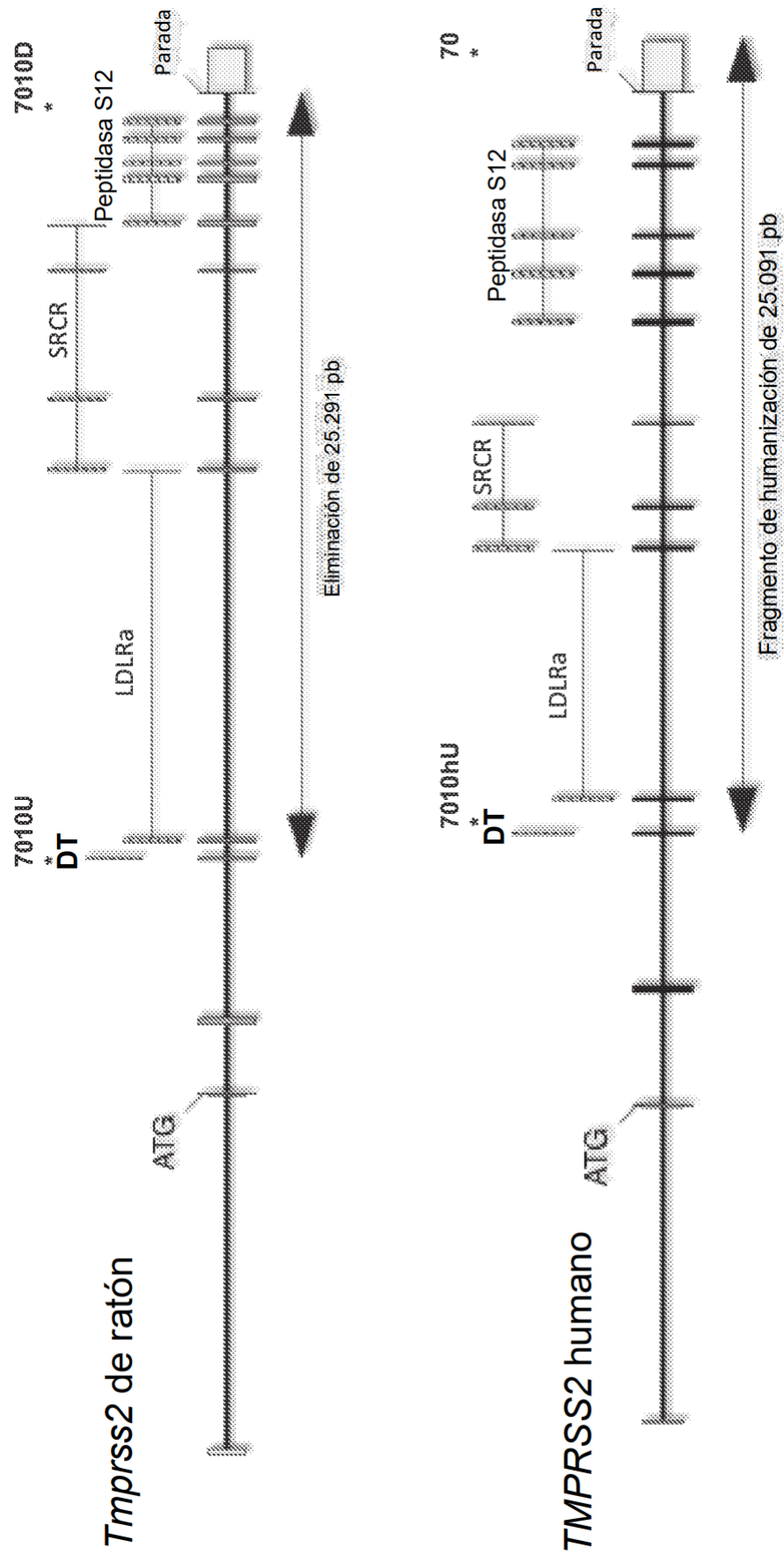
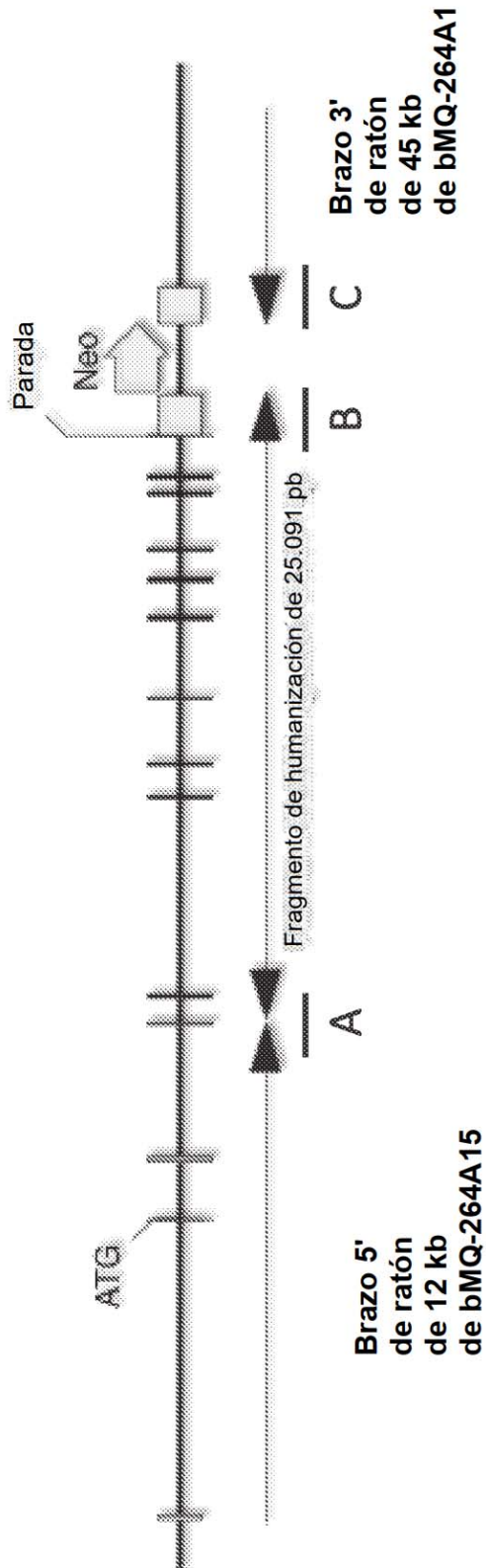
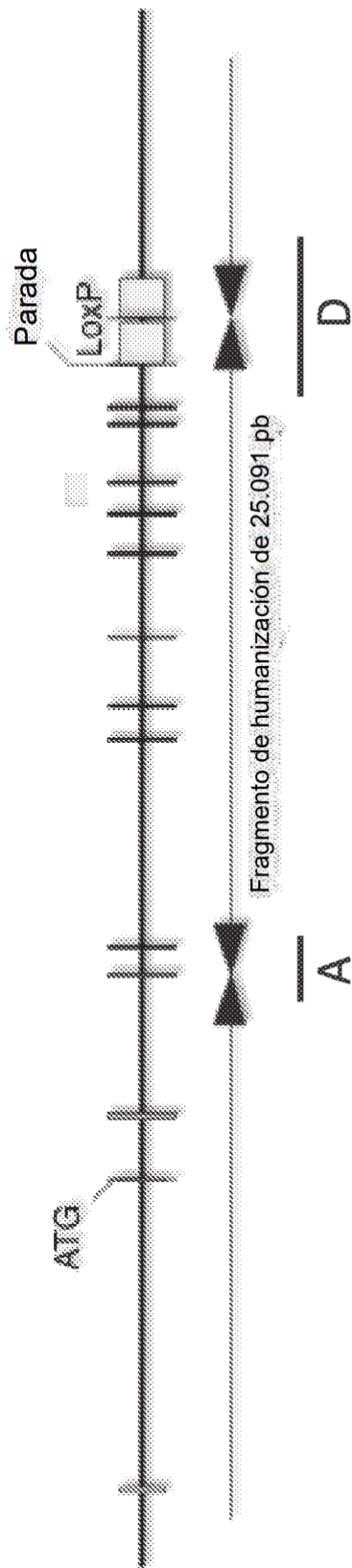


Figura 1A



A. 5' de ratón// 5' humana (SEQ ID NO:22)	AGCACCCCTC TCTCCGCAG AGTCTAAGAA ATCGCTGTGT TTAGCCCTCG CCCTGGGCAC TGTCTCAGG GGAGCTGCTG TGGCTGCTGT CTTCCTTTGG// AAGTTCAGTA AGTGCAGGGA GCCTCGATCC CACCATGTGC TCCTGCAGTC CCCAGTGCTC TGAGCCAGAC CCTGCTCTCT GGGCTATTGA GACCTCTGGA GGCCCTCCGT GAGGTTCTC TCTTACATAA CGAGGCTGTC TCTCTCCCT TCTCTTG
B. humana// <u>XhoI</u> // casete (<u>loxP</u>) (SEQ ID NO:23)	GGTCAGAGGA CCAAAGGTGA GGCAAGGCCA GACTTGGTGC TCCTGTGGTT// <u>CTCGAG</u> // <u>ATAACTTCG TATAATGTAT GCTATACGAA GTTAT</u> ATGCATGGCC TCCGCGCCGG GTTTTGGCGC CTCGCCGGG CGCCCCCTC CTCACGGCGA GCGCTGCCAC GTCAGACGAA GGGCGCAGCG AGCGTCCTGA
C. Casete (<u>loxP</u>)/ICEU// <u>NheI</u> // de ratón (SEQ ID NO:24)	ATTGTTTTCG CAAGTTCTAA TTCCATCAGA CCTCGACCTG CAGCCCCCTAG <u>ATAACTTCGI</u> <u>ATAATGTATG CTATACGAAG TTAI</u> /GCTAGTAACTATAACGGTCCCTAAGGTAGCGA// <u>GCTAGC</u> // TCCACGTGGC TTTGTCCCAG ACTTCCTTTG TCTTCAACAA CCTTCTGCAA

Figura 1B



A. 5' de ratón// 5' humana (SEQ ID NO:22)	AGCACCCCTC TCTTCCGCAG AGTCTAAGAA ATCGCTGTGT TTAGCCCTCG CCCTGGGCAC TGTCCTCAGG GGAGCTGCTG TGGCTGCTGT CTTGCTTTGG // AAGTTTCAGTA AGTGCAGGGA GCCTCGATCC CACCATGTGC TCCTGCAGTC CCCAGTGCTC TGAGCCAGAC CCTGCTCTCT GGGCTATTGA GACCTCTGGA GGCCCTCCGT GAGGTTCCCTC TCTTACATAA CGAGGCTGTC TCTCTCCCT TCTCTTG
D. humana// XhoI // (loxP) /ICEUI// NheI // de ratón (SEQ ID NO:25)	GGTCAGAGGA CCAAAGGTGA GGCAAGGCCA GACTTGGTGC TCCTGTGGTT// CTCGAG // ATAACTTCG TATAATGTAT GCTATACGAA GTTAT /GCTAGTAACTATAACGGTCCTAAGGTAGCGA // GCTAGC // TCCACGTGGC TTTGTCCCAG ACTTCCTTTG TCTTCAACAA CCTTCTGCAA

Figura 1C

	10	20	30	40	50	60
MALNSGSPPAIGPYENHGYPENPYPAQPTVVPTVEVHPAQYYPSVPQYAPRVLTQA						
MALNSGSPPGIGPCYENHGQSEHICPPRPVPAPNGYNLYPAQYYPSVPQYAPRITTQA						
MALNSGSPPGIGPCYENHGQSEHICPPRPVPAPNGYNLYPAQYYPSVPQYAPRITTQA						

	10	20	30	40	50	60
MALNSGSPPAIGPYENHGYPENPYPAQPTVVPTVEVHPAQYYPSVPQYAPRVLTQA						
MALNSGSPPGIGPCYENHGQSEHICPPRPVPAPNGYNLYPAQYYPSVPQYAPRITTQA						
MALNSGSPPGIGPCYENHGQSEHICPPRPVPAPNGYNLYPAQYYPSVPQYAPRITTQA						

	70	80	90	100	110	120
SNPVVCTQPKSPSGT	VCT	SKT	KAL	CIT	TLT	GLT
STSVIHTHPKS-	SGAL	CTSK	SKS	KL	CL	AL
STSVIHTHPKS-	SGAL	CTSK	SKS	KL	CL	AL

	70	80	90	100	110	120
SNPVVCTQPKSPSGT	VCT	SKT	KAL	CIT	TLT	GLT
STSVIHTHPKS-	SGAL	CTSK	SKS	KL	CL	AL
STSVIHTHPKS-	SGAL	CTSK	SKS	KL	CL	AL

	130	140	150	160	170	180
DSSGTCINPSNWC	DGVSHCPG	GEDENRCV	RLYGPNF	ILQVSSQR	KSWHPVCQ	DDWNENY
GSSGTCISSSLWC	DGVAHCPN	GEDENRCV	RLYGQSF	ILQVSSQR	KAWYPVCQ	DDWSESY
DSSGTCINPSNWC	DGVSHCPG	GEDENRCV	RLYGPNF	ILQVSSQR	KSWHPVCQ	DDWNENY
*****	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *	* * * * *

	130	140	150	160	170	180
DSSGTCINPSNWC	DGVSHCPG	GEDENRCV	RLYGPNF	ILQVSSQR	KSWHPVCQ	DDWNENY
GSSGTCISSSLWC	DGVAHCPN	GEDENRCV	RLYGQSF	ILQVSSQR	KAWYPVCQ	DDWSESY
DSSGTCINPSNWC	DGVSHCPG	GEDENRCV	RLYGPNF	ILQVSSQR	KSWHPVCQ	DDWNENY

[illegible][illegible]

	250	260	270	280	290	300																																																	
C	A	G	V	N	L	S	R	Q	S	R	I	V	G	E	S	A	L	P	G	A	W	P	Q	V	S	L	H	V	Q	N	V	H	V	C	G	S	I	I	T	P	E	W	I	V	T	A	A	H	C	V	E	K			
C	H	E	G	V	R	S	-	V	K	R	Q	S	R	I	V	G	L	N	A	S	P	G	D	W	P	Q	V	S	L	H	V	Q	N	V	H	V	C	G	S	I	I	T	P	E	W	I	V	T	A	A	H	C	V	E	E
C	H	A	C	G	V	N	L	S	R	Q	S	R	I	V	G	E	S	A	L	P	G	A	W	P	Q	V	S	L	H	V	Q	N	V	H	V	C	G	S	I	I	T	P	E	W	I	V	T	A	A	H	C	V	E	K	

	250	260	270	280	290	300																																																	
C	A	G	V	N	L	S	R	Q	S	R	I	V	G	E	S	A	L	P	G	A	W	P	Q	V	S	L	H	V	Q	N	V	H	V	C	G	S	I	I	T	P	E	W	I	V	T	A	A	H	C	V	E	K			
C	H	E	G	V	R	S	-	V	K	R	Q	S	R	I	V	G	L	N	A	S	P	G	D	W	P	Q	V	S	L	H	V	Q	N	V	H	V	C	G	S	I	I	T	P	E	W	I	V	T	A	A	H	C	V	E	E
C	H	A	C	G	V	N	L	S	R	Q	S	R	I	V	G	E	S	A	L	P	G	A	W	P	Q	V	S	L	H	V	Q	N	V	H	V	C	G	S	I	I	T	P	E	W	I	V	T	A	A	H	C	V	E	K	

Figura 1D

Alineación de proteínas Tmprss2

hTMPRSS2 310 320 330 340 350 360
PLNNPWHWTAFAGILRQSFMFYAGYQVEKVISHPNYDSKTKNNDIALMKLQKPLTFNDL
mTmprss2
PLSSPRYWTAFAGILRQSLMFYGSRRHQVEKVISHPNYDSKTKNNDIALMKLQKPLAFNDL
pro 7010 mutante
PLNNPWHWTAFAGILRQSFMFYAGYQVEKVISHPNYDSKTKNNDIALMKLQKPLTFNDL

hTMPRSS2 370 380 390 400 410 420
VKPVCLPNPGMMLQPEQLCWISGWGATEEKGKTSEVLNAAKVLLETQRCNSRYVDNLI
mTmprss2
VKPVCLPNPGMMLDLQECWISGWGATYEKGKTSVDLNAAAMVPLIEPSKNSKYIYNNLI
pro 7010 mutante
VKPVCLPNPGMMLQPEQLCWISGWGATEEKGKTSEVLNAAKVLLETQRCNSRYVDNLI

hTMPRSS2 430 440 450 460 470 480
TPAMICAGFLQGNVDSCQDGGPLVTSKNNIWWLIGDTSWGGCAKAYRPGVYGNVMVF
mTmprss2
TPAMICAGFLQGSVDSCQDGGPLVTLKNGIWWLIGDTSWGGCAKALRPGVYGNVTVF
pro 7010 mutante
TPAMICAGFLQGNVDSCQDGGPLVTSKNNIWWLIGDTSWGGCAKAYRPGVYGNVMVF

hTMPRSS2 490
TDWIYRQMRADG
mTmprss2
TDWIYQQMRANS
pro 7010 mutante
TDWIYRQMRADG

línea continua: DT
recuadro continuo: LDLRa
recuadro discontinuo: dominio SRCR
línea discontinua: peptidasa S1

Figura 1D (continuación)

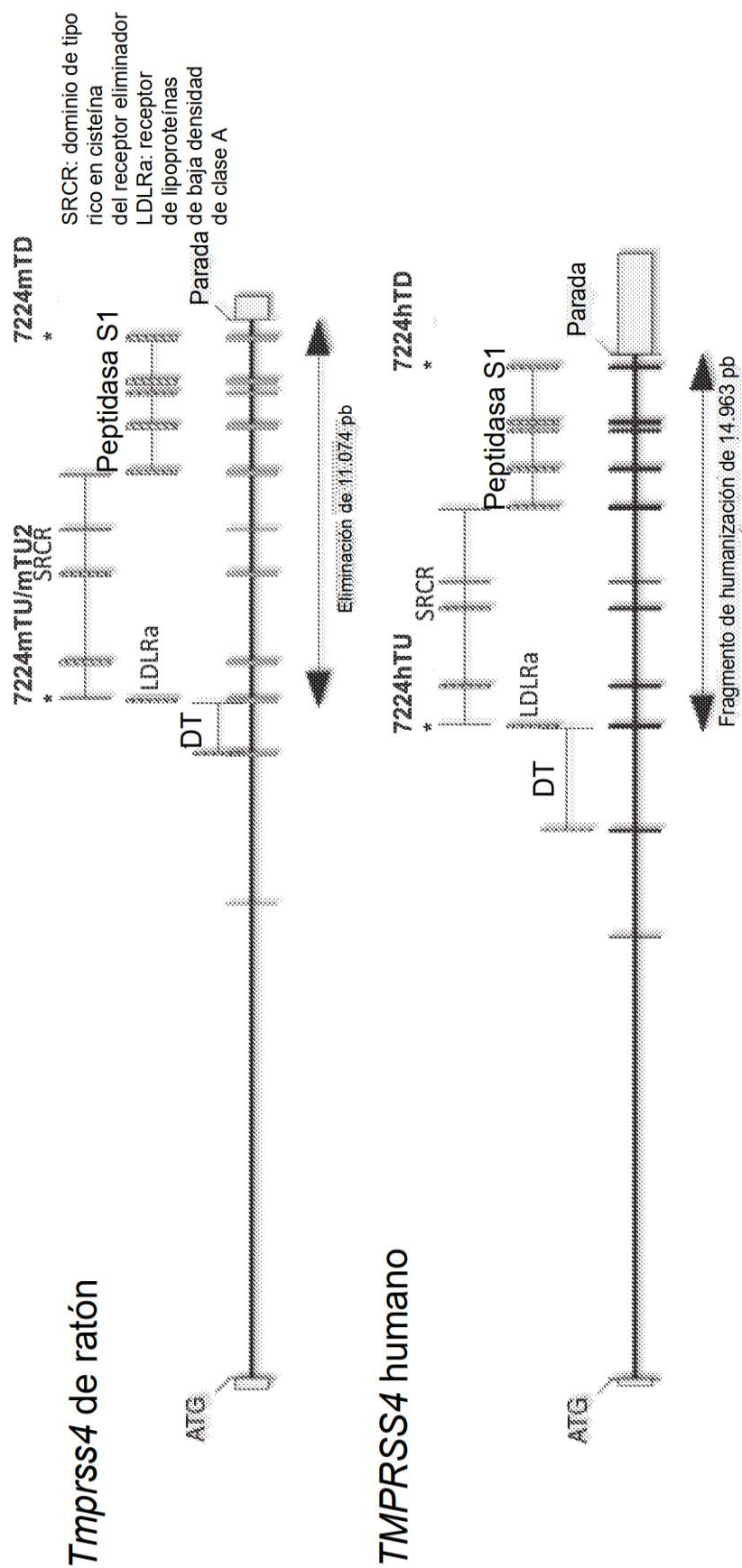
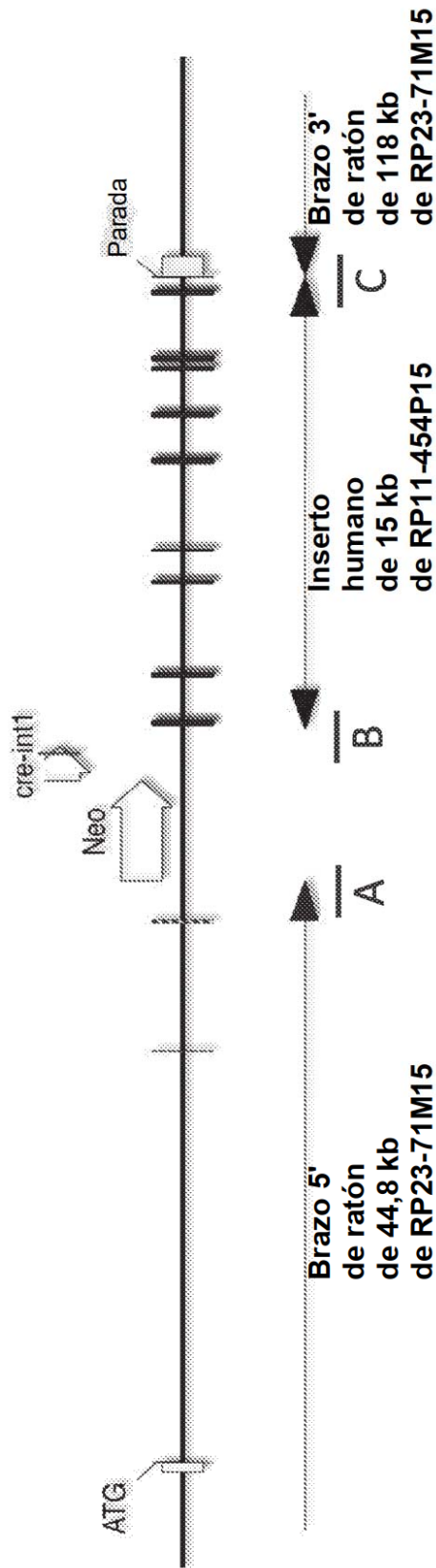


Figura 2A



A. De ratón// Sall-XhoI//casete (loxP) (SEQ ID NO:38)	CCAGTCAGGG ACACACATGC TCACACGCCC GCCCACC CGC ACACACTACA //GTCGAG //ATAACTTCGT ATAAIGTATG CTATACGAAG ITAI ATGCATGGCC TCCGCGCCGG GTTTTGGCGC
B. Casete (loxP)/ ICEUI// NheI // humana (SEQ ID NO:39)	ATTCTAGTTG TGGTTTGTC AAACATCATCA ATGTATCTTA TCATGCTGG AATAACTTCGTATATGCTATACGAAGTTAT/GCTAG / TAACATAACGGTCCTAAGGTAGCGA //GCTAGC// CAAGTCTGTG TGCTACCAAG TAGCAAACT GAGCCTGGAA CTCACACATG CGTGTCTGAG AGCCCAGCAC TATCGC
C. humana//de ratón (SEQ ID NO:40)	TAATCTGACT TTCTCTTCAT CCGTCTCTCT TATTCTAGGC TGAGCTGTAA // CGCTGCCGTC CCCCACATCC AGAAGCTGCT TCCCTTCAGA CCTACCTACG

Figura 2B

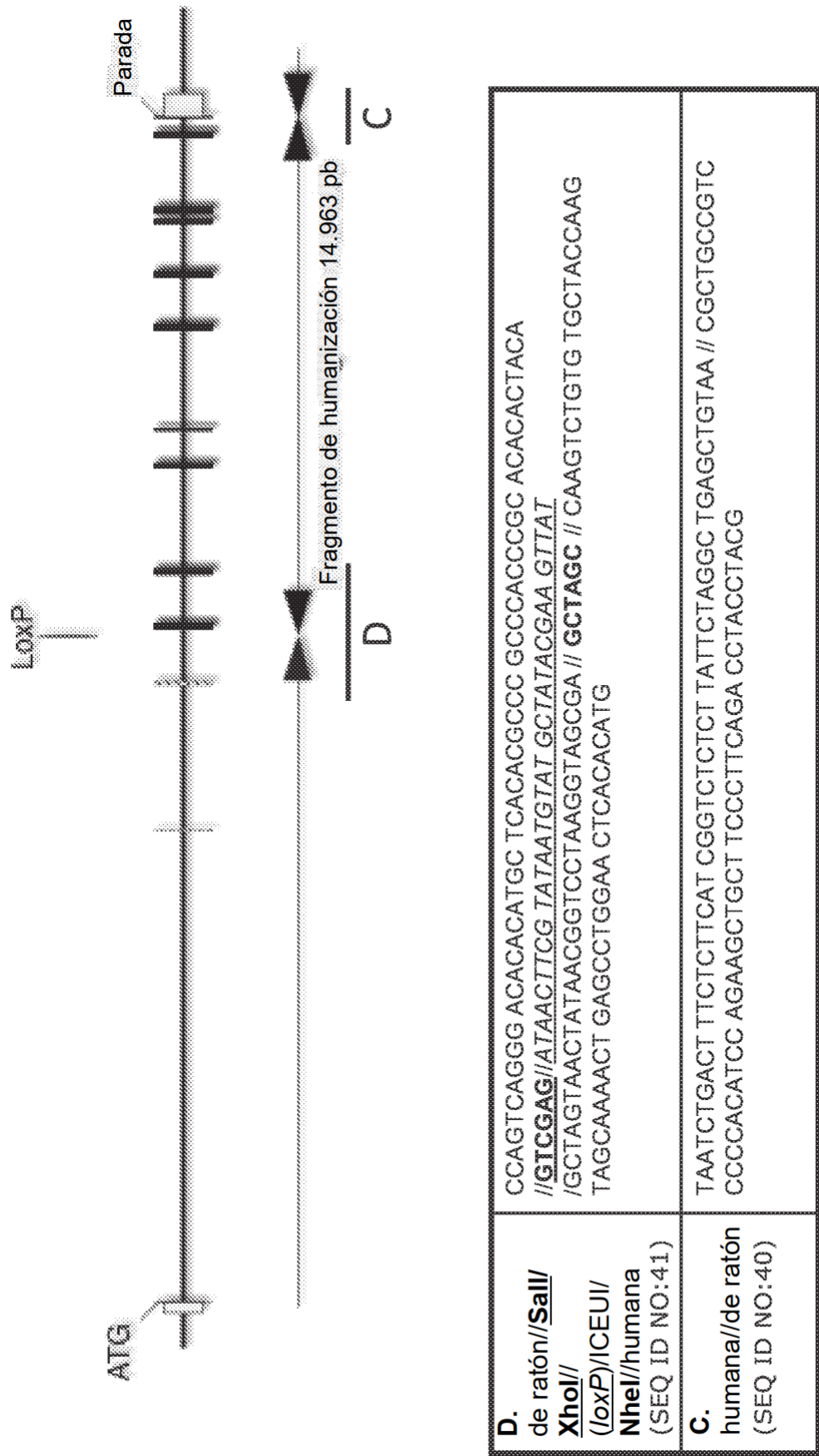


Figura 2C

hTMPRSS4
mTmprs4

hTMPRSS4
mTmprs4
pro 7224 mutante

hTMPRSS4
mTmprs4
pro 7224 mutante

hTMPRSS4
mTmprs4
pro 7224 mutante

Figura 2D

Alineación de proteínas Tmprss4

	310	320	330	340	350	360
hTMPRSS4	TFSGTVRPICLPFFDEELTPATPLWIIGWGFTKQNGKMSDILLQASVQVIDSTRCNADD					
mTmprss4	TFSGSVRPICLPFSDEVLPATPVVVIIGWGFTTEENGKMSDMLLQASVQVIDSTRCNAED					
pro 7224 mutante	TFSGTVRPICLPFFDEELTPATPLWIIGWGFTKQNGKMSDILLQASVQVIDSTRCNADD					
	.**	***	*****	*****	*****	*****
	370	380	390	400	410	420
hTMPRSS4	AYQGEVTEKMMCAGIPEGGVDTCCQGDSSGGPLMYQSDQWHVVGIVSWGCGGPSTPGVYT					
mTmprss4	AYEGEVTAEMLCAGTPQGGKDTCCQGDSSGGPLMYHSDKWQVVGI VSWGHC GGPSTPGVYT					
pro 7224 mutante	AYQGEVTEKMMCAGIPEGGVDTCCQGDSSGGPLMYQSDQWHVVGIVSWGCGGPSTPGVYT					
	.**	***	*****	*****	*****	*****
	430					
hTMPRSS4	KVSAYLNNWIYNVWKAEL					
mTmprss4	KVTAYLNNWIYNVRKSEM					
pro 7224 mutante	KVSAYLNNWIYNVWKAEL					
	.**	***				

línea continua: DT
recuadro continuo: LDLRa
recuadro discontinuo: dominio SRCR
línea discontinua: peptidasa S1

Figura 2D (continuación)

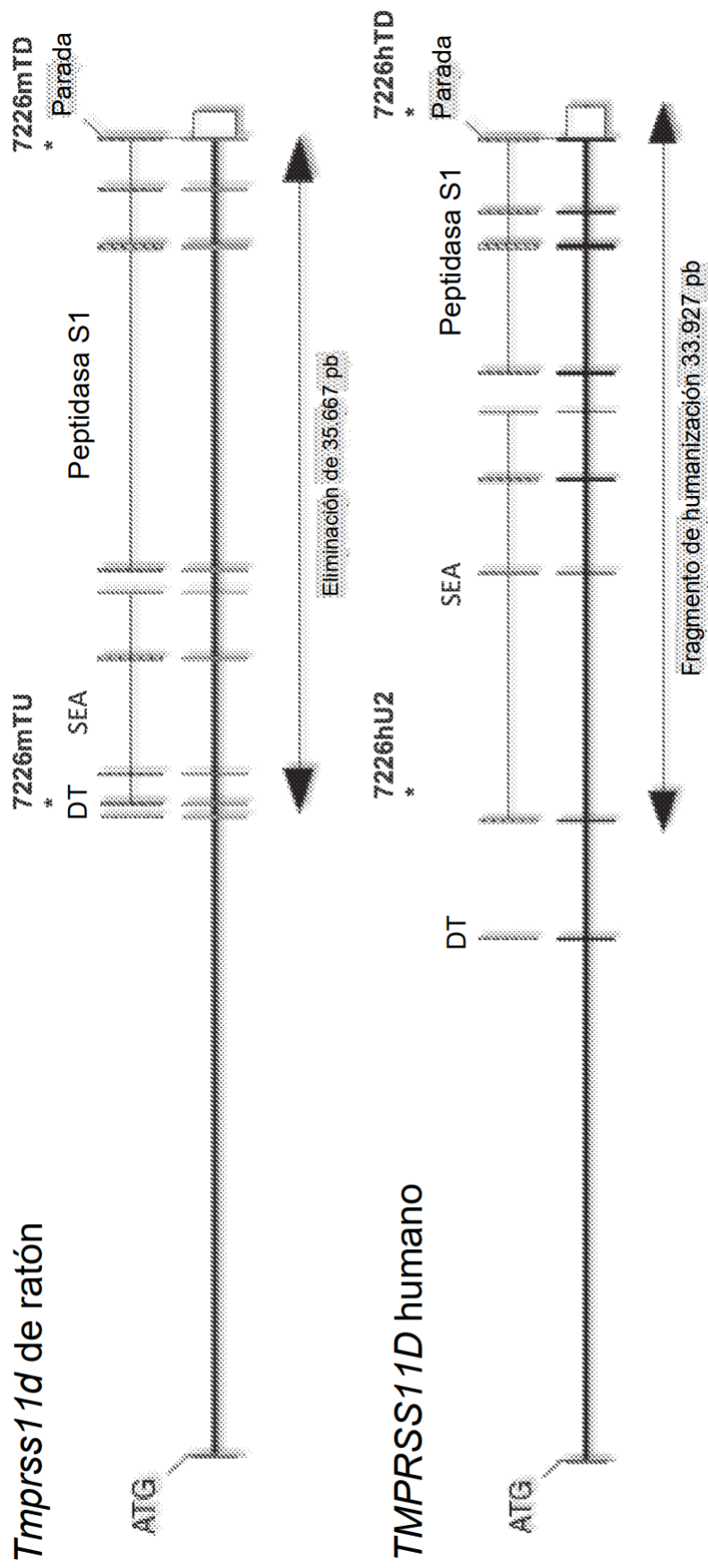
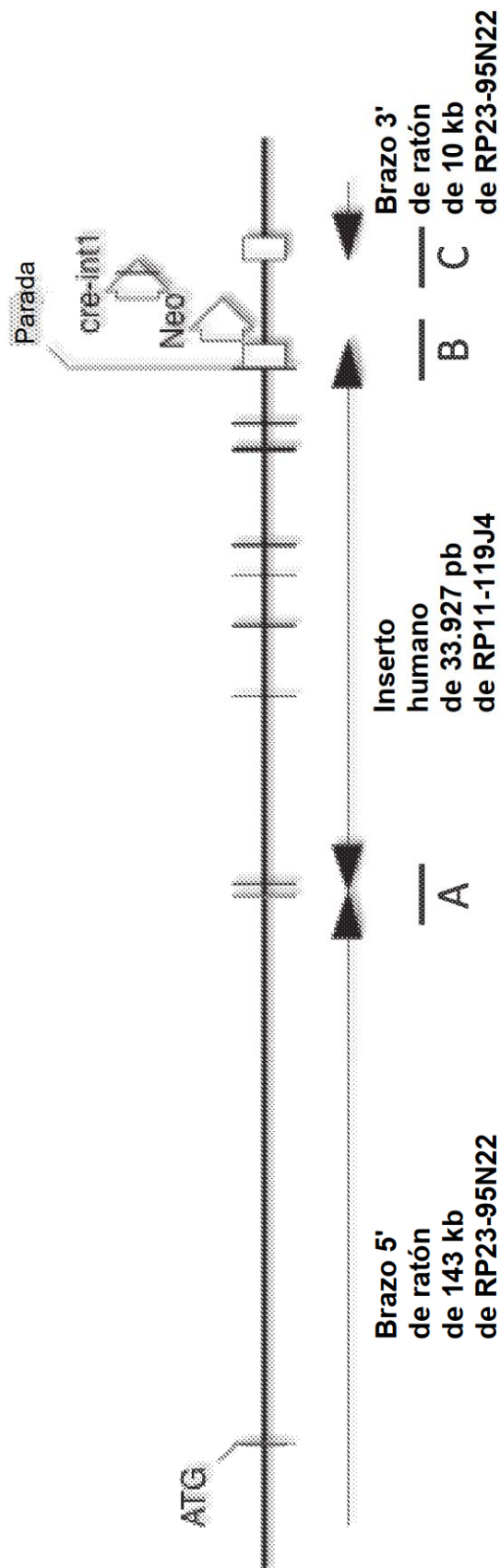
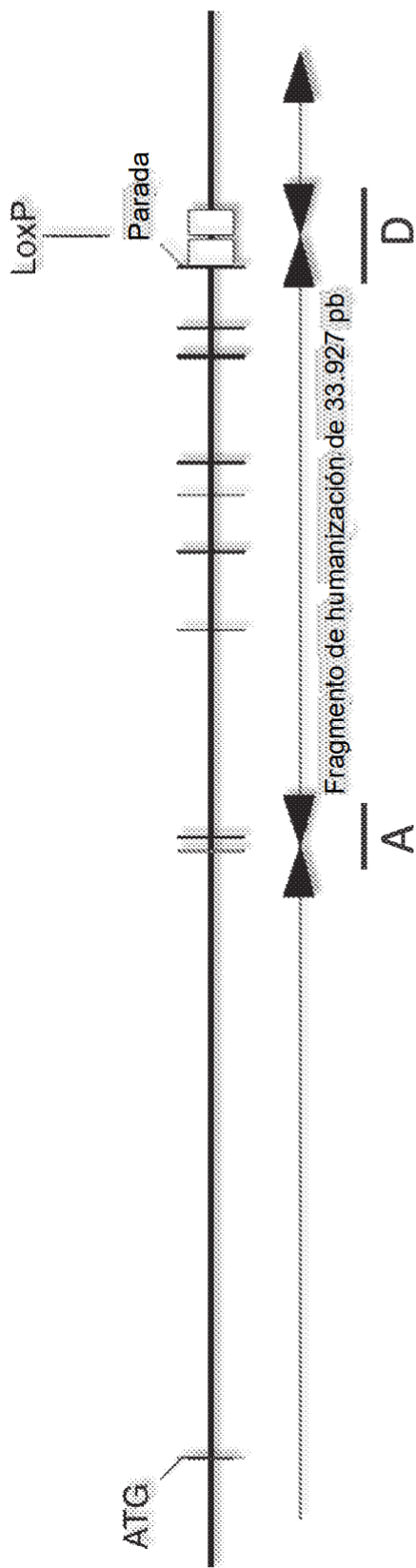


Figura 3A



A. de ratón// humana (SEQ ID NO:57)	GACCATTTTA AGGTTTTTGCT TGGTTGTTTT GGAGGGAGGG TGGTGCTTTG // CTAATGGTGA ATTACTAACT CCTCAATAAA GAATATTATT TGAATAAATT
B. humana// <u>XhoI</u> // casete (<u>loxP</u>) (SEQ ID NO:58)	GCTGCCTTTT AATCCAGCGC TATAATTGAG GCAAGCGTCC AGCTTGACACI// <u>CTCGAG</u> // <u>ATAACTTCG TATAATGTAT GCTATACGAA GTTAT</u> ATGCATGGCC TCCGCGCCGG GTTTTGGCGC CTCGCGCGG CGCCCCCTC CTCACGGCGA GCGCTGCCAC GTCAGACGAA GGCGGCAGCG AGCGTCCTGA
C. Casete (<u>loxP</u>)/ ICEU// <u>NheI</u> // de ratón (SEQ ID NO:59)	ATTGTTTTGC CAAGTTCTAA TTCCATCAGA CCTCGACCTG CAGCCCCCTAG <u>ATAACTTCGT</u> <u>ATAATGTATG CTATACGAAG TTAI</u> /GCTAGTAACATAACGGTCCTAAGGTAGCGA // <u>GCTAGC</u> //TGCAACCGAG GAAAAAACGT GCCATGAGGT CTCTGTATCC AAGTGTGACT

Figura 3B



A. de ratón// humana (SEQ ID NO:57)	GACCATTTTA AGGTTTTGCT TGGTTGTTTT GGAGGGAGGG TGGTGCTTTG // CTAATGGTGA ATTACTAACT CCTCAATAAA GAATATTATT TGAATAATT
D. humana//XhoI// (loxP)/ICEUI//NheI// de ratón (SEQ ID NO:60)	CCAGTCAGGG ACACACATGC TCACACGCC GCCACCCGC ACACACTACA //CTCGAG//ATAACTTCG TATAATGTAT GCTATACGAA GTTAT /GCTAGTAACTATAACGGTCCTAAGGTAGCGA // GCTAGC // TGCAACCGAG GAAAAAACGT GCCATGAGGT CTCTGTATCC AAGTGTGACT

Figura 3C

		Alineación de proteínas Tmprss11d									
		10	20	30	40	50	60				
hTMPRSS11D		MYRPARVTSTSRFLNPYVVC	FIVVAGV	VILAVT	IALLVYFLAFDQ	KSYFYRSSFQ	LLNVE				
mTmprss11d		MYRPRMLSPSRFFTPFAVAFV	VIITVGLLAM	AGLLIHFLAFD	KKAYFYHSSFQ	ILNVE					
pro 7226 mutante		MYRPRMLSPSRFFTPFAVAFV	VIITVGLLAM	AGLLIHFLAFD	KSYFYRSSFQ	ILNVE					
		***	. * ***	. * . * . * . *	. * . * . * . *	. * . * . * . *	. * . * . * . *				
hTMPRSS11D		YNSQLNSPATQEYRTL	SGRIESLI	TKTFKES	NLRNQFIRAHVAKLR	QDGGSVRADV	VVMKE				
mTmprss11d		YTEALNSPATHEYRTL	SERIEAMIT	DEFRGSS	LKSEFIRTHVVKL	KEGTGVADV	VVMKE				
pro 7226 mutante		YNSQLNSPATQEYRTL	SGRIESLI	TKTFKES	NLRNQFIRAHVAKLR	QDGGSVRADV	VVMKE				
		x . *****	*****	***	. * . * . * . *	. * . * . * . *	. * . * . * . *				
hTMPRSS11D		QFTRNNNGASMK	RIESVLRQ	MLNNSG	LEINPSTEIT	SLTDQ	AAANWL	INECGAGPDLI			
mTmprss11d		RSSKRNKRKVMK	TRIQSVLRR	-LSSSGN	LEIAPSNEIT	SLTDQ	TENVLT	QECGARPDLI			
pro 7226 mutante		QFTRNNNGASMK	RIESVLRQ	MLNNSG	LEINPSTEIT	SLTDQ	AAANWL	INECGAGPDLI			
		. . . * . *	. * . * . * . *	. * . * . * . *	. * . * . * . *	. * . * . * . *	. * . * . * . *				
hTMPRSS11D		TLSEQRILGGTEAE	EGSWPWQV	SLRLNNAH	HCGGLINN	MMWILTA	AHCFRS	NSNPRDWIA			
mTmprss11d		TLSEERIIGGMQAE	PGDWPQV	SLQLNNVH	HCGGALISNMWV	LTA	AHCFKSY	PNPQYWTA			
pro 7226 mutante		TLSEQRILGGTEAE	EGSWPWQV	SLRLNNAH	HCGGGLINN	MMWILTA	AHCFRS	NSNPRDWIA			
		***	. * . * . *	. * . * . * . *	. * . * . * . *	. * . * . * . *	. * . * . * . *				
hTMPRSS11D		TSGISTTFPKLR	MVRVRN	ILIHNNYK	SATHENDIA	LVRLNSV	TFTKDIH	SVCLPAATQNI			
mTmprss11d		TFGVSTMSPR	LRVRVRAI	LADGYSSV	TRDNDIAV	VQIDRSV	AFSRNIH	RVCLPAATQNI			
pro 7226 mutante		TSGISTTFPKLR	MVRVRN	ILIHNNYK	SATHENDIA	LVRLNSV	TFTKDIH	SVCLPAATQNI			
		* . * . *	. * . * . *	. * . * . * . *	. * . * . * . *	. * . * . * . *	. * . * . * . *				

Figura 3D

Alineación de proteínas Tmprss11d

	310	320	330	340	350	360
hTMPRSS11D	PPGSTAYVTGWGAQ	EYAGHTVPELRQ	QVRIISNDVCNAPH	SYNGAILSGMLCAG	VPQGG	
mTmprss11d	IPGSVAYVTGWGSL	TYGGNAVTLRQGE	VRIISSEECNTPAG	YSGSVLPGMLCAG	MRS	GA
pro 7226 mutante	PPGSTAYVTGWGAQ	EYAGHTVPELRQ	QVRIISNDVCNAPH	SYNGAILSGMLCAG	VPQGG	
	***	*****	*****	*****	*****	*
	*	..*	..*	..*	..*	*
	370	380	390	400	410	
hTMPRSS11D	VDACQGD	SGGFLVQEDSRRL	WFIVGIVSWG	DCQGLPKPGVYTR	VAYLDWIRQQTGI	
mTmprss11d	VDACQGD	SGGFLVQEDSRRL	WFVVGIVSWG	YQCGLPNKPGVYTR	VAYRNWIRQQTGI	
pro 7226 mutante	VDACQGD	SGGFLVQEDSRRL	WFIVGIVSWG	DCQGLPKPGVYTR	VAYLDWIRQQTGI	
	***	*****	*****	*****	*****	*
	*	..*	..*	..*	..*	*

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

Figura 3D (continuación)

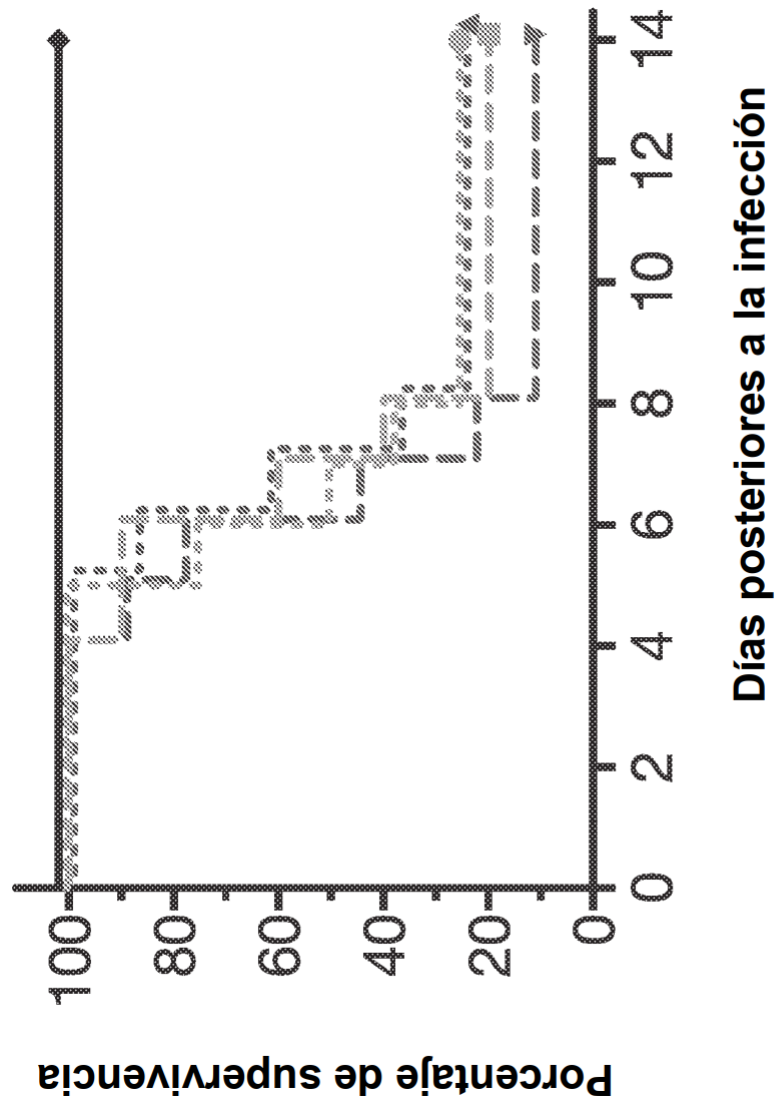


Figura 4