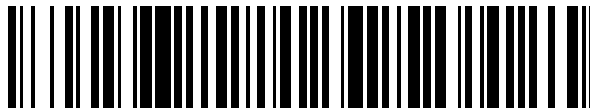


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 719 304**

51 Int. Cl.:

C12N 1/20 (2006.01)

C12N 9/04 (2006.01)

C12P 7/22 (2006.01)

C12P 1/04 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **07.08.2012 PCT/US2012/049842**

87 Fecha y número de publicación internacional: **14.02.2013 WO13022881**

96 Fecha de presentación y número de la solicitud europea: **07.08.2012 E 12822242 (9)**

97 Fecha y número de publicación de la concesión europea: **07.11.2018 EP 2742126**

54 Título: **Composiciones y métodos para la biosíntesis de vainillina o beta-D-glucósido de vainillina**

30 Prioridad:

08.08.2011 US 201161521090 P
10.08.2011 US 201161522096 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
09.07.2019

73 Titular/es:

**INTERNATIONAL FLAVORS & FRAGRANCES
INC. (50.0%)
521 West 57th Street
New York, NY 10019, US y
EVOLVA SA (50.0%)**

72 Inventor/es:

**HANSEN, JOERGEN;
HANSEN, ESBEN HALKJAER;
HONEY, POLUR;
SHERIDAN, JOSEPH M.;
HEAL, JONATHAN R. y
HAMILTON, WILLIAM D.O.**

74 Agente/Representante:

LEHMANN NOVO, María Isabel

Observaciones:

**Véase nota informativa (Remarks, Remarques o
Bemerkungen) en el folleto original publicado por
la Oficina Europea de Patentes**

ES 2 719 304 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

Composiciones y métodos para la biosíntesis de vainillina o beta-D-glucósido de vainillina

Introducción

5 Esta solicitud reivindica el beneficio de prioridad de la solicitud provisional de EE.UU. nº 61/521.090, presentada el 8 de agosto de 2011 y 61/522.096, presentada el 10 de agosto de 2011.

Antecedente de la invención

10 La vainillina es uno de los compuestos de sabor más importantes en el mundo con un mercado global de 180 millones de dólares. La vainillina natural se obtiene de las vainas de semillas curadas de la orquídea vainilla (*Vanilla planifolia*), pero la mayor parte de la vainilla mundial se sintetiza a partir de productos petroquímicos o ligninas de pasta de madera. La producción de vainilla natural a partir de la vaina de vainilla es un procedimiento lento y laborioso, que requiere la polinización manual de las flores y un procedimiento de curado de 1-6 meses de las vainas de vainilla verdes recogidas (Ramachandra & Ravishankar (2000) *J. Sci. Food Agric.* 80:289-304). La producción de 15 1 kilogramo (kg) de vainillina requiere aproximadamente 500 kg de vainilla, que corresponde a la polinización de aproximadamente 40.000 flores. Actualmente solo aproximadamente 0,25% (40 toneladas de 16.000) de vainillina vendidas anualmente procede de vainas de vainilla, mientras que la mayor parte del resto se sintetiza químicamente a partir de lignina o hidrocarburos fósiles, en particular guayacol. La vainilla producida de forma sintética se vende a aproximadamente 15 dólares por kg, en comparación con los precios de 1200-4000 dólares por kg de la vainilla natural (Walton, et al. (2003) *Phytochemistry* 63:505-515).

Resumen de la invención

20 Esta invención proporciona un método para producir vainillina o beta-D-glucósido de vainillina. El método de la invención implica las etapas de (a) proporcionar un hospedante recombinante capaz de producir vainillina, en donde dicho hospedante recombinante alberga un ácido nucleico heterólogo que codifica un polipéptido de enzima multifuncional arom (AROM) mutante; (b) cultivar dicho hospedante recombinante durante un tiempo suficiente para que dicho hospedante recombinante produzca vainillina y/o glucósido de vainillina; y (c) aislar la vainillina y/o 25 glucósido de vainillina de dicho hospedante recombinante o del líquido sobrenadante del cultivo, de modo que se produce vainillina y/o beta-D-glucósido de vainillina.

El polipéptido de AROM mutante tiene actividad de shikimato deshidrogenasa menor con respecto al polipéptido de AROM correspondiente que carece de la mutación. El polipéptido de AROM mutante tiene una o más mutaciones en el dominio 5, o carece del dominio 5.

30 Cualquiera de los polipéptidos descritos en la presente memoria puede incluir además un marcador de purificación, un péptido de tránsito al cloroplasto, un péptido de tránsito mitocondrial, un péptido de amiloplasto, péptido señal, o un marcador de secreción en el extremo N o C del polipéptido.

35 El hospedante es un microorganismo, p. ej., una levadura tal como *Saccharomyces cerevisiae*, *Schizosaccharomyces pombe* o *Escherichia coli*, una planta o célula de planta (p. ej., *Physcomitrella* o planta o célula de planta de tabaco).

40 Cualquiera de los mutantes descritos en la presente memoria puede incluir además un ácido nucleico heterólogo que codifica un polipéptido de catecol-O-metil transferasa (COMT) mutante, un gen que codifica una ácido carboxílico aromático reductasa (ACAR); una 3-deshidrosikimato deshidratasa (3DSD); una uridina 5'-difosfoglucosil transferasa (UGT); una fosfopanteteína transferasa (PPTasa); una AROM de tipo natural; y/o una O-metiltransferasas de tipo natural (OMT).

También se proporciona un ácido nucleico aislado que codifica un polipéptido de AROM mutante.

Breve descripción de los dibujos

45 La figura 1 es un esquema de la biosíntesis nueva de la vainillina (4) y describe los diferentes catabolitos de vainillina y productos secundarios metabólicos, es decir, ácido deshidrosikímico (1), ácido protocatéuico (2), aldehído protocatéuico (3), ácido vainílico (5), alcohol protocatéuico (6), alcohol de vainillilo (7), y β -D-glucósido de vainillina (8), encontrados en un organismo que expresa 3DSD, ACAR, OMT y UGT y una fosfopanteteína transferasa (PPTasa). Las flechas blancas muestran reacciones metabólicas primarias en levaduras; las flechas negras muestran reacciones enzimáticas introducidas por modificación genética metabólica; las flechas ralladas diagonalmente muestran reacciones metabólicas de levaduras inherentes indeseadas.

50 La figura 2 es una representación esquemática de la biosíntesis celular de los aminoácidos aromáticos tirosina, triptófano y fenilalanina usando un polipéptido de AROM de tipo natural (izquierda) y un polipéptido de AROM mutante que carece del dominio 5 (derecha).

La figura 3A muestra la producción de vainillina (VG) e isovainillina (IsoVG) en cepa de levadura que expresa diferentes polipéptidos de COMT mutantes.

La figura 3B y figura 3C muestra la producción de vainillina (VG) e isovainillina (IsoVG) por polipéptidos de COMT aislados de diferentes fuentes en comparación con OMT humana (Hs-OMT) y la OMT mutante L198Y humana (HS-OMT L198Y). *Phytophthora infestans* (PI), *Catharanthus roseus* (CR), *Yarrowia lipolytica* (YL), *Ciona intestinalis* números de acceso en GENBANK XP_002121420 y XP_002131313 (CI-1 y CI-2), *Capsasproa owczarzewski* (CO), *Chaetomium thermophilum* (CT), *Clavispora lusitanae* (CL), *Paracoccidioides* sp. 'Iutzii' Pb01 (PL), *Vanilla planifolia* (VP), *Coffea Arabica* (CA), *Rattus norvegicus* (RN), *Mus musculus* (MM), *Crenarchaeote* (CREN), *Mycobacterium vanbaaleeni* (MV), o *Schizosaccharomyces pombe* (SP).

La figura 4 muestra niveles de glucósido de vainillina, vainillina, glucósido de alcohol de vainillilo y alcohol de vainillilo en cepas de levaduras que expresan alcohol de vainillilo oxidasa (VAO) *Penicillium simplicissium* (PS) o *Rhodococcus jostii* (RJ), que se complementaban con alcohol de vainillilo 3 mM.

La figura 5 muestra los niveles de ácido vainílico, vainillina y glucósido de vainillina en cepas de levaduras *N. iowensis* ACAR o *N. crassa* ACAR y de *E. coli* que expresan fosfopanteteinilo transferasa (PPTasa) o PPTasa de *S. pombe*, cuando se cultivan en medio complementado con ácido vainílico 3 mM.

Descripción detallada de la invención

Esta invención se basa en el descubrimiento de que se pueden usar polipéptidos de AROM mutantes para aumentar la acumulación de 3-deshidroshikimato (3-DHS) en hospedantes recombinantes. El 3-DHS es un precursor para la biosíntesis de vainillina, y si está disponible más 3-DHS intracelular, se puede hacer más ácido protocatéquico en la primera etapa realizada de la ruta biosintética de la vainillina. Véase la figura 1. AROM es un complejo enzimático pentafuncional codificado en levaduras por el gen *ARO1*. El gen tiene 4764 pb de longitud y codifica un polipéptido correspondiente de 1588 aminoácidos de longitud. AROM lleva a cabo cinco conversiones enzimáticas consecutivas, es decir, conversión de DAHP (ácido 3-desoxi-D-arabino-heptulosónico 7-fosfato) en 3-DHQ (3-deshidroquinato), que se convierte en 3-DHS (ácido 3-deshidroshikímico), que se convierte en shikimato, que se convierte en shikimato-3-P (shikimato 3-fosfato), que se convierte en EPSP (5-enolpiruvilshikimato 3-fosfato), todas encaminadas a la biosíntesis celular de los aminoácidos aromáticos tirosina, triptófano y fenilalanina (véase la figura 2).

Las cinco funciones catalíticas de AROM residen en cinco dominios distintos del polipéptido que codifica *ARO1*. Los dominios funcionales de los polipéptidos AROM están en un orden diferente del orden en el que son necesarios en la conversión de cinco etapas de DAHP en EPSP (véase la figura 2). Por lo tanto, mientras el dominio 1 corresponde a la actividad catalítica 1, el dominio 2 corresponde a la actividad catalítica 5, el dominio 3 corresponde a la actividad catalítica 4, el dominio 4 corresponde a la actividad catalítica 2 y finalmente el dominio 5 corresponde a la actividad catalítica 3 (véase la figura 1).

Los hospedantes recombinantes que contienen 5 variantes de dominios del polipéptido AROM (p. ej., un polipéptido AROM que tiene una o más mutaciones dentro del dominio 5 o un polipéptido AROM en el que se elimina una parte del dominio 5) pueden tener niveles mayores de 3-DHS en comparación con un hospedante correspondiente que expresa un polipéptido AROM de tipo natural. Dichos hospedantes también pueden tener mayor producción de ácido protocatéquico.

Alternativamente o además, los polipéptidos de COMT mutante se pueden usar para la biosíntesis mejorada de beta-D-glucósido de vainillina. Por ejemplo, los polipéptidos de COMT mutante descritos en la presente memoria pueden tener una o más de las siguientes propiedades: mayor recambio; metilación preferencial en la posición meta (3'), en lugar de en la posición para (4') de modo que está favorecida la producción de vainillina frente a la de isovainillina; o mejor especificidad para los sustratos de la ruta de la vainillina, ácido protocatéquico y aldehído protocatéquico.

Por consiguiente, la invención proporciona polipéptidos de AROM mutantes y ácidos nucleicos que codifican dichos polipéptidos y el uso de los mismos en la biosíntesis de vainillina y/o glucósido de vainillina. El método incluye las etapas de proporcionar un hospedante recombinante capaz de producir vainillina en presencia de una fuente de carbono, en donde dicho hospedante recombinante alberga un ácido nucleico heterólogo que codifica un polipéptido de AROM mutante; cultivar dicho hospedante recombinante en presencia de la fuente de carbono; y purificar la vainillina y/o glucósido de vainillina de dicho hospedante recombinante o del líquido sobrenadante de cultivo.

Los hospedantes recombinantes descritos en la presente memoria se pueden usar en métodos para producir vainillina o glucósido de vainillina. Por ejemplo, si el hospedante recombinante es un microorganismo, el método puede incluir el crecimiento del microorganismo recombinante en un medio de cultivo en condiciones en las que se expresan los genes de la biosíntesis de la vainillina y/o glucósido de vainillina. El microorganismo recombinante se puede cultivar en un procedimiento discontinuo, semicontinuo o continuo o combinaciones de los mismos. Típicamente, el microorganismo recombinante se cultiva en un fermentador a una temperatura(s) definida en

presencia de una fuente de nutrientes adecuada, p. ej., una fuente de carbono, durante un periodo de tiempo deseado para producir una cantidad deseada de vainillina y/o glucósido de vainillina.

Las fuentes de carbono de uso en el presente método incluyen cualquier molécula que pueda ser metabolizada por la célula hospedante recombinante para facilitar el cultivo y/o producción de la vainillina y/o glucósido de vainillina.

5 Los ejemplos de fuentes de carbono adecuadas incluyen sacarosa (p. ej., como se encuentra en melazas), fructosa, xilosa, etanol, glicerol, glucosa, celulosa, almidón, celobiosa u otros polímeros que contienen glucosa. En realizaciones que usan levaduras como un hospedante, por ejemplo, son adecuadas fuentes de carbono tales como sacarosa, fructosa, xilosa, etanol, glicerol y glucosa. La fuente de carbono se puede proporcionar al organismo hospedante a lo largo del periodo de cultivo o, alternativamente, el organismo se puede cultivar durante un periodo de tiempo en presencia de otra fuente de energía, p. ej., proteína, y después proporcionar una fuente de carbono, solo durante la fase semicontinua.

15 En una realización, el microorganismo de este método alberga un ácido nucleico que codifica el polipéptido de AROM mutante y opcionalmente un polipéptido de COMT de tipo natural. En otra realización más, el microorganismo de este método alberga un ácido nucleico que codifica el polipéptido de AROM mutante y opcionalmente un polipéptido de COMT mutante. Dependiendo del microorganismo particular usado en el método, también pueden estar presentes y expresarse otros genes recombinantes que codifican una 3DSD, ACAR, UGT o PPTasa. Por lo tanto, el hospedante recombinante puede, en una realización, ser un microorganismo que alberga uno o más ácidos nucleicos heterólogos que codifican 3DSD, ACAR, UGT o PPTasa o un homólogo funcional de cualquiera de los mencionados que comparte al menos 80%, tal como al menos 90%, por ejemplo al menos 95% de la identidad de secuencia con los mismos. Los niveles de productos, sustratos y productos intermedios, p. ej., ácido deshidroshikímico, ácido protocatéquico, aldehído protocatéquico, vainillina, beta-D-glucósido de vainillina, se pueden determinar extrayendo muestras del medio de cultivo para el análisis de acuerdo con los métodos publicados.

25 Después de cultivar el microorganismo recombinante en cultivo durante el periodo de tiempo deseado, se puede entonces recuperar la vainillina y/o beta-D-glucósido de vainillina del cultivo usando diferentes técnicas conocidas en la técnica, p. ej., aislamiento y purificación por extracción, destilación a vacío y recristalización en múltiples etapas de soluciones acuosas y ultrafiltración (Böddeker, et al. (1997) *J. Membrane Sci.* 137:155-158; Borges da Silva, et al. (2009) *Chem. Eng. Des.* 87:1276-1292). Los procedimientos de extracción de dos fases, que usan compuestos de sulfhidrilo, tales como ditioneitol, ditioneitol, glutatión o L-cisteína (patente de EE.UU. n° 5.128.253), o soluciones alcalinas de KOH (documento WO 94/13614), se han usado en la recuperación de vainillina así como en su separación de otras sustancias aromáticas. La adsorción y pervaporación de vainillina de medios bioconvertidos usando membranas de copolímero de poliéter-poliamida también se ha descrito (Böddeker, et al. (1997) véase antes; Zucchi, et al. (1998) *J. Microbiol. Biotechnol.* 8:719-722). También se han usado resinas de adsorción macroporosas con armazones de poliestireno reticulado para recuperar la vainillina disuelta de soluciones acuosas (Zhang, et al. (2008) *Eur. Food Res. Technol.* 226:377-383). También se han evaluado las técnicas de ultrafiltración y contactores de membrana (MC) para recuperar vainillina (Zabkova, et al. (2007) *J. Membr. Sci.* 301:221-237; Scuibba, et al. (2009) *Desalination* 241:357-364). Alternativamente, se podrían usar técnicas convencionales tales como percolación o extracción con dióxido de carbono supercrítico y ósmosis inversa para la concentración. Si el hospedante recombinante es una planta o células de planta, la vainillina o glucósido de vainillina se pueden extraer del tejido vegetal usando diferentes técnicas conocidas en la técnica.

45 En algunas realizaciones, la vainillina o beta-D-glucósido de vainillina se puede producir usando células enteras que se alimentan con materias primas que contienen moléculas precursoras. Las materias primas se pueden alimentar durante el cultivo celular o después del cultivo celular. Las células enteras pueden estar en suspensión o inmovilizadas. Las células enteras pueden estar en caldo de fermentación o en un tampón de reacción. En algunas realizaciones puede ser necesario un agente de permeabilización para la transferencia eficaz de sustrato a las células.

50 En algunas realizaciones, la vainillina o beta-D-glucósido de vainillina se aísla y purifica hasta homogeneidad (p. ej., al menos 90%, 92%, 94%, 96% o 98% puro). En otras realizaciones, la vainillina o beta-D-glucósido de vainillina se aísla como un extracto de un hospedante recombinante. En relación con esto, la vainillina o beta-D-glucósido de vainillina se puede aislar, pero no necesariamente purificar hasta homogeneidad. De forma conveniente, la cantidad de vainillina o beta-D-glucósido de vainillina producida puede ser de aproximadamente 1 mg/l a aproximadamente 20.000 mg/l o mayor. Por ejemplo, se puede producir de aproximadamente 1 a aproximadamente 100 mg/l, de aproximadamente 30 a aproximadamente 100 mg/l, de aproximadamente 50 a aproximadamente 200 mg/l, de aproximadamente 100 a aproximadamente 500 mg/l, de aproximadamente 100 a aproximadamente 1,000 mg/l, de aproximadamente 250 a aproximadamente 5.000 mg/l, de aproximadamente 1.000 a aproximadamente 15.000 mg/l, o de aproximadamente 2.000 a aproximadamente 10.000 mg/l de vainillina o beta-D-glucósido de vainillina. En general, tiempos de cultivo más largos conducirán a mayores cantidades de producto. Por lo tanto, el microorganismo recombinante se puede cultivar durante de 1 día a 7 días, de 1 día a 5 días, de 3 días a 5 días, aproximadamente 3 días, aproximadamente 4 días o aproximadamente 5 días.

Se apreciará que los diferentes genes y módulos descritos en la presente memoria pueden estar presentes en dos o más microorganismos recombinantes más que en un solo microorganismo. Cuando se usa una pluralidad de microorganismos recombinantes, se pueden cultivar en un cultivo mixto para producir vainillina y/o glucósido de vainillina.

- 5 Los extractos de la vainillina o beta-D-glucósido de vainillina aislados y opcionalmente purificados son útiles en productos de consumo saborizantes tales como productos alimenticios, complementos dietéticos, productos nutracéuticos, composiciones farmacéuticas, composiciones higiénicas dentales y productos cosméticos.

10 La frase "producto alimenticio", como se usa en la presente memoria, incluye frutas, verduras, zumos, productos cárnicos tales como jamón, tocino y salchichas; productos de huevo, concentrados de frutas, gelatinas y productos similares a la gelatina tales como mermeladas, jaleas, conservas; productos lácteos tales como helado, crema agria y el sorbete; glaseados, jarabes incluyendo melazas; productos de maíz, trigo, centeno, soja, avena, arroz y cebada, corazones de nueces y productos de nueces, pasteles, galletas, confitería tales como caramelos, gomas, gotas con sabor a frutas y chocolates, chicles, mentas, cremas, glaseado, helados, tartas y panes, bebidas tales como café, té, refrescos gaseosos, tales como COKE y PEPSI, refrescos sin gas, zumos y otras bebidas de frutas, bebidas para el deporte tales como GATORADE, café, té, té helado, cola, bebidas alcohólicas, tales como cervezas, vinos y licores, y KOOL-AID.

15 Los productos alimenticios también incluyen condimentos tales como hierbas, especias y aderezos, potenciadores del sabor. Un producto alimenticio también incluye productos preparados, tales como edulcorantes dietéticos, edulcorantes líquidos, mezclas de sabores granulados que tras reconstitución con agua proporcionan bebidas no carbonatadas, mezclas de pudín instantáneo, café y té instantáneos, blanqueadores de café, mezclas de leche malteada, alimentos para mascotas, piensos para ganado, tabaco y materiales para aplicaciones de horneado, tales como mezclas para hornear en polvo para la preparación de panes, galletas, pasteles, tortitas, rosquillas. Los productos alimenticios también incluyen alimentos y bebidas de dieta o bajas en calorías que contiene poca o no contiene sacarosa. Otros ejemplos de productos alimenticios previstos de acuerdo con la presente invención se describen más adelante y a lo largo de la memoria descriptiva.

20 En otra realización, los productos alimenticios son frutas, verduras, zumos, productos cárnicos tales como jamón, tocino y salchichas; productos de huevo, concentrados de frutas, gelatinas y productos similares a la gelatina tales como mermeladas, jaleas, conservas; productos lácteos tales como helado, crema agria y el sorbete; glaseados, jarabes incluyendo melazas; productos de maíz, trigo, centeno, soja, avena, arroz y cebada, corazones de nueces y productos de nueces, pasteles, galletas, confitería tales como caramelos, gomas, gotas con sabor a frutas y chocolates, cremas, glaseado, helados, tartas y panes.

25 En otra realización, el producto consumible es una composición farmacéutica. Las composiciones preferidas son composiciones farmacéuticas que contienen vainillina y/o beta-D-glucósido de vainillina y uno o más excipientes farmacéuticamente aceptables. Estas composiciones farmacéuticas se pueden usar para formar fármacos farmacéuticos que contienen uno o más agentes activos que ejercen un efecto biológico. Como tal, la composición farmacéutica preferiblemente incluye además uno o más agentes activos que ejercen un efecto biológico. Dichos agentes activos incluyen agentes farmacéuticos y biológicos que tienen una actividad. Dichos agentes activos son bien conocidos en la técnica. Véase, p. ej., The Physician's Desk Reference. Dichas composiciones se pueden preparar de acuerdo con procedimientos conocidos en la técnica, por ejemplo, como se describe en *Remington's Pharmaceutical Sciences*, Mack Publishing Co., Easton, PA, EE.UU. En una realización, dicho agente activo incluye broncodilatadores, anorexígenos, antihistamínicos, suplementos nutricionales, laxantes, analgésicos, anestésicos, antiácidos, antagonistas del receptor H₂, anticolinérgicos, antidiarreicos, emolientes, antitusivos, antieméticos, antimicrobianos, antibacterianos, antifúngicos, antivirales, expectorantes, antiinflamatorios y mezclas de los mismos. En una realización, el agente activo es un antipirético o analgésico, p. ej., ibuprofeno, acetaminofeno o aspirina; laxantes, p. ej., fenolftaleína-dioctilsulfosuccinato de sodio; calmantes del apetito, p. ej., anfetaminas, fenilpropanolamina, hidrocloreto de fenilpropanolamina, o cafeína; antiácidos, p. ej., carbonato de calcio; antiasmáticos, p. ej., teofilina; antiurético, p. ej., hidrocloreto de difenoxilato; agentes activos contra la flatulencia, p. ej., simeticona; agentes para migraña, p. ej., tartrato de ergotamina; agentes psicofarmacológicos, p. ej., haloperidol; espasmolíticos o sedantes, p. ej., fenobarbitol; antihiperkinéticos, p. ej., metildopa o metilfenidato; tranquilizantes, p. ej., benzodiazepinas, hidroxinmeprobromatos o fenotiazinas; antihistamínicos, p. ej., astemizol, maleato de clorfeniramina, maleato de piridamina, succinato de doxlamina, maleato de bromofeniramina, citrato de feniltoloxamina, hidrocloreto de clorociclizina, maleato de feniramina, y tartrato de fenindamina; descongestionantes, p. ej., hidrocloreto de fenilpropanolamina, hidrocloreto de fenilefrina, hidrocloreto de pseudoefedrina, sulfato de pseudoefedrina, bitartrato de fenilpropanolamina y efedrina; bloqueantes de receptores beta, p. ej., propranolol; agentes para síndrome de abstinencia del alcohol, p. ej., disulfiram; antitusivos, p. ej., benzocaína, dextrometorfano, hidrobromuro de dextrometorfano, noscapina, citrato de carbapentano, e hidrocloreto de clofedianol; suplementos de flúor, p. ej., fluoruro sódico; antibióticos locales, p. ej., tetraciclina o cleocina; suplementos de corticosteroides, p. ej., prednisona o prednisolona; agentes contra la formación de bocio, p. ej., colchicina o alopurinol; antiépilépticos, p. ej., fenitoína de sodio; agentes contra la deshidratación, p. ej., suplementos de electrolitos; antisépticos, p. ej., cloruro de cetilpiridinio; AINE, p. ej., acetaminofeno, ibuprofeno, naproxeno, o sus sales; agentes activos gastrointestinales, p. ej., loperamida y famotidina; diferentes alcaloides, p. ej., fosfato de codeína, sulfato de codeína

o morfina; suplementos para oligoelementos, p. ej., cloruro de sodio, cloruro de zinc, carbonato de calcio, óxido de magnesio y otras sales de metales alcalinos y sales de metales alcalinotérreos; vitaminas; resinas de intercambio iónico, p. ej., colestiramina; sustancias reductoras del colesterol y que disminuyen los lípidos; antiarrítmicos, p. ej., N-acetilprocainamida; y expectorantes, p. ej., guaifenesina.

- 5 Sustancias activas que tienen un sabor particularmente desagradable incluyen agentes antibacterianos tales como ciprofloxacina, ofloxacina y pefloxacina; antiepilépticos tales como zonisamida; antibióticos macrólidos tales como eritromicina; antibióticos beta-lactámicos tales como penicilinas y cefalosporinas; sustancias activas psicotrópicas tales como clorpromazina; sustancias activas tales como sulpirina; y agentes activos contra las úlceras, tales como cimetidina.
- 10 Las composiciones farmacéuticas se administran a un sujeto en cualquier forma adecuada para lograr su propósito previsto. Sin embargo, preferiblemente la composición es una que se puede administrar por vía bucal u oral. Alternativamente, la composición farmacéutica puede ser un pulverizador oral o nasal. El sujeto es cualquier animal, tal como un ser humano. Otros animales adecuados incluyen caninos, felinos, perros, gatos, ganado, caballos, vacas, ovejas. Como se usa en la presente memoria, una composición veterinaria se refiere a una composición farmacéutica que es adecuada para animales no humanos. Dichas composiciones veterinarias son conocidas en la técnica.

En otra realización, la composición es una forma farmacéutica líquida para administración oral, que incluye emulsiones, soluciones, suspensiones, jarabes y elixires farmacéuticamente aceptables. Además de los compuestos activos, las formas farmacéuticas líquidas pueden contener diluyentes inertes usados normalmente en la técnica tales como, por ejemplo, agua u otros disolventes, agentes solubilizantes y emulsionantes tales como alcohol etílico, alcohol isopropílico, carbonato de etilo, acetato de etilo, alcohol bencilico, benzoato de bencilo, propilenglicol, 1,3-butilenglicol, dimetilformamida, aceites (en particular, aceites de semilla de algodón, cacahuete, maíz, germen, oliva, ricino y sésamo), glicerol, alcohol tetrahidrofurfurilo, polietilenglicoles y ésteres de ácido graso y sorbitán, y sus mezclas. Las suspensiones, además de los compuestos activos, pueden contener agentes de suspensión, por ejemplo, alcoholes de isoestearilo etoxilado, ésteres de sorbitol y sorbitán polioxietilénico, celulosa microcristalina, metahidróxido de aluminio, bentonita, agar-agar, y tragacanto y sus mezclas.

La composición farmacéutica puede estar en forma de un comprimido masticable. Los comprimidos masticables son conocidos en la técnica. Véanse, p. ej., las patentes de EE.UU. nº 4.684.534 y 6.060.078. Un comprimido masticable puede contener cualquier tipo de medicamento, preferiblemente un medicamento de sabor amargo, extractos de plantas naturales y otros compuestos orgánicos. Más preferiblemente, pueden estar contenidos en el núcleo vitaminas tales como vitamina A, vitamina B, vitamina B1, vitamina B2, vitamina B6, vitamina C, vitamina E y vitamina K; extractos de plantas naturales tales como extractos de sohgunjung-tang, extractos de Sipchundaebotang y extractos de Eleutherococcus senticosus; compuestos orgánicos tales como dimenhidrinato, meclazina, acetaminofeno, aspirina, fenilpropanolamina y cloruro de cetilpiridinio; o agentes gastrointestinales como el gel de hidróxido de aluminio seco, domperidona, azuleno soluble, L-glutamina e hidrotalcita.

La composición farmacéutica se puede disgregar por vía oral. Los comprimidos de disgregación oral son conocidos en la técnica. Véanse, p. ej., las patentes de EE.UU. nº 6.368.625 y 6.316.029.

La composición farmacéutica puede ser una forma farmacéutica sólida, que incluye vainillina o beta-D-glucósido de vainillina y un gránulo efervescente activado por agua y/o saliva, tal como uno que tiene una velocidad de efervescencia controlable. La composición efervescente puede comprender además un compuesto farmacéuticamente activo. Las composiciones farmacéuticas efervescentes son conocidas en la técnica. Véase, p. ej., la patente de EE.UU. nº 6.649.186. La composición efervescente se puede usar en aplicaciones farmacéuticas, veterinarias, hortícolas, domésticas, de alimentos, culinarias, plaguicidas, agrícolas, cosméticas, herbicidas, industriales, de limpieza, confitería y saborizantes. Las formulaciones que incorporan la composición efervescente que contiene vainillina o beta-D-glucósido de vainillina puede incluir además uno o más adyuvantes y/o ingredientes activos adicionales que se pueden elegir de los conocidos en la técnica, que incluyen sabores, diluyentes, colorantes, aglutinantes, carga, tensioactivo, disgregante, estabilizante, vehículos de compactación y disgregantes no efervescentes.

La composición farmacéutica puede ser una composición farmacéutica en forma de película o en forma de oblea. Dicha composición farmacéutica en forma de película o en forma de oblea se puede configurar, por ejemplo, como formas de administración de disgregación rápida, p. ej., formas de administración que se disgregan en un periodo de 1 segundo hasta 3 minutos, o como formas de administración que se disgregan lentamente, p. ej., formas de administración que se disgregan en un periodo de 3 a 15 minutos. Los tiempos de disgregación indicados se pueden ajustar a los intervalos mencionados antes usando, por ejemplo, polímeros formadores de matrices que tienen diferentes características de disgregación o solubilidad. Por lo tanto, mezclando los componentes poliméricos correspondientes, se puede ajustar el tiempo de disgregación. Además, se conocen disgregantes que "arrastran" agua dentro de la matriz y hacen que la matriz de rompa y abra desde dentro. Como consecuencia, algunas realizaciones de la invención incluyen dichos disgregantes con el fin de ajustar el tiempo de disgregación.

Los polímeros que son adecuados para usar en la composición farmacéutica en forma de película o en forma de oblea incluyen derivados de celulosa, poli(alcohol vinílico) (p. ej., MOWIOL), poliacrilatos, polivinilpirrolidona, éteres de celulosa, tales como etilcelulosa, así como poli(alcohol vinílico), poliuretano, polimetacrilatos, poli(metacrilatos de metilo) y derivados y copolimerizados de los polímeros mencionados antes.

- 5 En algunas realizaciones, el espesor total de la composición farmacéutica en forma de película o en forma de oblea es preferiblemente de 5 µm hasta 10 mm, preferiblemente de 30 µm a 2 mm, y con preferencia particular de 0,1 mm a 1 mm. Las preparaciones farmacéuticas pueden ser de forma redonda, ovalada, elíptica, triangular, cuadrangular o poligonal, pero también pueden tener cualquier forma redondeada.

- 10 La composición farmacéutica puede estar en forma de un aerosol. La composición en aerosol puede incluir además un agente farmacéuticamente activo. Las composiciones en aerosol son conocidas en la técnica. Véase, p. ej., la patente de EE.UU. nº 5.011.678. Como un ejemplo no limitante, una composición en aerosol puede incluir una cantidad médicamente eficaz de una sustancia farmacéuticamente activa, vainillina o beta-D-glucósido de vainillina y un propulsor biocompatible, tal como un propulsor de (hidro/fluoro)carbono.

- 15 En una realización, la composición farmacéutica es una composición nutricional. Los ejemplos de composiciones nutricionales que tienen un sabor indeseable incluyen productos de nutrición enteral para el tratamiento de la deficiencia nutricional, traumatismo, cirugía, enfermedad de Crohn, enfermedad renal, hipertensión, obesidad, para promover el rendimiento atlético, potenciación muscular o bienestar general o errores congénitos del metabolismo tales como fenilcetonuria. En particular, dichas formulaciones nutricionales pueden contener uno o más aminoácidos que tienen un sabor o regusto amargo o metálico. Dichos aminoácidos incluyen aminoácidos esenciales tales como un isómero L de la leucina, isoleucina, histidina, lisina, metionina, fenilalanina, treonina, triptófano, tirosina o valina.

- 20 En una realización, el producto consumible es una composición para la higiene dental que contiene vainillina y/o beta-D-glucósido de vainillina. Las composiciones para la higiene dental son conocidas en la técnica e incluyen pasta de dientes, enjuagues bucales, enjuague de placa, hilo dental, analgésicos dentales (tales como ANBESOL).

- 25 En otra realización, el producto consumible es un producto cosmético que contiene vainillina y/o beta-D-glucósido de vainillina. Por ejemplo, el producto cosmético puede ser una crema facial, barra de labios, brillo de labios. Otras composiciones adecuadas incluyen bálsamo de labios, tales como bálsamo de labios CHAPSTICK o BURT'S BEESWAX, que contienen además vainillina y/o beta-D-glucósido de vainillina.

Polipéptidos de enzima multifuncional arom (AROM)

- 30 Los ejemplos no limitantes de polipéptidos de AROM incluyen el polipéptido de *Saccharomyces cerevisiae* que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 4 (nº de acceso en GENBANK X06077); un polipéptido de *Schizosaccharomyces pombe* (nº de acceso en GENBANK NP_594681.1); un polipéptido de *Schizosaccharomyces japonicas* (nº de acceso en GENBANK XP_002171624); un polipéptido de *Neurospora crassa* (nº de acceso en GENBANK XP_956000); y un polipéptido de *Yarrowia lipolytica* (nº de acceso en GENBANK XP_505337).

- 35 La expresión "polipéptido de AROM" como se usa en la presente memoria se refiere a cualquier secuencia de aminoácidos que es al menos 80 por ciento (p. ej., al menos 85, 90, 95, 96, 97, 98, 99 o 100 por cien) idéntica a la secuencia expuesta en la SEQ ID NO: 4, y tiene actividad de 3-deshidroquinato deshidratasa, actividad de 3-deshidroquinato sintasa, actividad de 3-fosfoshikimato 1-carboxiviniltransferasa, actividad de shikimato 3-deshidrogenasa (NADP+) y actividad de shikimato quinasa.

- 40 De acuerdo con esta invención el polipéptido de AROM es un polipéptido de AROM mutante con actividad reducida de shikimato deshidrogenasa. Cuando se expresa en un hospedante recombinante, el polipéptido de AROM mutante dirige el flujo metabólico de la producción de aminoácidos a la producción de precursor de vainillina (figura 2). La actividad reducida de la shikimato deshidrogenasa se puede deducir de la acumulación de ácido deshidroshikímico en un hospedante recombinante que expresa un polipéptido de AROM mutante, usando LC-MS.

- 45 El polipéptido de AROM mutante descrito en la presente memoria puede tener una o más modificaciones en el dominio 5 (p. ej., una sustitución de uno o más aminoácidos. Por ejemplo, un polipéptido de AROM mutante puede tener una eliminación del dominio 5 entero, es decir, aminoácidos de 1305 a 1588 de la secuencia de aminoácidos en la SEQ ID NO: 4, de modo que el polipéptido de AROM mutante tiene actividad reducida de shikimato deshidrogenasa. Un polipéptido de AROM mutante de ejemplo que carece del dominio 5 se proporciona en la SEQ ID NO: 2. Los polipéptidos de AROM mutantes con al menos una sustitución de aminoácido en el dominio 5 incluyen los polipéptidos de AROM A1533P, P1500K, R1458W, V1349G, T1366G, I1387H, W1571V, T1392K, K1370L y A1441P como se exponen en la SEQ ID NO: 6, SEQ ID NO: 8, SEQ ID NO: 10, SEQ ID NO: 12, SEQ ID NO: 14, SEQ ID NO: 16, SEQ ID NO: 18, SEQ ID NO: 20, SEQ ID NO: 22 y SEQ ID NO: 24, respectivamente.

- 55 Las sustituciones de aminoácidos que son particularmente útiles se pueden encontrar, por ejemplo, en una o más posiciones que se alinean con la posición 1349, 1366, 1370, 1387, 1392, 1441, 1458, 1500, 1533 o 1571 de la secuencia de aminoácidos expuesta en la SEQ ID NO:4. Por ejemplo, un polipéptido de AROM modificado puede

tener una sustitución en una posición que se alinea con la posición 1370 o la posición 1392 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4.

Por ejemplo, un polipéptido de AROM modificado puede tener uno o más de los siguientes: un aminoácido distinto de valina (p. ej., una glicina) en una posición que se alinea con la posición 1349 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; un aminoácido distinto de treonina (p. ej., una glicina) en una posición que se alinea con la posición 1366 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; un aminoácido distinto de lisina (p. ej., leucina) en una posición que se alinea con la posición 1370 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; un aminoácido distinto de isoleucina (p. ej., histidina) en una posición que se alinea con la posición 1387 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; un aminoácido distinto de treonina (p. ej., lisina) en una posición que se alinea con la posición 1392 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; un aminoácido distinto de alanina (p. ej., prolina) en una posición que se alinea con la posición 1441 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; un aminoácido distinto de arginina (p. ej., triptófano) en una posición que se alinea con la posición 1458 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; un aminoácido distinto de prolina (p. ej., lisina) en una posición que se alinea con la posición 1500 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; un aminoácido distinto de alanina (p. ej., prolina) en una posición que se alinea con la posición 1533 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; o un aminoácido distinto de triptófano (p. ej., valina) en una posición que se alinea con la posición 1571 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4.

Polipéptidos de catecol-O-metil-transferasa (COMT)

El polipéptido de COMT de acuerdo con la invención, en algunas realizaciones puede ser una catecol-O-metiltransferasa. En otras realizaciones el polipéptido de COMT preferiblemente es una catecol-O-metiltransferasa. Más preferiblemente, un polipéptido de COMT de la invención es un polipéptido (COMT) mutante que tiene una metilación del hidroxilo en meta del aldehído protocatóquico, ácido protocatóquico y/o alcohol protocatóquico con respecto a la COMT de *Homo sapiens* que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 27.

Los ejemplos no limitantes de polipéptidos de COMT que se pueden mutar incluyen polipéptidos de COMT en la familia clasificada con el número EC 2.1.1.6, tal como el polipéptido de *Homo sapiens* (Hs) que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 27 (véase, también, n° de acceso en GENBANK NM_000754); un polipéptido de *Arabidopsis thaliana* que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 53 (n° de acceso en GENBANK AY062837); o un polipéptido de *Fragaria x ananassa* (fresa) que tiene la secuencia de aminoácidos expuesta en la SEQ ID NO: 54 (n° de acceso en GENBANK AF220491). El polipéptido de COMT humano existe en forma de varias variantes y el polipéptido de COMT puede ser cualquiera de esas variantes, sin embargo, en una realización preferida el polipéptido de COMT humano es de SEQ ID NO: 27 o SEQ ID NO: 55. Otros polipéptidos de COMT de mamífero adecuados útiles en esta invención incluyen los aislados de *Pan troglodytes* (n° de acceso en GENBANK XP_514984), *Macaca mulatta* (n° de acceso en GENBANK AFJ70145), *Equus caballus* (n° de acceso en GENBANK NP_001075303), *Canis lupus familiaris* (n° de acceso en GENBANK AAR20324), *Cricetulus griseus* (n° de acceso en GENBANK EGV97595), *Sus scrofa* (n° de acceso en GENBANK NP_001182259), y *Bos taurus* (n° de acceso en GENBANK NP_001095787). Otros polipéptidos de COMT de fuentes vegetales y microorganismos incluyen los aislados de *Rosa chinensis* (n° de acceso en GENBANK CAD29457), *Prunus dulcis* (n° de acceso en GENBANK CAA58218), *Gossypium hirsutum* (n° de acceso en GENBANK ACT32028), *Jatropha curcas* (n° de acceso en GENBANK ACT87981), *Eucalyptus camaldulensis* (ADB82906), *Candida orthopsilosis* (n° de acceso en GENBANK CCG25047), *Pichia stipitis* (n° de acceso en GENBANK ABN67921), y *Spathaspora passalidarum* (n° de acceso en GENBANK EGW29958). En ciertas realizaciones, el polipéptido de COMT se obtiene de *Phytophthora infestans* (n° de acceso en GENBANK XP_002899214), *Catharanthus roseus* (n° de acceso en GENBANK EGS21863), *Yarrowia lipolytica* (n° de acceso en GENBANK XP_500451), *Ciona intestinalis* (n° de acceso en GENBANK XP_002121420 o XP_002131313), *Capsasproa owczarzewski* (n° de acceso en GENBANK EFW46044), *Chaetomium therophilum* (n° de acceso en GENBANK EGS21863), *Clavisporea lusitanae* (n° de acceso en GENBANK XP_002899214), *Paracoccidioides* sp. 'lutzii' Pb01 (n° de acceso en GENBANK XP_002793380), *Vanilla planifolia* (SEQ ID NO:56), *Coffea Arabica* (n° de acceso en GENBANK AAN03726), *Rattus norvegicus* (n° de acceso en GENBANK NP_036663), *Mus musculus* (n° de acceso en GENBANK NP_031770), *Crenarchaeote* (n° de acceso en GENBANK ABZ07345), *Mycobacterium vanbaleeni* (n° de acceso en GENBANK ABM14078), o *Schizosaccharomyces pombe* (n° de acceso en GENBANK NP_001018770, que se ha mostrado que presenta la actividad de COMT deseada (figuras 3B y 3C).

La expresión "polipéptido de COMT" como se usa en la presente memoria se refiere a cualquier secuencia de aminoácidos que es al menos 80 por ciento (p. ej., al menos 85, 90, 95, 96, 97, 98, 99 o 100 por cien) idéntica a la secuencia de Hs COMT expuesta en la SEQ ID NO: 27 y tiene actividades enzimáticas de catecol-O-metiltransferasa del polipéptido de Hs COMT de tipo natural.

En una realización, la expresión "polipéptido de COMT mutante", como se usa en la presente memoria, se refiere a cualquier polipéptido que tiene una secuencia de aminoácidos que es al menos 80 por ciento, tal como al menos 85 por ciento, por ejemplo al menos 90 por ciento, tal como al menos 95 por ciento, por ejemplo al menos 96 por ciento, tal como al menos 97 por ciento, por ejemplo al menos 98 por ciento, tal como al menos 99 por ciento, idéntica a la

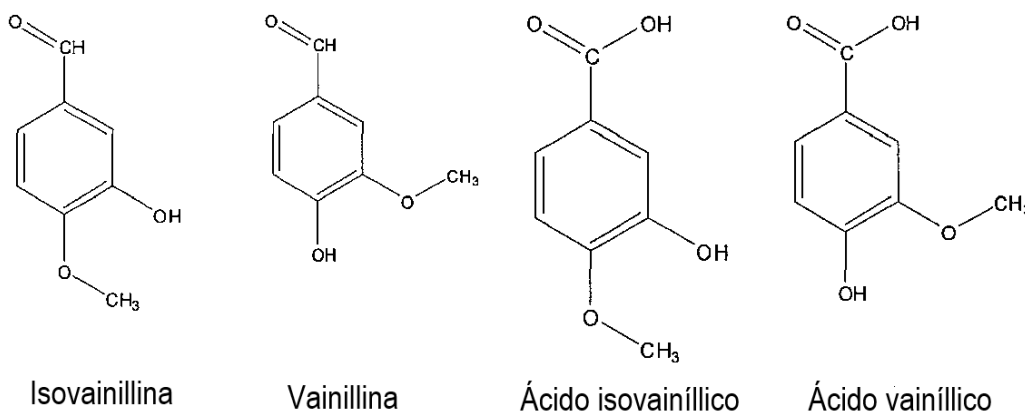
secuencia de Hs COMT expuesta en la SEQ ID NO: 27, y es capaz de catalizar la metilación del grupo -OH en la posición meta del ácido protocatéquico y/o aldehído protocatéquico, en donde la secuencia de aminoácidos de dicho polipéptido de COMT mutante difiere de la SEQ ID NO: 27 en al menos un aminoácido.

5 Los polipéptidos de COMT mutantes descritos en la presente memoria pueden tener una o más mutaciones (p. ej., una sustitución de uno o más aminoácidos, una eliminación de uno o más aminoácidos, inserciones de uno o más aminoácidos, o combinaciones de sustituciones, eliminaciones e inserciones), por ejemplo, en el sitio de unión del sustrato. Por ejemplo, un polipéptido de COMT mutante puede tener una o más sustituciones de aminoácidos en el sitio de unión del sustrato de la COMT humana.

10 En ciertas realizaciones, un "polipéptido de COMT mutante" difiere de la SEQ ID NO: 27 en solo uno o dos restos de aminoácidos, en donde las diferencias entre dichas proteínas mutante y de tipo natural están en el sitio de unión del sustrato.

15 Como se describe en la presente memoria, los polipéptidos de COMT mutante se pueden usar para mejorar la biosíntesis del glucósido de vainillina. Por ejemplo, los polipéptidos de COMT pueden tener una o más de las siguientes propiedades: mayor recambio; metilación preferencial en la posición meta (3'), más que en la posición para (4') de modo que está favorecida la producción de vainillina frente a la isovainillina; o mejor especificidad por los sustratos de la ruta de la vainillina, el ácido protocatéquico y aldehído protocatéquico. Los polipéptidos de COMT mutantes se pueden caracterizar por ensayos de metilación in vitro o se pueden caracterizar in vivo en un hospedante recombinante basándose en la producción de ácido vainílico, vainillina o glucósido de vainillina.

Las estructuras de isovainillina, vainillina, ácido isovainílico y ácido vainílico son las siguientes.



25 La Hs COMT de tipo natural carece de O-metilación regioselectiva del aldehído protocatéquico y el ácido protocatéquico, indicando que el sitio de unión de la Hs COMT no se une a estos sustratos en una orientación que permita la metilación regioselectiva deseada. Sin querer estar limitados por un mecanismo particular, el sitio activo de la Hs COMT está compuesto de la coenzima S-adenosil-metionina (SAM), que sirve como el donador de metilo, y el sustrato catecol, que contiene el hidroxilo que se va a metilar coordinado al Mg^{2+} y próximo a la Lys144. La O-metilación procede por un mecanismo de SN2, donde la Lys144 sirve como una base catalítica que desprotona el hidroxilo proximal para formar un oxianión que ataca un grupo metilo del sulfonio de la SAM. Véase, por ejemplo, Zheng & Bruice (1997) *J. Am. Chem. Soc.* 119(35):8137-8145; Kuhn & Kollman (2000) *J. Am. Chem. Soc.* 122(11):2586-2596; Roca, et al. (2003) *J. Am. Chem. Soc.* 125(25):7726-37.

30 En una realización un polipéptido de COMT mutante es capaz de catalizar la metilación de un grupo -OH del ácido protocatéquico, en donde dicha metilación da como resultado la generación de al menos 4 veces más ácido vainílico en comparación con el ácido isovainílico, preferiblemente al menos 5 veces más ácido vainílico en comparación con el ácido isovainílico, tal como al menos 10 veces más ácido vainílico en comparación con el ácido isovainílico, por ejemplo al menos 15 veces más ácido vainílico en comparación con el ácido isovainílico, tal como al menos 20 veces más ácido vainílico en comparación con el ácido isovainílico, por ejemplo al menos 25 veces más ácido vainílico en comparación con el ácido isovainílico, tal como al menos 30 veces más ácido vainílico en comparación con el ácido isovainílico; y tiene una secuencia de aminoácidos que difiere de la SEQ ID NO: 27 en al menos un aminoácido.

40 Además de las propiedades mencionadas antes, se prefiere además que un polipéptido de COMT mutante sea capaz de catalizar la metilación de un grupo -OH del ácido protocatéquico, en donde dicha metilación da como resultado la generación de al menos 4, 5, 10, 15, 20, 25 o 30 veces más vainillina en comparación con la isovainillina; y/o sea capaz de catalizar la metilación de un grupo -OH del alcohol protocatéquico, en donde dicha metilación da como resultado la generación de al menos 4, 5, 10, 15, 20, 25 o 30 veces más alcohol de vainillilo en comparación con el alcohol de isovainillilo.

Para determinar si un polipéptido de COMT mutante dado es capaz de catalizar la metilación de un grupo -OH del ácido protocatéutico, en donde dicha metilación da como resultado la generación de al menos X veces más ácido vainílico en comparación con el ácido isovainílico, se puede llevar a cabo un ensayo in vitro. En dicho ensayo, el ácido protocatéutico se incuba con un polipéptido de COMT mutante en presencia de un donador de metilo y posteriormente se determina el nivel de ácido isovainílico y ácido vainílico generado. Dicho donador de metilo puede ser, por ejemplo, S-adenosilmetionina. Más preferiblemente, esto se puede determinar generando un hospedante recombinante que alberga un ácido nucleico heterólogo que codifica el polipéptido de COMT mutante, en donde dicho hospedante recombinante además es capaz de producir ácido protocatéutico. Después de cultivar el hospedante recombinante, se puede determinar el nivel de ácido isovainílico y ácido vainílico generado. En relación con este método, se prefiere que dicho ácido nucleico heterólogo que codifica el polipéptido de COMT mutante que se va a ensayar esté operativamente unido a una región reguladora que permite la expresión en dicho hospedante recombinante. Además, se prefiere que el hospedante recombinante exprese al menos una 3DSD y al menos una ACAR, que pueden ser preferiblemente una de las 3DSD y ACAR descritas en la presente memoria. En realizaciones donde el hospedante recombinante expresa una ACAR capaz de catalizar la conversión de ácido vainílico en vainillina, entonces el método puede incluir también determinar el nivel de vainillina e isovainillina generado. El hospedante recombinante también puede expresar al menos una UGT capaz de catalizar la glucosilación de la vainillina e isovainillina, en cuyo caso se pueden determinar en su lugar los niveles de glucósido de vainillina y glucósido de isovainillina, respectivamente. Alternativamente, esto se puede determinar generando un hospedante recombinante que alberga un ácido nucleico heterólogo que codifica el polipéptido de COMT mutante que se va a ensayar, y alimentando ácido protocatéutico a dicho hospedante recombinante, seguido de la determinación del nivel generado de ácido vainílico y ácido isovainílico.

De forma similar, se puede usar un ensayo in vitro o una célula de hospedante recombinante para determinar si un polipéptido de COMT mutante es capaz de catalizar la metilación de un grupo -OH del aldehído protocatéutico, en donde dicha metilación da como resultado la generación de al menos X veces más vainillina en comparación con la isovainillina. Sin embargo, en este ensayo, se usa aldehído protocatéutico como material de partida y se determina el nivel de vainillina e isovainillina.

Igualmente, se puede usar un ensayo in vitro de una célula de hospedante recombinante para determinar si un polipéptido de COMT mutante es capaz de catalizar la metilación de un grupo -OH del alcohol protocatéutico, en donde dicha metilación da como resultado la generación de al menos X veces más alcohol de vainillilo en comparación con el alcohol de vainillilo. Sin embargo, en este ensayo, se usa aldehído protocatéutico como material de partida y se determina el nivel de vainillina e isovainillina.

El nivel de isovainillina y vainillina se pueden determinar por cualquier método adecuado útil para detectar estos compuestos, en donde dicho método puede distinguir entre la isovainillina y la vainillina. Dichos métodos incluyen, por ejemplo, HPLC. De forma similar, el nivel de ácido isovainílico, ácido vainílico, alcohol de isovainillilo y alcohol de vainillilo, se puede determinar usando cualquier método adecuado útil para detectar estos compuestos, en donde dicho método puede distinguir entre isovainillina y la vainillina. Dichos métodos incluyen, por ejemplo, HPLC.

Para sustratos de las dimensiones del aldehído protocatéutico y ácido protocatéutico, se encontró que el límite del sitio de unión del sustrato de la Hs COMT estaba formado por los siguientes restos hidrófobos: Trp38, Met40, Cys173, Pro174, Trp143 y Leu198. Un resto hidrófilo adicional que puede influir en la unión es Arg201. El hidroxilo del catecol que no es metilado, forma enlace de hidrógeno con la Glu199.

De acuerdo con este mecanismo, para que se produzca la metilación en la posición meta del aldehído protocatéutico y el ácido protocatéutico, estos sustratos deben unirse en una orientación que coloque el hidroxilo en meta de modo que esté coordinado con el Mg^{2+} y próximo a la Lys144, mientras que el hidroxilo en para está próximo a la Glu199. La falta de esta regioselectividad deseada observada en la Hs COMT de tipo natural sugiere que esta orientación de unión no ocurre de forma preferente. Uno o más aminoácidos pueden sustituir los restos del sitio de unión de la Hs COMT para permitir una orientación de unión de los sustratos que promueva la O-metilación en meta deseada del aldehído protocatéutico y el ácido protocatéutico.

Por lo tanto, un método para identificar los polipéptidos de COMT con especificidad por el sustrato mejorada proporciona metodología computacional para identificar mutaciones de restos que confieran O-metilación en meta mejorada por un polipéptido de COMT. El método incluye varias etapas diferentes. En particular, un método de identificación de mutaciones óptimas incluye las etapas de (a) seleccionar una estructura de proteína de un polipéptido de COMT; (b) acoplamiento de los sustratos ácido protocatéutico y aldehído protocatéutico a la estructura de proteína del polipéptido de COMT determinada en (a) para reducir las diferentes conformaciones que promueven la O-metilación en meta o para regioselectiva; (c) identificar mutaciones del sitio de unión próximas a los sustratos ácido protocatéutico y aldehído protocatéutico; (d) llevar a cabo análisis mutacional in silico en cada resto identificado en (c); (e) clasificar las mutaciones de restos candidatos para cada posición de (d) basado en conformaciones previstas del sustrato de la O-metilación en meta o para; y (f) seleccionar las mutaciones con mejor puntuación para cada resto candidato identificado en (c).

Como se ha descrito antes, el polipéptido de COMT puede estar en la familia de COMT clasificado con el número EC 2.1.1.6, catecol O-metil transferasa. Así pues, un experto en la técnica apreciará que, además de la Hs COMT, el método se puede aplicar a cualquier especie de COMT dentro de esta clasificación, en donde dichas proteínas tendrán restos del sitio de unión similares. Aunque se describe cada una de las etapas del método con más detalle con respecto a la Hs COMT, se apreciará que se pueden llevar a cabo etapas de métodos similares con otros polipéptidos de COMT.

En la etapa (a), se selecciona una estructura de proteína de COMT. Las estructuras de proteínas de Hs COMT están públicamente disponibles en el Banco de Datos de Proteínas y se puede evaluar la utilidad basándose en la resolución, inclusión de mutaciones y otras variaciones de secuencias que se pueden introducir para dar la cristalización. Otros factores que se pueden considerar cuando se selecciona una estructura incluyen si la estructura incluye o no un sustrato unido a la proteína, o la naturaleza del sustrato unido a la proteína. El código de estructura cristalina 3BWM (Banco de Datos de Proteínas RCSB) es una estructura particularmente útil para la Hs COMT, y se usó en los métodos descritos en la presente memoria. Un experto en la técnica apreciará que se puede usar otra estructura cristalina de Hs COMT como entradas para los procedimientos de modelización.

En la etapa (b), los sustratos de interés, p. ej., el ácido protocatéutico y aldehído protocatéutico, se acoplan a la estructura de proteína. Acoplamiento es un término para indicar un procedimiento computacional implementado en una variedad de algoritmos de predicción de una conformación probable de una molécula pequeña seleccionada en un sitio de unión de proteína. La técnica típicamente calcula una puntuación de la unión para proporcionar la base para evaluar la bondad de ajuste y la energía de unión prevista de la interacción. El programa de acoplamiento usado para obtener las conformaciones del sustrato usadas en el procedimiento actual es ProtoScreen (Haydon, et al. (2008) *Science* 321:1673-1675). El método crea puntuaciones de unión de acuerdo con el algoritmo de ProtoScore (Bhurruth-Alcor, et al. (2011) *Org. Biomol. Chem.* 9:1169-1188). Un experto en la técnica apreciará que se pueden usar otros programas de acoplamiento en esta etapa del método para identificar conformaciones de unión adecuadas de los sustratos ácido protocatéutico y aldehído protocatéutico en la Hs COMT (u otro miembro de la familia de COMT).

En la etapa (c) se lleva a cabo el análisis mutacional. Este puede incluir varias subetapas tales como (i) identificar un primer resto que se va a mutar; (ii) identificar un resto con el que se va a mutar de la lista de aminoácidos adecuados; (iii) para cada resto en (ii), buscar una biblioteca de rotámeros para los candidatos conformacionales; (iv) mutar el resto de (i) con cada una de las nuevas selecciones de rotámeros de la cadena lateral de restos de (iii); (v) minimizar la conformación del complejo de proteína; (vi) puntuar cada candidato a rotámero en (iv) con diferentes cálculos de energías de mutante-sustrato, mutante-proteína, mutante-disolvente y sustrato-disolvente con diferentes conformaciones del ácido o aldehído protocatéutico que permitan comparaciones de energía donde el sustrato se esté modificando en la posición meta o en la para; y (vii) clasificar las selecciones de rotámeros de (iii) y después dar la puntuación más alta a los mutantes de aminoácidos de (ii).

En la subetapa (i), se analiza a su vez cada uno de los restos que se va a mutar.

En la subetapa (ii), se selecciona un resto que se va a mutar de una lista determinística de aminoácidos disponibles. Por defecto esta lista son los 20 aminoácidos estándar que se encuentran en la naturaleza, pero podría incluir otros aminoácidos no naturales.

En la subetapa (iii), todos los rotámeros se identifican en una biblioteca de rotámeros que se corresponden con la identidad del aminoácido mutante. Una biblioteca de rotámeros es un conjunto precalculado de conformaciones preferidas de cadenas laterales de aminoácidos estándar en un formato 3D que se puede usar. Dichas bibliotecas se usan habitualmente en el trabajo de análisis estructural de proteínas, y están incluidas en la mayoría de los paquetes de modelización molecular disponibles en el mercado. Los conjuntos de rotámeros que se corresponden con la identidad del resto que muta se seleccionan para la subetapa (iv).

En la subetapa (iv), el resto de proteína seleccionado se intercambia por cada uno de los rotámeros identificados en (iii). Esto implica manipular la representación computacional de los átomos del resto de proteína de modo que se eliminen los átomos de la cadena lateral del resto de inicio antes de que el rotámero seleccionado entrante de (iii) se conecte a la posición del carbono alfa usando matemáticas vectoriales. Este método se repite para cada rotámero candidato para llegar a una lista de representaciones 3D de la proteína que varía cada una solo en las diferentes conformaciones del rotámero en la posición de un solo resto.

En la subetapa (v), los complejos de proteína-sustrato obtenidos en (iv) se someten a minimizaciones de campos de fuerza. En una realización particularmente útil, este campo de fuerza es AMBER99 con cargas AM1-BCC aplicadas al sustrato. En otro aspecto del método, se usa un término de solvatación de Born. La cadena principal de proteína se puede conectar usando restricciones de pared, mientras que las cadenas laterales permanecen sin restringir. Esto tiene el efecto de reducir el movimiento de proteína general, pero permitiendo explorar el efecto de la mutación de restos individuales en restos vecinos. Un experto en la técnica apreciará que varios paquetes de modelización molecular disponibles en el mercado son capaces de realizar estas tareas.

En la subetapa (vi), las conformaciones de proteína-complejo resultantes se someten a cálculos energéticos para determinar la viabilidad de las conformaciones mutantes individuales. Esto incluye el cálculo individual de la energía de interacción del resto de tipo natural con el sustrato, entorno de la proteína y disolvente y después llevar a cabo los mismos cálculos con la conformación de restos mutados. Los cálculos se determinan para ambas conformaciones (reacción en meta y para) del sustrato. Esta etapa del método identifica por lo tanto mutaciones que tienen energías de unión favorables para el sustrato en la colocación de reacción en meta, en comparación con la posición de activación para. Las energías se calculan usando términos basados en el campo de fuerzas. En una realización, el campo de fuerzas es AMBER99 y la energía de interacción entre el aminoácido y el entorno de proteína-sustrato se describe por la ecuación 1.

$$E_{no-unido} = E_{van-der-Waals} + E_{electrostático} \quad (\text{Ecuación 1})$$

donde

$$E_{van-der-Waals} = \sum_{\substack{\text{parejas} \\ \text{no unidas}}} \left(\frac{A_{ik}}{r_{ik}^{12}} - \frac{C_{ik}}{r_{ik}^6} \right)$$

$$E_{electrostático} = \sum_{\substack{\text{parejas} \\ \text{no unidas}}} \frac{q_i q_k}{D r_{ik}}$$

Un experto en la técnica apreciará que se pueden implementar otras ecuaciones similares para derivar energías de interacción adecuadas para estos análisis energéticos diferenciales de diferentes colocaciones de sustrato-aminoácido mutado.

En la subetapa (vii), los aminoácidos mutantes se seleccionan basándose en energías de unión favorables del sustrato unido en la colocación prevista de reacción en meta en comparación con la colocación prevista de reacción en para. Por lo tanto, estos cálculos determinan qué mutaciones de aminoácidos es probable que promuevan la O-metilación regioselectiva en meta.

En la etapa (iv), se repite la etapa (iii) hasta que resulta una lista de valores de energía para cada uno de los restos del sitio de unión que se van a mutar.

En la etapa (v), la lista de valores de energía derivada de la etapa (iv) se clasifica por la diferencia de energía entre cada mutación donde el sustrato está en la posición de reacción en meta en comparación con la posición de reacción en para. Las entradas en la parte superior de la lista representan cuando las mutaciones favorecen la O-metilación en meta frente a la O-metilación en para.

En la etapa (vi), se selecciona un número limitado de dichos candidatos de la etapa (v) basándose en los valores de energía. Las mutaciones no se seleccionan cuando (1) las mutaciones no son energéticamente favorables o (2) cuando no está previsto que las mutaciones alteren la regioselectividad.

Como se describe en la presente memoria, la aplicación del método descrito antes a la Hs COMT producía la identificación de un conjunto de mutaciones que se diseñan para mejorar la regioselectividad enzimática de la O-metilación. Las mutaciones se pueden usar independientemente o en combinación.

En la isoforma unida a membrana de Hs COMT, el número de restos equivalentes son los de Hs COMT solubles más cincuenta. Por lo tanto, los restos descritos o sustituidos en la Hs COMT también se deduce que se describen o sustituyen en el número de resto más cincuenta en la Hs COMT unida a membrana.

En una realización, un polipéptido de COMT mutante, (1) tiene una secuencia de aminoácidos que comparte al menos un porcentaje, tal como al menos 85 por ciento, por ejemplo al menos 90 por ciento, tal como al menos 95 por ciento, por ejemplo al menos 96 por ciento, tal como al menos 97 por ciento, por ejemplo al menos 98 por ciento, tal como al menos 99 por ciento de identidad de secuencia con la SEQ ID NO: 27, determinado a lo largo de la longitud entera de la SEQ ID NO:27; y (2) tiene al menos una sustitución de aminoácido en una posición que se alinea con las posiciones 198 a 199 de la SEQ ID NO: 27, que puede ser cualquiera de las sustituciones de aminoácidos descritas en la presente memoria más adelante; y (3) es capaz de catalizar la metilación de un grupo -OH del ácido protocatéquico, en donde dicha metilación da como resultado la generación de al menos 4, 5, 10, 15, 20, 25 o 30 veces más ácido vainílico en comparación con el ácido isovainílico. Además de estas características, dicho polipéptido de COMT mutante también puede ser capaz de catalizar la metilación de un grupo -OH del ácido protocatéquico, en donde dicha metilación da como resultado la generación de al menos 4, 5, 10, 15, 20, 25 o 30 veces más vainillina en comparación con la isovainillina; y/o ser capaz de catalizar la metilación de un grupo -OH del

alcohol protocatéquico, en donde dicha metilación da como resultado la generación de al menos 4, 5, 10, 15, 20, 25 o 30 veces más alcohol de vainillilo en comparación con el alcohol de isovainillilo.

Por lo tanto, el polipéptido de COMT mutante en una realización preferida puede tener una sustitución de aminoácido en la posición que se alinea con la posición 198 de la SEQ ID NO: 27. Por consiguiente, el polipéptido de COMT mutante puede ser un polipéptido de COMT mutante con las características indicadas antes, en donde dicha sustitución es una sustitución de la leucina en la posición que se alinea con la posición 198 de la SEQ ID NO: 27 por otro aminoácido que tiene un índice de hidropatía menor. Por ejemplo, el polipéptido de COMT mutante puede ser un polipéptido de COMT mutante con características como se han indicado antes, en donde dicha sustitución es una sustitución de la leucina en la posición que se alinea con la posición 198 de la SEQ ID NO: 27 por otro aminoácido que tiene un índice de hidropatía menor que 2. Por lo tanto, el polipéptido de COMT mutante puede ser un polipéptido de COMT mutante con características como se han indicado antes, en donde dicha sustitución es una sustitución de la leucina en la posición que se alinea con la posición 198 de la SEQ ID NO: 27 por una Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Lys, Met, Phe, Pro, Ser, Thr, Trp o Tyr, por ejemplo Ala, Arg, Asn, Asp, Glu, Gln, Gly, His, Lys, Met, Pro, Ser, Thr, Trp o Tyr. Sin embargo, preferiblemente dicha sustitución es una sustitución de la leucina en la posición que se alinea con la posición 198 de la SEQ ID NO: 27 por tirosina. La sustitución de la leucina que se alinea con la posición 198 de la SEQ ID NO: 27 por metionina aumentaba la regioselectividad de la O-metilación meta>para en el aldehído protocatéquico.

En otra realización preferida, el polipéptido de COMT mutante puede tener una sustitución de aminoácido en la posición que se alinea con la posición 199 de la SEQ ID NO: 27. Por consiguiente, el polipéptido de COMT mutante puede ser un polipéptido de COMT mutante con características como se han indicado antes, en donde dicha sustitución es una sustitución del ácido glutámico en la posición que se alinea con la posición 199 de la SEQ ID NO: 27 por otro aminoácido, que tiene carga neutra o positiva en la cadena lateral a pH 7,4. Por lo tanto, el polipéptido de COMT mutante puede ser un polipéptido de COMT mutante con características como se han indicado antes, en donde dicha sustitución es una sustitución del ácido glutámico en la posición que se alinea con la posición 199 de la SEQ ID NO: 27 por Ala, Arg, Asn, Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr o Val. Sin embargo, preferiblemente dicha sustitución es una sustitución del ácido glutámico en la posición que se alinea con la posición 199 de la SEQ ID NO: 27 por alanina o glutamina. La sustitución del ácido glutámico que se alinea con la posición 199 de la SEQ ID NO: 27 por alanina o glutamina aumentaba la regioselectividad de la O-metilación meta>para en el aldehído protocatéquico.

Por ejemplo, un polipéptido de COMT mutante puede tener una o más de las siguientes mutaciones: una sustitución de la leucina por un triptófano, tirosina, fenilalanina, ácido glutámico o arginina en una posición que se alinea con la posición 198 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27; una sustitución de la metionina por arginina, lisina o alanina en una posición que se alinea con la posición 40 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27, una sustitución del triptófano por una tirosina, lisina, histidina o arginina en una posición que se alinea con la posición 143 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27 una sustitución de la prolina por una isoleucina, arginina o tirosina en una posición que se alinea con la posición 174 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27; una sustitución del triptófano por una arginina o lisina en una posición que se alinea con la posición 38 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27; una sustitución de la cisteína por una fenilalanina, tirosina, ácido glutámico, triptófano o metionina en una posición que se alinea con la posición 173 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27; y/o una sustitución de arginina por una serina, ácido glutámico o ácido aspártico en una posición que se alinea con la posición 201 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27.

En una realización, un polipéptido de COMT mutante contiene sustitución de leucina por triptófano en una posición que se alinea con la posición 198. Esta mutación puede aumentar la regioselectividad de la O-metilación meta>para, para el ácido protocatéquico. La modelización del sitio de unión de la proteína de un polipéptido de COMT que contiene una mutación L198W, indica que puede ocurrir una colisión estérica entre el resto mutado y el sustrato. Esta colisión estérica no ocurre en la conformación de reacción en meta ya que el ácido carboxílico del sustrato está alejado de este resto.

En otra realización, el polipéptido de COMT mutante es un polipéptido de SEQ ID NO: 27, en donde el aminoácido en la posición 198 se ha sustituido por un aminoácido que tiene un índice de hidropatía menor que la leucina. Por ejemplo, el polipéptido de COMT mutante puede ser un polipéptido de SEQ ID NO: 27, en donde la leucina en la posición 198 se ha sustituido por un aminoácido que tiene un índice de hidropatía menor que 2. Por lo tanto, el polipéptido de COMT mutante puede ser un polipéptido de SEQ ID NO: 1, en donde la leucina en la posición 198 se ha sustituido por una Ala, Arg, Asn, Asp, Glu, Gln, Gly, His, Lys, Met, Pro, Ser, Thr, Trp o Tyr, preferiblemente Met o Tyr.

En otra realización preferida, el polipéptido de COMT mutante puede ser un polipéptido de SEQ ID NO: 27, en donde el aminoácido en la posición 199 se ha sustituido por otro aminoácido, que tiene carga neutra o positiva en la cadena lateral a pH 7,4. Por lo tanto, el polipéptido de COMT mutante puede ser un polipéptido de SEQ ID NO: 27, donde el ácido glutámico en la posición 199 se ha sustituido por Ala, Arg, Asn, Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr o Val, preferiblemente Ala o Gln.

En algunas realizaciones, un polipéptido de COMT mutante tiene dos o más mutaciones. Por ejemplo, se pueden mutar 2, 3, 4, 5, 6 o 7 de los restos en el sitio de unión del sustrato. Por ejemplo, en una realización, un polipéptido de COMT mutante puede tener una sustitución de metionina por una arginina o lisina en una posición que se alinea con la posición 40 de la secuencia de aminoácidos de SEQ ID NO: 27; una sustitución de triptófano por una tirosina o histidina en una posición que se alinea con la posición 143 de la secuencia de aminoácidos de SEQ ID NO: 27; una sustitución de prolina por una isoleucina en una posición que se alinea con la posición 174 de la secuencia de aminoácidos de SEQ ID NO: 27, y una sustitución de triptófano por arginina o lisina en la posición 38. Un polipéptido de COMT mutante también puede tener una sustitución de triptófano por lisina o arginina en una posición que se alinea con la posición 143 de la secuencia de aminoácidos de SEQ ID NO: 27 y una sustitución de prolina por una arginina o tirosina en la posición 174 de la SEQ ID NO: 27. Un polipéptido de COMT mutante también puede tener una sustitución de cisteína por fenilalanina, tirosina, ácido glutámico, triptófano o metionina en una posición que se alinea con la posición 173 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27, una sustitución de metionina por alanina en una posición que se alinea con la posición 40 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27, y una sustitución de la arginina por una serina, ácido glutámico o ácido aspártico en una posición que se alinea con la posición 201 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 27. También es posible que el polipéptido de COMT mutante tenga una sustitución de la leucina en una posición que se alinea con la posición 198 de la SEQ ID NO: 27, así como una sustitución del ácido glutámico en una posición que se alinea con la posición 199 de la SEQ ID NO: 27. Dichas sustituciones pueden ser cualquiera de las sustituciones descritas en la sección anterior, también es posible que el polipéptido de COMT mutante tenga una sustitución de la leucina en una posición que se alinea con la posición 198 de la SEQ ID NO: 27, así como una sustitución de la arginina en una posición que se alinea con la posición 201 de la SEQ ID NO: 27. Dichas sustituciones pueden ser cualquiera de las sustituciones descritas en esta sección anterior.

Porcentaje de identidad

Las identidades de secuencias dadas en la presente memoria son identidad de secuencia a lo largo de la longitud entera de la secuencia de referencia. Por consiguiente, la identidad de secuencia de la secuencia de aminoácidos proporcionada como la SEQ ID NO: 4 o SEQ ID NO: 27 en la presente memoria, es la identidad de secuencia a lo largo de la longitud entera de la SEQ ID NO: 4 o SEQ ID NO: 27, respectivamente.

El porcentaje de identidad se puede determinar como sigue. Una secuencia de referencia (p. ej., una secuencia de ácido nucleico o una secuencia de aminoácidos expuesta en la SEQ ID NO: 4 o SEQ ID NO: 27) se alinea con una o más secuencias candidatas usando el programa de ordenador ClustalW (versión 1.83, parámetros por defecto), que permite llevar a cabo alineamientos de secuencias de ácidos nucleicos o polipéptidos a lo largo de su longitud entera (alineamiento global). Chenna et al. (2003) *Nucleic Acids Res.* 31(13):3497-500.

ClustalW calcula el mejor emparejamiento entre una secuencia de referencia y una o más secuencias candidatas, y las alinea de modo que se pueden determinar las identidades, similitudes y diferencias. Se pueden insertar huecos de uno o más restos en una secuencia de referencia, una secuencia candidata, o ambas, para maximizar los alineamientos de las secuencias. Para el alineamiento por pares rápido de secuencias de ácidos nucleicos, se usan los siguientes parámetros por defecto: tamaño de palabra: 2; tamaño de ventana: 4; método de puntuación: porcentaje; número de diagonales superiores: 4; y penalización por hueco: 5. Para alineamientos múltiples de secuencias de ácidos nucleicos, se usan los siguientes parámetros: penalización por apertura de hueco: 10,0; penalización por extensión de hueco: 5,0; y transiciones de peso: sí. Para el alineamiento por pares rápido de secuencias de proteínas, se usan los siguientes parámetros: tamaño de palabra: 1; tamaño de ventana: 5; método de puntuación: porcentaje; número de diagonales superiores: 5; penalización por hueco: 3. Para alineamiento múltiple de secuencias de proteínas, se usan los siguientes parámetros: peso de matriz: blosum; penalización por apertura de hueco: 10,0; penalización por extensión de hueco: 0,05; huecos hidrófilos: seleccionado: restos hidrófilos: Gly, Pro, Ser, Asn, Asp, Gln, Glu, Arg y Lys; penalizaciones por hueco específico de resto: activado. El resultado de ClustalW es un alineamiento de secuencias que refleja la relación entre secuencias.

Para determinar el porcentaje de identidad de una secuencia de ácido nucleico o aminoácidos candidata con una secuencia de referencia, las secuencias se alinean usando ClustalW, el número de emparejamientos idénticos en el alineamiento se divide entre la longitud de la secuencia de referencia, y el resultado se multiplica por 100. Hay que indicar que el valor del porcentaje de identidad se puede redondear a la décima más cercana. Por ejemplo, 78,11, 78,12, 78,13 y 78,14 se pueden redondear a 78,1, mientras que 78,15, 78,16, 78,17, 78,18 y 78,19 se redondean a 78,2.

Sustituciones de aminoácidos

Las sustituciones de aminoácidos pueden ser conservadoras o no conservadoras. Las sustituciones de aminoácidos conservadoras sustituyen un aminoácido por un aminoácido de la misma clase, mientras que las sustituciones de aminoácidos no conservadoras sustituyen un aminoácido por un aminoácido de una clase diferente. Los ejemplos de sustituciones conservadoras incluyen sustituciones de aminoácidos dentro de los siguientes grupos: (1) glicina y alanina; (2) valina, isoleucina y leucina; (3) ácido aspártico y ácido glutámico; (4) asparagina, glutamina, serina y treonina; (5) lisina, histidina y arginina; y (6) fenilalanina y tirosina.

Las sustituciones de aminoácidos no conservadoras pueden sustituir un aminoácido de una clase por un aminoácido de una clase diferente. Las sustituciones no conservadoras pueden hacer un cambio sustancial en la carga o hidrofobicidad del producto génico. Las sustituciones de aminoácidos no conservadoras también pueden hacer un cambio sustancial en el volumen de la cadena lateral del resto, p. ej., sustitución de un resto de isoleucina por una alanina. Los ejemplos de sustituciones no conservadoras incluyen la sustitución de un aminoácido básico por un aminoácido no polar o un aminoácido ácido por un aminoácido polar. Un experto en la técnica apreciará que se pueden sustituir aminoácidos similares por los mutantes descritos en la presente memoria.

Ácidos nucleicos

Este documento también proporciona ácidos nucleicos que codifican los polipéptidos de AROM y COMT mutantes. Un "ácido nucleico aislado" se refiere a un ácido nucleico que se separa de otras moléculas de ácido nucleico que están presente en un genoma, incluyendo ácidos nucleicos que normalmente flanquean uno o ambos lados del ácido nucleico en un genoma. El término "aislado" como se usa en la presente memoria con respecto a ácidos nucleicos incluye cualquier secuencia de ácido nucleico que no es natural, puesto que dichas secuencias que no son naturales no se encuentran en la naturaleza y no tienen secuencias inmediatamente contiguas en un genoma que es natural.

Un ácido nucleico aislado puede ser, por ejemplo, una molécula de ADN, con la condición de que se elimine o esté ausente una de las secuencias de ácido nucleico que se encuentra flanqueando inmediatamente esa molécula de ADN en un genoma que es natural. Por lo tanto, un ácido nucleico aislado incluye, sin limitación, una molécula de ADN que existe como una molécula separada (p. ej., un ácido nucleico químicamente sintetizado, o un ADNc o fragmento de ADN genómico producido por PCR o tratamiento con endonucleasas de restricción) independiente de otras secuencias, así como ADN que se incorpora en un vector, un plásmido de replicación autónoma, un virus (p. ej., cualquier paramixovirus, retrovirus, lentivirus, adenovirus o herpesvirus), o en el ADN genómico de un procarionte o eucariota. Además, un ácido nucleico aislado puede incluir un ácido nucleico genéticamente modificado tal como una molécula de ADN que es parte de un ácido nucleico híbrido o de fusión. Un ácido nucleico que existe entre cientos de millones de otros ácidos nucleicos, por ejemplo, en bibliotecas de ADNc, genotecas o corte de gel que contienen una digestión por restricción de ADN genómico, no se considera un ácido nucleico aislado.

Los ácidos nucleicos que codifican polipéptidos de AROM o COMT se pueden modificar usando técnicas de clonación molecular comunes (p. ej., mutagénesis dirigida al sitio) para generar mutaciones en posiciones particulares del polipéptido codificado (p. ej., posiciones que se alinean con las posiciones 1349, 1366, 1370, 1387, 1392, 1441, 1458, 1500, 1533 o 1571 de la secuencia de aminoácidos de AROM expuesta en la SEQ ID NO: 4; o posiciones que se alinean con la posición 38, 40, 143, 173, 174, 198 o 201 de la forma soluble de la secuencia de aminoácidos de COMT humana expuesta en la SEQ ID NO: 27). Las moléculas de ácido nucleico pueden incluir una mutación de nucleótido única o más de una mutación, o más de un tipo de mutación. Se puede usar la reacción en cadena de la polimerasa (PCR) y técnicas de hibridación de ácidos nucleicos para identificar ácidos nucleicos que codifican polipéptidos de AROM que tienen secuencias de aminoácidos alteradas.

En algunas realizaciones, una molécula de ácido nucleico que codifica un polipéptido mutante de esta invención puede incluir una secuencia de marcador que codifica un "marcador" destinado a facilitar la posterior manipulación (p. ej., a facilitar la purificación o detección), secreción o localización del polipéptido codificado. Las secuencias de marcadores se pueden insertar en la secuencia de ácido nucleico que codifica el polipéptido de AROM o COMT de modo que el marcador codificado se localiza en el extremo carboxilo o amino del polipéptido de AROM o COMT. Ejemplos no limitantes de marcadores codificados incluyen la proteína verde fluorescente (GFP), glutatión S transferasa (GST), marcador HIS y marcador FLAG (Kodak, New Haven, CT). Otros ejemplos de marcadores incluyen un péptido de tránsito de cloroplasto, un péptido de tránsito mitocondrial, un péptido de amiloplasto, péptido señal o un marcador de secreción.

Los polipéptidos (mutantes o de tipo natural) de esta invención se pueden producir usando cualquier método. Por ejemplo, los polipéptidos se pueden producir por síntesis química. Alternativamente, los polipéptidos descritos en la presente memoria se pueden producir por tecnología recombinante convencional usando vectores de expresión heterólogos que codifican polipéptidos. Los vectores de expresión se pueden introducir en células hospedantes (p. ej., por transformación o transfección) para la expresión del polipéptido codificado, el cual después se puede purificar. Los sistemas de expresión que se pueden usar para la producción a pequeña o gran escala de polipéptidos incluyen, sin limitación, microorganismos tales como bacterias (p. ej., *E. coli* y *B. subtilis*) transformados con ADN bacteriófago recombinante, ADN plasmídico o vectores de expresión de ADN cósmido que contienen moléculas de ácido nucleico descritas en la presente memoria, y levaduras (p. ej., *S. cerevisiae* o *S. pombe*) transformadas con vectores de expresión de levaduras recombinantes que contienen las moléculas de ácido nucleico descritas en la presente memoria. Los sistemas de expresión útiles también incluyen sistemas de células de insectos infectadas con los vectores de expresión de virus recombinantes (p. ej., baculovirus) que contienen las moléculas de ácido nucleico descritas en la presente memoria, y sistemas de células vegetales infectadas con vectores de expresión de virus recombinantes (p. ej., virus del mosaico del tabaco) o transformadas con vectores de expresión de plásmidos recombinantes (p. ej., plásmido Ti) que contienen moléculas de ácido nucleico descritas en la presente memoria. Los polipéptidos de esta invención también se pueden producir usando sistemas de expresión de mamíferos que

albergan construcciones de expresión recombinantes que contienen promotores derivados del genoma de células de mamífero (p. ej., el promotor de metalotioneína) o de virus de mamíferos (p. ej., el promotor tardío de adenovirus y el promotor de citomegalovirus), junto con los ácidos nucleicos descritos en la presente memoria. Los polipéptidos de esta invención pueden tener un marcador N-terminal o C-terminal como se ha descrito antes.

5 Hospedantes recombinantes

Esta invención también caracteriza hospedantes recombinantes. Como se usa en la presente memoria, la expresión hospedante recombinante se pretende que se refiera a un hospedante, cuyo genoma se ha aumentado en al menos una secuencia de ADN incorporada. Dichas secuencias de ADN incluyen genes que no están presentes de forma natural, secuencias de ADN que normalmente no son transcritas en ARN o traducidas en una proteína ("expresadas"), y otros genes y secuencias de ADN que se desean introducir en el hospedante no recombinante. Se apreciará que típicamente el genoma de un hospedante recombinante descrito en la presente memoria se aumenta por la introducción estable de uno o más genes recombinantes. Sin embargo, también se pueden usar plásmidos o vectores autónomos o replicativos dentro del alcance de esta invención. Además, la presente invención se puede poner en práctica usando un plásmido o vector de número de copias bajo, p. ej., una sola copia, o número de copias alto (como se ilustra en la presente memoria).

En general, el ADN introducido no reside originalmente en el hospedante que es el receptor del ADN, pero está dentro del alcance de la invención aislar un segmento de ADN de un hospedante dado, y posteriormente introducir una o más copias adicionales de ese ADN en el mismo hospedante, p. ej., para potenciar la producción del producto de un gen o alterar el patrón de expresión de un gen. En algunos casos, el ADN introducido modificará o incluso sustituirá un gen endógeno o secuencia de ADN por, p. ej., recombinación homóloga o mutagénesis dirigida al sitio. Los hospedantes recombinantes adecuados incluyen microorganismos, células de plantas y plantas.

La expresión "gen recombinante" se refiere a un gen o secuencia de ADN que se introduce en un hospedante receptor, independientemente de si puede estar ya presente en dicho hospedante el mismo o similar gen o secuencia de ADN. "Introducido" o "aumentado" en este contexto, se sabe en la técnica que significa introducido o aumentado por la mano del hombre. Por lo tanto, un gen recombinante puede ser una secuencia de ADN de otra especie, o puede ser una secuencia de ADN que procede o está presente en la misma especie, pero se ha incorporado en un hospedante por métodos recombinantes para formar un hospedante recombinante. Se apreciará que un gen recombinante que se introduce en un hospedante puede ser idéntico a una secuencia de ADN que está normalmente presente en el hospedante que se va a transformar, y se introduce para proporcionar una o más copias adicionales del ADN para así permitir la sobreexpresión o expresión modificada del producto génico de ese ADN.

Un gen recombinante que codifica un polipéptido descrito en la presente memoria incluye la secuencia codificante para ese polipéptido, operativamente unida, en la orientación paralela, a una o más regiones reguladoras adecuadas para la expresión del polipéptido. Debido a que muchos microorganismos son capaces de expresar múltiples productos génicos a partir de un ARNm policistrónico, se pueden expresar múltiples polipéptidos bajo el control de una sola región reguladora para esos microorganismos, si se desea. Una secuencia codificante y una región reguladora se considera que están operativamente unidas cuando la región reguladora y la secuencia codificante están situadas de modo que la región reguladora es eficaz para regular la transcripción o traducción de la secuencia. Típicamente, el sitio de inicio de la traducción del marco de lectura traduccional de la secuencia codificante está situado entre uno y aproximadamente cincuenta nucleótidos en la dirección 3' de la región reguladora para un gen monocistrónico.

En muchos casos, la secuencia codificante de un polipéptido descrito en la presente memoria se identifica en una especie distinta del hospedante recombinante, es decir, es un ácido nucleico heterólogo. La expresión "ácido nucleico heterólogo" como se usa en la presente memoria, se refiere a un ácido nucleico introducido en un hospedante recombinante, en donde dicho ácido nucleico no está naturalmente presente en dicho hospedante. Por lo tanto, si el hospedante recombinante es un microorganismo, la secuencia codificante puede ser de microorganismos procariontes o eucariotes, de plantas o de animales. Sin embargo, en algún caso, la secuencia codificante es una secuencia que es originaria del hospedante y se está reintroduciendo en ese organismo. Una secuencia originaria se puede distinguir a menudo de la secuencia que se encuentra de forma natural por la presencia de secuencias no naturales unidas al ácido nucleico exógeno, p. ej., secuencias reguladoras no originarias que flanquean una secuencia originaria en una construcción de ácido nucleico recombinante. Además, los ácidos nucleicos exógenos transformados de forma estable típicamente se integran en posiciones distintas de la posición donde se encuentra la secuencia originaria.

"Región reguladora" se refiere a un ácido nucleico que tiene secuencias de nucleótidos que influyen en el inicio y velocidad de transcripción o traducción, y la estabilidad y/o movilidad de un producto de transcripción o traducción. Las regiones reguladoras incluyen, sin limitación, secuencias de promotores, secuencias de potenciadores, elementos de respuesta, sitios de reconocimiento de proteínas, elementos inducibles, secuencias de unión de proteínas, regiones no traducidas (UTRs) 5' y 3', sitios de inicio de la transcripción, secuencias de terminación, secuencias de poliadenilación, intrones y combinaciones de las mismas. Una región reguladora típicamente incluye al menos un promotor nuclear (basal). Una región reguladora también puede incluir al menos un elemento de

control, tal como una secuencia de potenciador, un elemento en la dirección 5' o región de activación en la dirección 5' (UAR). Una región reguladora se une operativamente a una secuencia codificante colocando la región reguladora y la secuencia codificante de modo que la región reguladora sea eficaz para regular la transcripción o traducción de la secuencia. Por ejemplo, para unir operativamente una secuencia codificante a una secuencia de promotor, el sitio de inicio de la traducción del marco de lectura traduccional de la secuencia codificante está colocada típicamente entre uno y aproximadamente cincuenta nucleótidos en la dirección 3' del promotor. Sin embargo, una región reguladora puede estar colocada como mucho aproximadamente 5.000 nucleótidos en la dirección 5' del sitio de inicio de la traducción, o aproximadamente 2.000 nucleótidos en la dirección 5' del sitio de inicio de la traducción.

La elección de las regiones reguladoras que se van a incluir depende de varios factores, que incluyen eficacia, seleccionabilidad, inducibilidad, nivel de expresión deseado y expresión preferencial durante ciertas etapas del cultivo. Para un experto en la técnica es una cuestión rutinaria modular la expresión de una secuencia codificante seleccionando adecuadamente y colocando regiones reguladoras con respecto a la secuencia codificante. Se entenderá que puede estar presente más de una región reguladora, p. ej., intrones, potenciadores, regiones de activación en dirección 5', terminadores de la transcripción y elementos inducibles.

Se pueden combinar uno o más genes, por ejemplo, uno o más ácidos nucleicos heterólogos en una construcción de ácido nucleico en "módulos" útiles para un aspecto discreto de la producción de vainillina y/o glucósido de vainillina. La combinación de una pluralidad de genes o ácidos nucleicos heterólogos en un módulo facilita el uso del módulo en una variedad de especies. Por ejemplo, un grupo de genes de vainillina se puede combinar de modo que cada secuencia codificante esté operativamente unida a una región reguladora separada, para formar un módulo de vainillina para la producción en organismos eucariotas. Alternativamente, el módulo puede expresar un mensaje policistrónico para la producción de vainillina y/o glucósido de vainillina en hospedantes procariotas tales como especies de *Rodobacter*, *E. coli*, *Bacillus* o *Lactobacillus*. Además de los genes útiles para la producción de vainillina o glucósido de vainillina, una construcción recombinante típicamente también contiene un origen de la replicación, y uno o más marcadores seleccionables para el mantenimiento de la construcción en especies adecuadas.

Se apreciará que, debido a la degeneración del código genético, una serie de ácidos nucleicos pueden codificar un polipéptido particular; es decir, para muchos aminoácidos, hay más de un triplete de nucleótidos que sirve como codón para el aminoácido. Por lo tanto, los codones en la secuencia codificante para un polipéptido dado se pueden modificar de modo que se obtenga la expresión óptima en un hospedante particular, usando tablas de sesgo de codones para ese hospedante (p. ej., microorganismo). Como ácidos nucleicos aislados, estas secuencias modificadas pueden existir como moléculas purificadas y se pueden incorporar en un vector o un virus para usar en módulos de construcción para construcciones de ácidos nucleicos recombinantes.

Los hospedantes recombinantes descritos en la presente memoria expresan polipéptidos de AROM mutantes. Por lo tanto, la presente invención se refiere a un hospedante recombinante que alberga un ácido nucleico heterólogo que codifica un polipéptido de AROM mutante, que puede ser cualquiera de los polipéptidos mutantes descritos en la presente memoria. En particular, la invención se refiere a un hospedante recombinante que alberga un ácido nucleico heterólogo que codifica dicho polipéptido de AROM mutante, en donde dicho ácido nucleico está operativamente unido a una región reguladora que permite la expresión en dicho hospedante recombinante.

Dichos hospedantes pueden incluir además genes adicionales o módulos biosintéticos para producir vainillina o glucósido de vainillina, mejorar la eficacia con la que la energía y fuentes de carbono se convierten en vainillina y su glucósido, y/o para potenciar la productividad del cultivo celular o planta. Dichos módulos biosintéticos adicionales pueden incluir uno o más de un gen que codifica un polipéptido de 3DSD, un gen que codifica una fosfopanteteinil transferasa (PPTasa) y un gen que codifica un polipéptido de UGT. Véase la figura 1. Estos genes pueden ser genes endógenos o genes recombinantes. Además, cuando la célula hospedante alberga un ácido nucleico heterólogo que codifica un polipéptido de AROM mutante, la célula hospedante puede incluir además un gen de OMT de tipo natural. Alternativamente, la célula hospedante puede albergar ácidos nucleicos heterólogos que codifican un polipéptido de COMT mutante y un polipéptido de AROM mutante, como se describe en la presente memoria. Además, el hospedante puede expresar además una enzima alcohol de vainillilo oxidasa (VAO).

Se conocen polipéptidos de 3DSD adecuados. Un polipéptido de 3DSD de acuerdo con la presente invención puede ser cualquier enzima con actividad de 3-deshidro-shikimato deshidratasa. Preferiblemente, el polipéptido de 3DSD es una enzima capaz de catalizar la conversión del 3-deshidro-shikimato en protocatecuato y H₂O. Un polipéptido de 3DSD de acuerdo con la presente invención preferiblemente es una enzima clasificada como EC 4.2.1.118. Por ejemplo, un polipéptido adecuado que tiene actividad de 3DSD incluye el polipéptido de 3DSD hecho por *Podospora pauciseta*, *Ustilago maydis*, *Rhodoicoccus jostii*, *Acinetobacter sp.*, *Aspergillus niger* o *Neurospora crassa*. Véanse los números de acceso en GENBANK CAD60599, XP_001905369.1, XP_761560.1, ABG93191.1, AAC37159.1 y XM_001392464. Por lo tanto, el hospedante recombinante puede incluir un ácido nucleico heterólogo que codifica el polipéptido de 3DSD de *Podospora anserina*, *Ustilago maydis*, *Rhodoicoccus jostii*, *Acinetobacter sp.*, *Aspergillus niger* o *Neurospora crassa* o un homólogo funcional de cualquiera de los anteriores que comparte al menos 80%, tal como al menos 85%, por ejemplo al menos 90%, tal como al menos 95%, por ejemplo al menos 98% de identidad de secuencia con el mismo.

Como se describe en la presente memoria, los polipéptidos de OMT de tipo natural adecuados son conocidos. Por ejemplo, un polipéptido de OMT de tipo natural adecuado incluye el OMT hecho por *H. sapiens*, *A. thaliana*, o *Fragaria x ananassa* (véanse, nº de acceso en GENBANK NM_000754, AY062837; y AF220491), así como polipéptidos de OMT aislados de una variedad de mamíferos, plantas o microorganismos.

- 5 Igualmente, los polipéptidos de AROM de tipo natural son conocidos. Por ejemplo, los polipéptidos de AROM de tipo natural adecuados incluyen AROM hecha por *S. cerevisiae*, *S. pombe*, *S. japonicas*, *N. crassa*, y *Y. Lipolytica*, véanse los nº de acceso en GENBANK X06077, NP_594681.1, XP_002171624 y XP_956000.

- Los polipéptidos de ACAR adecuados son conocidos. Los polipéptidos de ACAR pueden ser cualquier enzima que tenga actividad de ácido carboxílico aromático reductasa. Preferiblemente, el polipéptido de ACAR es una enzima capaz de catalizar la conversión de ácido protocatéuico en aldehído protocatéuico y/o la conversión de ácido vainillíco en vainillina. Un polipéptido de ACAR preferiblemente es una enzima clasificada como EC 1.2.1.30. Por ejemplo, un polipéptido de ACAR adecuado lo produce el género *Nocardia*. Véase, p. ej., nº de acceso en GENBANK AY495697. Por lo tanto, el hospedante recombinante puede incluir un ácido nucleico heterólogo que codifica el polipéptido de ACAR del género *Nocardia* o un homólogo funcional del mismo que comparte al menos 80%, tal como al menos 85%, por ejemplo al menos 90%, tal como al menos 95%, por ejemplo al menos 98% de identidad de secuencia con el mismo.

- Los polipéptidos de PPTasa adecuados son conocidos. Un polipéptido de PPTasa puede ser cualquier enzima capaz de catalizar la fosfopanteteinilación. Por ejemplo, un polipéptido de PPTasa adecuado es una enzima capaz de catalizar la fosfopanteteinilación de ACAR. Por ejemplo, un polipéptido de PPTasa adecuado es producido por *E. coli*, *Corynebacterium glutamicum* o *Nocardia farcinica*. Véanse los nº de acceso en GENBANK NP_601186, BAA35224 y YP_120266. Por lo tanto, el hospedante recombinante puede incluir un ácido nucleico heterólogo que codifica el polipéptido de PPTasa de *E. coli*, *C. glutamicum* o *N. farcinica* o un homólogo funcional de cualquiera de los mencionados antes que comparte al menos 80%, tal como al menos 85%, por ejemplo al menos 90%, tal como al menos 95%, por ejemplo al menos 98% de identidad de secuencia con el mismo.

- La glucosilación de la vainillina es particularmente útil. El β -D-glucósido de vainillina es la forma de almacenamiento de la vainillina que se encuentra en la vaina de vainilla. No es tóxico para la mayoría de los organismos, incluyendo levaduras, y tiene una mayor solubilidad en agua, en comparación con la vainillina. Además, la formación del β -D-glucósido de vainillina dirige lo más probablemente la biosíntesis hacia la producción de vainillina. UGT72E2 (Hansen, et al. (2009) *Appl. Environ. Microbiol.* 75:2765-27740) presentaba una alta especificidad de sustrato hacia la vainillina. En concordancia con esta observación, su expresión en la cepa de *S. cerevisiae* que produce vainillina daba como resultado que prácticamente toda la vainillina se convertía en β -D-glucósido de vainillina. La capacidad de convertir la vainillina en β -D-glucósido de vainillina in vivo es importante, porque la producción microbiana de vainillina no glucosilada más allá de la escala de 0,5-1 g/litro estaría dificultada por la toxicidad de la vainillina libre. La glucosilación sirve para evitar el efecto inhibitorio.

- Por consiguiente, el hospedante recombinante de esta invención también expresa un polipéptido de UGT. Un polipéptido de UGT puede ser cualquier UDP-Glucosa:Aglicón-Glucosiltransferasa. Preferiblemente, los polipéptidos de UGT pueden catalizar la glucosilación de la vainillina (es decir, para producir beta-D-glucósido de vainillina). Por lo tanto, el polipéptido de UGT puede ser una familia de 1 glicosiltransferasa. Los polipéptidos de UGT preferidos se clasifican como EC 2.4.1. Los polipéptidos de UGT adecuados incluyen UGT71C2, UGT72B1, UGT72E2, UGT84A2, UGT89B1, UGT85B1, y polipéptidos de arbutina sintasa. Véanse, p. ej., los nº de acceso en GENBANK AC0005496, NM_116337 y NM_126067. UGT72E2 de *A. thaliana* es particularmente útil (véase, p. ej., Hansen, et al. (2009) véase antes). Por lo tanto, el hospedante recombinante puede incluir un ácido nucleico heterólogo que codifica UGT71C2, UGT72B1, UGT72E2, UGT84A2, UGT89B1, UGT85B1, o la arbutina sintasa o un homólogo funcional de cualquiera de los mencionados antes que comparte al menos 80%, tal como al menos 85%, por ejemplo al menos 90%, tal como al menos 95%, por ejemplo al menos 98% de identidad de secuencia con el mismo. Se describen otros UGT útiles en el documento WO 01/40491.

- Como una realización adicional, una enzima VAO (EC 1.1.3.38) también puede ser expresada por células hospedantes para oxidar cualquier alcohol de vainillilo formado a vainillina. Las enzimas VAO son conocidas en la técnica e incluyen enzimas de hongos filamentosos tales como *Fusarium moniliformis* (nº de acceso en GENBANK AFJ11909) y *Penicillium simplicissimum* (nº de acceso en GENBANK P56216; Benen, et al. (1998) *J. Biol. Chem.* 273:7865-72) y bacterias tales como *Modestobacter marinus* (nº de acceso en GENBANK YP_006366868), *Rhodococcus jostii* (nº de acceso en GENBANK YP_703243.1) y *R. opacus* (nº de acceso en GENBANK EHI39392).

- En algunos casos, es deseable inhibir una o más funciones de un polipéptido endógeno con el fin de desviar productos intermedios metabólicos hacia la biosíntesis de vainillina o glucósido de vainillina. Por ejemplo, se puede reducir la actividad de piruvato descarboxilasa (PDC1) y/o de glutamato deshidrogenasa. En dichos casos, se puede incluir un ácido nucleico que inhibe la expresión del polipéptido o producto génico en una construcción recombinante que es transformada en la cepa. Alternativamente, se puede usar mutagénesis para generar mutantes en genes para los que se desea inhibir la función.

Son adecuados una serie de procariotas y eucariotas para usar en la construcción de microorganismos recombinantes descritos en la presente memoria, p. ej., bacterias gram negativas, bacterias gram positivas, levaduras y otros hongos. Se analiza primero una especie o cepa seleccionada para usar como una cepa de producción de vainillina o glucósido de vainillina para determinar que genes de producción son endógenos de la cepa y que genes no están presentes. Los genes para los que no está presente un equivalente endógeno en la cepa, se ensamblan en una o más construcciones recombinantes, que después se transforman en la cepa con el fin de suministrar la o las funciones que faltan.

Las especies procariotas y eucariotas de ejemplo se describen con más detalle más adelante. Sin embargo, se apreciará que pueden ser adecuadas otras especies. Por ejemplo, las especies adecuadas pueden ser de un género *Agaricus*, *Aspergillus*, *Bacillus*, *Candida*, *Corynebacterium*, *Escherichia*, *Fusarium*/*Gibberella*, *Kluyveromyces*, *Laetiporus*, *Lentinus*, *Phaffia*, *Phanerochaete*, *Pichia*, *Physcomitrella*, *Rhodoturula*, *Saccharomyces*, *Schizosaccharomyces*, *Sphaceloma*, *Xanthophyllomyces Yarrowia* y *Lactobacillus*. Las especies de ejemplo de dichos géneros incluyen *Lentinus tigrinus*, *Laetiporus sulphureus*, *Phanerochaete chrysosporium*, *Pichia pastoris*, *Physcomitrella patens*, *Rhodoturula glutinis* 32, *Rhodoturula mucilaginosa*, *Phaffia rhodozyma* UBV-AX, *Xanthophyllomyces dendrorhous*, *Fusarium fujikuroi*/*Gibberella fujikuroi*, *Candida utilis* y *Yarrowia lipolytica*. En algunas realizaciones, un microorganismo puede ser un ascomiceto tal como *Gibberella fujikuroi*, *Kluyveromyces lactis*, *Schizosaccharomyces pombe*, *Aspergillus niger* o *Saccharomyces cerevisiae*. En algunas realizaciones, un microorganismo puede ser un procariota tal como *Escherichia coli*, *Rhodobacter sphaeroides* o *Rhodobacter capsulatus*. Se apreciará que algunos microorganismos se pueden usar para cribar y ensayar genes de interés de una forma de alta capacidad, mientras que otros microorganismos con productividad o características de crecimiento deseadas se pueden usar para la producción a gran escala de beta-D-glucósido de vainillina.

Se describen ejemplos específicos de hospedantes recombinantes útiles en el documento WO 01/40491, así como en Hansen et al. (2009) *Appl. Environ. Microbiol.* 75:2765-2774 y Brochado, et al. (2010) *Microbial Cell Factories* 9:84, en donde el hospedante recombinante de acuerdo con esta invención contiene un ácido nucleico heterólogo que codifica un polipéptido de AROM mutante en lugar de los genes de OMT descritos en el documento WO 01/40491.

Un hospedante recombinante preferido para usar con la presente invención es *S. cerevisiae*, que se puede modificar genéticamente de forma recombinante como se describe en la presente memoria. *S. cerevisiae* es un organismo de bastidor ampliamente usado en biología sintética, y se puede usar como la plataforma de microorganismo recombinante. Hay bibliotecas de mutantes, plásmidos, modelos de ordenador detallados de metabolismo y otra información para *S. cerevisiae*, que permiten un diseño racional de diferentes módulos para potenciar el rendimiento del producto. Los métodos para hacer microorganismos recombinantes son conocidos. La cepa VG4 de *S. cerevisiae* (Brochado, et al. (2010) *Microb. Cell Fact.* 9:84) es particularmente útil. VG4 tiene el genotipo de *pdc1Δgdh1Δ↑GDH2*.

Las especies de *Aspergillus* tales como *A. oryzae*, *A. niger* y *A. sojae* son microorganismos ampliamente usados en la producción de alimento, y también se pueden usar como la plataforma de microorganismo recombinante. Por lo tanto, el hospedante recombinante puede ser *Aspergillus spp.* Están disponibles secuencias de nucleótidos para genomas de *A. nidulans*, *A. fumigatus*, *A. oryzae*, *A. clavatus*, *A. flavus*, *A. niger* y *A. terreus*, que permiten el diseño racional y la modificación de rutas endógenas para potenciar el flujo y aumentar el rendimiento del producto. Se han desarrollado modelos metabólicos para *Aspergillus*, así como estudios transcriptómicos y estudios proteómicos. *A. niger* se cultiva para la producción industrial de una serie de ingredientes de alimentos tales como ácido cítrico y ácido glucónico, y por lo tanto, especies como *A. niger* en general son adecuadas para la producción de ingredientes de alimentos tales como la vainillina y glucósido de vainillina.

E. coli, otro organismo plataforma ampliamente usado en biología sintética, también se puede usar como la plataforma de microorganismo recombinante. Por lo tanto, el hospedante recombinante puede ser *E. coli*. De forma similar a *Saccharomyces*, hay bibliotecas de mutantes, plásmidos, modelos de ordenador detallados de metabolismo y otra información para *E. coli*, que permiten un diseño racional de diferentes módulos para potenciar el rendimiento del producto. Se pueden usar métodos similares a los descritos antes para *Saccharomyces* para hacer microorganismos recombinantes de *E. coli*.

Se puede usar *Rhodobacter* como la plataforma de microorganismo recombinante. Por lo tanto, el hospedante recombinante puede ser *Rhodobacter spp.* De forma similar a *E. coli*, hay bibliotecas de mutantes así como vectores de plásmidos adecuados, que permiten el diseño racional de diferentes módulos para potenciar el rendimiento del producto. Se pueden usar métodos similares a los descritos antes para *E. coli* para hacer microorganismos recombinantes de *Rhodobacter*.

Los musgos *Physcomitrella*, cuando crecen en cultivo en suspensión, tienen características similares a cultivos de levaduras y otros cultivos fúngicos. Este género se está convirtiendo en un tipo importante de célula para la producción de metabolitos secundarios de plantas, que pueden ser difíciles de producir en otros tipos de células. Por lo tanto, el hospedante recombinante puede ser una *Physcomitrella spp.*

En algunas realizaciones, los ácidos nucleicos y polipéptidos descritos en la presente memoria se introducen en plantas o células de plantas para aumentar la producción general de vainillina o glucósido de vainillina. Por lo tanto, un hospedante recombinante puede ser una planta o una célula de planta que incluye al menos un ácido nucleico heterólogo descrito en la presente memoria. Una planta o célula de planta se puede transformar al tener un ácido nucleico heterólogo integrado en su genoma, es decir, se puede transformar establemente. Las células transformadas establemente típicamente retienen el ácido nucleico introducido con cada división celular. Una planta o célula de planta también se puede transformar de manera transitoria de modo que el ácido nucleico heterólogo no se integre en su genoma. Las células transformadas de forma transitoria típicamente pierden todo o una parte del ácido nucleico introducido con cada división celular de modo que el ácido nucleico introducido no se puede detectar en células hija después de un número suficiente de divisiones celulares. Las plantas y células de plantas transgénicas tanto transformadas de forma transitoria como transformadas establemente pueden ser útiles en los métodos descritos en la presente memoria.

Las células de plantas transgénicas usadas en los métodos descritos en la presente memoria pueden constituir parte o todo de una planta entera. Dichas plantas se pueden desarrollar de una forma adecuada para la especie en consideración, sea en una cámara de cultivo, un invernadero o en un campo. Las plantas transgénicas se pueden cruzar como se desee para un propósito particular, p. ej., para introducir un ácido nucleico heterólogo, por ejemplo una construcción de ácido nucleico recombinante en otras líneas, para transferir un ácido nucleico heterólogo a otras especies, o para la selección adicional de otras características deseables. Alternativamente, las plantas transgénicas se pueden propagar de forma vegetativa para las especies sensibles a dichas técnicas. Como se usa en la presente memoria, una planta transgénica también se refiere a la progenie de una planta transgénica inicial con la condición de que la progenie herede el transgén. Las semillas producidas por una planta transgénica se pueden cultivar y después autofecundar (o cruzar y autofecundar) para obtener semillas homocigotas para la construcción de ácido nucleico.

Las plantas transgénicas se pueden cultivar en un cultivo en suspensión, o cultivo tisular o de órgano. Para los propósitos de esta invención, se pueden usar técnicas de cultivo de tejido sólido y/o líquido. Cuando se usa medio sólido, las células de plantas transgénicas se pueden poner directamente sobre el medio o se pueden poner sobre un filtro que después se pone en contacto con el medio. Cuando se usa medio líquido, las células de plantas transgénicas se pueden poner sobre un dispositivo de flotación, p. ej., una membrana porosa que se pone en contacto con el medio líquido.

Cuando se usan células de plantas transformadas de forma transitoria, se puede incluir una secuencia indicadora que codifica un polipéptido indicador que tiene actividad de indicador, en el procedimiento de transformación y se puede llevar a cabo un ensayo de actividad o expresión de indicador en un tiempo adecuado después de la transformación. Un tiempo adecuado para llevar a cabo el ensayo típicamente es de aproximadamente 1-21 días después de la transformación, p. ej., aproximadamente 1-14 días, aproximadamente 1-7 días, o aproximadamente 1-3 días. El uso de ensayos transitorios es particularmente conveniente para el análisis rápido en diferentes especies, o para confirmar la expresión de un polipéptido heterólogo cuya expresión no se ha confirmado previamente en células receptoras particulares.

Las técnicas para introducir ácidos nucleicos en plantas monocotiledóneas y dicotiledóneas son conocidas en la técnica, e incluyen, sin limitación, transformación mediada por *Agrobacterium*, transformación mediada por vector vírico, electroporación y transformación con pistola de partículas; véanse las patentes de EE.UU. nº 5.538.880; 5.204.253; 6.329.571; y 6.013.863. Si una célula o tejido cultivado se usa como el tejido receptor para la transformación, las plantas se pueden regenerar a partir de cultivos transformados si se desea, por técnicas conocidas por los expertos en la técnica.

En una población de plantas transgénicas se puede cribar y/o seleccionar los miembros de la población que tengan un rasgo o fenotipo conferido por la expresión del transgén. Por ejemplo, se pueden cribar en una población de progenie de un solo suceso de transformación las plantas que tienen un nivel deseado de expresión de un polipéptido o ácido nucleico descrito en la presente memoria. Se pueden usar métodos físicos y bioquímicos para identificar niveles de expresión. Estos incluyen el análisis Southern o la amplificación por PCR para la detección de un polinucleótido; transferencias Northern, protección de S1 RNasa, extensión de cebador o amplificación por RT-PCR para detectar transcritos de ARN; ensayos enzimáticos para detectar actividad de enzimas o ribozimas de polipéptidos y polinucleótidos; y electroforesis en gel de proteínas, transferencias Western, inmunoprecipitación e inmunoensayos ligados a enzimas para detectar polipéptidos. También se pueden usar otras técnicas tales como hibridación in situ, tinción enzimática e inmunotinción, para detectar la presencia o expresión de polipéptidos y/o ácidos nucleicos. Se conocen métodos para llevar a cabo todas las técnicas mencionadas.

Como una alternativa, se puede cribar en una población de plantas con sucesos de transformación independientes las plantas que tienen un rasgo deseado, tal como producción de glucósido de vainillina. La selección y/o cribado se puede llevar a cabo en una o más generaciones y/o en más de una localización geográfica. En algunos casos, las plantas transgénicas se pueden cultivar y seleccionar en condiciones que inducen un fenotipo deseado o que es necesario de otra forma para producir un fenotipo deseado en una planta transgénica. Además, la selección y/o cribado se puede aplicar durante una etapa del desarrollo particular en el que se espera que la planta presente el

fenotipo. La selección y/o cribado se puede llevar a cabo para seleccionar las plantas transgénicas que tienen una diferencia estadísticamente significativa en el nivel de vainillina o beta-D-glucósido de vainillina con respecto a una planta de control que carece del transgén.

Homólogos funcionales

5 Los homólogos funcionales de los polipéptidos descritos antes también son adecuados para usar en la producción de vainillina o glucósido de vainillina en un hospedante recombinante. Por lo tanto, el hospedante recombinante puede incluir uno o más ácidos nucleicos heterólogos que codifican homólogos funcionales del polipéptido descrito antes y/o un ácido nucleico heterólogo que codifica un polipéptido de COMT o AROM mutante como se describe en la presente memoria. Un homólogo funcional es un polipéptido que tiene similitud de secuencia con un polipéptido de referencia, y que lleva una o más funciones bioquímicas o fisiológicas del polipéptido de referencia. Un homólogo funcional y el polipéptido de referencia pueden ser polipéptidos naturales, y la similitud de secuencia se puede deber a sucesos convergentes o divergentes. Como tales, los homólogos funcionales a veces se denominan en la bibliografía homólogos, ortólogos o parálogos. Las variantes de los homólogos funcionales naturales, tales como polipéptidos codificados por mutantes de una secuencia codificante de tipo natural, pueden ser ellas mismas homólogos funcionales. Los homólogos funcionales también se pueden crear por mutagénesis dirigida al sitio de la secuencia codificante de un polipéptido, o por combinación de dominios de secuencias codificantes de diferentes polipéptidos naturales ("intercambio de dominios"). Las técnicas para modificar genes que codifican polipéptidos de AROM y/o COMT funcionales descritos en la presente memoria son conocidas e incluyen, entre otras, técnicas de evolución dirigida, técnicas de mutagénesis dirigida al sitio y técnicas de mutagénesis aleatoria, y pueden ser útiles para aumentar la actividad específica de un polipéptido, alterar la especificidad de sustrato, alterar los niveles de expresión, alterar la localización subcelular o modificar las interacciones polipéptido:polipéptido de una forma deseada. Dichos polipéptidos modificados se consideran homólogos funcionales. La expresión "homólogo funcional" a veces se aplica al ácido nucleico que codifica un polipéptido funcionalmente homólogo.

25 Los homólogos funcionales se pueden identificar por análisis de alineamientos de nucleótidos y polipéptidos. Por ejemplo, realizar una búsqueda en una base de datos de secuencias de nucleótidos o polipéptidos puede identificar homólogos de polipéptidos de AROM o COMT, polipéptidos de 3DSD, ACAR, PPTasa o UGT. Los análisis de secuencia pueden implicar análisis por BLAST, Reciprocal BLAST o PSI-BLAST de bases de datos no redundantes que usan una secuencia de aminoácidos de AROM o COMT, 3DSD, ACAR, una PPTasa, o secuencia de aminoácidos de UGT como la secuencia de referencia. La secuencia de aminoácidos a veces se deduce de la secuencia de nucleótidos. Los polipéptidos en la base de datos que tienen más de 40% de identidad de secuencia son candidatos para la evaluación adicional de la idoneidad como polipéptido para la biosíntesis de vainillina o glucósido de vainillina. La similitud de secuencias de aminoácidos permite sustituciones de aminoácidos conservadoras, tales como sustitución de un resto hidrófobo por otro o sustitución de un resto polar por otro. Si se desea, se puede llevar a cabo la inspección manual de dichos candidatos con el fin de estrechar el número de candidatos que se van a evaluar más. La inspección manual se puede llevar a cabo seleccionando los candidatos que parece que tienen dominios presentes en polipéptidos de AROM o COMT o polipéptidos de biosíntesis de vainillina, p. ej., dominios funcionales conservados.

40 Las regiones conservadas se pueden identificar localizando una región dentro de la secuencia de aminoácidos primaria de un polipéptido que es una secuencia repetida, forma alguna estructura secundaria (p. ej., hélices y láminas beta), establece dominios cargados positivos o negativos, o representa un motivo o dominio de proteína. Véase, p. ej., la base de datos Pfam que describe secuencias consenso para una variedad de motivos y dominios de proteínas. La información incluida en la base de datos Pfam se describe en Sonnhammer, et al. (1998) *Nucl. Acids Res.* 26:320-322; Sonnhammer et al. (1997) *Proteins* 28:405-420; y Bateman et al. (1999) *Nucl. Acids Res.* 27:260-262. Las regiones conservadas también se pueden determinar alineando secuencias de los mismos polipéptidos o relacionados de especies estrechamente relacionadas. Especies estrechamente relacionadas preferiblemente son de la misma familia. En algunas realizaciones, es adecuado el alineamiento de secuencias de dos especies diferentes.

50 Típicamente, los polipéptidos que presentan al menos aproximadamente 40% de identidad de secuencia de aminoácidos son útiles para identificar regiones conservadas. Las regiones conservadas de polipéptidos relacionados presentan al menos 45% de identidad de secuencia de aminoácidos (p. ej., al menos 50%, al menos 60%, al menos 70%, al menos 80% o al menos 90% de identidad de secuencia de aminoácidos). En algunas realizaciones, una región conservada presenta al menos 92%, 94%, 96%, 98% o 99% de identidad de secuencia de aminoácidos. La identidad de secuencia se puede determinar como se ha expuesto antes.

La invención se describirá además en los siguientes ejemplos.

55 Ejemplo 1: Cepa indicadora de levadura para la producción de glucósido de vainillina a partir de glucosa

Se creó una cepa de levadura genéticamente estable productora de glucósido de vainillina a partir de glucosa como se describe en Brochado, et al. (2010) *Microb. Cell Fact.* 9:84, es decir, cepa VG4 con una eliminación de gen de PDC1 (piruvato descarboxilasa) y GDH1 (glutamato deshidrogenasa), y que sobreexpresaba GDH2. Además, la

cepa albergaba una construcción de expresión que contiene una PPTasa integrada en la región interlocus ECM3 del genoma de levadura. La expresión de la secuencia que codifica la PPTasa de *Corynebacterium glutamicum* se controló por el promotor TPI1 de levadura (Hansen, et al. (2009) *Appl. Environ. Microbiol.* 75(9):2765-74. Epub 2009 Mar 13). La cepa resultante se designó V12.

5 Ejemplo 2: Construcción de una AROM que carece de dominio 5

Se aislaron los 3912 pb más cercanos a 5' del gen *ARO1* de levadura, que incluyen todos los dominios funcionales excepto el dominio 5 (que tiene la actividad de shikimato deshidrogenasa), por amplificación por PCR de ADN genómico preparado a partir de la cepa de *S. cerevisiae* S288C, usando polimerasa de PCR correctora de errores. El fragmento de ADN resultante se subclonó en el vector de pTOPO y se secuenció para confirmar la secuencia de ADN. La secuencia de ácido nucleico y la correspondiente secuencia de aminoácidos se presentan en la SEQ ID NO: 1 y SEQ ID NO: 2, respectivamente. Este fragmento se sometió a digestión de restricción con *SpeI* y *SalI* y se clonó en los correspondientes sitios de restricción en el vector de expresión de levadura de alto número de copias p426-GPD (un vector basado en 2 μ), a partir del cual se puede expresar el gen insertado por el promotor fuerte GPD1 de levadura constitutivo. El plásmido resultante se denominó pVAN133.

15 Ejemplo 3: AROM de levadura con sustituciones de un solo aminoácido en el dominio 5

Todos los polipéptidos de AROM mutantes descritos en este ejemplo son polipéptidos de SEQ ID NO: 4, en donde un aminoácido se ha sustituido por otro aminoácido. Los polipéptidos de AROM mutantes se nombran como sigue: XnnnY, donde nnn indica la posición en la SEQ ID NO:4 del aminoácido, el cual se sustituye, X es el código de una letra del aminoácido en la posición nnn en la SEQ ID NO: 4 e Y es el código de una letra del aminoácido que sustituye a X. A modo de ejemplo A1533P se refiere a un polipéptido de AROM mutante de SEQ ID NO: 4, donde la alanina en la posición 1533 se sustituye por una prolina.

Se aislaron los 4764 pb completos del gen *ARO1* de levadura por amplificación por PCR de ADN genómico preparado a partir de la cepa de *S. cerevisiae* S288C, usando polimerasa de PCR correctora de errores. El fragmento de ADN resultante se subclonó en el vector de pTOPO y se secuenció para confirmar la secuencia de ADN. La secuencia de ácido nucleico y la correspondiente secuencia de aminoácidos se presentan en la SEQ ID NO: 3 y SEQ ID NO: 4, respectivamente. Este fragmento se sometió a digestión de restricción con *SpeI* y *SalI* y se clonó en los correspondientes sitios de restricción en el vector de expresión de levadura de bajo número de copias p416-TEF (un vector basado en CEN-ARS), a partir del cual se puede expresar el gen a partir del promotor fuerte TEF. El plásmido resultante se denominó pVAN183.

30 El plásmido pVAN183 se usó para hacer 10 mutantes de *ARO1* con dominio 5 diferentes, usando el kit de mutagénesis dirigida al sitio QUICKCHANGE II (Agilent Technologies). Con referencia a la SEQ ID NO: 4, los mutantes contenían las siguientes sustituciones de aminoácidos: A1533P, P1500K, R1458W, V1349G, T1366G, I1387H, W1571V, T1392K, K1370L y A1441P.

35 Después de confirmar la secuencia de estos genes de AROM mutantes, los plásmidos de expresión que contenían las sustituciones A1533P, P1500K, R1458W, V1349G, T1366G, I1387H, W1571V, T1392K, K1370L y A1441P se designaron pVAN368-pVAN377, respetivamente. La secuencia de ácido nucleico y las correspondientes secuencias de aminoácidos de los mutantes de AROM se dan en la tabla 1.

Tabla 1

Mutante	Plásmido	Secuencia de ácido nucleico	Secuencia de aminoácidos
A1533P	pVAN368	SEQ ID NO: 5	SEQ ID NO: 6
P1500K	pVAN369	SEQ ID NO: 7	SEQ ID NO: 8
R1458W	pVAN370	SEQ ID NO: 9	SEQ ID NO: 10
V1349G	pVAN371	SEQ ID NO: 11	SEQ ID NO: 12
T1366G	pVAN372	SEQ ID NO: 13	SEQ ID NO: 14
I1387H	pVAN373	SEQ ID NO: 15	SEQ ID NO: 16
W1571V	pVAN374	SEQ ID NO: 17	SEQ ID NO: 18
T1392K	pVAN375	SEQ ID NO: 19	SEQ ID NO: 20
K1370L	pVAN376	SEQ ID NO: 21	SEQ ID NO: 22

A1441P	pVAN377	SEQ ID NO: 23	SEQ ID NO: 24
--------	---------	---------------	---------------

Ejemplo 4: Proteína de fusión de AROM y deshidratasa 3DHS de levadura

Se aislaron los 3951 pb más cercanos a 5' del gen *ARO1* de levadura, que incluyen todos los dominios funcionales excepto el dominio 5 con la actividad de shikimato deshidrogenasa, por amplificación por PCR de ADN genómico preparado a partir de la cepa de *S. cerevisiae* S288C, usando polimerasa de PCR correctora de errores. El fragmento de ADN resultante se subclonó en el vector de pTOPO y se secuenció para confirmar la secuencia de ADN. Con el fin de fusionar este fragmento con el gen de la 3-deshidroschikimato deshidratasa (3DSD) de la ruta de la vainillina, se insertó el gen de 3DSD de *Podospora pauciseta* (Hansen, et al. (2009) véase antes) en los sitios *XmaI-EcoRI* del vector de expresión p426-GPD, y después el fragmento de *ARO1* clonado se liberó y se insertó en los sitios *SpeI-XmaI* de la construcción resultante. El gen de fusión final se expresa a partir del promotor fuerte GPD1 de levadura constitutivo. El plásmido resultante se denominó pVAN132. La secuencia de ácido nucleico y la correspondiente secuencia de aminoácidos de esta proteína de fusión se presentan en la SEQ ID NO: 25 y SEQ ID NO: 26, respectivamente.

Ejemplo 5: Expresión de enzimas AROM mutantes o de fusión en levaduras que ya biosintetizan glucósido de vainillina.

Cada uno de los plásmidos descritos en los ejemplos 2, 3 y 4 se introdujo en la cepa de levadura V12 por transformación, usando el protocolo de transformación de acetato de litio, que da como resultado las siguientes cepas de levaduras: V12-Aro1-1 (que contiene el plásmido pVAN133), V12-Aro1-2 (que contiene el plásmido pVAN132), V12-1-3 (que contiene el plásmido pVAN183), V12-Aro1-4 (que contiene el plásmido pVAN368), V12-Aro1-5 (que contiene el plásmido pVAN369), V12-Aro1-6 (que contiene el plásmido pVAN370), V12-Aro1-7 (que contiene el plásmido pVAN371), V12-Aro1-8 (que contiene el plásmido pVAN372), V12-Aro1-9 (que contiene el plásmido pVAN373), V12-Aro1-10 (que contiene el plásmido pVAN374), V12-Aro1-11 (que contiene el plásmido pVAN375), V12-Aro1-12 (que contiene el plásmido pVAN376) and V12-Aro1-13 (que contiene el plásmido pVAN377).

Las cepas de levadura V12-Aro1-1, V12-Aro1-2, V12-1-3, V12-Aro1-4, V12-Aro1-5, V12-Aro1-6, V12-Aro1-7, V12-Aro1-8, V12-Aro1-9, V12-Aro1-10, V12-Aro1-11, V12-Aro1-12, y V12-Aro1-13 se cultivaron como cultivos de 200 ml en matraces de agitación Erlenmeyer de 500 ml usando medio de crecimiento SC (completo sintético) sin aminoácidos aromáticos, a 30°C con revoluciones moderadas (150 rpm) durante 72 horas. Se tomaron muestras a las 48 horas, y se determinó el contenido de glucósido de vainillina. Se incluyó como controles la cepa de levadura V12 que contiene los vectores vacíos p416-TEF o p426-GPD. La producción de glucósido de vainillina (VG) en las cepas de control (que contienen los plásmidos vacíos p416- y p426-GPD) era típicamente aproximadamente 250 mg/l. Se espera que la expresión de AROM con dominio 5 truncado aumente la producción de VG (cepa V12-Aro1-1) y se espera que la fusión física adicional de esta AROM truncada con la primera enzima comprometida en la ruta de la vainillina heteróloga (*Podospora pauciseta* 3DSD) de como resultado un aumento adicional en la producción del glucósido de vainillina (cepa V12-Aro1-2).

De las versiones mutantes de AROM, en las que se cambiaban aminoácidos individuales del dominio 5, los T1392K (cepa V12-Aro1-9) y K1370L (cepa V12-Aro1-10) pueden ser útiles para la mayor producción de VG. Por ejemplo, se puede observar un aumento de aproximadamente 30-35% en la producción de VG para T1392K (cepa V12-Aro1-9) y K1370L (cepa V12-Aro1-10).

Este ejemplo demuestra que mediante la sobreexpresión del polipéptido de AROM mutante con menor actividad de shikimato deshidrogenasa, se puede aumentar la concentración celular de 3-DHS suficientemente para tener una función en el rendimiento de la valoración final en la ruta heteróloga. Este ejemplo también demuestra que la fusión de la primera enzima de la ruta heteróloga, es decir, 3DHD, con una enzima AROM truncada da como resultado un aumento del flujo en la ruta heteróloga, obteniendo canalización del sustrato. Finalmente, los experimentos descritos en la presente memoria indicaban que cambiando aminoácidos discretos en el dominio de AROM que metaboliza naturalmente 3DHS, el compuesto necesario para la producción de vainillina, se puede aumentar la cantidad de 3DHS disponible para la biosíntesis de vainillina.

Ejemplo 6: Mutantes de COMT

Todos los polipéptidos de COMT mutantes descritos en este ejemplo son polipéptidos de SEQ ID NO: 27, en donde un aminoácido se ha sustituido por otro aminoácido. Los polipéptidos de COMT mutantes se nombran como sigue: XnnnY, donde nnn indica la posición en la SEQ ID NO: 27 del aminoácido, el cual se sustituye, X es el código de una letra del aminoácido en la posición nnn en la SEQ ID NO: 27 e Y es el código de una letra del aminoácido que sustituye a X. A modo de ejemplo L198Y se refiere a un polipéptido de COMT mutante de SEQ ID NO: 27, donde la leucina en la posición 198 se sustituye por un triptófano.

ES 2 719 304 T3

Los ácidos nucleicos que codifican los polipéptidos de COMT mutantes se construyeron por PCR usando cebadores que contienen los codones deseados. Las PCR se hicieron bien como una PCR individual o usando la PCR de extensión con solapamiento de secuencias (SOE) por procedimientos convencionales. Esta etapa la pueden hacer, por ejemplo, proveedores comerciales, tales como Life Technologies. Los cebadores usados se dan en la tabla 2.

5 TABLA 2

Cebador	Secuencia	SEQ ID NO:
HSOMT_F	CGTAGCATGCAGTCTAGAAAAATGGGTGACACTAAGGAG C	28
HSOMT_R	GACGACGTTAGTGACAgaaattcTTATGGACCAGCTTCAG AACCTG	29
HSOMT2_F	CGTAGCATGCAGTCTAGAAAAATGG	30
HSOMT2_R	GACGACGTTAGTGACAgaaattc	31
W38NRK_F	CTATTGACACTTATTGTGAGCAAAAGGAGnrKGCTATGA ACGTTG	32
W38NYK_F	CTATTGACACTTATTGTGAGCAAAAGGAGnyKGCTATGA ACGTTG	33
W38_R	CTCCTTTTGCTCACAATAAGTGCAATAG	34
M40NRK_F	GACACTTATTGTGAGCAAAAGGAGTGGGCTnrKAACGTT GGTGAC	35
M40NRK_R	CACTTATTGTGAGCAAAAGGAGTGGGCTnyKAACGTTGG TGAC	36
M40_R	CTCCTTTTGCTCACAATAAGTGTC	37
W143NRK_F	CTTTGGACATGGTTTTCTTGGACCATnrKAAGGACAGAT ATTTGCC	38
W143NYK_F	CTTTGGACATGGTTTTCTTGGACCATnyKAAGGACAGAT ATTTGCC	39
W143_R	ATGGTCCAAGAAAACCATGTCCAAAG	40
C173NRK_F	GTA CTGTTTTGTTAGCTGACAACGTTATTnrKccaGGTG CTCCAGACTTCTTG	41
C173NYK_F	GTA CTGTTTTGTTAGCTGACAACGTTATTnyKccaGGTG CTCCAGACTTCTTG	42
C173_R	AATAACGTTGTCAGCTAACAAAACAGTAC	43
P174NRK_F	CTGTTTTGTTAGCTGACAACGTTATTtgnrKGGTGCTC CAGACTTC	44
P174NYK_F	CTGTTTTGTTAGCTGACAACGTTATTtgnnyKGGTGCTC CAGACTTC	45
P174_R	ACAAATAACGTTGTCAGCTAACAAAACAG	46
L198NRK_R	atttagaattcTTATGGACCAGCTTCAGAACCTGGACCC TTATATATAGCCTTCTCCAAACCGTCAACAACCTCTCTA TATTCMYNGAAAGATTGATAATGAG	47

Cebador	Secuencia	SEQ ID NO:
L198NYK_R	atttagaattcTATGGACCAGCTTCAGAACCTGGACCCT TATATATAGCCTTCTCCAAACCGTCAACAACCTCTCTAT ATTCMRNGAAAGATTGATAATGAG	48
E199NRK_R	atttagaattcTTATGGACCAGCTTCAGAACCTGGACCC TTATATATAGCCTTCTCCAAACCGTCAACAACCTCTCTA TAMYNCAAGAAAGATTGATAATG	49
E199NYK_R	atttagaattcTTATGGACCAGCTTCAGAACCTGGACCC TTATATATAGCCTTCTCCAAACCGTCAACAACCTCTCTA TAMRNCAAGAAAGATTGATAATG	50
R201NRK_R	atttagaattcTTATGGACCAGCTTCAGAACCTGGACCC TTATATATAGCCTTCTCCAAACCGTCAACAACCTCMYNA TATTCGAAGAAAGATTG	51
R201NYK_R	atttagaattcTTATGGACCAGCTTCAGAACCTGGACCC TTATATATAGCCTTCTCCAAACCGTCAACAACCTCMRNA TATTCGAAGAAAGATTG	52

En las secuencias de cebadores proporcionadas antes: M puede ser A o C; R puede ser A o G; W puede ser A o T; S puede ser G o C; Y puede ser C o T; K puede ser G o T; V puede ser A, G o C; H puede ser A, C o T; D puede ser A, G o T; B puede ser G, C o T; y N puede ser A, G, C o T.

- 5 Se incluyeron sitios de restricción para *EcoRI/XbaI* en los cebadores para facilitar la clonación de los productos de la PCR en el vector de expresión de levadura centromérico p416-TEF (Mumberg, et al. (1995) *Gene* 156 (1):119-22). Los plásmidos resultantes se transformaron en una cepa de levadura EFSC2055 (Genotipo: Mata his3D1 leu2D0 met15D0 ura3D0 adh6::LEU2 bgl1::KanMX4 PTP11::3DSD[AurC]::(HsOMT::MET15[NatMX])::ACAR[HphMX]::UGT7 2E2[HIS3] ECM3::(CorPPTase-SchAP4). Esta cepa de levadura se basa en la cepa VG4 (Brochado, et al. (2010) *Microbial. Cell Factories* 9:84), e incluye además una alteración de HsOMT con el marcador MET15, y tiene dos genes adicionales integrados, en concreto *Corynebacterium glutamicum* PPTase (nº de acceso en la base de datos NCBI NP_601186) y *S. cerevisiae* HAP4 (nº de acceso en la base de datos NCBI Z28109).

15 La HsOMT se alteró por amplificación por PCR de *S. cerevisiae* MET15 (marcador de selección auxotrófico metionina) usando cebadores con colas de 70 pb homólogos al extremo frontal y posterior de HsOMT, respectivamente. Después la cepa de levadura se transformó con el producto de la PCR, dando como resultado transformantes que no tenían actividad de HsOMT y capaces de crecer en placas no complementadas con metionina. La cepa resultante se designó EFSC2055,

20 Los plásmidos que expresan CorPPTasa y SchAP4 se transformaron en la cepa predecesora EFSC2055 usando el protocolo convencional de transformación de levaduras con acetato de litio/PEG y los transformantes se seleccionaron en placas de SC-uracilo. Los transformantes se ensayaron por cultivo en cultivos de 3 ml durante 72 horas en medio Delft complementado con 8% de melazas de remolacha azucarera.

25 Se analizaron los cultivos usando HPLC-UV para cuantificar el glucósido de vainillina/glucósido de isovainillina y productos relacionados. El análisis por HPLC se llevó a cabo con un sistema AGILENT 1100 series con bomba binaria y una columna Phenomenex Synergi Polar-RP 2,5 u 100 Å 100x2,00 mm, que separa los precursores y la isovainillina y vainillina. Se ejecutó un gradiente plano con agua/acetonitrilo + ácido trifluoroacético al 0,1%. Se llevó a cabo un programa de 8,9 minutos + 1,1 minutos postejecución como se presenta en la tabla 3.

TABLA 3

Tiempo	% de Acetonitrilo	Caudal ml/min
0	5	0,5
0,7	5	0,5
5,7	27	0,5

6,2	100	0,5
6,6	100	0,7
7,8	100	1,0
8,1	100	1,0
8,6	5	0,8
8,9	5	0,6

5 El glucósido de vainillina y glucósido de isovainillina se cuantificaron integrando el área de los picos de HPLC y comparando los mismos con una curva patrón. Los resultados de este análisis se muestran en la figura 3A. Las células de levadura que expresaban Hs-OMT de SEQ ID NO: 27 (denominadas Hs-COMT wt) producían isovainillina y vainillina en una relación 1:3, mientras que el mutante L198Y producía isovainillina y vainillina en una relación de aproximadamente 1:125. La relación exacta era difícil de determinar ya que la producción de isovainillina estaba en el límite de detección.

Además de las mutaciones analizadas en la figura 3A, buena especificidad y baja producción de isovainillina en los mutantes L198C, L198N, L198D, L198F y L198E.

10 Ejemplo 7: Reducción de alcohol de vainillilo

A modo de ilustración, los genes de VAO de *P. simplicissium* (nº de acceso en GENBANK P56216) y *R. jostii* (nº de acceso en GENBANK YP_703243.1) se aislaron y se clonaron en un vector de expresión de levaduras. Los vectores de expresión se transformaron posteriormente en una cepa de levadura que expresa glucosiltransferasa. Se ensayó en las cepas transformadas la actividad de VAO por cultivo de la levadura durante 48 horas en medio complementado con alcohol de vainillilo 3 mM. Los resultados de este análisis se presentan en la figura 4. Las enzimas VAO tanto de *P. simplicissium* como *R. jostii* presentaban actividad en levaduras. Cuando las enzimas VAO se analizaron en una cepa capaz de producir glucósido de vainillina, hubo una reducción en la acumulación de alcohol de vainillilo durante la fermentación del glucósido de vainillina.

Ejemplo 8: Gen de ACAR de *Neurospora crassa*

20 Como alternativa a una proteína ACAR (EC 1.2.1.30) de *Nocardia iowensis* (Hansen, et al. ((2009) *Appl. Environ. Microbiol.* 75:2765-74), se investigó el uso de una enzima ACAR de *Neurospora crassa* (Gross & Zenk (1969) *Eur. J. Biochem.* 8:413-9; US 6.372.461) en levaduras, ya que *Neurospora* (moho común del pan) es un organismo GRAS. Se aisló un gen de *N. crassa* (GENBANK XP_955820) con homología con la ACAR de *Nocardia iowensis* y se clonó en un vector de expresión de levadura. El vector se transformó en una cepa de levadura que expresa una PPtasa, se seleccionaron cepas según la presencia del gen de ACAR, y la levadura seleccionada se cultivó durante 72 horas en medio complementado con ácido vainílico 3 mM para demostrar la actividad de ACAR. Los resultados de este análisis se presentan en la figura 5. Se encontró que la enzima ACAR de *N. crassa* presenta una actividad mayor en levadura que la ACAR de *N. iowensis*. Por lo tanto, en algunas realizaciones del método descrito en la presente memoria, se usa una enzima ACAR de *N. crassa* en la producción de vainillina o glucósido de vainillina.

30 Además de las proteínas ACAR de *N. iowensis* o *N. crassa* ACAR, está contemplado que se pueden usar otras proteínas ACAR, incluyendo las aisladas de *Nocardia brasiliensis* (nº de acceso en GENBANK EHY26728), *Nocardia farcinica* (nº de acceso en GENBANK BAD56861), *Podospira anserina* (nº de acceso en GENBANK CAP62295), o *Sordaria macropora* (nº de acceso en GENBANK CCC14931), con identidad de secuencia significativa con la proteína ACAR de *N. iowensis* o *N. crassa*.

35

LISTA DE SECUENCIAS

- <110> International Flavors & Fragrances Inc.
 Evolva SA
 Hansen, Joergen
 5 Hansen, Esben H.
 Sompalli, Honey Polur
 Sheridan, Joseph M.
 Heal, Jonathan R.
 Hamilton, William D.O.
- 10 <120> COMPOSICIONES Y MÉTODOS PARA LA BIOSÍNTESIS DE VAINILLINA O BETA-D-GLUCÓSIDO DE VAINILLINA
 <130> IFF0063WO
 <150> US 61/521,090
 < 151> 2011-08-08
- 15 <150> US 61/522,096
 < 151> 2011-08-10
 <160> 56
 <170> PatentIn versión 3.5
- 20 <210> 1
 < 211> 3912
 < 212> ADN
 < 213> Saccharomyces cerevisiae
- <400> 1
 atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac 60
 attcatgacc atttggttga aaccataaatt aaacattgtc cttcttcgac atacgttatt 120
 tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag 180
 gcttctttgc cagaaggctc tcgtttactt acttatgttg ttaaaccagg tgagacaagt 240
 aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt 300
 gatacggtta tggtagcgat cgggtggtgt gttattggtg acatgattgg gttcgttgca 360
 tctacattta tgagaggtgt tcgtggtgtc caagtaccaa catccttatt ggcaatggtc 420
 gattcctoca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt 480
 gcattttggc aaccaaaatt tgccttgta gatattaaat ggctagaaac gttagccaag 540
 agagagtta tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac 600
 gaatttacta gattagaatc aaacgcttcg ttgttcttaa atgttgtaa tggggcaaaa 660
 aatgtcaagg ttaccaatca attgacaaac gagattgacg agatatacga tacagatatt 720
 gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa 780
 gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct 840
 attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc 900
 attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa 960

25

ES 2 719 304 T3

gttgcacgtc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa 1020
 tggtttaaag agctaaccct acataagaaa acaccattgg atatcttatt gaagaaaatg 1080
 agtattgaca agaaaaacga gggttccaaa aagaaggtagg tcattttaga aagtattggt 1140
 aagtgctatg gtgactccgc tcaatttggt agcgatgaag acctgagatt tattctaaca 1200
 gatgaaacct tcgtttacct cttcaaggac atccctgctg atcaacagaa agttgttatc 1260
 cccctggtt ctaagtccat ctccaatcgt gctttaattc ttgctgccct cgggtaaggt 1320
 caatgtaaaa tcaagaactt attacattct gatgatacta aacatatgtt aaccgctggt 1380
 catgaattga aagggtctac gatatcatgg gaagataatg gtgagacggg agtgggtgaa 1440
 ggacatggtg gttccacatt gtcagcttgt gctgaccctt tataatctagg taatgcaggt 1500
 actgcatcta gatttttgac ttccttggtc gccttggtca attctacttc aagccaaaag 1560
 tataatggtt taactggtaa cgcaagaatg caacaaagac caattgctcc tttggtcgat 1620
 tctttgctg ctaatggtac taaaattgag tacttgaata atgaaggttc cctgccaatc 1680
 aaagtttata ctgattcggg attcaaagggt ggtagaattg aattagctgc tacagtttct 1740
 tctcagtacg tatcctctat cttgatgtgt gcccatacgt ctgaagaacc tgtaactttg 1800
 gctcttggtg gtggttaagcc aatctctaaa ttgtacgtcg atatgacaat aaaaatgatg 1860
 gaaaaattcg gtatcaatgt tgaacttct actacagaac cttacactta ttatattcca 1920
 aaggacatt atattaacct atcagaatac gtcattgaaa gtgatgcctc aagtgtaca 1980
 taccattgg ccttcgccc aatgactggg actaccgtaa cggttccaaa cattggtttt 2040
 gagtcttac aagggtgatg cagatttgca agagatgtct tgaaacctat gggttgtaaa 2100
 ataactcaa cggcaacttc aactactggt tgggtcctc ctgtaggtac tttaaagcca 2160
 ttaaaacatg ttgatatgga gccaatgact gatgcgttct taactgcatg tgttgttgcc 2220
 gctatttcgc acgacagtga tccaaattct gcaaatacaa ccaccattga aggtattgca 2280
 aaccagcgtg tcaaagagtg taacagaatt ttggccatgg ctacagagct cgccaaattt 2340
 ggcgtcaaaa ctacagaatt accagatggg attcaagtcc atggtttaa ctcgataaaa 2400
 gatttgaagg ttccttccga ctctctgga cctgtcggtg tatgcacata tgatgatcat 2460
 cgtgtggcca tgagtttctc gcttctgca ggaatggtaa attctcaaaa tgaacgtgac 2520
 gaagttgcta atoctgtaag aatacttgaa agacattgta ctggtaaaac ctggcctggc 2580
 tgggtgggatg tgttacattc cgaactaggt gccaaattag atgggtgcaga acctttagag 2640
 tgcacatcca aaaagaactc aaagaaaagc gttgtcatta ttggcatgag agcagctggc 2700
 aaaactacta taagtaaatg gtgcgcatcc gctctgggtt acaaattagt tgacctagac 2760
 gagctggtt agcaacagca taacaatcaa agtgtaaac aatttgttgt ggagaacggg 2820

ES 2 719 304 T3

tgggagaagt tccgtgagga agaaacaaga attttcaagg aagttattca aaattacggc 2880
 gatgatggat atgttttctc aacaggtggc ggtattggtg aaagcgctga gtctagaaaa 2940
 gccttaaaag attttgcctc atcaggtgga tacgttttac acttacatag ggatattgag 3000
 gagacaattg tctttttaca aagtgatcct tcaagacctg cctatgtgga agaaattcgt 3060
 gaagtttgga acagaagga ggggtggtat aaagaatgct caaatcttc tttctttgct 3120
 cctcattgct ccgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
 gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
 ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
 gaggtctgag aggtcagagt agaccatttg gctaactact ctgctgattt cgtgagtaaa 3360
 cagttatcta tattgcgtaa agccactgac agtattccta tcatttttac tgtgcgaacc 3420
 atgaagcaag gtggcaactt tcttgatgaa gagttcaaaa ccttgagaga gctatacgat 3480
 attgccttga agaatggtgt tgaattcctt gacttagaac taactttacc tactgatatc 3540
 caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
 caaggattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
 cttgatgtgg atgttgtaaa atttgtgggt acggctgta atttcgaaga taatttgaga 3720
 ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagtaa tatgacttct 3780
 aaaggtagca tttctcgtgt tttgaataat gtttaacac ctgtgacatc agatttattg 3840
 cctaactccg ctgcccttg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
 atgggaggtt ga 3912

<210> 2

<211> 1303

<212> PRT

5 <213> *Saccharomyces cerevisiae*

<400> 2

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
 1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
 20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
 35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
 50 55 60

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

ES 2 719 304 T3

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

ES 2 719 304 T3

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
 325 330 335

Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
 340 345 350

Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly
 355 360 365

Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly
 370 375 380

Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr
 385 390 395 400

Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln
 405 410 415

Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu
 420 425 430

Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu
 435 440 445

His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys
 450 455 460

Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu
 465 470 475 480

Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495

Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510

Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525

Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540

Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560

Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala
 565 570 575

ES 2 719 304 T3

Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
580 585 590

Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
595 600 605

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
725 730 735

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
805 810 815

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
820 825 830

ES 2 719 304 T3

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
885 890 895

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu
900 905 910

Gly Tyr Lys Leu Val Asp Leu Asp Glu Leu Phe Glu Gln Gln His Asn
915 920 925

Asn Gln Ser Val Lys Gln Phe Val Val Glu Asn Gly Trp Glu Lys Phe
930 935 940

Arg Glu Glu Glu Thr Arg Ile Phe Lys Glu Val Ile Gln Asn Tyr Gly
945 950 955 960

Asp Asp Gly Tyr Val Phe Ser Thr Gly Gly Gly Ile Val Glu Ser Ala
965 970 975

Glu Ser Arg Lys Ala Leu Lys Asp Phe Ala Ser Ser Gly Gly Tyr Val
980 985 990

Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
995 1000 1005

Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
1010 1015 1020

Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
1025 1030 1035

Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
1040 1045 1050

Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
1055 1060 1065

Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe

ES 2 719 304 T3

1070		1075		1080
Asp Asp Leu Thr Glu Gln Thr		Glu Asn Leu Thr Pro		Ile Cys Tyr
1085		1090		1095
Gly Cys Glu Ala Val Glu Val		Arg Val Asp His Leu		Ala Asn Tyr
1100		1105		1110
Ser Ala Asp Phe Val Ser Lys		Gln Leu Ser Ile Leu		Arg Lys Ala
1115		1120		1125
Thr Asp Ser Ile Pro Ile Ile		Phe Thr Val Arg Thr		Met Lys Gln
1130		1135		1140
Gly Gly Asn Phe Pro Asp Glu		Glu Phe Lys Thr Leu		Arg Glu Leu
1145		1150		1155
Tyr Asp Ile Ala Leu Lys Asn		Gly Val Glu Phe Leu		Asp Leu Glu
1160		1165		1170
Leu Thr Leu Pro Thr Asp Ile		Gln Tyr Glu Val Ile		Asn Lys Arg
1175		1180		1185
Gly Asn Thr Lys Ile Ile Gly		Ser His His Asp Phe		Gln Gly Leu
1190		1195		1200
Tyr Ser Trp Asp Asp Ala Glu		Trp Glu Asn Arg Phe		Asn Gln Ala
1205		1210		1215
Leu Thr Leu Asp Val Asp Val		Val Lys Phe Val Gly		Thr Ala Val
1220		1225		1230
Asn Phe Glu Asp Asn Leu Arg		Leu Glu His Phe Arg		Asp Thr His
1235		1240		1245
Lys Asn Lys Pro Leu Ile Ala		Val Asn Met Thr Ser		Lys Gly Ser
1250		1255		1260
Ile Ser Arg Val Leu Asn Asn		Val Leu Thr Pro Val		Thr Ser Asp
1265		1270		1275
Leu Leu Pro Asn Ser Ala Ala		Pro Gly Gln Leu Thr		Val Ala Gln
1280		1285		1290
Ile Asn Lys Met Tyr Thr Ser		Met Gly Gly		
1295		1300		

<210> 3
 <211> 4767
 <212> ADN
 <213> Saccharomyces cerevisiae

ES 2 719 304 T3

<400> 3
atggtgcagt tagccaaagt cccaattcota ggaaatgata ttatccacgt tgggtataac 60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt 120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag 180
gcttctttgc cagaaggctc tcgtttactt acttatggtg ttaaaccagg tgagacaagt 240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtgggaag atgtactcgt 300
gatacggtta tggtagcgat cgggtgggtg gttattggtg acatgattgg gttcgttgca 360
tctacattta tgagaggtgt tcgtggtgtc caagtaccaa catccttatt ggcaatggtc 420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt 480
gcattttggc aacccaaatt tgccttcta gatattaaat ggctagaaac gttagccaag 540
agagagtta tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac 600
gaatttacta gattagaatc aaacgcttcg ttgttcttaa atggtggtta tggggcaaaa 660
aatgtcaagg ttaccaatca attgacaaac gagattgacg agatatcgaa tacagatatt 720
gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa 780
gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct 840
attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc 900
attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa 960
gttgcacgtc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa 1020
tggtttaaag agctaactt acataagaaa acaccattgg atatcttatt gaagaaaatg 1080
agtattgaca agaaaaacga gggttccaaa aagaaggtgg tcattttaga aagtattggt 1140
aagtgctatg gtgactccgc tcaatttggt agcgaatgaag acctgagatt tattctaaca 1200
gatgaaacct tcgtttacct cttcaaggac atccctgctg atcaacagaa agttgttacc 1260
ccccctggtt ctaagtccat ctccaatcgt gctttaaattc ttgctgccct cggatgaagg 1320
caatgtaaaa tcaagaactt attacattct gatgatacta aacatatgtt aaccgctggt 1380
catgaattga aaggtgctac gatatcatgg gaagataatg gtgagacggt agtgggtgaa 1440
ggacatggtg gttccacatt gtcagcttgt gctgaccocct tatatctagg taatgcaggt 1500
actgcatota gatTTTTGAC ttccttggct gccttgggtca attctacttc aagccaaaag 1560
tatatcgttt taactggtaa cgcaagaatg caacaaagac caattgctcc tttggtcgat 1620
tctttgcgtg ctaatggtag taaaattgag tacttgaata atgaaggttc cctgccaatc 1680
aaagtttata ctgattcggg attcaaaggg ggtagaattg aattagctgc tacagtttct 1740

ES 2 719 304 T3

tctcagtag	tatcctctat	cttgatgtgt	gccccatag	ctgaagaacc	tgtaactttg	1800
gctcctgttg	gtggttaagcc	aatctctaaa	ttgtacgtcg	atatgacaat	aaaaatgatg	1860
gaaaaattcg	gtatcaatgt	tgaaacttct	actacagaac	cttacctta	ttatattcca	1920
aagggacatt	atattaaccc	atcagaatac	gtcattgaaa	gtgatgcctc	aagtgtaca	1980
taccattgg	ccttcgccgc	aatgactgg	actaccgtaa	cggttcctaaa	cattggtttt	2040
gagtcgttac	aaggtgatgc	cagatttgca	agagatgtct	tgaaacctat	gggttgtaaa	2100
ataactcaaa	cggaacttc	aactactggt	togggtcctc	ctgtaggtac	tttaaagcca	2160
ttaaaacatg	ttgatatgga	gccaatgact	gatgcgttct	taactgcatg	tggtgtggc	2220
gctatttcgc	acgacagtga	tccaaattct	gcaaatacaa	ccaccattga	aggtattgca	2280
aaccagcgtg	tcaaagatg	taacagaatt	ttggccatgg	ctacagagct	cgccaaattt	2340
ggcgtcaaaa	ctacagaatt	accagatgg	attcaagtcc	atggtttaa	ctcgataaaa	2400
gatttgaag	ttccttccga	ctctctgga	cctgtcggtg	tatgcacata	tgatgatcat	2460
cgtgtggcca	tgagtttctc	gcttcttgca	ggaatggtaa	attctcaaaa	tgaacgtgac	2520
gaagttgcta	atcctgtaag	aatacttgaa	agacattgta	ctggtaaac	ctggcctggc	2580
tggtgggatg	tgttacattc	cgaactaggt	gccaaattag	atggtgcaga	accttagag	2640
tgacatcca	aaaagaactc	aaagaaaagc	gttgtcatta	ttggcatgag	agcagctggc	2700
aaaactacta	taagtaaatg	gtgcgcatcc	gctctgggtt	acaaattagt	tgacctagac	2760
gagctgtttg	agcaacagca	taacaatcaa	agtgttaaac	aatttgttgt	ggagaacggt	2820
tgggagaagt	tccgtgagga	agaaacaaga	atcttcaagg	aagttattca	aaattacggc	2880
gatgatggat	atgttttctc	aacaggtggc	ggtattgttg	aaagcgtga	gtctagaaaa	2940
gccttaaaag	atcttgcctc	atcaggtgga	tacgttttac	acttacatag	ggatattgag	3000
gagacaattg	tctttttaca	aagtgatcct	tcaagacctg	cctatgtgga	agaaattcgt	3060
gaagtttga	acagaaggga	gggttggtat	aaagaatgct	caaatttctc	tttctttgct	3120
cctcattgct	ccgcagaagc	tgagttccaa	gctctaagaa	gatcgtttag	taagtacatt	3180
gcaaccatta	caggtgtcag	agaaatagaa	attccaagcg	gaagatctgc	ctttgtgtgt	3240
ttaacctttg	atgacttaac	tgaacaaact	gagaatttga	ctccaatctg	ttatggttgt	3300
gaggctgtag	aggtcagagt	agaccatttg	gctaattact	ctgctgattt	cgtgagtaaa	3360
cagttatcta	tattgcgtaa	agccactgac	agtattccta	tcatttttac	tgtgcgaacc	3420
atgaagcaag	gtggcaactt	tctgatgaa	gagttcaaaa	ccttgagaga	gctatacgat	3480
attgccttga	agaatggtgt	tgaattcctt	gacttagaac	taactttacc	tactgatata	3540
caatatgagg	ttattaacaa	aaggggcaac	accaagatca	ttggttccca	tcatgacttc	3600

ES 2 719 304 T3

caaggattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga 3720
ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
aaaggtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
cctaactcgg ctgcccctgg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
atgggaggtg tcgagcctaa ggaactgttt gttgttgaa agccaattgg ccaactctaga 3960
tcgccaattt tacataacac tggctatgaa attttaggtt tacctcacia gttcgataaa 4020
tttgaaactg aatccgcaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt 4080
ggcggtgctg cagtcacaat tctctgaaa ttagatataa tgcagtacat ggatgaattg 4140
actgatgctg ctaaagttat tgggtgctgta aacacagtta taccattggg taacaagaag 4200
ttaaagggtg ataataccga ctgggttaggt atccgtaatg ccttaattaa caatggcgtt 4260
cccgaatatg ttggtcatac cgctggtttg gttatcggtg caggtggcac ttctagagcc 4320
gccctttacg ccttgcacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaatagag 4440
tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaagg 4560
gcccatgctg cttttgtacc aaccttattg gaagccgcat acaaaccaag cgttactccc 4620
gttatgacaa tttcacaga caaatatcaa tggcacggtt tccctggatc acaaatgtta 4680
gtacaccaag gtgtagctca gtttgaaaag tggacaggat tcaagggccc tttcaaggcc 4740
atttttgatg ccgttacgaa agagtag 4767

<210> 4

< 211> 1588

< 212> PRT

5 < 213> Saccharomyces cerevisiae

<400> 4

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
50 55 60

ES 2 719 304 T3

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

ES 2 719 304 T3

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
 325 330 335
 Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
 340 345 350
 Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly
 355 360 365
 Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly
 370 375 380
 Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr
 385 390 395 400
 Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln
 405 410 415
 Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu
 420 425 430
 Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu
 435 440 445
 His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys
 450 455 460
 Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu
 465 470 475 480
 Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495
 Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510
 Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525
 Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540
 Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560
 Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala

ES 2 719 304 T3

565 570 575

Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
580 585 590

Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
595 600 605

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
725 730 735

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
805 810 815

ES 2 719 304 T3

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
 820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
 835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
 850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
 865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
 885 890 895

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu
 900 905 910

Gly Tyr Lys Leu Val Asp Leu Asp Glu Leu Phe Glu Gln Gln His Asn
 915 920 925

Asn Gln Ser Val Lys Gln Phe Val Val Glu Asn Gly Trp Glu Lys Phe
 930 935 940

Arg Glu Glu Glu Thr Arg Ile Phe Lys Glu Val Ile Gln Asn Tyr Gly
 945 950 955 960

Asp Asp Gly Tyr Val Phe Ser Thr Gly Gly Gly Ile Val Glu Ser Ala
 965 970 975

Glu Ser Arg Lys Ala Leu Lys Asp Phe Ala Ser Ser Gly Gly Tyr Val
 980 985 990

Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
 995 1000 1005

Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
 1010 1015 1020

Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
 1025 1030 1035

Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
 1040 1045 1050

Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
 1055 1060 1065

ES 2 719 304 T3

Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe
 1070 1075 1080

Asp Asp Leu Thr Glu Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr
 1085 1090 1095

Gly Cys Glu Ala Val Glu Val Arg Val Asp His Leu Ala Asn Tyr
 1100 1105 1110

Ser Ala Asp Phe Val Ser Lys Gln Leu Ser Ile Leu Arg Lys Ala
 1115 1120 1125

Thr Asp Ser Ile Pro Ile Ile Phe Thr Val Arg Thr Met Lys Gln
 1130 1135 1140

Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val
 1220 1225 1230

Asn Phe Glu Asp Asn Leu Arg Leu Glu His Phe Arg Asp Thr His
 1235 1240 1245

Lys Asn Lys Pro Leu Ile Ala Val Asn Met Thr Ser Lys Gly Ser
 1250 1255 1260

Ile Ser Arg Val Leu Asn Asn Val Leu Thr Pro Val Thr Ser Asp
 1265 1270 1275

Leu Leu Pro Asn Ser Ala Ala Pro Gly Gln Leu Thr Val Ala Gln
 1280 1285 1290

Ile Asn Lys Met Tyr Thr Ser Met Gly Gly Ile Glu Pro Lys Glu
 1295 1300 1305

ES 2 719 304 T3

Leu Phe Val Val Gly Lys Pro Ile Gly His Ser Arg Ser Pro Ile
 1310 1315 1320

Leu His Asn Thr Gly Tyr Glu Ile Leu Gly Leu Pro His Lys Phe
 1325 1330 1335

Asp Lys Phe Glu Thr Glu Ser Ala Gln Leu Val Lys Glu Lys Leu
 1340 1345 1350

Leu Asp Gly Asn Lys Asn Phe Gly Gly Ala Ala Val Thr Ile Pro
 1355 1360 1365

Leu Lys Leu Asp Ile Met Gln Tyr Met Asp Glu Leu Thr Asp Ala
 1370 1375 1380

Ala Lys Val Ile Gly Ala Val Asn Thr Val Ile Pro Leu Gly Asn
 1385 1390 1395

Lys Lys Phe Lys Gly Asp Asn Thr Asp Trp Leu Gly Ile Arg Asn
 1400 1405 1410

Ala Leu Ile Asn Asn Gly Val Pro Glu Tyr Val Gly His Thr Ala
 1415 1420 1425

Gly Leu Val Ile Gly Ala Gly Gly Thr Ser Arg Ala Ala Leu Tyr
 1430 1435 1440

Ala Leu His Ser Leu Gly Cys Lys Lys Ile Phe Ile Ile Asn Arg
 1445 1450 1455

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys

ES 2 719 304 T3

1535

1540

1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
1580 1585

<210> 5

< 211> 4767

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Ácido nucleico sintético

<400> 5

atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac	60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt	120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag	180
gcttctttgc cagaaggctc tcgtttactt acttatgttg ttaaaccagg tgagacaagt	240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt	300
gatacggtta tggtagcgat cgggtggtgt gttattggtg acatgattgg gttcgttgca	360
tctacattta tgagaggtgt tcgtgttgtc caagtaccaa catccttatt ggcaatggtc	420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt	480
gcattttggc aacccaaatt tgtccttgta gatattaaat ggctagaaac gttagccaag	540
agagagttha tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac	600
gaatttacta gattagaatc aaacgcttcg ttgttcttaa atggtgttaa tggggcaaaa	660
aatgtcaagg ttaccaatca attgacaaac gagattgacg agatatcgaa tacagatatt	720
gaagctatgt tggatcatac atataagtha gttcttgaga gtattaaggt caaagcggaa	780
gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct	840
attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc	900
attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa	960
gttgcacgtc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa	1020
tggtttaaag agctaacctt acataagaaa acaccattgg atatcttatt gaagaaaatg	1080
agtattgaca agaaaaacga gggttccaaa agaaggtgg tcattttaga aagtattggt	1140
aagtgctatg gtgactccgc tcaatttggt agcgaatgaag acctgagatt tattctaaca	1200

10

ES 2 719 304 T3

gatgaaaccc	tcgtttacc	cttcaaggac	atccctgctg	atcaacagaa	agttgttatac	1260
ccccctggtt	ctaagtccat	ctccaatcgt	gctttaattc	ttgctgccct	cggngaaggt	1320
caatgtaaaa	tcaagaactt	attacattct	gatgatacta	aacatatgtt	aaccgctggt	1380
catgaattga	aaggtgctac	gatatcatgg	gaagataatg	gtgagacggt	agtggtggaa	1440
ggacatggtg	gttccacatt	gtcagcttgt	gctgaccctt	tatatctagg	taatgcaggt	1500
actgcatcta	gatttttgac	ttccttggct	gccttggcca	attctacttc	aagccaaaag	1560
tatatcgttt	taactggtaa	cgcaagaatg	caacaaagac	caattgctcc	tttggctgat	1620
tctttgogtg	ctaattgtac	taaaattgag	tacttgaata	atgaaggttc	cctgccaatc	1680
aaagtttata	ctgattcggg	attcaaaggt	ggtagaattg	aattagctgc	tacagtttct	1740
tctcagtacg	tatcctctat	cttgatgtgt	gccccatacg	ctgaagaacc	tgtaactttg	1800
gctcttggtg	gtggtaaacc	aatctctaaa	ttgtacgtcg	atatgacaat	aaaaatgatg	1860
gaaaaattcg	gtatcaatgt	tgaaaacttct	actacagaac	cttacctta	ttatattcca	1920
aagggacatt	atattaaccc	atcagaatac	gtcattgaaa	gtgatgcctc	aagtgcatac	1980
taccattggg	ccttcgccgc	aatgactggg	actaccgtaa	cggttccaaa	cattggtttt	2040
gagtcgttac	aaggtgatgc	cagatttgca	agagatgtct	tgaaacctat	gggttgtaaa	2100
ataactcaaa	cggcaacttc	aactactggt	togggtcctc	ctgtaggtac	tttaaagcca	2160
ttaaaacatg	ttgatatgga	gccaatgact	gatgcgttct	taactgcatg	tgttgttgcc	2220
gctatttcgc	acgacagtga	tccaaattct	gcaaatacaa	ccaccattga	aggtattgca	2280
aaccagcgtg	tcaaagagtg	taacagaatt	ttggccatgg	ctacagagct	cgccaaattt	2340
ggcgtcaaaa	ctacagaatt	accagatggg	attcaagtcc	atggtttaaa	ctcgataaaa	2400
gatttgaagg	ttccttccga	ctcttctgga	cctgtcggtg	tatgcacata	tgatgatcat	2460
cgtgtggcca	tgagtttctc	gcttcttgca	ggaatggtaa	attctcaaaa	tgaactgac	2520
gaagttgcta	atcctgtaag	aatacttgaa	agacattgta	ctggtaaaac	ctggcctggc	2580
tggtgggatg	tgttacattc	cgaactaggt	gccaaattag	atggtgcaga	acctttagag	2640
tgcacatcca	aaaagaactc	aaagaaaagc	gttgtcatta	ttggcatgag	agcagctggc	2700
aaaactacta	taagtaaag	gtgcgcatcc	gctctgggtt	acaaattagt	tgacctagac	2760
gagctgtttg	agcaacagca	taacaatcaa	agtgttaaac	aatttgttgt	ggagaacggt	2820
tgggagaagt	tcogtgagga	agaaacaaga	atcttcaagg	aagttattca	aaattacggc	2880
gatgatggat	atgttttctc	aacaggtggc	ggtattgttg	aaagcgtga	gtctagaaaa	2940
gccttaaaag	atcttgcctc	atcaggtgga	tacgttttac	acttacatag	ggatattgag	3000
gagacaattg	tctttttaca	aagtgatcct	tcaagacctg	cctatgtgga	agaaattcgt	3060

ES 2 719 304 T3

gaagtttga acagaagga ggggtggtat aaagaatgct caaatcttc tttctttgct 3120
cctcattgct cgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
gaggctgtag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
cagttatcta tattgcgtaa agccactgac agtatttcta tcatttttac tgtgcgaacc 3420
atgaagcaag gtggcaactt tcctgatgaa gagttcaaaa ccttgagaga gctatacgat 3480
attgccttga agaatggtgt tgaattcctt gacttagaac taactttacc tactgatatc 3540
caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
caaggattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga 3720
ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
aaagtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
cctaactcgc ctgcccctgg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
atgggaggtc tcgagcctaa ggaactgttt gttgttgtaa agccaattgg ccactctaga 3960
tcgccaatth tacataacac tggctatgaa attttaggtt tacctcacia gttcgataaa 4020
tttgaaactg aatcgcaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt 4080
ggcgggtgctg cagtcacaat tcctctgaaa ttagatataa tgcagtacat ggatgaattg 4140
actgatgctg ctaaagttat tgggtgctgta aacacagtta taccattggg taacaagaag 4200
tttaaggggtg ataataccga ctgggttaggt atccgtaatg cettaattaa caatggcgtt 4260
cccgaatatg ttggtcatac cgctggtttg gttatcggtg caggtggcac ttctagagcc 4320
gccctttacg ccttgcacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaaatagag 4440
tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaaggt 4560
gccccatgctg cttttgtacc aaccttattg gaagccccat acaaaccaag cgttactccc 4620
gttatgacaa tttcacaaga caaatatcaa tggcacggtt tccttgatc acaaatgtta 4680
gtacaccaag gtgtagctca gtttgaaaag tggacaggat tcaagggccc tttcaaggcc 4740
atthttgatg ccgttacgaa agagtag 4767

<210> 6

< 211> 1588

< 212> PRT

5 < 213> Secuencia artificial

<220>

< 223> Polipéptido sintético

ES 2 719 304 T3

<400> 6

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
 1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
 20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
 35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
 50 55 60

Glu Gly Ser Arg Leu Leu Thr Tyr Val Val Lys Pro Gly Glu Thr Ser
 65 70 75 80

Lys Ser Arg Glu Thr Lys Ala Gln Leu Glu Asp Tyr Leu Leu Val Glu
 85 90 95

Gly Cys Thr Arg Asp Thr Val Met Val Ala Ile Gly Gly Gly Val Ile
 100 105 110

Gly Asp Met Ile Gly Phe Val Ala Ser Thr Phe Met Arg Gly Val Arg
 115 120 125

Val Val Gln Val Pro Thr Ser Leu Leu Ala Met Val Asp Ser Ser Ile
 130 135 140

Gly Gly Lys Thr Ala Ile Asp Thr Pro Leu Gly Lys Asn Phe Ile Gly
 145 150 155 160

Ala Phe Trp Gln Pro Lys Phe Val Leu Val Asp Ile Lys Trp Leu Glu
 165 170 175

Thr Leu Ala Lys Arg Glu Phe Ile Asn Gly Met Ala Glu Val Ile Lys
 180 185 190

Thr Ala Cys Ile Trp Asn Ala Asp Glu Phe Thr Arg Leu Glu Ser Asn
 195 200 205

Ala Ser Leu Phe Leu Asn Val Val Asn Gly Ala Lys Asn Val Lys Val
 210 215 220

Thr Asn Gln Leu Thr Asn Glu Ile Asp Glu Ile Ser Asn Thr Asp Ile

ES 2 719 304 T3

225		230		235		240
Glu Ala Met Leu Asp His Thr Tyr Lys Leu Val Leu Glu Ser Ile Lys		245		250		255
Val Lys Ala Glu Val Val Ser Ser Asp Glu Arg Glu Ser Ser Leu Arg		260		265		270
Asn Leu Leu Asn Phe Gly His Ser Ile Gly His Ala Tyr Glu Ala Ile		275		280		285
Leu Thr Pro Gln Ala Leu His Gly Glu Cys Val Ser Ile Gly Met Val			290	295		300
Lys Glu Ala Glu Leu Ser Arg Tyr Phe Gly Ile Leu Ser Pro Thr Gln			305	310	315	320
Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser		325		330		335
Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro		340		345		350
Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly		355		360		365
Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly		370		375		380
Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr			385	390	395	400
Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln		405		410		415
Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu		420		425		430
Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu		435		440		445
His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys		450		455		460
Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu			465	470	475	480

ES 2 719 304 T3

Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495

Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510

Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525

Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540

Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560

Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala
 565 570 575

Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
 580 585 590

Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
 595 600 605

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
 610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
 625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
 645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
 660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
 675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
 690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
 705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
 725 730 735

ES 2 719 304 T3

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
 740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
 755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
 770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
 785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
 805 810 815

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
 820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
 835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
 850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
 865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
 885 890 895

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu
 900 905 910

Gly Tyr Lys Leu Val Asp Leu Asp Glu Leu Phe Glu Gln Gln His Asn
 915 920 925

Asn Gln Ser Val Lys Gln Phe Val Val Glu Asn Gly Trp Glu Lys Phe
 930 935 940

Arg Glu Glu Glu Thr Arg Ile Phe Lys Glu Val Ile Gln Asn Tyr Gly
 945 950 955 960

Asp Asp Gly Tyr Val Phe Ser Thr Gly Gly Gly Ile Val Glu Ser Ala
 965 970 975

Glu Ser Arg Lys Ala Leu Lys Asp Phe Ala Ser Ser Gly Gly Tyr Val
 980 985 990

ES 2 719 304 T3

Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
 995 1000 1005

 Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
 1010 1015 1020

 Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
 1025 1030 1035

 Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
 1040 1045 1050

 Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
 1055 1060 1065

 Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe
 1070 1075 1080

 Asp Asp Leu Thr Glu Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr
 1085 1090 1095

 Gly Cys Glu Ala Val Glu Val Arg Val Asp His Leu Ala Asn Tyr
 1100 1105 1110

 Ser Ala Asp Phe Val Ser Lys Gln Leu Ser Ile Leu Arg Lys Ala
 1115 1120 1125

 Thr Asp Ser Ile Pro Ile Ile Phe Thr Val Arg Thr Met Lys Gln
 1130 1135 1140

 Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

 Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

 Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

 Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

 Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

 Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val

ES 2 719 304 T3

1220		1225		1230		
Asn Phe	Glu Asp	Asn Leu	Arg Leu	Glu His Phe	Arg Asp	Thr His
1235			1240		1245	
Lys Asn	Lys Pro	Leu Ile	Ala Val	Asn Met Thr	Ser Lys	Gly Ser
1250			1255		1260	
Ile Ser	Arg Val	Leu Asn	Asn Val	Leu Thr Pro	Val Thr	Ser Asp
1265			1270		1275	
Leu Leu	Pro Asn	Ser Ala	Ala Pro	Gly Gln Leu	Thr Val	Ala Gln
1280			1285		1290	
Ile Asn	Lys Met	Tyr Thr	Ser Met	Gly Gly Ile	Glu Pro	Lys Glu
1295			1300		1305	
Leu Phe	Val Val	Gly Lys	Pro Ile	Gly His Ser	Arg Ser	Pro Ile
1310			1315		1320	
Leu His	Asn Thr	Gly Tyr	Glu Ile	Leu Gly Leu	Pro His	Lys Phe
1325			1330		1335	
Asp Lys	Phe Glu	Thr Glu	Ser Ala	Gln Leu Val	Lys Glu	Lys Leu
1340			1345		1350	
Leu Asp	Gly Asn	Lys Asn	Phe Gly	Gly Ala Ala	Val Thr	Ile Pro
1355			1360		1365	
Leu Lys	Leu Asp	Ile Met	Gln Tyr	Met Asp Glu	Leu Thr	Asp Ala
1370			1375		1380	
Ala Lys	Val Ile	Gly Ala	Val Asn	Thr Val Ile	Pro Leu	Gly Asn
1385			1390		1395	
Lys Lys	Phe Lys	Gly Asp	Asn Thr	Asp Trp Leu	Gly Ile	Arg Asn
1400			1405		1410	
Ala Leu	Ile Asn	Asn Gly	Val Pro	Glu Tyr Val	Gly His	Thr Ala
1415			1420		1425	
Gly Leu	Val Ile	Gly Ala	Gly Gly	Thr Ser Arg	Ala Ala	Leu Tyr
1430			1435		1440	
Ala Leu	His Ser	Leu Gly	Cys Lys	Lys Ile Phe	Ile Ile	Asn Arg
1445			1450		1455	

ES 2 719 304 T3

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Pro
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys
 1535 1540 1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
 1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
 1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
 1580 1585

<210> 7

<211> 4767

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Ácido nucleico sintético

<400> 7

atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac	60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt	120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag	180
gcttctttgc cagaaggctc tcgtttactt acttatggtg ttaaaccagg tgagacaagt	240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt	300
gatacggtta tggtagcgat cgggtggtgt gttattggtg acatgattgg gttcgttgca	360
tctacattta tgagaggtgt tcgtggttgc caagtaccaa catccttatt ggcaatggtc	420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt	480
gcattttggc aacccaaaatt tgtccttcta gatattaaat ggctagaaac gttagccaag	540

10

ES 2 719 304 T3

agagagttta tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac 600
 gaatttacta gattagaatc aaacgcttcg ttgttcttaa atgttgtaa tggggcaaaa 660
 aatgtcaagg ttaccaatca attgacaaa gagattgacg agatatacga tacagatatt 720
 gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa 780
 gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct 840
 attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc 900
 attggtatgg ttaaagaggg ggaattatcc cgttatttcg gtattctctc ccctacccaa 960
 gttgcacgtc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa 1020
 tggtttaaag agctaaccct acataagaaa acaccattgg atatcttatt gaagaaaatg 1080
 agtattgaca agaaaaacga gggttccaaa aagaaggtgg tcattttaga aagtattggt 1140
 aagtgctatg gtgactccgc tcaatttggt agcgaagaag acctgagatt tattctaaca 1200
 gatgaaacct tcgtttacc cttcaaggac atccctgctg atcaacagaa agttgttacc 1260
 cccctcggtt ctaagtccat ctccaatcgt gctttaattc ttgctgccct cggatgaagg 1320
 caatgtaaaa tcaagaactt attacattct gatgatacta aacatatggt aaccgctggt 1380
 catgaattga aagtgctac gatatcatgg gaagataatg gtgagacggg agtggggaa 1440
 ggacatggg gttccacatt gtcagcttgt gctgaccctt tatatctagg taatgcagg 1500
 actgcatcta gatttttgac ttccttggtt gccttggtca atttacttc aagccaaaag 1560
 tataatggtt taactggtaa cgcaagaatg caacaaagac caattgctcc tttggatgat 1620
 tctttgctg ctaatgggtac taaaattgag taactgaata atgaaggttc cctgccaatc 1680
 aaagtttata ctgattcggg attcaaaggg ggtagaattg aattagctgc tacagtttct 1740
 tctcagtacg tatcctctat cttgatgtgt gcccatacgt ctgaagaacc tgtaactttg 1800
 gctcttggtg gtggttaagc aatctctaaa ttgtacgtcg atatgacaat aaaaatgatg 1860
 gaaaaattcg gtatcaatgt tgaaacttct actacagaac cttacactta ttatattcca 1920
 aaggacatt atattaacc atcagaatac gtcattgaaa gtgatgcctc aagtgctaca 1980
 taccatttgg ccttcgccgc aatgactggg actaccgtaa cggttccaaa cattggtttt 2040
 gagtcgttac aagtgatgc cagatttgca agagatgtct tgaaacctat gggttgtaa 2100
 ataactcaa cggcaacttc aactactggt tcgggtcctc ctgtaggtac tttaaagcca 2160
 ttaaaacatg ttgatatgga gccaatgact gatgcgttct taactgcatg gttgttgcc 2220
 gctatttcgc acgacagtga tccaaattct gcaaatatac ccaccattga aggtattgca 2280
 aaccagcgtg tcaaagagtg taacagaatt ttggccatgg ctacagagct cgcctaaatt 2340
 ggcgtcaaaa ctacagaatt accagatggg attcaagtc atggtttaaa ctcgataaaa 2400
 gatttgaagg ttccttcga ctctctgga cctgtcggg tatgcacata tgatgatcat 2460

ES 2 719 304 T3

cgtgtggcca tgagtttctc gcttcttgca ggaatggtaa attctcaaaa tgaacgtgac 2520
 gaagttgcta atcctgtaag aatacttgaa agacattgta ctggtaaaac ctggcctggc 2580
 tgggtgggatg tgttacattc cgaactaggt gccaaattag atgggtgcaga acctttagag 2640
 tgcacatcca aaaagaactc aaagaaaagc gttgtcatta ttggcatgag agcagctggc 2700
 aaaactacta taagtaaatg gtgcgcatcc gctctggggtt acaaattagt tgacctagac 2760
 gagctgtttg agcaacagca taacaatcaa agtggttaaac aatttgttgt ggagaacggc 2820
 tgggagaagt tccgtgagga agaaacaaga attttcaagg aagttattca aaattacggc 2880
 gatgatggat atgttttctc aacaggtggc ggtattgttg aaagcgtga gtctagaaaa 2940
 gccttaaaag attttgccctc atcaggtgga tacgttttac acttacatag ggatattgag 3000
 gagacaattg tctttttaca aagtgatcct tcaagacctg cctatgtgga agaaattcgt 3060
 gaagtttgga acagaaggga ggggtggtat aaagaatgct caaatttctc tttctttgct 3120
 cctcattgct ccgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
 gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
 ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
 gaggtctgag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
 cagttatcta tattgcgtaa agccactgac agtattccta tcatttttac tgtgcgaacc 3420
 atgaagcaag gtggcaactt tcctgatgaa gagttcaaaa ccttgagaga gctatacgat 3480
 attgccttga agaatggtgt tgaattcctt gacttagaac taactttacc tactgatatc 3540
 caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
 caaggattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
 cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga 3720
 ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
 aaaggtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
 cctaactccg ctgcccctgg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
 atgggaggta tcgagcctaa ggaactggtt gttgttgtaa agccaattgg ccaactetaga 3960
 tcgccaatth tacataacac tggctatgaa attttaggtt tacctcacia gttcgataaa 4020
 tttgaaactg aatccgcaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt 4080
 ggcggtgctg cagtcacaat tcctctgaaa ttagatataa tgcagtacat ggatgaattg 4140
 actgatgctg ctaaagttat tgggtctgta aacacagtta taccattggg taacaagaag 4200
 tttaaaggtg ataataccga ctggttaggt atccgtaatg cettaattaa caatggcgtt 4260
 cccgaatatg ttggtcatac cgctggtttg gttatcggtg caggtggcac ttctagagcc 4320

ES 2 719 304 T3

gccctttacg ccttgacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
 tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaatagag 4440
 tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtaaaa 4500
 gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaaggt 4560
 gcccatgctg cttttgtacc aaccttattg gaagccgcat acaaaccaag cgttactccc 4620
 gttatgacaa tttcacaaga caaatatcaa tggcacgttg tccctggatc acaaatgta 4680
 gtacaccaag gtgtagctca gtttgaagaa tggacaggat tcaagggccc tttcaaggcc 4740
 atttttgatg ccgttacgaa agagtag 4767

<210> 8

<211> 1588

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 8

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
 1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
 20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
 35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
 50 55 60

Glu Gly Ser Arg Leu Leu Thr Tyr Val Val Lys Pro Gly Glu Thr Ser
 65 70 75 80

Lys Ser Arg Glu Thr Lys Ala Gln Leu Glu Asp Tyr Leu Leu Val Glu
 85 90 95

Gly Cys Thr Arg Asp Thr Val Met Val Ala Ile Gly Gly Gly Val Ile
 100 105 110

Gly Asp Met Ile Gly Phe Val Ala Ser Thr Phe Met Arg Gly Val Arg
 115 120 125

Val Val Gln Val Pro Thr Ser Leu Leu Ala Met Val Asp Ser Ser Ile
 130 135 140

10

ES 2 719 304 T3

Gly Gly Lys Thr Ala Ile Asp Thr Pro Leu Gly Lys Asn Phe Ile Gly
 145 150 155 160

Ala Phe Trp Gln Pro Lys Phe Val Leu Val Asp Ile Lys Trp Leu Glu
 165 170 175

Thr Leu Ala Lys Arg Glu Phe Ile Asn Gly Met Ala Glu Val Ile Lys
 180 185 190

Thr Ala Cys Ile Trp Asn Ala Asp Glu Phe Thr Arg Leu Glu Ser Asn
 195 200 205

Ala Ser Leu Phe Leu Asn Val Val Asn Gly Ala Lys Asn Val Lys Val
 210 215 220

Thr Asn Gln Leu Thr Asn Glu Ile Asp Glu Ile Ser Asn Thr Asp Ile
 225 230 235 240

Glu Ala Met Leu Asp His Thr Tyr Lys Leu Val Leu Glu Ser Ile Lys
 245 250 255

Val Lys Ala Glu Val Val Ser Ser Asp Glu Arg Glu Ser Ser Leu Arg
 260 265 270

Asn Leu Leu Asn Phe Gly His Ser Ile Gly His Ala Tyr Glu Ala Ile
 275 280 285

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

Lys Glu Ala Glu Leu Ser Arg Tyr Phe Gly Ile Leu Ser Pro Thr Gln
 305 310 315 320

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
 325 330 335

Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
 340 345 350

Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly
 355 360 365

Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly
 370 375 380

Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr
 385 390 395 400

ES 2 719 304 T3

Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln
 405 410 415
 Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu
 420 425 430
 Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu
 435 440 445
 His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys
 450 455 460
 Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu
 465 470 475 480
 Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495
 Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510
 Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525
 Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540
 Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560
 Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala
 565 570 575
 Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
 580 585 590
 Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
 595 600 605
 Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
 610 615 620
 Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
 625 630 635 640
 Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
 645 650 655

ES 2 719 304 T3

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
725 730 735

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
805 810 815

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
885 890 895

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu

ES 2 719 304 T3

Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val
 1220 1225 1230

Asn Phe Glu Asp Asn Leu Arg Leu Glu His Phe Arg Asp Thr His
 1235 1240 1245

Lys Asn Lys Pro Leu Ile Ala Val Asn Met Thr Ser Lys Gly Ser
 1250 1255 1260

Ile Ser Arg Val Leu Asn Asn Val Leu Thr Pro Val Thr Ser Asp
 1265 1270 1275

Leu Leu Pro Asn Ser Ala Ala Pro Gly Gln Leu Thr Val Ala Gln
 1280 1285 1290

Ile Asn Lys Met Tyr Thr Ser Met Gly Gly Ile Glu Pro Lys Glu
 1295 1300 1305

Leu Phe Val Val Gly Lys Pro Ile Gly His Ser Arg Ser Pro Ile
 1310 1315 1320

Leu His Asn Thr Gly Tyr Glu Ile Leu Gly Leu Pro His Lys Phe
 1325 1330 1335

Asp Lys Phe Glu Thr Glu Ser Ala Gln Leu Val Lys Glu Lys Leu
 1340 1345 1350

Leu Asp Gly Asn Lys Asn Phe Gly Gly Ala Ala Val Thr Ile Pro
 1355 1360 1365

Leu Lys Leu Asp Ile Met Gln Tyr Met Asp Glu Leu Thr Asp Ala
 1370 1375 1380

ES 2 719 304 T3

Ala Lys Val Ile Gly Ala Val Asn Thr Val Ile Pro Leu Gly Asn
 1385 1390 1395

Lys Lys Phe Lys Gly Asp Asn Thr Asp Trp Leu Gly Ile Arg Asn
 1400 1405 1410

Ala Leu Ile Asn Asn Gly Val Pro Glu Tyr Val Gly His Thr Ala
 1415 1420 1425

Gly Leu Val Ile Gly Ala Gly Gly Thr Ser Arg Ala Ala Leu Tyr
 1430 1435 1440

Ala Leu His Ser Leu Gly Cys Lys Lys Ile Phe Ile Ile Asn Arg
 1445 1450 1455

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Lys Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys
 1535 1540 1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
 1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
 1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
 1580 1585

<210> 9

< 211> 4767

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Ácido nucleico sintético

ES 2 719 304 T3

<400> 9
atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac 60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt 120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag 180
gcttctttgc cagaaggctc tcgtttactt acttatggtt ttaaaccagg tgagacaagt 240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt 300
gatacggtta tggtagcgat cggtaggtgt gttattggtg acatgattgg gttcgttgca 360
tctacattta tgagaggtgt tcgtggtgtc caagtaccaa catccttatt ggcaatggtc 420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt 480
gcattttggc aaccaaatt tgccttgta gatattaaat ggctagaaac gttagccaag 540
agagagtta tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac 600
gaatttacta gattagaatc aaacgcttcg ttgttcttaa atggtgtaa tggggcaaaa 660
aatgtcaagg ttaccaatca attgacaaac gagattgacg agatatacga tacagatatt 720
gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa 780
gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct 840
attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc 900
attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa 960
gttgcacgtc tatccaagat tttggttgc tacgggttgc ctgtttcgcc tgatgagaaa 1020
tggtttaaag agctaacctt acataagaaa acaccattgg atatcttatt gaagaaaatg 1080
agtattgaca agaaaaacga gggttccaaa aagaagggtg tcattttaga aagtattggt 1140
aagtgctatg gtgactccgc tcaatttgtt agcgatgaag acctgagatt tattctaaca 1200
gatgaaacc tcgtttacc cttcaaggac atccctgctg atcaacagaa agttgttatc 1260
ccccctggtt ctaagtccat ctccaatcgt gctttaattc ttgctgccct cggatgaagg 1320
caatgtaaaa tcaagaactt attacattct gatgatacta aacatatgtt aaccgctgtt 1380
catgaattga aagggtgctac gatatcatgg gaagataatg gtgagacggt agtgggtgaa 1440
ggacatggtg gttccacatt gtcagcttgt gctgaccctt tatatctagg taatgcagg 1500
actgcatcta gatTTTTGAC ttccttggct gccttggca attctacttc aagccaaaag 1560
tatatcgttt taactggtaa cgcaagaatg caacaaagac caattgctcc tttggtcgat 1620
tctttgcgtg ctaatggtac taaaattgag tacttgaata atgaaggttc cctgccaatc 1680
aaagtttata ctgattcgggt attcaaagggt ggtagaattg aattagctgc tacagtttct 1740
tctcagtacg tatcctctat cttgatgtgt gcccatacgt ctgaagaacc tgtaactttg 1800

ES 2 719 304 T3

gctcttggtg gtggttaagcc aatctctaaa ttgtacgtcg atatgacaat aaaaatgatg 1860
gaaaaattcg gtatcaatgt tgaaacttct actacagaac cttacactta ttatattcca 1920
aagggacatt atattaaccc atcagaatac gtcattgaaa gtgatgcctc aagtgtctaca 1980
taccatttgg ccttcgccc aatgactggg actaccgtaa cggttccaaa cattggtttt 2040
gagtcgttac aaggtgatgc cagatttgca agagatgtct tgaaacctat gggttgtaaa 2100
ataactcaaa cggcaacttc aactactggt tcgggtcctc ctgtaggtac tttaaagcca 2160
ttaaaacatg ttgatatgga gccaatgact gatgcgttct taactgcatg tgttggtgcc 2220
gctatttcgc acgacagtga tccaaattct gcaaatacaa ccaccattga aggtattgca 2280
aaccagcgtg tcaaagagtg taacagaatt ttggccatgg ctacagagct cgccaaattt 2340
ggcgtcaaaa ctacagaatt accagatggg attcaagtcc atggtttaaa ctcgataaaa 2400
gatttgaagg ttccttcgca ctctctgga cctgtcggg tatgcacata tgatgatcat 2460
cgtgtggcca tgagtttctc gcttcttgca ggaatgtaa attctcaaaa tgaacgtgac 2520
gaagttgcta atcctgtaag aatacttgaa agacattgta ctggtaaaac ctggcctggc 2580
tgggtgggatg tgttacattc cgaactaggt gccaaattag atgggtgcaga acctttagag 2640
tgcaatcca aaaagaactc aaagaaaagc gttgtcatta ttggcatgag agcagctggc 2700
aaaactacta taagtaaag gtgcgcatcc gctctgggtt acaaattagt tgacctagac 2760
gagctgtttg agcaacagca taacaatcaa agtgttaaac aatttgttgt ggagaacggt 2820
tgggagaagt tccgtgagga agaaacaaga attttcaagg aagttattca aaattacggc 2880
gatgatggat atgttttctc aacaggtggc ggtattgttg aaagcgtga gtctagaaaa 2940
gccttaaaag attttgcctc atcaggtgga tacgttttac acttacatag ggatattgag 3000
gagacaattg tctttttaca aagtgatect tcaagacctg cctatgtgga agaaattcgt 3060
gaagtttggg acagaagggg ggggtgggat aaagaatgct caaatttctc tttctttgct 3120
cctcattgct ccgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
gaggctgtag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
cagttatcta tattcgttaa agccactgac agtattccta tcatttttac tgtgcgaacc 3420
atgaagcaag gtggcaactt tectgatgaa gagttcaaaa cettgagaga gctatacgat 3480
attgccttga agaattggtgt tgaattcctt gacttagaac taactttacc tactgatatc 3540
caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
caaggattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcaaga taatttgaga 3720

ES 2 719 304 T3

ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
aaaggtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
cctaactccg ctgccctcg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
atgggaggtta tgcagcctaa ggaactgttt gttgttgaa agccaattgg ccaactctaga 3960
tcgccaattt tacataacac tggctatgaa attttaggtt tacctcaciaa gttcgataaa 4020
tttgaaactg aatccgcaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt 4080
ggcggtgctg cagtcacaat tcctctgaaa ttagatataa tgcagtacat ggatgaattg 4140
actgatgctg ctaaagttat tgggtgctgta aacacagtta taccattggg taacaagaag 4200
tttaagggtg ataataccga ctgggttaggt atccgtaatg ccttaattaa caatggcgtt 4260
cccgaatatg ttggtcatac cgctggtttg gttatcggtg caggtggcac ttctagagcc 4320
gccctttacg ccttgcacag tttaggttgc aaaaagatct tcataatcaa ctggacaact 4380
tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaatagag 4440
tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaagg 4560
gcccatgctg cttttgtacc aaccttattg gaagccgcat acaaaccaag cgttactccc 4620
gttatgacaa tttcacaaga caaatatcaa tggcacggtg tccttgatc acaaatgtta 4680
gtacaccaag gtgtagctca gtttgaaaag tggacaggat tcaagggcc tttcaaggcc 4740
atttttgatg ccgttacgaa agagtag 4767

<210> 10

< 211> 1588

< 212> PRT

5 < 213> Secuencia artificial

<220>

< 223> Polipéptido sintético

<400> 10

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
50 55 60

10

ES 2 719 304 T3

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

ES 2 719 304 T3

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
 325 330 335

Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
 340 345 350

Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly
 355 360 365

Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly
 370 375 380

Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr
 385 390 395 400

Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln
 405 410 415

Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu
 420 425 430

Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu
 435 440 445

His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys
 450 455 460

Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu
 465 470 475 480

Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495

Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510

Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525

Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540

Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560

Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala

ES 2 719 304 T3

565 570 575

Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
580 585 590

Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
595 600 605

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
725 730 735

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
805 810 815

ES 2 719 304 T3

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
 820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
 835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
 850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
 865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
 885 890

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu
 900 905 910

Gly Tyr Lys Leu Val Asp Leu Asp Glu Leu Phe Glu Gln Gln His Asn
 915 920 925

Asn Gln Ser Val Lys Gln Phe Val Val Glu Asn Gly Trp Glu Lys Phe
 930 935 940

Arg Glu Glu Glu Thr Arg Ile Phe Lys Glu Val Ile Gln Asn Tyr Gly
 945 950 955 960

Asp Asp Gly Tyr Val Phe Ser Thr Gly Gly Gly Ile Val Glu Ser Ala
 965 970 975

Glu Ser Arg Lys Ala Leu Lys Asp Phe Ala Ser Ser Gly Gly Tyr Val
 980 985 990

Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
 995 1000 1005

Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
 1010 1015 1020

Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
 1025 1030 1035

Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
 1040 1045 1050

Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
 1055 1060 1065

ES 2 719 304 T3

Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe
 1070 1075 1080

Asp Asp Leu Thr Glu Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr
 1085 1090 1095

Gly Cys Glu Ala Val Glu Val Arg Val Asp His Leu Ala Asn Tyr
 1100 1105 1110

Ser Ala Asp Phe Val Ser Lys Gln Leu Ser Ile Leu Arg Lys Ala
 1115 1120 1125

Thr Asp Ser Ile Pro Ile Ile Phe Thr Val Arg Thr Met Lys Gln
 1130 1135 1140

Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val
 1220 1225 1230

Asn Phe Glu Asp Asn Leu Arg Leu Glu His Phe Arg Asp Thr His
 1235 1240 1245

Lys Asn Lys Pro Leu Ile Ala Val Asn Met Thr Ser Lys Gly Ser
 1250 1255 1260

Ile Ser Arg Val Leu Asn Asn Val Leu Thr Pro Val Thr Ser Asp
 1265 1270 1275

Leu Leu Pro Asn Ser Ala Ala Pro Gly Gln Leu Thr Val Ala Gln
 1280 1285 1290

Ile Asn Lys Met Tyr Thr Ser Met Gly Gly Ile Glu Pro Lys Glu
 1295 1300 1305

ES 2 719 304 T3

Leu Phe Val Val Gly Lys Pro Ile Gly His Ser Arg Ser Pro Ile
 1310 1315 1320

Leu His Asn Thr Gly Tyr Glu Ile Leu Gly Leu Pro His Lys Phe
 1325 1330 1335

Asp Lys Phe Glu Thr Glu Ser Ala Gln Leu Val Lys Glu Lys Leu
 1340 1345 1350

Leu Asp Gly Asn Lys Asn Phe Gly Gly Ala Ala Val Thr Ile Pro
 1355 1360 1365

Leu Lys Leu Asp Ile Met Gln Tyr Met Asp Glu Leu Thr Asp Ala
 1370 1375 1380

Ala Lys Val Ile Gly Ala Val Asn Thr Val Ile Pro Leu Gly Asn
 1385 1390 1395

Lys Lys Phe Lys Gly Asp Asn Thr Asp Trp Leu Gly Ile Arg Asn
 1400 1405 1410

Ala Leu Ile Asn Asn Gly Val Pro Glu Tyr Val Gly His Thr Ala
 1415 1420 1425

Gly Leu Val Ile Gly Ala Gly Gly Thr Ser Arg Ala Ala Leu Tyr
 1430 1435 1440

Ala Leu His Ser Leu Gly Cys Lys Lys Ile Phe Ile Ile Asn Trp
 1445 1450 1455

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys

ES 2 719 304 T3

1535

1540

1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
 1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
 1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
 1580 1585

<210> 11

< 211> 4767

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Ácido nucleico sintético

<400> 11

atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac	60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt	120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag	180
gcttctttgc cagaaggctc tcgtttactt acttatgttg ttaaaccagg tgagacaagt	240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt	300
gatacggtta tggtagcgat cgggtggtgt gttattggtg acatgattgg gttcgttgca	360
tctacattta tgagaggtgt tcgtgttgtc caagtaacaa catccttatt ggcaatggtc	420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt	480
gcattttggc aacccaaatt tgtccttgta gatattaaat ggctagaaac gttagccaag	540
agagagttha tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac	600
gaatttacta gattagaatc aaacgcttcg ttgttcttaa atggtgttaa tggggcaaaa	660
aatgtcaagg ttaccaatca attgacaac gagattgacg agatatcgaa tacagatatt	720
gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa	780
gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct	840
attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc	900
attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa	960
gttgcacgctc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa	1020
tggtttaaag agctaacctt acataagaaa acaccattgg atatcttatt gaagaaaatg	1080
agtattgaca agaaaaacga gggttccaaa agaaggtgg tcattttaga aagtattggt	1140
aagtgctatg gtgactccgc tcaatttggt agcgatgaag acctgagatt tattctaaca	1200

10

ES 2 719 304 T3

gatgaaaccc	tcgtttacct	cttcaaggac	atccctgctg	atcaacagaa	agttgttatac	1260
ccccctggtt	ctaagtccat	ctccaatcgt	gctttaattc	ttgctgccct	cggngaaggt	1320
caatgtaaaa	tcaagaactt	attacattct	gatgatacta	aacatatggt	aaccgctggt	1380
catgaattga	aagggtgctac	gatatcatgg	gaagataatg	gtgagacggt	agtgggtggaa	1440
ggacatggtg	gttccacatt	gtcagcttgt	gctgaccctt	tatatctagg	taatgcaggt	1500
actgcatcta	gatttttgac	ttccttggct	gccttggcca	attctacttc	aagccaaaag	1560
tatatcgttt	taactggtaa	cgcaagaatg	caacaaagac	caattgctcc	tttggctgat	1620
tctttgogtg	ctaattgtac	taaaattgag	tacttgaata	atgaagggtc	cctgccaatc	1680
aaagtttata	ctgattcggg	attcaaaggt	ggtagaattg	aattagctgc	tacagtttct	1740
tctcagtacg	tatcctctat	cttgatgtgt	gccccatacg	ctgaagaacc	tgtaactttg	1800
gctcttggtg	gtggtaaacc	aatctctaaa	ttgtacgtcg	atatgacaat	aaaaatgatg	1860
gaaaaattcg	gtatcaatgt	tgaaaacttct	actacagaac	cttacctta	ttatattcca	1920
aagggacatt	atattaaccc	atcagaatac	gtcattgaaa	gtgatgcctc	aagtgcatac	1980
taccattggg	ccttcgccgc	aatgactggg	actaccgtaa	cggttccaaa	cattggtttt	2040
gagtcgttac	aaggtgatgc	cagatttgca	agagatgtct	tgaaacctat	gggttgtaaa	2100
ataactcaaa	cggcaacttc	aactactggt	togggtcctc	ctgtaggtac	tttaaagcca	2160
ttaaaacatg	ttgatatgga	gccaatgact	gatgcgttct	taactgcatg	tgttgttgcc	2220
gctatttcgc	acgacagtga	tccaaattct	gcaaatacaa	ccaccattga	aggtattgca	2280
aaccagcgtg	tcaaagagtg	taacagaatt	ttggccatgg	ctacagagct	cgccaaattt	2340
ggcgtcaaaa	ctacagaatt	accagatggg	attcaagtcc	atggtttaaa	ctcgataaaa	2400
gatttgaagg	ttccttccga	ctcttctgga	cctgtcggtg	tatgcacata	tgatgatcat	2460
cgtgtggcca	tgagtttctc	gcttcttgca	ggaatggtaa	attctcaaaa	tgaactgac	2520
gaagttgcta	atcctgtaag	aatacttgaa	agacattgta	ctggtaaaac	ctggcctggc	2580
tggtgggatg	tgttacattc	cgaactaggt	gccaaattag	atggtgcaga	acctttagag	2640
tgacatcca	aaaagaactc	aaagaaaagc	gttgtcatta	ttggcatgag	agcagctggc	2700
aaaactacta	taagtaaag	gtgcgcaccc	gctctgggtt	acaaattagt	tgacctagac	2760
gagctgtttg	agcaacagca	taacaatcaa	agtgttaaac	aatttgttgt	ggagaacggt	2820
tgggagaagt	tccgtgagga	agaaacaaga	atcttcaagg	aagttattca	aaattacggc	2880
gatgatggat	atgttttctc	aacaggtggc	ggtattgttg	aaagcgtga	gtctagaaaa	2940
gccttaaaag	atcttgcctc	atcaggtgga	tacgttttac	acttacatag	ggatattgag	3000
gagacaattg	tctttttaca	aagtgatcct	tcaagacctg	cctatgtgga	agaaattcgt	3060

ES 2 719 304 T3

gaagtttga acagaagga ggggtggtat aaagaatgct caaatcttc tttctttgct 3120
 cctcattgct cgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
 gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
 ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
 gaggtctgtag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
 cagttatcta tattgcgtaa agccactgac agtatttcta tcatttttac tgtgcgaacc 3420
 atgaagcaag gtggcaactt tcctgatgaa gagttcaaaa ccttgagaga gctatacgat 3480
 attgccttga agaatggtgt tgaattcctt gacttagaac taactttacc tactgatatc 3540
 caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
 caagattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
 cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga 3720
 ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
 aaagtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
 cctaactcgc ctgcccctgg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
 atgggaggtc tcgagcctaa ggaactgttt gttgttgaa agccaattgg ccactctaga 3960
 tcgccaattt tacataacac tggctatgaa attttaggtt tacctcacia gttcgataaa 4020
 tttgaaactg aatcgcaca attgggaaaa gaaaaacttt tggacggaaa caagaacttt 4080
 ggcggtgctg cagtcacaat tcctctgaaa ttagatataa tgcagtacat ggatgaattg 4140
 actgatgctg ctaaagttat tgggtgctgta aacacagtta taccattggg taacaagaag 4200
 tttaaaggtg ataataccga ctggttaggt atccgtaatg cettaattaa caatggcgtt 4260
 cccgaatatg ttggtcatac cgctggtttg gttatcggtg caggtggcac ttctagagcc 4320
 gccctttacg ccttgcacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
 tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaatagag 4440
 tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
 gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaaggt 4560
 gcccatgctg cttttgtacc aaccttattg gaagccgcat acaaaccaag cgttactccc 4620
 gttatgacaa tttcacaaga caaatatcaa tggcacggtt tccttgatc acaaatgtta 4680
 gtacaccaag gtgtagctca gtttgaaaag tggacaggat tcaagggccc tttcaaggcc 4740
 atttttgatg ccgttacgaa agagtag 4767

<210> 12
 < 211> 1588
 < 212> PRT

5 < 213> Secuencia artificial

<220>
 < 223> Polipéptido sintético

ES 2 719 304 T3

<400> 12

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
 1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
 20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
 35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
 50 55 60

Glu Gly Ser Arg Leu Leu Thr Tyr Val Val Lys Pro Gly Glu Thr Ser
 65 70 75 80

Lys Ser Arg Glu Thr Lys Ala Gln Leu Glu Asp Tyr Leu Leu Val Glu
 85 90 95

Gly Cys Thr Arg Asp Thr Val Met Val Ala Ile Gly Gly Gly Val Ile
 100 105 110

Gly Asp Met Ile Gly Phe Val Ala Ser Thr Phe Met Arg Gly Val Arg
 115 120 125

Val Val Gln Val Pro Thr Ser Leu Leu Ala Met Val Asp Ser Ser Ile
 130 135 140

Gly Gly Lys Thr Ala Ile Asp Thr Pro Leu Gly Lys Asn Phe Ile Gly
 145 150 155 160

Ala Phe Trp Gln Pro Lys Phe Val Leu Val Asp Ile Lys Trp Leu Glu
 165 170 175

Thr Leu Ala Lys Arg Glu Phe Ile Asn Gly Met Ala Glu Val Ile Lys
 180 185 190

Thr Ala Cys Ile Trp Asn Ala Asp Glu Phe Thr Arg Leu Glu Ser Asn
 195 200 205

Ala Ser Leu Phe Leu Asn Val Val Asn Gly Ala Lys Asn Val Lys Val
 210 215 220

Thr Asn Gln Leu Thr Asn Glu Ile Asp Glu Ile Ser Asn Thr Asp Ile

ES 2 719 304 T3

Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495

Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510

Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525

Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540

Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560

Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala
 565 570 575

Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
 580 585 590

Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
 595 600 605

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
 610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
 625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
 645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
 660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
 675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
 690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
 705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
 725 730 735

ES 2 719 304 T3

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

ES 2 719 304 T3

Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
 995 1000 1005

 Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
 1010 1015 1020

 Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
 1025 1030 1035

 Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
 1040 1045 1050

 Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
 1055 1060 1065

 Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe
 1070 1075 1080

 Asp Asp Leu Thr Glu Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr
 1085 1090 1095

 Gly Cys Glu Ala Val Glu Val Arg Val Asp His Leu Ala Asn Tyr
 1100 1105 1110

 Ser Ala Asp Phe Val Ser Lys Gln Leu Ser Ile Leu Arg Lys Ala
 1115 1120 1125

 Thr Asp Ser Ile Pro Ile Ile Phe Thr Val Arg Thr Met Lys Gln
 1130 1135 1140

 Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

 Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

 Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

 Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

 Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

 Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val

ES 2 719 304 T3

1220		1225		1230	
Asn Phe	Glu Asp	Asn Leu	Arg	Leu Glu His Phe	Arg Asp Thr His
1235			1240	1245	
Lys Asn	Lys Pro	Leu Ile	Ala	Val Asn Met Thr	Ser Lys Gly Ser
1250			1255	1260	
Ile Ser	Arg Val	Leu Asn	Asn	Val Leu Thr Pro	Val Thr Ser Asp
1265			1270	1275	
Leu Leu	Pro Asn	Ser Ala	Ala	Pro Gly Gln Leu	Thr Val Ala Gln
1280			1285	1290	
Ile Asn	Lys Met	Tyr Thr	Ser	Met Gly Gly Ile	Glu Pro Lys Glu
1295			1300	1305	
Leu Phe	Val Val	Gly Lys	Pro	Ile Gly His Ser	Arg Ser Pro Ile
1310			1315	1320	
Leu His	Asn Thr	Gly Tyr	Glu	Ile Leu Gly Leu	Pro His Lys Phe
1325			1330	1335	
Asp Lys	Phe Glu	Thr Glu	Ser	Ala Gln Leu Gly	Lys Glu Lys Leu
1340			1345	1350	
Leu Asp	Gly Asn	Lys Asn	Phe	Gly Gly Ala Ala	Val Thr Ile Pro
1355			1360	1365	
Leu Lys	Leu Asp	Ile Met	Gln	Tyr Met Asp Glu	Leu Thr Asp Ala
1370			1375	1380	
Ala Lys	Val Ile	Gly Ala	Val	Asn Thr Val Ile	Pro Leu Gly Asn
1385			1390	1395	
Lys Lys	Phe Lys	Gly Asp	Asn	Thr Asp Trp Leu	Gly Ile Arg Asn
1400			1405	1410	
Ala Leu	Ile Asn	Asn Gly	Val	Pro Glu Tyr Val	Gly His Thr Ala
1415			1420	1425	
Gly Leu	Val Ile	Gly Ala	Gly	Gly Thr Ser Arg	Ala Ala Leu Tyr
1430			1435	1440	
Ala Leu	His Ser	Leu Gly	Cys	Lys Lys Ile Phe	Ile Ile Asn Arg
1445			1450	1455	

ES 2 719 304 T3

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys
1535 1540 1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
1580 1585

<210> 13

< 211> 4767

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Ácido nucleico sintético

<400> 13

atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac 60

attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt 120

tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag 180

gcttctttgc cagaaggctc tcgtttactt acttatgttg ttaaaccagg tgagacaagt 240

aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt 300

gatacggtta tggtagcgat cgggtggtgt gttattggtg acatgattgg gttcgttgca 360

tctacattta tgagaggtgt tcgtgttgtc caagtaccaa catccttatt ggcaatggtc 420

gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt 480

gcattttggc aacccaaaatt tgtccttgta gatattaaat ggctagaaac gttagccaag 540

10

ES 2 719 304 T3

agagagttta tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac 600
 gaatttacta gattagaatc aaacgcttcg ttgttcttaa atgttgtaa tggggcaaaa 660
 aatgtcaagg ttaccaatca attgacaaa gagattgacg agatatacga tacagatatt 720
 gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa 780
 gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct 840
 attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc 900
 attggtatgg ttaaagaggg ggaattatcc cgttatttcg gtattctctc ccctacccaa 960
 gttgcacgtc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa 1020
 tggtttaaag agctaaccct acataagaaa acaccattgg atatcttatt gaagaaaatg 1080
 agtattgaca agaaaaacga gggttccaaa aagaaggtgg tcattttaga aagtattggt 1140
 aagtgctatg gtgactccgc tcaatttggt agcgaagaag acctgagatt tattctaaca 1200
 gatgaaacct tcgtttacc cttcaaggac atccctgctg atcaacagaa agttgttacc 1260
 cccccgtgt ctaagtccat ctccaatcgt gctttaattc ttgctgccct cggatgaagg 1320
 caatgtaaaa tcaagaactt attacattct gatgatacta aacatatggt aaccgctgtt 1380
 catgaattga aagtgctac gatatcatgg gaagataatg gtgagacggg agtggggaa 1440
 ggacatggg gttccacatt gtcagcttgt gctgaccctt tatatctagg taatgcagg 1500
 actgcatcta gatttttgac ttccttggct gccttggta atttacttc aagccaaaag 1560
 tataatggtt taactggtaa cgcaagaatg caacaaagac caattgctcc tttggtgat 1620
 tctttgctg ctaatggtag taaaattgag taactgaata atgaaggttc cctgccaatc 1680
 aaagtttata ctgattcggg attcaaagg ggtagaattg aattagctgc tacagtttct 1740
 tctcagtacg tatcctctat cttgatgtgt gcccatacgt ctgaagaacc tgtaactttg 1800
 gctcttggtg gtgtaagcc aatctctaaa ttgtacgtcg atatgacaat aaaaatgatg 1860
 gaaaaattcg gtatcaatgt tgaaacttct actacagaac cttacactta ttatattcca 1920
 aaggacatt atattaacc atcagaatac gtcattgaaa gtgatgcctc aagtgctaca 1980
 taccattgg ccttcgccgc aatgactggg actaccgtaa cggttccaaa cattggtttt 2040
 gagtcgttac aagtgatgc cagatttgca agagatgtct tgaaacctat gggttgtaa 2100
 ataactcaa cggcaacttc aactactggt tcgggtcctc ctgtaggtac tttaaagcca 2160
 ttaaaacatg ttgatatgga gccaatgact gatgcgttct taactgcatg tgttgttgc 2220
 gctatttcgc acgacagtga tccaaattct gcaaatata ccaccattga aggtattgca 2280
 aaccagcgtg tcaaagagt taacagaatt ttggccatgg ctacagagct cgccaaattt 2340
 ggcgtcaaaa ctacagaatt accagatggg attcaagtc atggtttaaa ctcgataaaa 2400
 gatttgaagg ttccttcga ctctctgga cctgtcggg tatgcacata tgatgatcat 2460

ES 2 719 304 T3

cggtgggcca tgagtttctc gcttcttgca ggaatggtaa attctcaaaa tgaacgtgac	2520
gaagttgcta atcctgtaag aatacttgaa agacattgta ctggtaaaac ctggcctggc	2580
tgggtgggatg tgttacattc cgaactaggt gccaaattag atgggtgcaga acctttagag	2640
tgcacatcca aaaagaactc aaagaaaagc gttgtcatta ttggcatgag agcagctggc	2700
aaaactacta taagtaaatg gtgcgcatcc gctctgggtt acaaattagt tgacctagac	2760
gagctgtttg agcaacagca taacaatcaa agtggttaaac aatttgttgt ggagaacggt	2820
tgggagaagt tccgtgagga agaaacaaga attttcaagg aagttattca aaattacggc	2880
gatgatggat atgttttctc aacaggtggc ggtattgttg aaagcgctga gtctagaaaa	2940
gccttaaaag attttgccctc atcaggtgga tacgttttac acttacatag ggatattgag	3000
gagacaattg tctttttaca aagtgatcct tcaagacctg cctatgtgga agaaattcgt	3060
gaagtttgga acagaaggga ggggtggtat aaagaatgct caaatttctc tttctttgct	3120
cctcattgct ccgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt	3180
gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt	3240
ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt	3300
gaggctgtag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa	3360
cagttatcta tattgcgtaa agccactgac agtattccta tcatttttac tgtgcgaacc	3420
atgaagcaag gtggcaactt tcctgatgaa gagttcaaaa ccttgagaga gctatacgat	3480
attgccttga agaatggtgt tgaattcctt gacttagaac taactttacc tactgatatc	3540
caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc	3600
caaggattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact	3660
cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga	3720
ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct	3780
aaaggtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg	3840
cctaactccg ctgcccctgg ccaattgaca gtagcacaaa ttaacaagat gtatacatct	3900
atgggaggta tcgagcctaa ggaactgttt gttgttgtaa agccaattgg ccactctaga	3960
tcgccaatth tacataacac tggctatgaa attttaggtt tacctcacia gttcgataaa	4020
tttgaaactg aatccgcaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt	4080
ggcgtgctg cagtcaggat tcctctgaaa ttagatataa tgcagtacat ggatgaattg	4140
actgatgctg ctaaagttat tgggtctgta aacacagtta taccattggg taacaagaag	4200
tttaaggggtg ataataccga ctggttaggt atccgtaatg cettaattaa caatggcgtt	4260
cccgaatatg ttggtcatac cgctggtttg gttatcggtg caggtggcac ttctagagcc	4320

ES 2 719 304 T3

gccctttacg ccttgacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
 tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaatagag 4440
 tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
 gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaaggt 4560
 gcccatgctg cttttgtacc aaccttattg gaagccgcat acaaaccaag cgttactccc 4620
 gttatgacaa tttcacaaga caaatatcaa tggcacgttg tccctggatc acaaatgtta 4680
 gtacaccaag gtgtagctca gtttgaaaag tggacaggat tcaagggccc tttcaaggcc 4740
 atttttgatg ccgttacgaa agagtag 4767

<210> 14

<211> 1588

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 14

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
 1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
 20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
 35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
 50 55 60

Glu Gly Ser Arg Leu Leu Thr Tyr Val Val Lys Pro Gly Glu Thr Ser
 65 70 75 80

Lys Ser Arg Glu Thr Lys Ala Gln Leu Glu Asp Tyr Leu Leu Val Glu
 85 90 95

Gly Cys Thr Arg Asp Thr Val Met Val Ala Ile Gly Gly Gly Val Ile
 100 105 110

Gly Asp Met Ile Gly Phe Val Ala Ser Thr Phe Met Arg Gly Val Arg
 115 120 125

Val Val Gln Val Pro Thr Ser Leu Leu Ala Met Val Asp Ser Ser Ile
 130 135 140

10

ES 2 719 304 T3

Gly Gly Lys Thr Ala Ile Asp Thr Pro Leu Gly Lys Asn Phe Ile Gly
 145 150 155 160

Ala Phe Trp Gln Pro Lys Phe Val Leu Val Asp Ile Lys Trp Leu Glu
 165 170 175

Thr Leu Ala Lys Arg Glu Phe Ile Asn Gly Met Ala Glu Val Ile Lys
 180 185 190

Thr Ala Cys Ile Trp Asn Ala Asp Glu Phe Thr Arg Leu Glu Ser Asn
 195 200 205

Ala Ser Leu Phe Leu Asn Val Val Asn Gly Ala Lys Asn Val Lys Val
 210 215 220

Thr Asn Gln Leu Thr Asn Glu Ile Asp Glu Ile Ser Asn Thr Asp Ile
 225 230 235 240

Glu Ala Met Leu Asp His Thr Tyr Lys Leu Val Leu Glu Ser Ile Lys
 245 250 255

Val Lys Ala Glu Val Val Ser Ser Asp Glu Arg Glu Ser Ser Leu Arg
 260 265 270

Asn Leu Leu Asn Phe Gly His Ser Ile Gly His Ala Tyr Glu Ala Ile
 275 280 285

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

Lys Glu Ala Glu Leu Ser Arg Tyr Phe Gly Ile Leu Ser Pro Thr Gln
 305 310 315 320

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
 325 330 335

Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
 340 345 350

Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly
 355 360 365

Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly
 370 375 380

Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr
 385 390 395 400

ES 2 719 304 T3

Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln
 405 410 415
 Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu
 420 425 430
 Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu
 435 440 445
 His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys
 450 455 460
 Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu
 465 470 475 480
 Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495
 Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510
 Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525
 Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540
 Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560
 Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala
 565 570 575
 Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
 580 585 590
 Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
 595 600 605
 Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
 610 615 620
 Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
 625 630 635 640
 Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
 645 650 655

ES 2 719 304 T3

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
725 730 735

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
805 810 815

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
885 890 895

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu

ES 2 719 304 T3

Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val
 1220 1225 1230

Asn Phe Glu Asp Asn Leu Arg Leu Glu His Phe Arg Asp Thr His
 1235 1240 1245

Lys Asn Lys Pro Leu Ile Ala Val Asn Met Thr Ser Lys Gly Ser
 1250 1255 1260

Ile Ser Arg Val Leu Asn Asn Val Leu Thr Pro Val Thr Ser Asp
 1265 1270 1275

Leu Leu Pro Asn Ser Ala Ala Pro Gly Gln Leu Thr Val Ala Gln
 1280 1285 1290

Ile Asn Lys Met Tyr Thr Ser Met Gly Gly Ile Glu Pro Lys Glu
 1295 1300 1305

Leu Phe Val Val Gly Lys Pro Ile Gly His Ser Arg Ser Pro Ile
 1310 1315 1320

Leu His Asn Thr Gly Tyr Glu Ile Leu Gly Leu Pro His Lys Phe
 1325 1330 1335

Asp Lys Phe Glu Thr Glu Ser Ala Gln Leu Val Lys Glu Lys Leu
 1340 1345 1350

Leu Asp Gly Asn Lys Asn Phe Gly Gly Ala Ala Val Arg Ile Pro
 1355 1360 1365

Leu Lys Leu Asp Ile Met Gln Tyr Met Asp Glu Leu Thr Asp Ala
 1370 1375 1380

ES 2 719 304 T3

Ala Lys Val Ile Gly Ala Val Asn Thr Val Ile Pro Leu Gly Asn
 1385 1390 1395

Lys Lys Phe Lys Gly Asp Asn Thr Asp Trp Leu Gly Ile Arg Asn
 1400 1405 1410

Ala Leu Ile Asn Asn Gly Val Pro Glu Tyr Val Gly His Thr Ala
 1415 1420 1425

Gly Leu Val Ile Gly Ala Gly Gly Thr Ser Arg Ala Ala Leu Tyr
 1430 1435 1440

Ala Leu His Ser Leu Gly Cys Lys Lys Ile Phe Ile Ile Asn Arg
 1445 1450 1455

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys
 1535 1540 1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
 1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
 1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
 1580 1585

<210> 15

< 211> 4767

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Ácido nucleico sintético

ES 2 719 304 T3

<400> 15
atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac 60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt 120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag 180
gcttctttgc cagaaggctc tcgtttactt acttatggtt ttaaaccagg tgagacaagt 240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt 300
gatacggtta tggtagcgat cggtaggtgt gttattggtg acatgattgg gttcgttgca 360
tctacattta tgagaggtgt tcgtggtgtc caagtaccaa catccttatt ggcaatggtc 420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt 480
gcattttggc aacccaaatt tgccttcta gatattaaat ggctagaaac gttagccaag 540
agagagttta tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac 600
gaatttacta gattagaatc aaacgcttcg ttgttcttaa atggtgtaa tggggcaaaa 660
aatgtcaagg ttaccaatca attgacaaac gagattgacg agatatacga tacagatatt 720
gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa 780
gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct 840
attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc 900
attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa 960
gttgcacgtc tatccaagat tttggttgc tacgggttgc ctgtttcgcc tgatgagaaa 1020
tggtttaaag agctaacctt acataagaaa acaccattgg atatcttatt gaagaaaatg 1080
agtattgaca agaaaaacga gggttccaaa aagaagggtg tcattttaga aagtattggt 1140
aagtgtatg gtgactccgc tcaatttgtt agcgatgaag acctgagatt tattctaaca 1200
gatgaaacc tcgtttacc cttcaaggac atccctgctg atcaacagaa agttgttatc 1260
ccccctggtt ctaagtccat ctccaatcgt gctttaattc ttgctgccct cggatgaagg 1320
caatgtaaaa tcaagaactt attacattct gatgatacta aacatatgtt aaccgctgtt 1380
catgaattga aagggtgctac gatatcatgg gaagataatg gtgagacggt agtgggtgaa 1440
ggacatggtg gttccacatt gtcagcttgt gctgaccctt tatatctagg taatgcagg 1500
actgcatcta gatTTTTGAC ttccttggct gccttggca attctacttc aagccaaaag 1560
tatatcgttt taactggtaa cgcaagaatg caacaaagac caattgctcc tttggtcgat 1620
tctttgcgtg ctaatggtac taaaattgag tacttgaata atgaaggttc cctgccaatc 1680
aaagtttata ctgattcgggt attcaaagggt ggtagaattg aattagctgc tacagtttct 1740
tctcagtacg tatcctctat cttgatgtgt gcccatacgt ctgaagaacc tgtaactttg 1800

ES 2 719 304 T3

gctcttggtg gtggttaagcc aatctctaaa ttgtacgtcg atatgacaat aaaaatgatg 1860
gaaaaattcg gtatcaatgt tgaaacttct actacagaac cttacactta ttatattcca 1920
aaggacatt atattaaccc atcagaatac gtcattgaaa gtgatgcctc aagtgtctaca 1980
taccattgg ccttcgccc aatgactggg actaccgtaa cggttccaaa cattggtttt 2040
gagtcgttac aaggtgatgc cagatttgca agagatgtct tgaaacctat gggttgtaaa 2100
ataactcaaa cggcaacttc aactactggt tcgggtcctc ctgtaggtac tttaaagcca 2160
ttaaaacatg ttgatatgga gccaatgact gatgcgttct taactgcatg tgttggtgcc 2220
gctatttcgc acgacagtga tccaaattct gcaaatacaa ccaccattga aggtattgca 2280
aaccagcgtg tcaaagagtg taacagaatt ttggccatgg ctacagagct cgccaaattt 2340
ggcgtcaaaa ctacagaatt accagatggg attcaagtcc atggtttaaa ctcgataaaa 2400
gatttgaagg ttccttcgca ctctctgga cctgtcggg tatgcacata tgatgatcat 2460
cgtgtggcca tgagtttctc gcttcttgca ggaatgtaa attctcaaaa tgaacgtgac 2520
gaagttgcta atcctgtaag aatacttgaa agacattgta ctggtaaaac ctggcctggc 2580
tgggtgggatg tgttacattc cgaactaggt gccaaattag atgggtgcaga accttagag 2640
tgcaatcca aaaagaactc aaagaaaagc gttgtcatta ttggcatgag agcagctggc 2700
aaaactacta taagtaaatg gtgcgcatcc gctctgggtt acaaattagt tgacctagac 2760
gagctgtttg agcaacagca taacaatcaa agtgttaaac aatttgttgt ggagaacggt 2820
tgggagaagt tccgtgagga agaaacaaga attttcaagg aagttattca aaattacggc 2880
gatgatggat atgttttctc aacaggtggc ggtattgttg aaagcgtga gtctagaaaa 2940
gccttaaaag attttgcctc atcaggtgga tacgttttac acttacetag ggatattgag 3000
gagacaattg tctttttaca aagtgatect tcaagacctg cctatgtgga agaaattcgt 3060
gaagtttga acagaagggg ggggtgggat aaagaatgct caaatttctc tttctttgct 3120
cctcattgct ccgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
gaggctgtag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
cagttatcta tattcgtaa agccactgac agtattccta tcatttttac tgtgcgaacc 3420
atgaagcaag gtggcaactt tectgatgaa gagttcaaaa cettgagaga gctatacgat 3480
attgccttga agaattggtg tgaattcctt gacttagaac taactttacc tactgatatc 3540
caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
caaggattat actcctggga cgacgtgaa tgggaaaaca gattcaatca agcgttaact 3660
cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcaaga taatttgaga 3720

ES 2 719 304 T3

ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
aaaggtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
cctaactccg ctgccctg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
atgggaggtg tgcagcctaa ggaactgttt gttgttgaa agccaattgg ccaactctaga 3960
tcgccaatth tacataacac tggctatgaa attttaggtt tacctcaciaa gttcgataaa 4020
tttgaaactg aatccgcaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt 4080
ggcgggtgctg cagtcacaat tcctctgaaa ttagatataa tgcagtacat ggatgaattg 4140
actgatgctg ctaaagttca tgggtgctgta aacacagtta taccattggg taacaagaag 4200
tttaagggtg ataataccga ctgggttaggt atccgtaatg ccttaattaa caatggcgtt 4260
cccgaatatg ttggtcatac cgctggttg gttatcgggtg caggtggcac ttctagagcc 4320
gccctttacg ccttgcacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaatagag 4440
tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaagg 4560
gcccatgctg cttttgtacc aaccttattg gaagccgcat acaaaccaag cgttactccc 4620
gttatgacaa tttcacaaga caaatatcaa tggcacggtg tccttgatc acaaatgtta 4680
gtacaccaag gtgtagctca gtttgaaaag tggacaggat tcaagggcc tttcaaggcc 4740
atthttgatg ccgttacgaa agagtag 4767

<210> 16

< 211> 1588

< 212> PRT

5 < 213> Secuencia artificial

<220>

< 223> Polipéptido sintético

<400> 16

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
50 55 60

10

ES 2 719 304 T3

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

ES 2 719 304 T3

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
 325 330 335

Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
 340 345 350

Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly
 355 360 365

Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly
 370 375 380

Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr
 385 390 395 400

Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln
 405 410 415

Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu
 420 425 430

Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu
 435 440 445

His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys
 450 455 460

Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu
 465 470 475 480

Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495

Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510

Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525

Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540

Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560

Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala

ES 2 719 304 T3

565 570 575

Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
580 585 590

Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
595 600 605

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
725 730 735

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
805 810 815

ES 2 719 304 T3

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
 820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
 835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
 850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
 865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
 885 890 895

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu
 900 905 910

Gly Tyr Lys Leu Val Asp Leu Asp Glu Leu Phe Glu Gln Gln His Asn
 915 920 925

Asn Gln Ser Val Lys Gln Phe Val Val Glu Asn Gly Trp Glu Lys Phe
 930 935 940

Arg Glu Glu Glu Thr Arg Ile Phe Lys Glu Val Ile Gln Asn Tyr Gly
 945 950 955 960

Asp Asp Gly Tyr Val Phe Ser Thr Gly Gly Gly Ile Val Glu Ser Ala
 965 970 975

Glu Ser Arg Lys Ala Leu Lys Asp Phe Ala Ser Ser Gly Gly Tyr Val
 980 985 990

Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
 995 1000 1005

Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
 1010 1015 1020

Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
 1025 1030 1035

Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
 1040 1045 1050

Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
 1055 1060 1065

ES 2 719 304 T3

Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe
 1070 1075 1080

Asp Asp Leu Thr Glu Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr
 1085 1090 1095

Gly Cys Glu Ala Val Glu Val Arg Val Asp His Leu Ala Asn Tyr
 1100 1105 1110

Ser Ala Asp Phe Val Ser Lys Gln Leu Ser Ile Leu Arg Lys Ala
 1115 1120 1125

Thr Asp Ser Ile Pro Ile Ile Phe Thr Val Arg Thr Met Lys Gln
 1130 1135 1140

Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val
 1220 1225 1230

Asn Phe Glu Asp Asn Leu Arg Leu Glu His Phe Arg Asp Thr His
 1235 1240 1245

Lys Asn Lys Pro Leu Ile Ala Val Asn Met Thr Ser Lys Gly Ser
 1250 1255 1260

Ile Ser Arg Val Leu Asn Asn Val Leu Thr Pro Val Thr Ser Asp
 1265 1270 1275

Leu Leu Pro Asn Ser Ala Ala Pro Gly Gln Leu Thr Val Ala Gln
 1280 1285 1290

Ile Asn Lys Met Tyr Thr Ser Met Gly Gly Ile Glu Pro Lys Glu
 1295 1300 1305

ES 2 719 304 T3

Leu Phe Val Val Gly Lys Pro Ile Gly His Ser Arg Ser Pro Ile
 1310 1315 1320

Leu His Asn Thr Gly Tyr Glu Ile Leu Gly Leu Pro His Lys Phe
 1325 1330 1335

Asp Lys Phe Glu Thr Glu Ser Ala Gln Leu Val Lys Glu Lys Leu
 1340 1345 1350

Leu Asp Gly Asn Lys Asn Phe Gly Gly Ala Ala Val Thr Ile Pro
 1355 1360 1365

Leu Lys Leu Asp Ile Met Gln Tyr Met Asp Glu Leu Thr Asp Ala
 1370 1375 1380

Ala Lys Val His Gly Ala Val Asn Thr Val Ile Pro Leu Gly Asn
 1385 1390 1395

Lys Lys Phe Lys Gly Asp Asn Thr Asp Trp Leu Gly Ile Arg Asn
 1400 1405 1410

Ala Leu Ile Asn Asn Gly Val Pro Glu Tyr Val Gly His Thr Ala
 1415 1420 1425

Gly Leu Val Ile Gly Ala Gly Gly Thr Ser Arg Ala Ala Leu Tyr
 1430 1435 1440

Ala Leu His Ser Leu Gly Cys Lys Lys Ile Phe Ile Ile Asn Arg
 1445 1450 1455

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys

ES 2 719 304 T3

1535

1540

1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
 1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
 1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
 1580 1585

<210> 17

< 211> 4767

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Ácido nucleico sintético

<400> 17

```

atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac      60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt      120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag      180
gcttctttgc cagaaggctc tcgtttactt acttatgttg ttaaaccagg tgagacaagt      240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt      300
gatacggtta tggtagcgat cggtggtggt gttattggtg acatgattgg gttcgttgca      360
tctacattta tgagaggtgt tcgtgttgtc caagtaccaa catccttatt ggcaatggtc      420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt      480
gcattttggc aacccaaatt tgtccttgta gatattaaat ggctagaaac gttagccaag      540
agagagttha tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac      600
gaatttacta gattagaatc aaacgcttcg ttgttcttaa atggtgttaa tggggcaaaa      660
aatgtcaagg ttaccaatca attgacaaac gagattgacg agatatcgaa tacagatatt      720
gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa      780
gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct      840
attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc      900
attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa      960
gttgcacgtc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa     1020
tggtttaaag agctaacctt acataagaaa acaccattgg atatcttatt gaagaaaatg     1080
agtattgaca agaaaaacga gggttccaaa agaaggtgg tcattttaga aagtattggt     1140
aagtgctatg gtgactccgc tcaatttggt agcgatgaag acctgagatt tattctaaca     1200
    
```

10

ES 2 719 304 T3

gatgaaaccc	tcgtttacct	cttcaaggac	atccctgctg	atcaacagaa	agttgttatac	1260
ccccctggtt	ctaagtccat	ctccaatcgt	gctttaattc	ttgctgccct	cggngaaggt	1320
caatgtaaaa	tcaagaactt	attacattct	gatgatacta	aacatatgtt	aaccgctggt	1380
catgaattga	aagggtgctac	gatatcatgg	gaagataatg	gtgagacggt	agtgggtggaa	1440
ggacatggtg	gttccacatt	gtcagcttgt	gctgacccct	tatatctagg	taatgcaggt	1500
actgcatcta	gatttttgac	ttccttggct	gccttggcca	attctacttc	aagccaaaag	1560
tatatcgttt	taactggtaa	cgcaagaatg	caacaaagac	caattgctcc	tttggctgat	1620
tctttgogtg	ctaattgtac	taaaattgag	tacttgaata	atgaaggttc	cctgccaatc	1680
aaagtttata	ctgattcggg	attcaaaggt	ggtagaattg	aattagctgc	tacagtttct	1740
tctcagtacg	tatcctctat	cttgatgtgt	gccccatacg	ctgaagaacc	tgtaactttg	1800
gctcttggtg	gtggaagcc	aatctctaaa	ttgtacgtcg	atatgacaat	aaaaatgatg	1860
gaaaaattcg	gtatcaatgt	tgaaacttct	actacagaac	cttacctta	ttatattcca	1920
aagggacatt	atattaaccc	atcagaatac	gtcattgaaa	gtgatgcctc	aagtgcatac	1980
taccattggg	ccttcgccgc	aatgactggg	actaccgtaa	cggttccaaa	cattggtttt	2040
gagtcgttac	aaggtgatgc	cagatttgca	agagatgtct	tgaaacctat	gggttgtaaa	2100
ataactcaaa	cggcaacttc	aactactggt	togggtcctc	ctgtaggtac	tttaaagcca	2160
ttaaaacatg	ttgatatgga	gccaatgact	gatgcgttct	taactgcatg	tgttgttgcc	2220
gctatttcgc	acgacagtga	tccaaattct	gcaaatacaa	ccaccattga	aggtattgca	2280
aaccagcgtg	tcaaagagtg	taacagaatt	ttggccatgg	ctacagagct	cgccaaattt	2340
ggcgtcaaaa	ctacagaatt	accagatggg	attcaagtcc	atggtttaaa	ctcgataaaa	2400
gatttgaagg	ttccttccga	ctcttctgga	cctgtcggtg	tatgcacata	tgatgatcat	2460
cgtgtggcca	tgagtttctc	gcttcttgca	ggaatggtaa	attctcaaaa	tgaactgac	2520
gaagttgcta	atcctgtaag	aatacttgaa	agacattgta	ctggtaaaac	ctggcctggc	2580
tggtgggatg	tgttacattc	cgaactaggt	gccaaattag	atggtgcaga	acctttagag	2640
tgcacatcca	aaaagaactc	aaagaaaagc	gttgtcatta	ttggcatgag	agcagctggc	2700
aaaactacta	taagtaaag	gtgcgcatcc	gctctggggt	acaaattagt	tgacctagac	2760
gagctgtttg	agcaacagca	taacaatcaa	agtgttaaac	aatttgttgt	ggagaacggt	2820
tgggagaagt	tccgtgagga	agaaacaaga	atcttcaagg	aagttattca	aaattacggc	2880
gatgatggat	atgttttctc	aacaggtggc	ggtattgttg	aaagcgtga	gtctagaaaa	2940
gccttaaaag	atcttgcctc	atcaggtgga	tacgttttac	acttacatag	ggatattgag	3000
gagacaattg	tctttttaca	aagtgatcct	tcaagacctg	cctatgtgga	agaaattcgt	3060

ES 2 719 304 T3

gaagtttga acagaagga ggggtggtat aaagaatgct caaatcttc tttctttgct 3120
 cctcattgct cgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
 gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
 ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
 gaggtctgtag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
 cagttatcta tattggttaa agccactgac agtatttcta tcatttttac tgtgcgaacc 3420
 atgaagcaag gtggcaactt tcctgatgaa gagttcaaaa ccttgagaga gctatacgat 3480
 attgccttga agaatggtgt tgaattcctt gacttagaac taactttacc tactgatatc 3540
 caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
 caagattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
 cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga 3720
 ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
 aaagtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
 cctaactcgc ctgcccctgg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
 atgggaggtc tcgagcctaa ggaactgttt gttgttgaa agccaattgg ccactctaga 3960
 tcgccaattt tacataacac tggctatgaa attttaggtt tacctcacia gttcgataaa 4020
 tttgaaactg aatcgcaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt 4080
 ggcggtgctg cagtcacaat tcctctgaaa ttagatataa tgcagtacat ggatgaattg 4140
 actgatgctg ctaaagttat tgggtgctgta aacacagtta taccattggg taacaagaag 4200
 tttaaaggtg ataataccga ctggttaggt atccgtaatg cettaattaa caatggcgtt 4260
 cccgaatatg ttggtcatac cgctggtttg gttatcggtg caggtggcac ttctagagcc 4320
 gccctttacg ccttgcacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
 tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaaatagag 4440
 tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
 gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaaggt 4560
 gcccatgctg cttttgtacc aaccttattg gaagccgcat acaaaccaag cgttactccc 4620
 gttatgacaa tttcacaaga caaatatcaa tggcacggtt tccttgatc acaaatgtta 4680
 gtacaccaag gtgtagctca gtttgaagag gttacaggat tcaagggccc tttcaaggcc 4740
 atttttgatg ccgttacgaa agagtag 4767

<210> 18
 < 211> 1588
 < 212> PRT

5 < 213> Secuencia artificial

<220>
 < 223> Polipéptido sintético

ES 2 719 304 T3

<400> 18

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
 1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
 20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
 35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
 50 55 60

Glu Gly Ser Arg Leu Leu Thr Tyr Val Val Lys Pro Gly Glu Thr Ser
 65 70 75 80

Lys Ser Arg Glu Thr Lys Ala Gln Leu Glu Asp Tyr Leu Leu Val Glu
 85 90 95

Gly Cys Thr Arg Asp Thr Val Met Val Ala Ile Gly Gly Gly Val Ile
 100 105 110

Gly Asp Met Ile Gly Phe Val Ala Ser Thr Phe Met Arg Gly Val Arg
 115 120 125

Val Val Gln Val Pro Thr Ser Leu Leu Ala Met Val Asp Ser Ser Ile
 130 135 140

Gly Gly Lys Thr Ala Ile Asp Thr Pro Leu Gly Lys Asn Phe Ile Gly
 145 150 155 160

Ala Phe Trp Gln Pro Lys Phe Val Leu Val Asp Ile Lys Trp Leu Glu
 165 170 175

Thr Leu Ala Lys Arg Glu Phe Ile Asn Gly Met Ala Glu Val Ile Lys
 180 185 190

Thr Ala Cys Ile Trp Asn Ala Asp Glu Phe Thr Arg Leu Glu Ser Asn
 195 200 205

Ala Ser Leu Phe Leu Asn Val Val Asn Gly Ala Lys Asn Val Lys Val
 210 215 220

Thr Asn Gln Leu Thr Asn Glu Ile Asp Glu Ile Ser Asn Thr Asp Ile

ES 2 719 304 T3

Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495

Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510

Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525

Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540

Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560

Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala
 565 570 575

Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
 580 585 590

Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
 595 600 605

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
 610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
 625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
 645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
 660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
 675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
 690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
 705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
 725 730 735

ES 2 719 304 T3

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

ES 2 719 304 T3

Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
 995 1000 1005

 Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
 1010 1015 1020

 Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
 1025 1030 1035

 Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
 1040 1045 1050

 Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
 1055 1060 1065

 Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe
 1070 1075 1080

 Asp Asp Leu Thr Glu Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr
 1085 1090 1095

 Gly Cys Glu Ala Val Glu Val Arg Val Asp His Leu Ala Asn Tyr
 1100 1105 1110

 Ser Ala Asp Phe Val Ser Lys Gln Leu Ser Ile Leu Arg Lys Ala
 1115 1120 1125

 Thr Asp Ser Ile Pro Ile Ile Phe Thr Val Arg Thr Met Lys Gln
 1130 1135 1140

 Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

 Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

 Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

 Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

 Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

 Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val

ES 2 719 304 T3

1220		1225		1230
Asn Phe	Glu Asp	Asn Leu	Arg Leu	Glu His Phe
1235			1240	Arg Asp Thr His
				1245
Lys Asn	Lys Pro	Leu Ile	Ala Val	Asn Met Thr
1250			1255	Ser Lys Gly Ser
				1260
Ile Ser	Arg Val	Leu Asn	Asn Val	Leu Thr Pro
1265			1270	Val Thr Ser Asp
				1275
Leu Leu	Pro Asn	Ser Ala	Ala Pro	Gly Gln Leu
1280			1285	Thr Val Ala Gln
				1290
Ile Asn	Lys Met	Tyr Thr	Ser Met	Gly Gly Ile
1295			1300	Glu Pro Lys Glu
				1305
Leu Phe	Val Val	Gly Lys	Pro Ile	Gly His Ser
1310			1315	Arg Ser Pro Ile
				1320
Leu His	Asn Thr	Gly Tyr	Glu Ile	Leu Gly Leu
1325			1330	Pro His Lys Phe
				1335
Asp Lys	Phe Glu	Thr Glu	Ser Ala	Gln Leu Val
1340			1345	Lys Glu Lys Leu
				1350
Leu Asp	Gly Asn	Lys Asn	Phe Gly	Gly Ala Ala
1355			1360	Val Thr Ile Pro
				1365
Leu Lys	Leu Asp	Ile Met	Gln Tyr	Met Asp Glu
1370			1375	Leu Thr Asp Ala
				1380
Ala Lys	Val Ile	Gly Ala	Val Asn	Thr Val Ile
1385			1390	Pro Leu Gly Asn
				1395
Lys Lys	Phe Lys	Gly Asp	Asn Thr	Asp Trp Leu
1400			1405	Gly Ile Arg Asn
				1410
Ala Leu	Ile Asn	Asn Gly	Val Pro	Glu Tyr Val
1415			1420	Gly His Thr Ala
				1425
Gly Leu	Val Ile	Gly Ala	Gly Gly	Thr Ser Arg
1430			1435	Ala Ala Leu Tyr
				1440
Ala Leu	His Ser	Leu Gly	Cys Lys	Lys Ile Phe
1445			1450	Ile Ile Asn Arg
				1455

ES 2 719 304 T3

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys
 1535 1540 1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
 1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Val Thr Gly Phe Lys Gly Pro Phe
 1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
 1580 1585

<210> 19

< 211> 4767

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Ácido nucleico sintético

<400> 19

atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac	60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt	120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag	180
gcttctttgc cagaaggctc tcgtttactt acctatggtg ttaaaccagg tgagacaagt	240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt	300
gatacggtta tggtagcgat cgggtggtgt gttattggtg acatgattgg gttcgttgca	360
tctacattta tgagaggtgt tcgtggttgc caagtaccaa catccttatt ggcaatggtc	420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt	480
gcattttggc aacccaaaatt tgtccttcta gatattaaat ggctagaaac gttagccaag	540

10

ES 2 719 304 T3

agagagttta tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac 600
 gaatttacta gattagaatc aaacgcttcg ttgttcttaa atgttgtaa tggggcaaaa 660
 aatgtcaagg ttaccaatca attgacaaa gagattgacg agatatacga tacagatatt 720
 gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa 780
 gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct 840
 attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc 900
 attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa 960
 gttgcacgtc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa 1020
 tggtttaaag agctaaccct acataagaaa acaccattgg atatcttatt gaagaaaatg 1080
 agtattgaca agaaaaacga gggttccaaa aagaaggtgg tcattttaga aagtattggt 1140
 aagtgctatg gtgactccgc tcaatttggt agcgaagaag acctgagatt tattctaaca 1200
 gatgaaaccc tcgtttacc cttcaaggac atccctgctg atcaacagaa agttgttacc 1260
 cccctcggtt ctaagtccat ctccaatcgt gctttaattc ttgctgccct cggatgaagg 1320
 caatgtaaaa tcaagaactt attacattct gatgatacta aacatatggt aaccgctggt 1380
 catgaattga aagtgctac gatatcatgg gaagataatg gtgagacggg agtggggaa 1440
 ggacatggg gttccacatt gtcagcttgt gctgaccctt tatatctagg taatgcagg 1500
 actgcatcta gatttttgac ttccttggtt gccttggtca atttacttc aagccaaaag 1560
 tataatggtt taactggtaa cgcaagaatg caacaaagac caattgctcc tttggtgat 1620
 tctttgctg ctaatggtag taaaattgag taactgaata atgaaggttc cctgccaatc 1680
 aaagtttata ctgattcggg attcaaaggg ggtagaattg aattagctgc tacagtttct 1740
 tctcagtacg tatcctctat cttgatgtgt gcccatacgt ctgaagaacc tgtaactttg 1800
 gctcttggtg gtgtaagcc aatctctaaa ttgtacgtcg atatgacaat aaaaatgatg 1860
 gaaaaattcg gtatcaatgt tgaaacttct actacagaac cttacactta ttatattcca 1920
 aaggacatt atattaaccc atcagaatac gtcattgaaa gtgatgcctc aagtgctaca 1980
 taccattgg ccttcgccgc aatgactggg actaccgtaa cggttccaaa cattggtttt 2040
 gagtcgttac aagtgatgc cagatttgca agagatgtct tgaaacctat gggttgtaa 2100
 ataactcaa cggcaacttc aactactggt tcgggtctc ctgtaggtac ttaaagcca 2160
 ttaaacaatg ttgatatgga gccaatgact gatgcgttct taactgcatg gttgttgcc 2220
 gctatttcgc acgacagtga tccaaattct gcaaatatac ccaccattga aggtattgca 2280
 aaccagcgtg tcaaagagt taacagaatt ttggccatgg ctacagagct cgccaaattt 2340
 ggcgtcaaaa ctacagaatt accagatggg attcaagtc atggtttaaa ctcgataaaa 2400
 gatttgaagg ttccttcga ctctctgga cctgtcggg tatgcacata tgatgatcat 2460

ES 2 719 304 T3

cgtgtggcca tgagtttctc gcttcttgca ggaatggtaa attctcaaaa tgaacgtgac 2520
 gaagttgcta atcctgtaag aatacttgaa agacattgta ctggtaaaac ctggcctggc 2580
 tgggtgggatg tgttacattc cgaactaggt gccaaattag atgggtgcaga acctttagag 2640
 tgcacatcca aaaagaactc aaagaaaagc gttgtcatta ttggcatgag agcagctggc 2700
 aaaactacta taagtaaatg gtgcgcaccc gctctgggtt acaaattagt tgacctagac 2760
 gagctgtttg agcaacagca taacaatcaa agtggttaaac aatttgttgt ggagaacggc 2820
 tgggagaagt tccgtgagga agaaacaaga attttcaagg aagttattca aaattacggc 2880
 gatgatggat atgttttctc aacaggtggc ggtattgttg aaagcgtga gtctagaaaa 2940
 gccttaaaag attttgccctc atcaggtgga tacgttttac acttacatag ggatattgag 3000
 gagacaattg tctttttaca aagtgatcct tcaagacctg cctatgtgga agaaattcgt 3060
 gaagtttggg acagaaggga ggggtgggat aaagaatgct caaatttctc tttctttgct 3120
 cctcattgct ccgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
 gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
 ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
 gaggtctgag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
 cagttatcta tattgcgtaa agccactgac agtattccta tcatttttac tgtgcgaacc 3420
 atgaagcaag gtggcaactt tcctgatgaa gagttcaaaa ccttgagaga gctatacgat 3480
 attgccttga agaatggtgt tgaattcctt gacttagaac taactttacc tactgatatc 3540
 caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
 caaggattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
 cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga 3720
 ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
 aaaggtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
 cctaactccg ctgcccctgg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
 atgggaggta tcgagcctaa ggaactggtt gttgttgtaa agccaattgg ccactctaga 3960
 tcgccaatth tacataacac tggctatgaa attttaggtt tacctcacia gttcgataaa 4020
 tttgaaactg aatccgcaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt 4080
 ggcggtgctg cagtcacaat tcctctgaaa ttagatataa tgcagtacat ggatgaattg 4140
 actgatgctg ctaaagttat tgggtctgta aacaaagtta taccattggg taacaagaag 4200
 tttaaaggtg ataataccga ctggttaggt atccgtaatg cettaattaa caatggcgtt 4260
 cccgaatatg ttggtcatac cgctggtttg gttatcggtg caggtggcac ttctagagcc 4320

ES 2 719 304 T3

gccctttacg ccttgacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
 tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaatagag 4440
 tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
 gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaaggt 4560
 gcccatgctg cttttgtacc aaccttattg gaagccgcat acaaaccaag cgttactccc 4620
 gttatgacaa ttccacaaga caaatatcaa tggcacgttg tccctggatc acaaatgta 4680
 gtacaccaag gtgtagctca gtttgaaaag tggacaggat tcaagggccc tttcaaggcc 4740
 attttgatg ccgttacgaa agagtag 4767

<210> 20

<211> 1588

<212> PRT

5 <213> Secuencia artificial

<220>

<223> Polipéptido sintético

<400> 20

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
 1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
 20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
 35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
 50 55 60

Glu Gly Ser Arg Leu Leu Thr Tyr Val Val Lys Pro Gly Glu Thr Ser
 65 70 75 80

Lys Ser Arg Glu Thr Lys Ala Gln Leu Glu Asp Tyr Leu Leu Val Glu
 85 90 95

Gly Cys Thr Arg Asp Thr Val Met Val Ala Ile Gly Gly Gly Val Ile
 100 105 110

Gly Asp Met Ile Gly Phe Val Ala Ser Thr Phe Met Arg Gly Val Arg
 115 120 125

Val Val Gln Val Pro Thr Ser Leu Leu Ala Met Val Asp Ser Ser Ile
 130 135 140

10

ES 2 719 304 T3

Gly Gly Lys Thr Ala Ile Asp Thr Pro Leu Gly Lys Asn Phe Ile Gly
 145 150 155 160

Ala Phe Trp Gln Pro Lys Phe Val Leu Val Asp Ile Lys Trp Leu Glu
 165 170 175

Thr Leu Ala Lys Arg Glu Phe Ile Asn Gly Met Ala Glu Val Ile Lys
 180 185 190

Thr Ala Cys Ile Trp Asn Ala Asp Glu Phe Thr Arg Leu Glu Ser Asn
 195 200 205

Ala Ser Leu Phe Leu Asn Val Val Asn Gly Ala Lys Asn Val Lys Val
 210 215 220

Thr Asn Gln Leu Thr Asn Glu Ile Asp Glu Ile Ser Asn Thr Asp Ile
 225 230 235 240

Glu Ala Met Leu Asp His Thr Tyr Lys Leu Val Leu Glu Ser Ile Lys
 245 250 255

Val Lys Ala Glu Val Val Ser Ser Asp Glu Arg Glu Ser Ser Leu Arg
 260 265 270

Asn Leu Leu Asn Phe Gly His Ser Ile Gly His Ala Tyr Glu Ala Ile
 275 280 285

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

Lys Glu Ala Glu Leu Ser Arg Tyr Phe Gly Ile Leu Ser Pro Thr Gln
 305 310 315 320

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
 325 330 335

Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
 340 345 350

Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly
 355 360 365

Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly
 370 375 380

Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr
 385 390 395 400

ES 2 719 304 T3

Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln
 405 410 415
 Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu
 420 425 430
 Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu
 435 440 445
 His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys
 450 455 460
 Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu
 465 470 475 480
 Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495
 Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510
 Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525
 Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540
 Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560
 Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala
 565 570 575
 Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
 580 585 590
 Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
 595 600 605
 Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
 610 615 620
 Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
 625 630 635 640
 Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
 645 650 655

ES 2 719 304 T3

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
725 730 735

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
805 810 815

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
885 890 895

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu

ES 2 719 304 T3

Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val
 1220 1225 1230

Asn Phe Glu Asp Asn Leu Arg Leu Glu His Phe Arg Asp Thr His
 1235 1240 1245

Lys Asn Lys Pro Leu Ile Ala Val Asn Met Thr Ser Lys Gly Ser
 1250 1255 1260

Ile Ser Arg Val Leu Asn Asn Val Leu Thr Pro Val Thr Ser Asp
 1265 1270 1275

Leu Leu Pro Asn Ser Ala Ala Pro Gly Gln Leu Thr Val Ala Gln
 1280 1285 1290

Ile Asn Lys Met Tyr Thr Ser Met Gly Gly Ile Glu Pro Lys Glu
 1295 1300 1305

Leu Phe Val Val Gly Lys Pro Ile Gly His Ser Arg Ser Pro Ile
 1310 1315 1320

Leu His Asn Thr Gly Tyr Glu Ile Leu Gly Leu Pro His Lys Phe
 1325 1330 1335

Asp Lys Phe Glu Thr Glu Ser Ala Gln Leu Val Lys Glu Lys Leu
 1340 1345 1350

Leu Asp Gly Asn Lys Asn Phe Gly Gly Ala Ala Val Thr Ile Pro
 1355 1360 1365

Leu Lys Leu Asp Ile Met Gln Tyr Met Asp Glu Leu Thr Asp Ala
 1370 1375 1380

ES 2 719 304 T3

Ala Lys Val Ile Gly Ala Val Asn Lys Val Ile Pro Leu Gly Asn
 1385 1390 1395

Lys Lys Phe Lys Gly Asp Asn Thr Asp Trp Leu Gly Ile Arg Asn
 1400 1405 1410

Ala Leu Ile Asn Asn Gly Val Pro Glu Tyr Val Gly His Thr Ala
 1415 1420 1425

Gly Leu Val Ile Gly Ala Gly Gly Thr Ser Arg Ala Ala Leu Tyr
 1430 1435 1440

Ala Leu His Ser Leu Gly Cys Lys Lys Ile Phe Ile Ile Asn Arg
 1445 1450 1455

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys
 1535 1540 1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
 1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
 1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
 1580 1585

<210> 21

<211> 4767

<212> ADN

5 <213> Secuencia artificial

<220>

<223> Ácido nucleico sintético

ES 2 719 304 T3

<400> 21

atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac	60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt	120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag	180
gcttctttgc cagaaggctc tcgtttactt acttatggtg ttaaaccagg tgagacaagt	240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt	300
gatacggtta tggtagcgat cggtaggtgt gttattggtg acatgattgg gttcgttgca	360
tctacattta tgagaggtgt tcgtggtgtc caagtaccaa catccttatt ggcaatggtc	420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt	480
gcattttggc aaccaaatt tgccttcta gatattaaat ggctagaaac gttagccaag	540
agagagtta tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac	600
gaatttacta gattagaatc aaacgcttcg ttgttcttaa atggtgtaa tggggcaaaa	660
aatgtcaagg ttaccaatca attgacaaac gagattgacg agatatcgaa tacagatatt	720
gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa	780
gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct	840
attggtcatg cttatgaagc tataactaacc ccacaagcat tacatggtga atgtgtgtcc	900
attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa	960
gttgcacgtc tatccaagat tttggttgc tacgggttgc ctgtttcgcc tgatgagaaa	1020
tggtttaaag agctaacctt acataagaaa acaccattgg atatcttatt gaagaaaatg	1080
agtattgaca agaaaaacga gggttccaaa aagaaggtgg tcattttaga aagtattggt	1140
aagtgctatg gtgactccgc tcaatttgtt agcgatgaag acctgagatt tattctaaca	1200
gatgaaacc tcgtttacc cttcaaggac atccctgctg atcaacagaa agttgttatc	1260
ccccctggtt ctaagtccat ctccaatcgt gctttaattc ttgctgccct cgggtgaagg	1320
caatgtaaaa tcaagaactt attacattct gatgatacta aacatatgtt aaccgctgtt	1380
catgaattga aagggtgctac gatatcatgg gaagataatg gtgagacggt agtgggtgaa	1440
ggacatggtg gttccacatt gtcagcttgt gctgaccctt tatatctagg taatgcagg	1500
actgcatcta gatTTTTgac ttccttggct gccttggca attctacttc aagccaaaag	1560
tatatcgttt taactggtaa cgcaagaatg caacaaagac caattgctcc tttggtcgat	1620
tctttgcgtg ctaatggtac taaaattgag tacttgaata atgaaggttc cctgccaatc	1680
aaagtttata ctgattcgggt attcaaagggt ggtagaattg aattagctgc tacagtttct	1740
tctcagtacg tatcctctat cttgatgtgt gcccatacgt ctgaagaacc tgtaactttg	1800

ES 2 719 304 T3

gctcttggtg gtggttaagcc aatctctaaa ttgtacgtcg atatgacaat aaaaatgatg 1860
gaaaaattcg gtatcaatgt tgaaacttct actacagaac cttacactta ttatattcca 1920
aaggacatt atattaaccc atcagaatac gtcattgaaa gtgatgcctc aagtgtctaca 1980
taccattgg ccttcgccc aatgactggg actaccgtaa cggttccaaa cattggtttt 2040
gagtcgttac aaggtgatgc cagatttgca agagatgtct tgaaacctat gggttgtaaa 2100
ataactcaaa cggcaacttc aactactggt tcgggtcctc ctgtaggtac tttaaagcca 2160
ttaaaacatg ttgatatgga gccaatgact gatgcgttct taactgcatg tgttggtgcc 2220
gctatttcgc acgacagtga tccaaattct gcaaatacaa ccaccattga aggtattgca 2280
aaccagcgtg tcaaagagtg taacagaatt ttggccatgg ctacagagct cgccaaattt 2340
ggcgtcaaaa ctacagaatt accagatggg attcaagtcc atggtttaaa ctcgataaaa 2400
gatttgaagg ttccttcgca ctctctgga cctgtcggg tatgcacata tgatgatcat 2460
cgtgtggcca tgagtttctc gcttcttgca ggaatgtaa attctcaaaa tgaacgtgac 2520
gaagttgcta atcctgtaag aatacttgaa agacattgta ctggtaaaac ctggcctggc 2580
tgggtgggatg tgttacattc cgaactaggt gccaaattag atgggtcaga accttagag 2640
tgcaatcca aaaagaactc aaagaaaagc gttgtcatta ttggcatgag agcagctggc 2700
aaaactacta taagtaaatg gtgcgcatcc gctctgggtt acaaattagt tgacctagac 2760
gagctgtttg agcaacagca taacaatcaa agtgttaaac aatttgttgt ggagaacggt 2820
tgggagaagt tccgtgagga agaaacaaga attttcaagg aagttattca aaattacggc 2880
gatgatggat atgttttctc aacaggtggc ggtattgttg aaagcgtga gtctagaaaa 2940
gccttaaaag attttgcctc atcaggtgga tacgttttac acttacatag ggatattgag 3000
gagacaattg tctttttaca aagtgatect tcaagacctg cctatgtgga agaaattcgt 3060
gaagtttggg acagaagggg ggggtgggat aaagaatgct caaatttctc tttctttgct 3120
cctcattgct ccgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
gaggctgtag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
cagttatcta tattcgtaa agccactgac agtattccta tcatttttac tgtgcgaacc 3420
atgaagcaag gtggcaactt tectgatgaa gagttcaaaa cettgagaga gctatacgat 3480
attgccttga agaattggtg tgaattcctt gacttagaac taactttacc tactgatatc 3540
caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
caaggattat actcctggga cgacgtgaa tgggaaaaca gattcaatca agcgttaact 3660
cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga 3720

ES 2 719 304 T3

ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
aaaggtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
cctaactccg ctgccctcg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
atgggaggtg tgcagcctaa ggaactgttt gttgttgaa agccaattgg ccaactctaga 3960
tcgccaattt tacataacac tggctatgaa attttaggtt tacctcaciaa gttcgataaa 4020
tttgaaactg aatccgcaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt 4080
ggcggtgctg cagtcacaat tcctctgtta ttagatataa tgcagtacat ggatgaattg 4140
actgatgctg ctaaagttat tgggtgctgta aacacagtta taccattggg taacaagaag 4200
tttaagggtg ataataccga ctgggttaggt atccgtaatg ccttaattaa caatggcgtt 4260
cccgaatatg ttggtcatac cgctggttg gttatcggtg caggtggcac ttctagagcc 4320
gccctttacg ccttgcacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaatagag 4440
tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaagg 4560
gcccatgctg cttttgtacc aaccttattg gaagccgat acaaaccaag cgttactccc 4620
gttatgacaa tttcacaaga caaatatcaa tggcacggtt tccttgatc acaaatgtta 4680
gtacaccaag gtgtagctca gtttgaaaag tggacaggat tcaagggcc tttcaaggcc 4740
atttttgatg ccgttacgaa agagtag 4767

<210> 22

< 211> 1588

< 212> PRT

5 < 213> Secuencia artificial

<220>

< 223> Polipéptido sintético

<400> 22

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
50 55 60

10

ES 2 719 304 T3

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

ES 2 719 304 T3

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
 325 330 335

Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
 340 345 350

Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly
 355 360 365

Ser Lys Lys Lys Val Val Ile Leu Glu Ser Ile Gly Lys Cys Tyr Gly
 370 375 380

Asp Ser Ala Gln Phe Val Ser Asp Glu Asp Leu Arg Phe Ile Leu Thr
 385 390 395 400

Asp Glu Thr Leu Val Tyr Pro Phe Lys Asp Ile Pro Ala Asp Gln Gln
 405 410 415

Lys Val Val Ile Pro Pro Gly Ser Lys Ser Ile Ser Asn Arg Ala Leu
 420 425 430

Ile Leu Ala Ala Leu Gly Glu Gly Gln Cys Lys Ile Lys Asn Leu Leu
 435 440 445

His Ser Asp Asp Thr Lys His Met Leu Thr Ala Val His Glu Leu Lys
 450 455 460

Gly Ala Thr Ile Ser Trp Glu Asp Asn Gly Glu Thr Val Val Val Glu
 465 470 475 480

Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495

Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510

Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525

Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540

Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560

Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala

ES 2 719 304 T3

565 570 575

Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
580 585 590

Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
595 600 605

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
725 730 735

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
805 810 815

ES 2 719 304 T3

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
 820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
 835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
 850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
 865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
 885 890 895

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu
 900 905 910

Gly Tyr Lys Leu Val Asp Leu Asp Glu Leu Phe Glu Gln Gln His Asn
 915 920 925

Asn Gln Ser Val Lys Gln Phe Val Val Glu Asn Gly Trp Glu Lys Phe
 930 935 940

Arg Glu Glu Glu Thr Arg Ile Phe Lys Glu Val Ile Gln Asn Tyr Gly
 945 950 955 960

Asp Asp Gly Tyr Val Phe Ser Thr Gly Gly Gly Ile Val Glu Ser Ala
 965 970 975

Glu Ser Arg Lys Ala Leu Lys Asp Phe Ala Ser Ser Gly Gly Tyr Val
 980 985 990

Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
 995 1000 1005

Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
 1010 1015 1020

Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
 1025 1030 1035

Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
 1040 1045 1050

Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
 1055 1060 1065

ES 2 719 304 T3

Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe
 1070 1075 1080

Asp Asp Leu Thr Glu Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr
 1085 1090 1095

Gly Cys Glu Ala Val Glu Val Arg Val Asp His Leu Ala Asn Tyr
 1100 1105 1110

Ser Ala Asp Phe Val Ser Lys Gln Leu Ser Ile Leu Arg Lys Ala
 1115 1120 1125

Thr Asp Ser Ile Pro Ile Ile Phe Thr Val Arg Thr Met Lys Gln
 1130 1135 1140

Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val
 1220 1225 1230

Asn Phe Glu Asp Asn Leu Arg Leu Glu His Phe Arg Asp Thr His
 1235 1240 1245

Lys Asn Lys Pro Leu Ile Ala Val Asn Met Thr Ser Lys Gly Ser
 1250 1255 1260

Ile Ser Arg Val Leu Asn Asn Val Leu Thr Pro Val Thr Ser Asp
 1265 1270 1275

Leu Leu Pro Asn Ser Ala Ala Pro Gly Gln Leu Thr Val Ala Gln
 1280 1285 1290

Ile Asn Lys Met Tyr Thr Ser Met Gly Gly Ile Glu Pro Lys Glu
 1295 1300 1305

ES 2 719 304 T3

Leu Phe Val Val Gly Lys Pro Ile Gly His Ser Arg Ser Pro Ile
 1310 1315 1320

Leu His Asn Thr Gly Tyr Glu Ile Leu Gly Leu Pro His Lys Phe
 1325 1330 1335

Asp Lys Phe Glu Thr Glu Ser Ala Gln Leu Val Lys Glu Lys Leu
 1340 1345 1350

Leu Asp Gly Asn Lys Asn Phe Gly Gly Ala Ala Val Thr Ile Pro
 1355 1360 1365

Leu Leu Leu Asp Ile Met Gln Tyr Met Asp Glu Leu Thr Asp Ala
 1370 1375 1380

Ala Lys Val Ile Gly Ala Val Asn Thr Val Ile Pro Leu Gly Asn
 1385 1390 1395

Lys Lys Phe Lys Gly Asp Asn Thr Asp Trp Leu Gly Ile Arg Asn
 1400 1405 1410

Ala Leu Ile Asn Asn Gly Val Pro Glu Tyr Val Gly His Thr Ala
 1415 1420 1425

Gly Leu Val Ile Gly Ala Gly Gly Thr Ser Arg Ala Ala Leu Tyr
 1430 1435 1440

Ala Leu His Ser Leu Gly Cys Lys Lys Ile Phe Ile Ile Asn Arg
 1445 1450 1455

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys

ES 2 719 304 T3

1535

1540

1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
 1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
 1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
 1580 1585

<210> 23

< 211> 4767

< 212> ADN

5 < 213> Secuencia artificial

<220>

< 223> Ácido nucleico sintético

<400> 23

atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac	60
attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt	120
tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag	180
gcttctttgc cagaaggctc tcgtttactt acttatgttg ttaaaccagg tgagacaagt	240
aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt	300
gatacggtta tggtagcgat cgggtggtgt gttattggtg acatgattgg gttcgttgca	360
tctacattta tgagaggtgt tcgtgttgtc caagtaccaa catccttatt ggcaatggtc	420
gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt	480
gcattttggc aacccaaatt tgtccttgta gatattaaat ggctagaaac gttagccaag	540
agagagttha tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac	600
gaatttacta gattagaatc aaacgcttcg ttgttcttaa atggtgttaa tggggcaaaa	660
aatgtcaagg ttaccaatca attgacaac gagattgacg agatatcgaa tacagatatt	720
gaagctatgt tggatcatac atataagtha gttcttgaga gtattaaggt caaagcggaa	780
gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct	840
attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc	900
attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa	960
gttgcacgtc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa	1020
tggtttaaag agctaacctt acataagaaa acaccattgg atatcttatt gaagaaaatg	1080
agtattgaca agaaaaacga gggttccaaa agaaggtgg tcattttaga aagtattggt	1140
aagtgctatg gtgactccgc tcaatttggt agcgatgaag acctgagatt tattctaaca	1200

10

ES 2 719 304 T3

gatgaaaccc	tcgtttacc	cttcaaggac	atccctgctg	atcaacagaa	agttgttatac	1260
ccccctggtt	ctaagtccat	ctccaatcgt	gctttaattc	ttgctgccct	cggngaaggt	1320
caatgtaaaa	tcaagaactt	attacattct	gatgatacta	aacatatgtt	aaccgctggt	1380
catgaattga	aaggtgctac	gatatcatgg	gaagataatg	gtgagacggt	agtgggtgaa	1440
ggacatggtg	gttccacatt	gtcagcttgt	gctgaccctt	tatatctagg	taatgcaggt	1500
actgcatcta	gatttttgac	ttccttggct	gccttggcca	attctacttc	aagccaaaag	1560
tatatcgttt	taactggtaa	cgcaagaatg	caacaaagac	caattgctcc	tttggctgat	1620
tctttgogtg	ctaattgtac	taaaattgag	tacttgaata	atgaaggttc	cctgccaatc	1680
aaagtttata	ctgattcggg	attcaaaggt	ggtagaattg	aattagctgc	tacagtttct	1740
tctcagtacg	tatcctctat	cttgatgtgt	gccccatacg	ctgaagaacc	tgtaactttg	1800
gctcttggtg	gtggaagcc	aatctctaaa	ttgtacgtcg	atatgacaat	aaaaatgatg	1860
gaaaaattcg	gtatcaatgt	tgaaacttct	actacagaac	cttacctta	ttatattcca	1920
aagggacatt	atattaaccc	atcagaatac	gtcattgaaa	gtgatgcctc	aagtgcatac	1980
taccattggg	ccttcgccgc	aatgactggg	actaccgtaa	cggttccaaa	cattggtttt	2040
gagtcgttac	aaggtgatgc	cagatttgca	agagatgtct	tgaaacctat	gggttgtaaa	2100
ataactcaaa	cggcaacttc	aactactggt	togggtcctc	ctgtaggtac	tttaaagcca	2160
ttaaaacatg	ttgatatgga	gccaatgact	gatgcgttct	taactgcatg	tgttgttgcc	2220
gctatttcgc	acgacagtga	tccaaattct	gcaaatacaa	ccaccattga	aggtattgca	2280
aaccagcgtg	tcaaagagtg	taacagaatt	ttggccatgg	ctacagagct	cgccaaattt	2340
ggcgtcaaaa	ctacagaatt	accagatggg	attcaagtcc	atggtttaaa	ctcgataaaa	2400
gatttgaagg	ttccttccga	ctcttctgga	cctgtcggtg	tatgcacata	tgatgatcat	2460
cgtgtggcca	tgagtttctc	gcttcttgca	ggaatggtaa	attctcaaaa	tgaactgac	2520
gaagttgcta	atcctgtaag	aatacttgaa	agacattgta	ctggtaaaac	ctggcctggc	2580
tggtgggatg	tgttacattc	cgaactaggt	gccaaattag	atggtgcaga	acctttagag	2640
tgcacatcca	aaaagaactc	aaagaaaagc	gttgtcatta	ttggcatgag	agcagctggc	2700
aaaactacta	taagtaaag	gtgcgcatcc	gctctggggt	acaaattagt	tgacctagac	2760
gagctgtttg	agcaacagca	taacaatcaa	agtgttaaac	aatttgttgt	ggagaacggt	2820
tgggagaagt	tcogtgagga	agaaacaaga	atcttcaagg	aagttattca	aaattacggc	2880
gatgatggat	atgttttctc	aacaggtggc	ggtattgttg	aaagcgtga	gtctagaaaa	2940
gccttaaaag	atcttgcctc	atcaggtgga	tacgttttac	acttacatag	ggatattgag	3000
gagacaattg	tctttttaca	aagtgatcct	tcaagacctg	cctatgtgga	agaaattcgt	3060

ES 2 719 304 T3

gaagtttga acagaagga ggggtggtat aaagaatgct caaatcttc tttctttgct 3120
 cctcattgct cgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
 gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
 ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
 gaggtctgtag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
 cagttatcta tattggttaa agccactgac agtatttcta tcatttttac tgtgcgaacc 3420
 atgaagcaag gtggcaactt tcctgatgaa gagttcaaaa ccttgagaga gctatacgat 3480
 attgccttga agaatggtgt tgaattcctt gacttagaac taactttacc tactgatatc 3540
 caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
 caagattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
 cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga 3720
 ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
 aaagtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
 cctaactcgc ctgcccctgg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
 atgggaggtc tcgagcctaa ggaactgttt gttgttgtaa agccaattgg ccactctaga 3960
 tcgccaattt tacataacac tggctatgaa attttaggtt tacctcacia gttcgataaa 4020
 tttgaaactg aatcgcgaca attggtgaaa gaaaaacttt tggacggaaa caagaacttt 4080
 ggcggtgctg cagtcacaat tcctctgaaa ttagatataa tgcagtacat ggatgaattg 4140
 actgatgctg ctaaagttat tgggtgctgta aacacagtta taccattggg taacaagaag 4200
 tttaaaggtg ataataccga ctggttaggt atccgtaatg cettaattaa caatggcgtt 4260
 cccgaatatg ttggtcatac cgctggtttg gttatcggtg caggtggcac ttctagagcc 4320
 ccactttacg ccttgcacag tttaggttgc aaaaagatct tcataatcaa caggacaact 4380
 tcgaaattga agccattaat agagtcactt ccatctgaat tcaacattat tggaaatagag 4440
 tccactaaat ctatagaaga gattaaggaa cacgttggcg ttgctgtcag ctgtgtacca 4500
 gccgacaaac cattagatga cgaactttta agtaagctgg agagattcct tgtgaaaggt 4560
 gcccatgctg cttttgtacc aaccttattg gaagccgcat acaaaccaag cgttactccc 4620
 gttatgacaa tttcacaaga caaatatcaa tggcacggtt tccttgatc acaaatgtta 4680
 gtacaccaag gtgtagctca gtttgaaaag tggacaggat tcaagggccc tttcaaggcc 4740
 atttttgatg ccgttacgaa agagtag 4767

<210> 24
 < 211> 1588
 < 212> PRT

5 < 213> Secuencia artificial

<220>
 < 223> Polipéptido sintético

ES 2 719 304 T3

<400> 24

Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
 1 5 10 15

Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
 20 25 30

Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
 35 40 45

Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
 50 55 60

Glu Gly Ser Arg Leu Leu Thr Tyr Val Val Lys Pro Gly Glu Thr Ser
 65 70 75 80

Lys Ser Arg Glu Thr Lys Ala Gln Leu Glu Asp Tyr Leu Leu Val Glu
 85 90 95

Gly Cys Thr Arg Asp Thr Val Met Val Ala Ile Gly Gly Gly Val Ile
 100 105 110

Gly Asp Met Ile Gly Phe Val Ala Ser Thr Phe Met Arg Gly Val Arg
 115 120 125

Val Val Gln Val Pro Thr Ser Leu Leu Ala Met Val Asp Ser Ser Ile
 130 135 140

Gly Gly Lys Thr Ala Ile Asp Thr Pro Leu Gly Lys Asn Phe Ile Gly
 145 150 155 160

Ala Phe Trp Gln Pro Lys Phe Val Leu Val Asp Ile Lys Trp Leu Glu
 165 170 175

Thr Leu Ala Lys Arg Glu Phe Ile Asn Gly Met Ala Glu Val Ile Lys
 180 185 190

Thr Ala Cys Ile Trp Asn Ala Asp Glu Phe Thr Arg Leu Glu Ser Asn
 195 200 205

Ala Ser Leu Phe Leu Asn Val Val Asn Gly Ala Lys Asn Val Lys Val
 210 215 220

Thr Asn Gln Leu Thr Asn Glu Ile Asp Glu Ile Ser Asn Thr Asp Ile

ES 2 719 304 T3

Gly His Gly Gly Ser Thr Leu Ser Ala Cys Ala Asp Pro Leu Tyr Leu
 485 490 495

Gly Asn Ala Gly Thr Ala Ser Arg Phe Leu Thr Ser Leu Ala Ala Leu
 500 505 510

Val Asn Ser Thr Ser Ser Gln Lys Tyr Ile Val Leu Thr Gly Asn Ala
 515 520 525

Arg Met Gln Gln Arg Pro Ile Ala Pro Leu Val Asp Ser Leu Arg Ala
 530 535 540

Asn Gly Thr Lys Ile Glu Tyr Leu Asn Asn Glu Gly Ser Leu Pro Ile
 545 550 555 560

Lys Val Tyr Thr Asp Ser Val Phe Lys Gly Gly Arg Ile Glu Leu Ala
 565 570 575

Ala Thr Val Ser Ser Gln Tyr Val Ser Ser Ile Leu Met Cys Ala Pro
 580 585 590

Tyr Ala Glu Glu Pro Val Thr Leu Ala Leu Val Gly Gly Lys Pro Ile
 595 600 605

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
 610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
 625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
 645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
 660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
 675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
 690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
 705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
 725 730 735

ES 2 719 304 T3

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
 740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
 755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
 770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
 785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
 805 810 815

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
 820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
 835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
 850 855 860

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
 865 870 875 880

Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
 885 890 895

Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu
 900 905 910

Gly Tyr Lys Leu Val Asp Leu Asp Glu Leu Phe Glu Gln Gln His Asn
 915 920 925

Asn Gln Ser Val Lys Gln Phe Val Val Glu Asn Gly Trp Glu Lys Phe
 930 935 940

Arg Glu Glu Glu Thr Arg Ile Phe Lys Glu Val Ile Gln Asn Tyr Gly
 945 950 955 960

Asp Asp Gly Tyr Val Phe Ser Thr Gly Gly Gly Ile Val Glu Ser Ala
 965 970 975

Glu Ser Arg Lys Ala Leu Lys Asp Phe Ala Ser Ser Gly Gly Tyr Val
 980 985 990

ES 2 719 304 T3

Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
 995 1000 1005

 Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
 1010 1015 1020

 Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
 1025 1030 1035

 Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
 1040 1045 1050

 Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
 1055 1060 1065

 Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe
 1070 1075 1080

 Asp Asp Leu Thr Glu Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr
 1085 1090 1095

 Gly Cys Glu Ala Val Glu Val Arg Val Asp His Leu Ala Asn Tyr
 1100 1105 1110

 Ser Ala Asp Phe Val Ser Lys Gln Leu Ser Ile Leu Arg Lys Ala
 1115 1120 1125

 Thr Asp Ser Ile Pro Ile Ile Phe Thr Val Arg Thr Met Lys Gln
 1130 1135 1140

 Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
 1145 1150 1155

 Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
 1160 1165 1170

 Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
 1175 1180 1185

 Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
 1190 1195 1200

 Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
 1205 1210 1215

 Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val

ES 2 719 304 T3

1220		1225		1230		
Asn Phe	Glu Asp	Asn Leu	Arg Leu	Glu His Phe	Arg Asp	Thr His
1235			1240		1245	
Lys Asn	Lys Pro	Leu Ile	Ala Val	Asn Met Thr	Ser Lys	Gly Ser
1250			1255		1260	
Ile Ser	Arg Val	Leu Asn	Asn Val	Leu Thr Pro	Val Thr	Ser Asp
1265			1270		1275	
Leu Leu	Pro Asn	Ser Ala	Ala Pro	Gly Gln Leu	Thr Val	Ala Gln
1280			1285		1290	
Ile Asn	Lys Met	Tyr Thr	Ser Met	Gly Gly Ile	Glu Pro	Lys Glu
1295			1300		1305	
Leu Phe	Val Val	Gly Lys	Pro Ile	Gly His Ser	Arg Ser	Pro Ile
1310			1315		1320	
Leu His	Asn Thr	Gly Tyr	Glu Ile	Leu Gly Leu	Pro His	Lys Phe
1325			1330		1335	
Asp Lys	Phe Glu	Thr Glu	Ser Ala	Gln Leu Val	Lys Glu	Lys Leu
1340			1345		1350	
Leu Asp	Gly Asn	Lys Asn	Phe Gly	Gly Ala Ala	Val Thr	Ile Pro
1355			1360		1365	
Leu Lys	Leu Asp	Ile Met	Gln Tyr	Met Asp Glu	Leu Thr	Asp Ala
1370			1375		1380	
Ala Lys	Val Ile	Gly Ala	Val Asn	Thr Val Ile	Pro Leu	Gly Asn
1385			1390		1395	
Lys Lys	Phe Lys	Gly Asp	Asn Thr	Asp Trp Leu	Gly Ile	Arg Asn
1400			1405		1410	
Ala Leu	Ile Asn	Asn Gly	Val Pro	Glu Tyr Val	Gly His	Thr Ala
1415			1420		1425	
Gly Leu	Val Ile	Gly Ala	Gly Gly	Thr Ser Arg	Ala Pro	Leu Tyr
1430			1435		1440	
Ala Leu	His Ser	Leu Gly	Cys Lys	Lys Ile Phe	Ile Ile	Asn Arg
1445			1450		1455	

ES 2 719 304 T3

Thr Thr Ser Lys Leu Lys Pro Leu Ile Glu Ser Leu Pro Ser Glu
 1460 1465 1470

Phe Asn Ile Ile Gly Ile Glu Ser Thr Lys Ser Ile Glu Glu Ile
 1475 1480 1485

Lys Glu His Val Gly Val Ala Val Ser Cys Val Pro Ala Asp Lys
 1490 1495 1500

Pro Leu Asp Asp Glu Leu Leu Ser Lys Leu Glu Arg Phe Leu Val
 1505 1510 1515

Lys Gly Ala His Ala Ala Phe Val Pro Thr Leu Leu Glu Ala Ala
 1520 1525 1530

Tyr Lys Pro Ser Val Thr Pro Val Met Thr Ile Ser Gln Asp Lys
 1535 1540 1545

Tyr Gln Trp His Val Val Pro Gly Ser Gln Met Leu Val His Gln
 1550 1555 1560

Gly Val Ala Gln Phe Glu Lys Trp Thr Gly Phe Lys Gly Pro Phe
 1565 1570 1575

Lys Ala Ile Phe Asp Ala Val Thr Lys Glu
 1580 1585

<210> 25
 < 211> 5064
 < 212> ADN
 < 213> Secuencia artificial

5

<220>
 < 223> Ácido nucleico sintético

<400> 25
 atggtgcagt tagccaaagt cccaattcta ggaaatgata ttatccacgt tgggtataac 60
 attcatgacc atttggttga aaccataatt aaacattgtc cttcttcgac atacgttatt 120
 tgcaatgata cgaacttgag taaagttcca tactaccagc aattagtcct ggaattcaag 180
 gcttctttgc cagaaggctc tcgtttactt acttatgttg ttaaaccagg tgagacaagt 240
 aaaagtagag aaaccaaagc gcagctagaa gattatcttt tagtggaagg atgtactcgt 300
 gatacggtta tggtagcgat cgggtggtgt gttattggtg acatgattgg gttcgttgca 360
 tctacattta tgagaggtgt tcgtgttgtc caagtaccaa catccttatt ggcaatggtc 420
 gattcctcca ttggtggtaa aactgctatt gacactcctc taggtaaaaa ctttattggt 480
 gcattttggc aacccaaaatt tgtccttcta gatattaaat ggctagaaac gttagccaag 540

10

ES 2 719 304 T3

agagagttta tcaatgggat ggcagaagtt atcaagactg cttgtatttg gaacgctgac 600
 gaatttacta gattagaatc aaacgcttcg ttgttcttaa atgttgtaa tggggcaaaa 660
 aatgtcaagg ttaccaatca attgacaaa gagattgacg agatatacga tacagatatt 720
 gaagctatgt tggatcatac atataagtta gttcttgaga gtattaaggt caaagcggaa 780
 gttgtctctt cggatgaacg tgaatccagt ctaagaaacc ttttgaactt cggacattct 840
 attggtcatg cttatgaagc tatactaacc ccacaagcat tacatggtga atgtgtgtcc 900
 attggtatgg ttaaagaggc ggaattatcc cgttatttcg gtattctctc ccctacccaa 960
 gttgcacgtc tatccaagat tttggttgcc tacgggttgc ctgtttcgcc tgatgagaaa 1020
 tggtttaaag agctaaccct acataagaaa acaccattgg atatcttatt gaagaaaatg 1080
 agtattgaca agaaaaacga gggttccaaa aagaaggtgg tcattttaga aagtattggt 1140
 aagtgctatg gtgactccgc tcaatttggt agcgaagaag acctgagatt tattctaaca 1200
 gatgaaacct tcgtttacc cttcaaggac atccctgctg atcaacagaa agttgttacc 1260
 cccctcggtt ctaagtccat ctccaatcgt gctttaattc ttgctgccct cggtgaaggt 1320
 caatgtaaaa tcaagaactt attacattct gatgatacta aacatatggt aaccgctggt 1380
 catgaattga aagtgctac gatatcatgg gaagataatg gtgagacggg agtggggaa 1440
 ggacatggtg gttccacatt gtcagcttgt gctgaccctt tatatctagg taatgcaggt 1500
 actgcatcta gatttttgac ttccttggct gccttggtca atttacttc aagccaaaag 1560
 tataatggtt taactggtaa cgcaagaatg caacaaagac caattgctcc tttggtgat 1620
 tctttgctg ctaatggtag taaaattgag taactgaata atgaaggttc cctgccaatc 1680
 aaagtttata ctgattcggg attcaaagggt ggtagaattg aattagctgc tacagtttct 1740
 tctcagtacg tatcctctat cttgatgtgt gcccatacgt ctgaagaacc tgtaactttg 1800
 gctcttggtg gtggttaagc aatctctaaa ttgtacgtcg atatgacaat aaaaatgatg 1860
 gaaaaattcg gtatcaatgt tgaaacttct actacagaac cttacactta ttatattcca 1920
 aaggacatt atattaacc atcagaatac gtcattgaaa gtgatgcctc aagtgtaca 1980
 taccatttg ccttcgccgc aatgactggg actaccgtaa cggttccaaa cattggtttt 2040
 gagtcgttac aagtgatgc cagatttgca agagatgtct tgaaacctat gggttgtaa 2100
 ataactcaa cggcaacttc aactactggt tcgggtcctc ctgtaggtac tttaaagcca 2160
 ttaaaacatg ttgatatgga gccaatgact gatgcgttct taactgcatg gttgttgcc 2220
 gctatttcgc acgacagtga tccaaattct gcaaatatac ccaccattga aggtattgca 2280
 aaccagcgtg tcaaagagt taacagaatt ttggccatgg ctacagagct cgccaaattt 2340
 ggcgtcaaaa ctacagaatt accagatggg attcaagtc atggtttaaa ctcgataaaa 2400
 gatttgaagg ttccttcga ctctctgga cctgtcggg tatgcacata tgatgatcat 2460

ES 2 719 304 T3

cgtgtggcca tgagtttctc gcttcttgca ggaatggtaa attctcaaaa tgaacgtgac 2520
 gaagttgcta atcctgtaag aatacttgaa agacattgta ctggtaaaac ctggcctggc 2580
 tgggtgggatg tgttacattc cgaactaggt gccaaattag atgggtgcaga acctttagag 2640
 tgcacatcca aaaagaactc aaagaaaagc gttgtcatta ttggcatgag agcagctggc 2700
 aaaactacta taagtaaatg gtgcgcatcc gctctgggtt acaaattagt tgacctagac 2760
 gagctgtttg agcaacagca taacaatcaa agtggttaaac aatttgttgt ggagaacggc 2820
 tgggagaagt tccgtgagga agaaacaaga attttcaagg aagttattca aaattacggc 2880
 gatgatggat atgttttctc aacaggtggc ggtattgttg aaagcgtga gtctagaaaa 2940
 gccttaaaag attttgccctc atcaggtgga tacgttttac acttacatag ggatattgag 3000
 gagacaattg tctttttaca aagtgatcct tcaagacctg cctatgtgga agaaattcgt 3060
 gaagtttgga acagaaggga ggggtggtat aaagaatgct caaatttctc tttctttgct 3120
 cctcattgct ccgcagaagc tgagttccaa gctctaagaa gatcgtttag taagtacatt 3180
 gcaaccatta caggtgtcag agaaatagaa attccaagcg gaagatctgc ctttgtgtgt 3240
 ttaacctttg atgacttaac tgaacaaact gagaatttga ctccaatctg ttatggttgt 3300
 gaggtctgag aggtcagagt agaccatttg gctaattact ctgctgattt cgtgagtaaa 3360
 cagttatcta tattgcgtaa agccactgac agtattccta tcatttttac tgtgcgaacc 3420
 atgaagcaag gtggcaactt tcctgatgaa gagttcaaaa ccttgagaga gctatacgat 3480
 attgccttga agaatggtgt tgaattcctt gacttagaac taactttacc tactgatatc 3540
 caatatgagg ttattaacaa aaggggcaac accaagatca ttggttccca tcatgacttc 3600
 caaggattat actcctggga cgacgctgaa tgggaaaaca gattcaatca agcgttaact 3660
 cttgatgtgg atgttgtaaa atttgtgggt acggctgtta atttcgaaga taatttgaga 3720
 ctggaacact ttagggatac acacaagaat aagcctttaa ttgcagttaa tatgacttct 3780
 aaaggtagca tttctcgtgt tttgaataat gttttaacac ctgtgacatc agatttattg 3840
 cctaactccg ctgcccctgg ccaattgaca gtagcacaaa ttaacaagat gtatacatct 3900
 atgggaggta tcgagcctaa ggaactgttt gttgttgtaa agccaattgg ccccgggaaa 3960
 atgccttcca aactcgccat cacttccatg tcacttggcc ggtgttatgc cggccactcc 4020
 ttcaccaacta agctcgatat ggcccggaaa tatggctatc aaggcctaga gctcttccac 4080
 gaggacttgg ctgatgtagc ctatcgtctc tccggagaga ccccttcccc atgtggcccg 4140
 tccccagcag cccagctctc ggetgcccgt caaatctctc gcatgtgcca agtcagaaac 4200
 attgaaatcg tctgcctcca gcccttcagc cagtacgacg gcctactcga ccgcgaggag 4260
 cacgagcgcc gtctggagca gctcgagttc tggatcgagc tcgcccacga gcttgacaca 4320

ES 2 719 304 T3

gacattatcc aaatccccgc caactttctc cccgccgagg aagtaactga ggacatttcg 4380
ctcatcgtct cggaccttca agaagtggcc gacatgggcc tgcaggccaa cccacccatc 4440
cgctttgtct acgaggctct gtgctggagc actcgtgtcg acaacttgga gcgtagctgg 4500
gaggtggtgc agaggggtgaa caggcccaac tttggcgtgt gcctggacac tttcaacatt 4560
gcggggcggg tatatgctga tccgacggtt gcctctggcc gcacccccaa cgcggaggaa 4620
gcgatacggg agtcgattgc gcgtctcgtt gaaagggtcg atgtcagcaa ggtcttttat 4680
gtgcaggttg tggacgctga gaagttgaag aagccgctgg tgccgggtca tcggttttat 4740
gacccggagc agccggcgag gatgagctgg tcaaggaact gcaggttatt ctacggggag 4800
aaggacagag gggcgtatct gcccgtaag gagattgcct gggccttctt caacgggctc 4860
ggattcaggg gttgggtcag tctggagctc ttcaacagaa gaatgtcggg cacaggcttt 4920
ggggtgcccg aggagctggc caggagaggg gccgtgtcgt gggcaaagct ggtgagggac 4980
atgaagatca ctgttgattc accaacacaa caacaagcca cacagcagcc catcaggatg 5040
ctgtcgtcgt cagcggcttt gtaa 5064

<210> 26
<211> 1687
<212> PRT
5 <213> Secuencia artificial

<220>
<223> Polipéptido sintético

<400> 26
Met Val Gln Leu Ala Lys Val Pro Ile Leu Gly Asn Asp Ile Ile His
1 5 10 15
Val Gly Tyr Asn Ile His Asp His Leu Val Glu Thr Ile Ile Lys His
20 25 30
Cys Pro Ser Ser Thr Tyr Val Ile Cys Asn Asp Thr Asn Leu Ser Lys
35 40 45
Val Pro Tyr Tyr Gln Gln Leu Val Leu Glu Phe Lys Ala Ser Leu Pro
50 55 60
Glu Gly Ser Arg Leu Leu Thr Tyr Val Val Lys Pro Gly Glu Thr Ser
65 70 75 80
Lys Ser Arg Glu Thr Lys Ala Gln Leu Glu Asp Tyr Leu Leu Val Glu
85 90 95
Gly Cys Thr Arg Asp Thr Val Met Val Ala Ile Gly Gly Gly Val Ile
100 105 110

ES 2 719 304 T3

Gly Asp Met Ile Gly Phe Val Ala Ser Thr Phe Met Arg Gly Val Arg
 115 120 125

Val Val Gln Val Pro Thr Ser Leu Leu Ala Met Val Asp Ser Ser Ile
 130 135 140

Gly Gly Lys Thr Ala Ile Asp Thr Pro Leu Gly Lys Asn Phe Ile Gly
 145 150 155 160

Ala Phe Trp Gln Pro Lys Phe Val Leu Val Asp Ile Lys Trp Leu Glu
 165 170 175

Thr Leu Ala Lys Arg Glu Phe Ile Asn Gly Met Ala Glu Val Ile Lys
 180 185 190

Thr Ala Cys Ile Trp Asn Ala Asp Glu Phe Thr Arg Leu Glu Ser Asn
 195 200 205

Ala Ser Leu Phe Leu Asn Val Val Asn Gly Ala Lys Asn Val Lys Val
 210 215 220

Thr Asn Gln Leu Thr Asn Glu Ile Asp Glu Ile Ser Asn Thr Asp Ile
 225 230 235 240

Glu Ala Met Leu Asp His Thr Tyr Lys Leu Val Leu Glu Ser Ile Lys
 245 250 255

Val Lys Ala Glu Val Val Ser Ser Asp Glu Arg Glu Ser Ser Leu Arg
 260 265 270

Asn Leu Leu Asn Phe Gly His Ser Ile Gly His Ala Tyr Glu Ala Ile
 275 280 285

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

Lys Glu Ala Glu Leu Ser Arg Tyr Phe Gly Ile Leu Ser Pro Thr Gln
 305 310 315 320

Val Ala Arg Leu Ser Lys Ile Leu Val Ala Tyr Gly Leu Pro Val Ser
 325 330 335

Pro Asp Glu Lys Trp Phe Lys Glu Leu Thr Leu His Lys Lys Thr Pro
 340 345 350

Leu Asp Ile Leu Leu Lys Lys Met Ser Ile Asp Lys Lys Asn Glu Gly

ES 2 719 304 T3

355	360	365																			
Ser	Lys	Lys	Lys	Val	Val	Ile	Leu	Glu	Ser	Ile	Gly	Lys	Cys	Tyr	Gly						
370						375					380										
Asp	Ser	Ala	Gln	Phe	Val	Ser	Asp	Glu	Asp	Leu	Arg	Phe	Ile	Leu	Thr						
385					390					395					400						
Asp	Glu	Thr	Leu	Val	Tyr	Pro	Phe	Lys	Asp	Ile	Pro	Ala	Asp	Gln	Gln						
				405					410					415							
Lys	Val	Val	Ile	Pro	Pro	Gly	Ser	Lys	Ser	Ile	Ser	Asn	Arg	Ala	Leu						
			420					425					430								
Ile	Leu	Ala	Ala	Leu	Gly	Glu	Gly	Gln	Cys	Lys	Ile	Lys	Asn	Leu	Leu						
		435					440					445									
His	Ser	Asp	Asp	Thr	Lys	His	Met	Leu	Thr	Ala	Val	His	Glu	Leu	Lys						
	450					455					460										
Gly	Ala	Thr	Ile	Ser	Trp	Glu	Asp	Asn	Gly	Glu	Thr	Val	Val	Val	Glu						
465					470					475					480						
Gly	His	Gly	Gly	Ser	Thr	Leu	Ser	Ala	Cys	Ala	Asp	Pro	Leu	Tyr	Leu						
				485					490					495							
Gly	Asn	Ala	Gly	Thr	Ala	Ser	Arg	Phe	Leu	Thr	Ser	Leu	Ala	Ala	Leu						
			500					505					510								
Val	Asn	Ser	Thr	Ser	Ser	Gln	Lys	Tyr	Ile	Val	Leu	Thr	Gly	Asn	Ala						
		515					520						525								
Arg	Met	Gln	Gln	Arg	Pro	Ile	Ala	Pro	Leu	Val	Asp	Ser	Leu	Arg	Ala						
	530					535					540										
Asn	Gly	Thr	Lys	Ile	Glu	Tyr	Leu	Asn	Asn	Glu	Gly	Ser	Leu	Pro	Ile						
545					550					555					560						
Lys	Val	Tyr	Thr	Asp	Ser	Val	Phe	Lys	Gly	Gly	Arg	Ile	Glu	Leu	Ala						
				565					570					575							
Ala	Thr	Val	Ser	Ser	Gln	Tyr	Val	Ser	Ser	Ile	Leu	Met	Cys	Ala	Pro						
			580					585					590								
Tyr	Ala	Glu	Glu	Pro	Val	Thr	Leu	Ala	Leu	Val	Gly	Gly	Lys	Pro	Ile						
		595					600					605									

ES 2 719 304 T3

Ser Lys Leu Tyr Val Asp Met Thr Ile Lys Met Met Glu Lys Phe Gly
610 615 620

Ile Asn Val Glu Thr Ser Thr Thr Glu Pro Tyr Thr Tyr Tyr Ile Pro
625 630 635 640

Lys Gly His Tyr Ile Asn Pro Ser Glu Tyr Val Ile Glu Ser Asp Ala
645 650 655

Ser Ser Ala Thr Tyr Pro Leu Ala Phe Ala Ala Met Thr Gly Thr Thr
660 665 670

Val Thr Val Pro Asn Ile Gly Phe Glu Ser Leu Gln Gly Asp Ala Arg
675 680 685

Phe Ala Arg Asp Val Leu Lys Pro Met Gly Cys Lys Ile Thr Gln Thr
690 695 700

Ala Thr Ser Thr Thr Val Ser Gly Pro Pro Val Gly Thr Leu Lys Pro
705 710 715 720

Leu Lys His Val Asp Met Glu Pro Met Thr Asp Ala Phe Leu Thr Ala
725 730 735

Cys Val Val Ala Ala Ile Ser His Asp Ser Asp Pro Asn Ser Ala Asn
740 745 750

Thr Thr Thr Ile Glu Gly Ile Ala Asn Gln Arg Val Lys Glu Cys Asn
755 760 765

Arg Ile Leu Ala Met Ala Thr Glu Leu Ala Lys Phe Gly Val Lys Thr
770 775 780

Thr Glu Leu Pro Asp Gly Ile Gln Val His Gly Leu Asn Ser Ile Lys
785 790 795 800

Asp Leu Lys Val Pro Ser Asp Ser Ser Gly Pro Val Gly Val Cys Thr
805 810 815

Tyr Asp Asp His Arg Val Ala Met Ser Phe Ser Leu Leu Ala Gly Met
820 825 830

Val Asn Ser Gln Asn Glu Arg Asp Glu Val Ala Asn Pro Val Arg Ile
835 840 845

Leu Glu Arg His Cys Thr Gly Lys Thr Trp Pro Gly Trp Trp Asp Val
850 855 860

ES 2 719 304 T3

Leu His Ser Glu Leu Gly Ala Lys Leu Asp Gly Ala Glu Pro Leu Glu
 865 870 875 880
 Cys Thr Ser Lys Lys Asn Ser Lys Lys Ser Val Val Ile Ile Gly Met
 885 890 895
 Arg Ala Ala Gly Lys Thr Thr Ile Ser Lys Trp Cys Ala Ser Ala Leu
 900 905 910
 Gly Tyr Lys Leu Val Asp Leu Asp Glu Leu Phe Glu Gln Gln His Asn
 915 920 925
 Asn Gln Ser Val Lys Gln Phe Val Val Glu Asn Gly Trp Glu Lys Phe
 930 935 940
 Arg Glu Glu Glu Thr Arg Ile Phe Lys Glu Val Ile Gln Asn Tyr Gly
 945 950 955 960
 Asp Asp Gly Tyr Val Phe Ser Thr Gly Gly Gly Ile Val Glu Ser Ala
 965 970 975
 Glu Ser Arg Lys Ala Leu Lys Asp Phe Ala Ser Ser Gly Gly Tyr Val
 980 985 990
 Leu His Leu His Arg Asp Ile Glu Glu Thr Ile Val Phe Leu Gln Ser
 995 1000 1005
 Asp Pro Ser Arg Pro Ala Tyr Val Glu Glu Ile Arg Glu Val Trp
 1010 1015 1020
 Asn Arg Arg Glu Gly Trp Tyr Lys Glu Cys Ser Asn Phe Ser Phe
 1025 1030 1035
 Phe Ala Pro His Cys Ser Ala Glu Ala Glu Phe Gln Ala Leu Arg
 1040 1045 1050
 Arg Ser Phe Ser Lys Tyr Ile Ala Thr Ile Thr Gly Val Arg Glu
 1055 1060 1065
 Ile Glu Ile Pro Ser Gly Arg Ser Ala Phe Val Cys Leu Thr Phe
 1070 1075 1080
 Asp Asp Leu Thr Glu Gln Thr Glu Asn Leu Thr Pro Ile Cys Tyr
 1085 1090 1095
 Gly Cys Glu Ala Val Glu Val Arg Val Asp His Leu Ala Asn Tyr
 1100 1105 1110

ES 2 719 304 T3

Ser Ala Asp Phe Val Ser Lys Gln Leu Ser Ile Leu Arg Lys Ala
1115 1120 1125

Thr Asp Ser Ile Pro Ile Ile Phe Thr Val Arg Thr Met Lys Gln
1130 1135 1140

Gly Gly Asn Phe Pro Asp Glu Glu Phe Lys Thr Leu Arg Glu Leu
1145 1150 1155

Tyr Asp Ile Ala Leu Lys Asn Gly Val Glu Phe Leu Asp Leu Glu
1160 1165 1170

Leu Thr Leu Pro Thr Asp Ile Gln Tyr Glu Val Ile Asn Lys Arg
1175 1180 1185

Gly Asn Thr Lys Ile Ile Gly Ser His His Asp Phe Gln Gly Leu
1190 1195 1200

Tyr Ser Trp Asp Asp Ala Glu Trp Glu Asn Arg Phe Asn Gln Ala
1205 1210 1215

Leu Thr Leu Asp Val Asp Val Val Lys Phe Val Gly Thr Ala Val
1220 1225 1230

Asn Phe Glu Asp Asn Leu Arg Leu Glu His Phe Arg Asp Thr His
1235 1240 1245

Lys Asn Lys Pro Leu Ile Ala Val Asn Met Thr Ser Lys Gly Ser
1250 1255 1260

Ile Ser Arg Val Leu Asn Asn Val Leu Thr Pro Val Thr Ser Asp
1265 1270 1275

Leu Leu Pro Asn Ser Ala Ala Pro Gly Gln Leu Thr Val Ala Gln
1280 1285 1290

Ile Asn Lys Met Tyr Thr Ser Met Gly Gly Ile Glu Pro Lys Glu
1295 1300 1305

Leu Phe Val Val Gly Lys Pro Ile Gly Pro Gly Lys Met Pro Ser
1310 1315 1320

Lys Leu Ala Ile Thr Ser Met Ser Leu Gly Arg Cys Tyr Ala Gly
1325 1330 1335

His Ser Phe Thr Thr Lys Leu Asp Met Ala Arg Lys Tyr Gly Tyr

ES 2 719 304 T3

1340		1345		1350
Gln Gly 1355	Leu Glu Leu Phe	His 1360	Glu Asp Leu Ala	Asp Val Ala Tyr 1365
Arg Leu 1370	Ser Gly Glu Thr	Pro 1375	Ser Pro Cys Gly	Pro Ser Pro Ala 1380
Ala Gln 1385	Leu Ser Ala Ala	Arg 1390	Gln Ile Leu Arg	Met Cys Gln Val 1395
Arg Asn 1400	Ile Glu Ile Val	Cys 1405	Leu Gln Pro Phe	Ser Gln Tyr Asp 1410
Gly Leu 1415	Leu Asp Arg Glu	Glu 1420	His Glu Arg Arg	Leu Glu Gln Leu 1425
Glu Phe 1430	Trp Ile Glu Leu	Ala 1435	His Glu Leu Asp	Thr Asp Ile Ile 1440
Gln Ile 1445	Pro Ala Asn Phe	Leu 1450	Pro Ala Glu Glu	Val Thr Glu Asp 1455
Ile Ser 1460	Leu Ile Val Ser	Asp 1465	Leu Gln Glu Val	Ala Asp Met Gly 1470
Leu Gln 1475	Ala Asn Pro Pro	Ile 1480	Arg Phe Val Tyr	Glu Ala Leu Cys 1485
Trp Ser 1490	Thr Arg Val Asp	Thr 1495	Trp Glu Arg Ser	Trp Glu Val Val 1500
Gln Arg 1505	Val Asn Arg Pro	Asn 1510	Phe Gly Val Cys	Leu Asp Thr Phe 1515
Asn Ile 1520	Ala Gly Arg Val	Tyr 1525	Ala Asp Pro Thr	Val Ala Ser Gly 1530
Arg Thr 1535	Pro Asn Ala Glu	Glu 1540	Ala Ile Arg Lys	Ser Ile Ala Arg 1545
Leu Val 1550	Glu Arg Val Asp	Val 1555	Ser Lys Val Phe	Tyr Val Gln Val 1560
Val Asp 1565	Ala Glu Lys Leu	Lys 1570	Lys Pro Leu Val	Pro Gly His Arg 1575

ES 2 719 304 T3

Phe Tyr Asp Pro Glu Gln Pro Ala Arg Met Ser Trp Ser Arg Asn
 1580 1585 1590

Cys Arg Leu Phe Tyr Gly Glu Lys Asp Arg Gly Ala Tyr Leu Pro
 1595 1600 1605

Val Lys Glu Ile Ala Trp Ala Phe Phe Asn Gly Leu Gly Phe Glu
 1610 1615 1620

Gly Trp Val Ser Leu Glu Leu Phe Asn Arg Arg Met Ser Asp Thr
 1625 1630 1635

Gly Phe Gly Val Pro Glu Glu Leu Ala Arg Arg Gly Ala Val Ser
 1640 1645 1650

Trp Ala Lys Leu Val Arg Asp Met Lys Ile Thr Val Asp Ser Pro
 1655 1660 1665

Thr Gln Gln Gln Ala Thr Gln Gln Pro Ile Arg Met Leu Ser Leu
 1670 1675 1680

Ser Ala Ala Leu
 1685

<210> 27

<211> 221

<212> PRT

5 <213> Homo sapiens

<400> 27

Met Gly Asp Thr Lys Glu Gln Arg Ile Leu Asn His Val Leu Gln His
 1 5 10 15

Ala Glu Pro Gly Asn Ala Gln Ser Val Leu Glu Ala Ile Asp Thr Tyr
 20 25 30

Cys Glu Gln Lys Glu Trp Ala Met Asn Val Gly Asp Lys Lys Gly Lys
 35 40 45

Ile Val Asp Ala Val Ile Gln Glu His Gln Pro Ser Val Leu Leu Glu
 50 55 60

Leu Gly Ala Tyr Cys Gly Tyr Ser Ala Val Arg Met Ala Arg Leu Leu
 65 70 75 80

Ser Pro Gly Ala Arg Leu Ile Thr Ile Glu Ile Asn Pro Asp Cys Ala
 85 90 95

ES 2 719 304 T3

Ala Ile Thr Gln Arg Met Val Asp Phe Ala Gly Val Lys Asp Lys Val
 100 105 110

Thr Leu Val Val Gly Ala Ser Gln Asp Ile Ile Pro Gln Leu Lys Lys
 115 120 125

Lys Tyr Asp Val Asp Thr Leu Asp Met Val Phe Leu Asp His Trp Lys
 130 135 140

Asp Arg Tyr Leu Pro Asp Thr Leu Leu Leu Glu Glu Cys Gly Leu Leu
 145 150 155 160

Arg Lys Gly Thr Val Leu Leu Ala Asp Asn Val Ile Cys Pro Gly Ala
 165 170 175

Pro Asp Phe Leu Ala His Val Arg Gly Ser Ser Cys Phe Glu Cys Thr
 180 185 190

His Tyr Gln Ser Phe Leu Glu Tyr Arg Glu Val Val Asp Gly Leu Glu
 195 200 205

Lys Ala Ile Tyr Lys Gly Pro Gly Ser Glu Ala Gly Pro
 210 215 220

<210> 28

< 211> 40

5 < 212> ADN

< 213> Secuencia artificial

<220>

< 223> Oligonucleótido sintético

<400> 28

10 cgtagcatgc agtctagaaa aatgggtgac actaaggagc 40

<210> 29

< 211> 45

< 212> ADN

< 213> Secuencia artificial

15 <220>

< 223> Oligonucleótido sintético

<400> 29

gacgacgta gtcacagaat tcttatggac cagcttcaga acctg 45

<210> 30

20 < 211> 25

< 212> ADN

< 213> Secuencia artificial

<220>

< 223> Oligonucleótido sintético

25 <400> 30

cgtagcatgc agtctagaaa aatgg 25

<210> 31

< 211> 22

< 212> ADN

30 < 213> Secuencia artificial

<220>
 < 223> Oligonucleótido sintético

<400> 31
 gacgacgta gtgacagaat tc 22

5 <210> 32
 < 211> 45
 < 212> ADN
 < 213> Secuencia artificial

10 <220>
 < 223> Oligonucleótido sintético

 <220>
 < 221> misc_feature
 < 222> (30)..(30)
 < 223> n es a, c, g, o t

15 <400> 32
 ctattgacac ttattgtgag caaaaggagn rkgctatgaa cgttg 45

 <210> 33
 < 211> 45
 < 212> ADN

20 < 213> Secuencia artificial

 <220>
 < 223> Oligonucleótido sintético

 <220>
 < 221> misc_feature
 < 222> (30)..(30)
 < 223> n es a, c, g, o t

25 <400> 33
 ctattgacac ttattgtgag caaaaggagn ykgctatgaa cgttg 45

 <210> 34
 < 211> 29
 < 212> ADN

30 < 213> Secuencia artificial

 <220>
 < 223> Oligonucleótido sintético

35 <400> 34
 ctccttttgc tcacaataag tgtcaatag 29

 <210> 35
 < 211> 45
 < 212> ADN

40 < 213> Secuencia artificial

 <220>
 < 223> Oligonucleótido sintético

 <220>
 < 221> misc_feature
 < 222> (31)..(31)
 < 223> n es a, c, g, o t

45 <400> 35
 gacactatt gtgagcaaaa ggagtgggct nrkaacgttg gtgac 45

 <210> 36
 < 211> 43

50

< 212> ADN
 < 213> Secuencia artificial

 <220>
 < 223> Oligonucleótido sintético

 5 <220>
 < 221> misc_feature
 < 222> (29)..(29)
 < 223> n es a, c, g, o t

 <400> 36
 10 cacttattgt gagcaaaagg agtgggctny kaacgttggt gac 43

 <210> 37
 < 211> 24
 < 212> ADN
 < 213> Secuencia artificial

 15 <220>
 < 223> Oligonucleótido sintético

 <400> 37
 ctccttttgc tcacaataag tgtc 24

 <210> 38
 20 < 211> 46
 < 212> ADN
 < 213> Secuencia artificial

 <220>
 < 223> Oligonucleótido sintético

 25 <220>
 < 221> misc_feature
 < 222> (27)..(27)
 < 223> n es a, c, g, o t

 <400> 38
 30 ctttgacat ggttttcttg gaccatnrka aggacagata ttgccc 46

 <210> 39
 < 211> 46
 < 212> ADN
 < 213> Secuencia artificial

 35 <220>
 < 223> Oligonucleótido sintético

 <220>
 < 221> misc_feature
 < 222> (27)..(27)
 40 < 223> n es a, c, g, o t

 <400> 39
 ctttgacat ggttttcttg gaccatnyka aggacagata ttgccc 46

 <210> 40
 < 211> 26
 45 < 212> ADN
 < 213> Secuencia artificial

 <220>
 < 223> Oligonucleótido sintético

 <400> 40
 50 atggccaag aaaacatgt ccaaag 26

<210> 41
 < 211> 53
 < 212> ADN
 < 213> Secuencia artificial

5 <220>
 < 223> Oligonucleótido sintético

<220>
 < 221> misc_feature
 < 222> (30)..(30)

10 < 223> n es a, c, g, o t

<400> 41
 gtactgtttt gtagctgac aacgttattn rkccaggtgc tccagacttc ttg 53

<210> 42
 < 211> 53
 < 212> ADN
 < 213> Secuencia artificial

15 <220>
 < 223> Oligonucleótido sintético

<220>
 < 221> misc_feature
 < 222> (30)..(30)

20 < 223> n es a, c, g, o t

<400> 42
 gtactgtttt gtagctgac aacgttattn ykccaggtgc tccagacttc ttg 53

25 <210> 43
 < 211> 29
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> Oligonucleótido sintético

30 <400> 43
 aataacgttg tcagctaaca aaacagtac 29

<210> 44
 < 211> 47
 < 212> ADN
 < 213> Secuencia artificial

35 <220>
 < 223> Oligonucleótido sintético

<220>
 < 221> misc_feature
 < 222> (30)..(30)

40 < 223> n es a, c, g, o t

<400> 44
 ctgttttgtt agctgacaac gttattgtn rkggtgctcc agacttc 47

45 <210> 45
 < 211> 47
 < 212> ADN
 < 213> Secuencia artificial

<220>
 < 223> Oligonucleótido sintético

50

<220>
 < 221> misc_feature
 < 222> (30)..(30)
 < 223> n es a, c, g, o t

5 <400> 45
 ctgtttgtt agctgacaac gttattgtn ykgtgctcc agacttc 47
 <210> 46
 < 211> 29
 < 212> ADN
 10 < 213> Secuencia artificial
 <220>
 < 223> Oligonucleótido sintético
 <400> 46
 acaaataacg ttgtcagcta acaaaacag 29

15 <210> 47
 < 211> 103
 < 212> ADN
 < 213> Secuencia artificial
 <220>
 20 < 223> Oligonucleótido sintético
 <220>
 < 221> misc_feature
 < 222> (86)..(86)
 < 223> n es a, c, g, o t

25 <400> 47
 atttagaatt cttatggacc agcttcagaa cctggaccct tatatatagc cttctccaaa 60
 ccgtcaaca cctctctata ttcmyngaaa gattgataat gag 103
 <210> 48
 < 211> 102
 < 212> ADN
 30 < 213> Secuencia artificial
 <220>
 < 223> Oligonucleótido sintético
 <220>
 < 221> misc_feature
 35 < 222> (85)..(85)
 < 223> n es a, c, g, o t
 <400> 48
 atttagaatt ctatggacca gcttcagaac ctggaccctt atatatagcc ttctccaaac 60
 cgtcaacaac ctctctatat tcmrngaaag attgataatg ag 102

40 <210> 49
 < 211> 101
 < 212> ADN
 < 213> Secuencia artificial
 <220>
 < 223> Oligonucleótido sintético

45 <220>
 < 221> misc_feature
 < 222> (83)..(83)
 < 223> n es a, c, g, o t

ES 2 719 304 T3

<400> 49
 atttagaatt cttatggacc agcttcagaa cctggaccct tatatatagc cttctccaaa 60
 ccgtcaacaa cctctctata myncaagaaa gattgataat g 101
 <210> 50
 < 211> 101
 5 < 212> ADN
 < 213> Secuencia artificial
 <220>
 < 223> Oligonucleótido sintético
 <220>
 10 < 221> misc_feature
 < 222> (83)..(83)
 < 223> n es a, c, g, o t
 <400> 50
 atttagaatt cttatggacc agcttcagaa cctggaccct tatatatagc cttctccaaa 60
 ccgtcaacaa cctctctata mrncaagaaa gattgataat g 101
 15 <210> 51
 < 211> 95
 < 212> ADN
 < 213> Secuencia artificial
 <220>
 20 < 223> Oligonucleótido sintético
 <220>
 < 221> misc_feature
 < 222> (77)..(77)
 < 223> n es a, c, g, o t
 25 <400> 51
 atttagaatt cttatggacc agcttcagaa cctggaccct tatatatagc cttctccaaa 60
 ccgtcaacaa cctcmynata ttccaagaaa gattg 95
 <210> 52
 < 211> 95
 < 212> ADN
 30 < 213> Secuencia artificial
 <220>
 < 223> Oligonucleótido sintético
 <220>
 < 221> misc_feature
 35 < 222> (77)..(77)
 < 223> n es a, c, g, o t
 <400> 52
 atttagaatt cttatggacc agcttcagaa cctggaccct tatatatagc cttctccaaa 60
 ccgtcaacaa cctcmrnata ttccaagaaa gattg 95
 <210> 53
 40 < 211> 363
 < 212> PRT
 < 213> Arabidopsis thaliana

ES 2 719 304 T3

<400> 53

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

Asp Glu Ala Ala Leu Phe Ala Met Gln Leu Ala Ser Ala Ser Val Leu
 20 25 30

Pro Met Ala Leu Lys Ser Ala Leu Glu Leu Asp Leu Leu Glu Ile Met
 35 40 45

Ala Lys Asn Gly Ser Pro Met Ser Pro Thr Glu Ile Ala Ser Lys Leu
 50 55 60

Pro Thr Lys Asn Pro Glu Ala Pro Val Met Leu Asp Arg Ile Leu Arg
 65 70 75 80

Leu Leu Thr Ser Tyr Ser Val Leu Thr Cys Ser Asn Arg Lys Leu Ser
 85 90 95

Gly Asp Gly Val Glu Arg Ile Tyr Gly Leu Gly Pro Val Cys Lys Tyr
 100 105 110

Leu Thr Lys Asn Glu Asp Gly Val Ser Ile Ala Ala Leu Cys Leu Met
 115 120 125

Asn Gln Asp Lys Val Leu Met Glu Ser Trp Tyr His Leu Lys Asp Ala
 130 135 140

Ile Leu Asp Gly Gly Ile Pro Phe Asn Lys Ala Tyr Gly Met Ser Ala
 145 150 155 160

Phe Glu Tyr His Gly Thr Asp Pro Arg Phe Asn Lys Val Phe Asn Asn
 165 170 175

Gly Met Ser Asn His Ser Thr Ile Thr Met Lys Lys Ile Leu Glu Thr
 180 185 190

Tyr Lys Gly Phe Glu Gly Leu Thr Ser Leu Val Asp Val Gly Gly Gly
 195 200 205

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

Gly Ile Asn Phe Asp Leu Pro His Val Ile Glu Asp Ala Pro Ser His

ES 2 719 304 T3

85 90 95

Leu Pro Asp Gly Lys Val Glu Arg Leu Tyr Cys Leu Gly Pro Val Cys
 100 105 110

Lys Phe Leu Thr Lys Asn Glu Asp Gly Val Ser Ile Ala Ala Leu Cys
 115 120 125

Leu Met Asn Gln Asp Lys Val Leu Val Glu Ser Trp Tyr His Leu Lys
 130 135 140

Asp Ala Val Leu Asp Gly Gly Ile Pro Phe Asn Lys Ala Tyr Gly Met
 145 150 155 160

Thr Ala Phe Asp Tyr His Gly Thr Asp Pro Arg Phe Asn Lys Val Phe
 165 170 175

Asn Lys Gly Met Ala Asp His Ser Thr Ile Thr Met Lys Lys Ile Leu
 180 185 190

Glu Thr Tyr Lys Gly Phe Glu Gly Leu Lys Ser Ile Val Asp Val Gly
 195 200 205

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

Ile Lys Gly Ile Asn Phe Asp Leu Pro His Val Ile Glu Asp Ala Pro
 225 230 235 240

Gln Tyr Pro Gly Val Gln His Val Gly Gly Asp Met Phe Val Ser Val
 245 250 255

Pro Lys Gly Asn Ala Ile Phe Met Lys Trp Ile Cys His Asp Trp Ser
 260 265 270

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

Asp Asp Gly Lys Val Ile Leu Ala Glu Cys Ile Leu Pro Val Ala Pro
 290 295 300

Asp Thr Ser Leu Ala Thr Lys Gly Val Val His Met Asp Val Ile Met
 305 310 315 320

Leu Ala His Asn Pro Gly Gly Lys Glu Arg Thr Glu Gln Glu Phe Glu
 325 330 335

Ala Leu Ala Lys Gly Ser Gly Phe Gln Gly Ile Arg Val Cys Cys Asp
 340 345 350

Ala Phe Asn Thr Tyr Val Ile Glu Phe Leu Lys Lys Ile
 355 360 365

<210> 55
 <211> 271

5

ES 2 719 304 T3

< 212> PRT

< 213> Homo sapiens

<400> 55

Met Pro Glu Ala Pro Pro Leu Leu Leu Ala Ala Val Leu Leu Gly Leu
1 5 10 15

Val Leu Leu Val Val Leu Leu Leu Leu Arg His Trp Gly Trp Gly
20 25 30

Leu Cys Leu Ile Gly Trp Asn Glu Phe Ile Leu Gln Pro Ile His Asn
35 40 45

Leu Leu Met Gly Asp Thr Lys Glu Gln Arg Ile Leu Asn His Val Leu
50 55 60

Gln His Ala Glu Pro Gly Asn Ala Gln Ser Val Leu Glu Ala Ile Asp
65 70 75 80

Thr Tyr Cys Glu Gln Lys Glu Trp Ala Met Asn Val Gly Asp Lys Lys
85 90 95

Gly Lys Ile Val Asp Ala Val Ile Gln Glu His Gln Pro Ser Val Leu
100 105 110

Leu Glu Leu Gly Ala Tyr Cys Gly Tyr Ser Ala Val Arg Met Ala Arg
115 120 125

Leu Leu Ser Pro Gly Ala Arg Leu Ile Thr Ile Glu Ile Asn Pro Asp
130 135 140

Cys Ala Ala Ile Thr Gln Arg Met Val Asp Phe Ala Gly Val Lys Asp
145 150 155 160

Lys Val Thr Leu Val Val Gly Ala Ser Gln Asp Ile Ile Pro Gln Leu
165 170 175

Lys Lys Lys Tyr Asp Val Asp Thr Leu Asp Met Val Phe Leu Asp His
180 185 190

5

Trp Lys Asp Arg Tyr Leu Pro Asp Thr Leu Leu Leu Glu Glu Cys Gly
195 200 205

Leu Leu Arg Lys Gly Thr Val Leu Leu Ala Asp Asn Val Ile Cys Pro
210 215 220

Gly Ala Pro Asp Phe Leu Ala His Val Arg Gly Ser Ser Cys Phe Glu
225 230 235 240

Cys Thr His Tyr Gln Ser Phe Leu Glu Tyr Arg Glu Val Val Asp Gly
245 250 255

Leu Glu Lys Ala Ile Tyr Lys Gly Pro Gly Ser Glu Ala Gly Pro
260 265 270

ES 2 719 304 T3

<210> 56
 < 211> 263
 < 212> PRT
 < 213> Vanilla planifolia

5 <400> 56
 Met Ala Thr Thr Val Ala Thr Ala Thr Arg Ala Thr Glu Asn Lys Thr
 1 5 10 15
 Gln Thr Glu Glu Asn Ser Gln Asn Gly Gly Gln Gln Thr Gly His Gln
 20 25 30
 Glu Ile Gly His Lys Ser Leu Leu Lys Ser Asp Ala Leu Tyr Gln Tyr
 35 40 45
 Ile Leu Glu Thr Ser Val Tyr Pro Arg Glu Pro Glu Cys Leu Lys Glu
 50 55 60
 Leu Arg Glu Ile Thr Ala Lys His Pro Trp Asn Leu Met Thr Thr Ser
 65 70 75 80
 Ala Asp Glu Gly Gln Phe Leu Gly Met Leu Leu Lys Leu Ile Asn Ala
 85 90 95
 Lys Asn Thr Met Glu Ile Gly Val Phe Thr Gly Tyr Ser Leu Leu Ala
 100 105 110
 Thr Ala Leu Ala Leu Pro Asp Asp Gly Lys Ile Leu Ala Met Asp Ile
 115 120 125
 Asn Arg Glu Asn Tyr Glu Leu Gly Leu Pro Leu Ile Gln Lys Ala Gly
 130 135 140
 Val Ala His Lys Ile Asp Phe Arg Glu Gly Pro Ala Leu Pro Val Leu
 145 150 155 160
 Asp Glu Leu Met Lys Asp Glu Ser Lys His Gly Ser Phe Asp Phe Ile
 165 170 175
 Phe Val Asp Ala Asp Lys Asp Asn Tyr Leu Asn Tyr His Gln Arg Ile
 180 185 190
 Ile Asp Leu Val Lys Val Gly Gly Val Ile Gly Tyr Asp Asn Thr Leu
 195 200 205
 Trp Asn Gly Ala Val Val Leu Pro Pro Asp Ala Pro Met Arg Lys Tyr
 210 215 220
 Ile Arg Tyr Tyr Arg Asp Phe Val Ile Glu Leu Asn Lys Glu Leu Ala
 225 230 235 240
 Ala Asp Pro Arg Ile Glu Ile Cys Gln Leu Pro Val Gly Asp Gly Ile
 245 250 255
 Thr Leu Cys Arg Arg Val Lys
 260

REIVINDICACIONES

1. Un método para producir vainillina y/o beta-D-glucósido de vainillina que comprende
- 5 (A) proporcionar un hospedante recombinante capaz de producir vainillina, en donde dicho hospedante recombinante es un microorganismo, planta o célula de planta y alberga un ácido nucleico heterólogo que codifica un polipéptido de la enzima multifuncional Arom (AROM) mutante, cuyo polipéptido de AROM mutante tiene una secuencia de aminoácidos que comparte al menos 80 por ciento de identidad de secuencia con la SEQ ID NO: 4, determinado a lo largo de la longitud entera de la SEQ ID NO: 4, presenta actividad de 3-deshidroquinato deshidratasa, actividad de 3-deshidroquinato sintasa, actividad de 3-fosfoshikimato 1-carboxiviniltransferasa, actividad de shikimato quinasa, menor actividad de shikimato 3-deshidrogenasa (NADP+), y comprende una o más de las siguientes mutaciones:
- 10 (a) un aminoácido distinto de valina en una posición que se alinea con la posición 1349 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 15 (b) un aminoácido distinto de treonina en una posición que se alinea con la posición 1366 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (c) un aminoácido distinto de lisina en una posición que se alinea con la posición 1370 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (d) un aminoácido distinto de isoleucina en una posición que se alinea con la posición 1387 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 20 (e) un aminoácido distinto de treonina en una posición que se alinea con la posición 1392 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (f) un aminoácido distinto de alanina en una posición que se alinea con la posición 1441 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 25 (g) un aminoácido distinto de arginina en una posición que se alinea con la posición 1458 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (h) un aminoácido distinto de prolina en una posición que se alinea con la posición 1500 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (i) un aminoácido distinto de alanina en una posición que se alinea con la posición 1533 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; y
- 30 (j) un aminoácido distinto de triptófano en una posición que se alinea con la posición 1571 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; o
- tiene un aminoácido como se expone en la SEQ ID NO: 2;
- (B) cultivar dicho hospedante recombinante durante un tiempo suficiente para que dicho hospedante recombinante produzca vainillina y/o glucósido de vainillina; y
- 35 (C) aislar la vainillina y/o glucósido de vainillina de dicho hospedante recombinante o del líquido sobrenadante del cultivo, produciendo así vainillina y/o beta-D-glucósido de vainillina.
2. El método de la reivindicación 1, en donde el polipéptido de AROM mutante comprende una o más de las siguientes mutaciones:
- 40 (a) un resto de glicina en una posición que se alinea con la posición 1349 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (b) un resto de glicina en una posición que se alinea con la posición 1366 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (c) un resto de leucina en una posición que se alinea con la posición 1370 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 45 (d) un resto de histidina en una posición que se alinea con la posición 1387 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;

- (c) un resto de leucina en una posición que se alinea con la posición 1370 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (d) un resto de histidina en una posición que se alinea con la posición 1387 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 5 (e) un resto de lisina en una posición que se alinea con la posición 1392 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (f) un resto de prolina en una posición que se alinea con la posición 1441 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 10 (g) un resto de triptófano en una posición que se alinea con la posición 1458 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (h) un resto de lisina en una posición que se alinea con la posición 1500 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (i) un resto de prolina en una posición que se alinea con la posición 1533 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; y
- 15 (j) un resto de valina en una posición que se alinea con la posición 1571 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4.
3. El método de la reivindicación 1 o 2, en donde dicho hospedante recombinante alberga además un ácido nucleico heterólogo que codifica un polipéptido de catecol-O-metil transferasa (COMT) mutante; en donde el polipéptido de COMT mutante tiene una secuencia de aminoácidos que comparte al menos 80 por ciento de identidad de secuencia con la SEQ ID NO: 27, determinado a lo largo de la longitud entera de la SEQ ID NO: 27; y que tiene una secuencia de aminoácidos que difiere de la SEQ ID NO: 27 en al menos un resto de aminoácido.
- 20 4. El método de la reivindicación 3, en donde el polipéptido de COMT mutante comprende Ala, Arg, Asn, Asp, Cys, Glu, Gln, Gly, His, Lys, Met, Phe, Pro, Ser, Thr, Trp y Tyr en la posición 198 de la SEQ ID NO: 27.
5. El método de la reivindicación 3 o 4, en donde el polipéptido de COMT mutante comprende Ala, Arg, Asn, Cys, Gln, Gly, His, Ile, Leu, Lys, Met, Phe, Pro, Ser, Thr, Trp, Tyr y Val en la posición 199 de la SEQ ID NO: 27.
- 25 6. El método de una cualquiera de las reivindicaciones 1 a 5, en donde el polipéptido de AROM mutante o polipéptido de COMT mutante comprende además un marcador de purificación, un péptido de tránsito al cloroplasto, un péptido de tránsito mitocondrial, un péptido de amiloplasto, péptido señal, o un marcador de secreción en el extremo N o C de dicho polipéptido.
- 30 7. Un polipéptido de la enzima multifuncional Arom (AROM) mutante aislado, cuyo polipéptido de AROM mutante tiene una secuencia de aminoácidos que comparte al menos 80 por ciento de identidad de secuencia con la SEQ ID NO: 4, determinado a lo largo de la longitud entera de la SEQ ID NO: 4, presenta actividad de 3-deshidroquinato deshidratasa, actividad de 3-deshidroquinato sintasa, actividad de 3-fosfoshikimato 1-carboxiviniltransferasa, actividad de shikimato quinasa, menor actividad de shikimato 3-deshidrogenasa (NADP+), y comprende una o más de las siguientes mutaciones:
- 35 (a) un aminoácido distinto de valina en una posición que se alinea con la posición 1349 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (b) un aminoácido distinto de treonina en una posición que se alinea con la posición 1366 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 40 (c) un aminoácido distinto de lisina en una posición que se alinea con la posición 1370 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (d) un aminoácido distinto de isoleucina en una posición que se alinea con la posición 1387 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 45 (e) un aminoácido distinto de treonina en una posición que se alinea con la posición 1392 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (f) un aminoácido distinto de alanina en una posición que se alinea con la posición 1441 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (g) un aminoácido distinto de arginina en una posición que se alinea con la posición 1458 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;

- (h) un aminoácido distinto de prolina en una posición que se alinea con la posición 1500 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (i) un aminoácido distinto de alanina en una posición que se alinea con la posición 1533 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; y
- 5 (j) un aminoácido distinto de triptófano en una posición que se alinea con la posición 1571 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; o
- tiene un aminoácido como se expone en la SEQ ID NO: 2.
8. El polipéptido de AROM mutante aislado de la reivindicación 7, en donde el polipéptido de AROM mutante comprende una o más de las siguientes mutaciones:
- 10 (a) un resto de glicina en una posición que se alinea con la posición 1349 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (b) un resto de glicina en una posición que se alinea con la posición 1366 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 15 (c) un resto de leucina en una posición que se alinea con la posición 1370 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (d) un resto de histidina en una posición que se alinea con la posición 1387 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (e) un resto de lisina en una posición que se alinea con la posición 1392 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 20 (f) un resto de prolina en una posición que se alinea con la posición 1441 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (g) un resto de triptófano en una posición que se alinea con la posición 1458 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- 25 (h) un resto de lisina en una posición que se alinea con la posición 1500 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4;
- (i) un resto de prolina en una posición que se alinea con la posición 1533 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4; y
- (j) un resto de valina en una posición que se alinea con la posición 1571 de la secuencia de aminoácidos expuesta en la SEQ ID NO: 4.
- 30 9. El polipéptido de AROM mutante aislado de la reivindicación 7 u 8, en donde el polipéptido de AROM mutante comprende además un marcador de purificación, un péptido de tránsito al cloroplasto, un péptido de tránsito mitocondrial, un péptido de amiloplasto, péptido señal, o un marcador de secreción en el extremo N o C de dicho polipéptido.
- 35 10. Un ácido nucleico aislado que codifica el polipéptido de AROM mutante de una cualquiera de las reivindicaciones 7 a 9.
11. Un hospedante recombinante, en donde dicho hospedante recombinante es un microorganismo, planta o célula de planta y comprende un ácido nucleico heterólogo que codifica el polipéptido de AROM mutante de una cualquiera de las reivindicaciones 7 a 9.
- 40 12. El hospedante recombinante de la reivindicación 11; o el método de una cualquiera de las reivindicaciones 1 a 6, en donde el hospedante recombinante comprende además un ácido nucleico que codifica un polipéptido de 3-deshidrosihikimato deshidratasa (3DSD), un ácido nucleico que codifica un polipéptido de ácido carboxílico aromático reductasa (ACAR), un ácido nucleico que codifica un polipéptido de fosfopanteteína transferasa (PPTasa), un ácido nucleico que codifica un polipéptido de uridina 5'-difosfoglucosil transferasa (UGT) y/o un ácido nucleico que codifica un alcohol de vainillilo oxidasa (VAO).
- 45 13. El hospedante recombinante de la reivindicación 11 o 12; o el método de una cualquiera de las reivindicaciones 1 a 6, en donde dicho hospedante recombinante alberga un ácido nucleico heterólogo que codifica un polipéptido de AROM mutante de acuerdo con una cualquiera de las reivindicaciones 7 a 9, y además comprende un gen que codifica una O-metiltransferasa (OMT) de tipo natural.

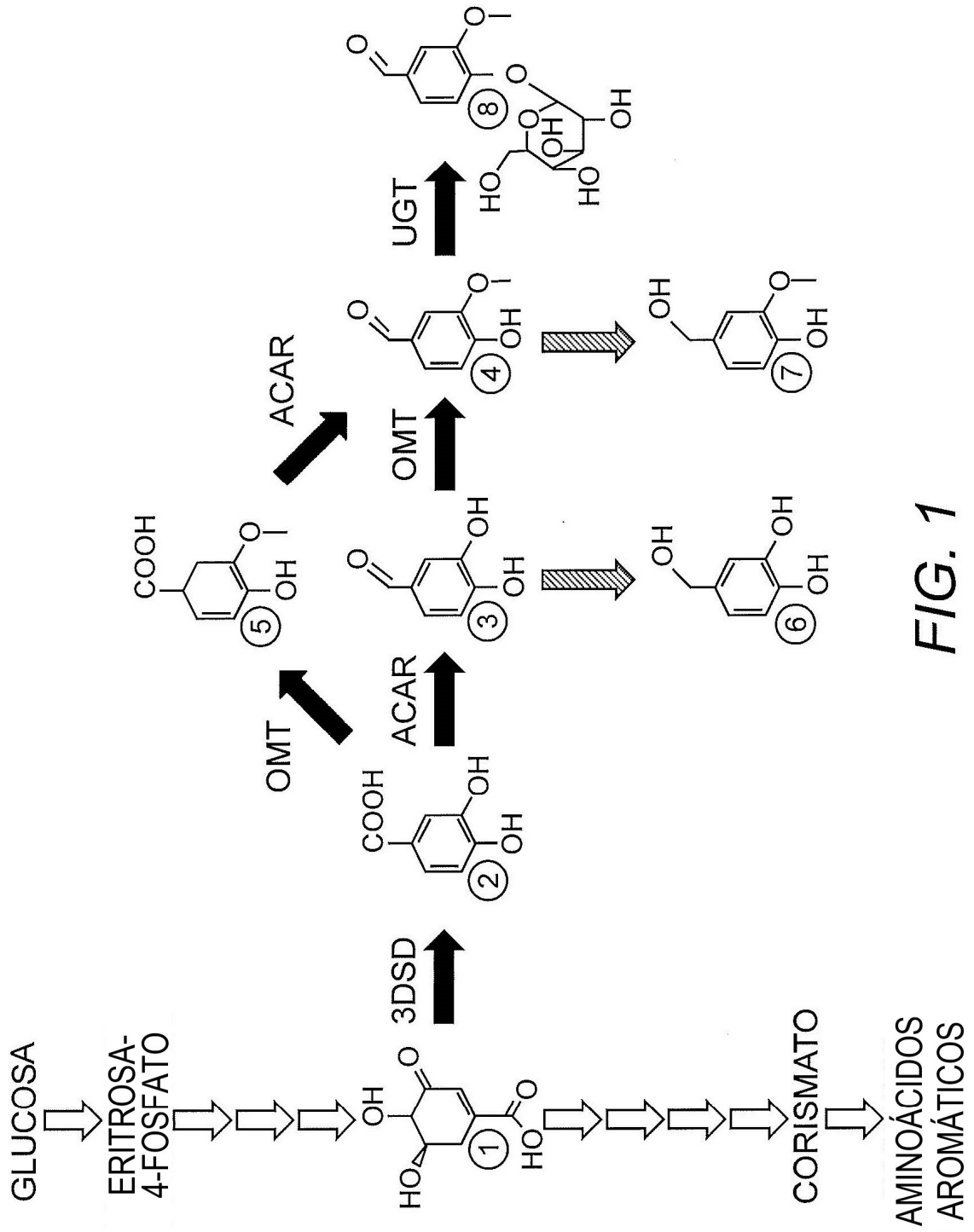


FIG. 1

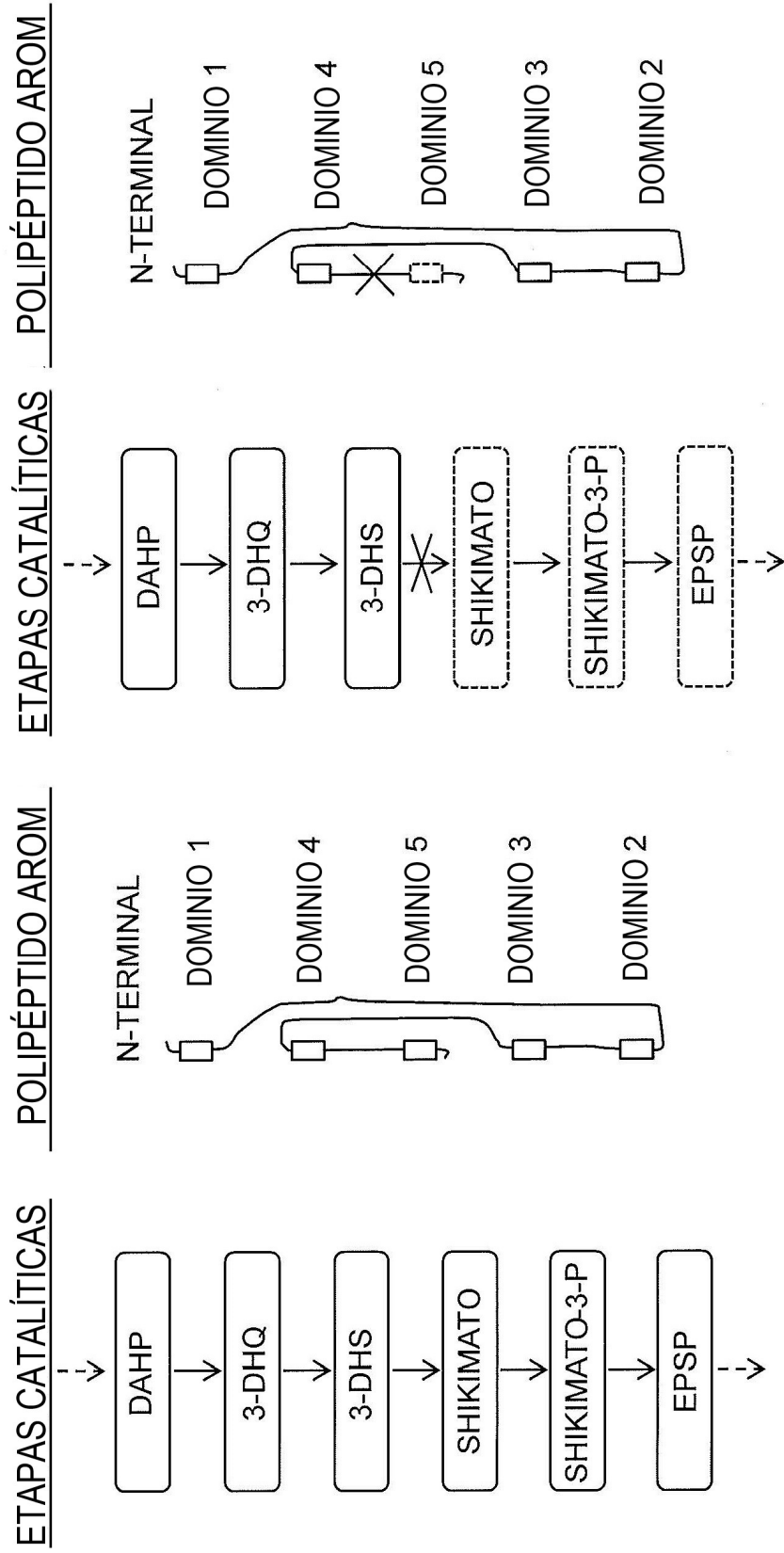


FIG. 2

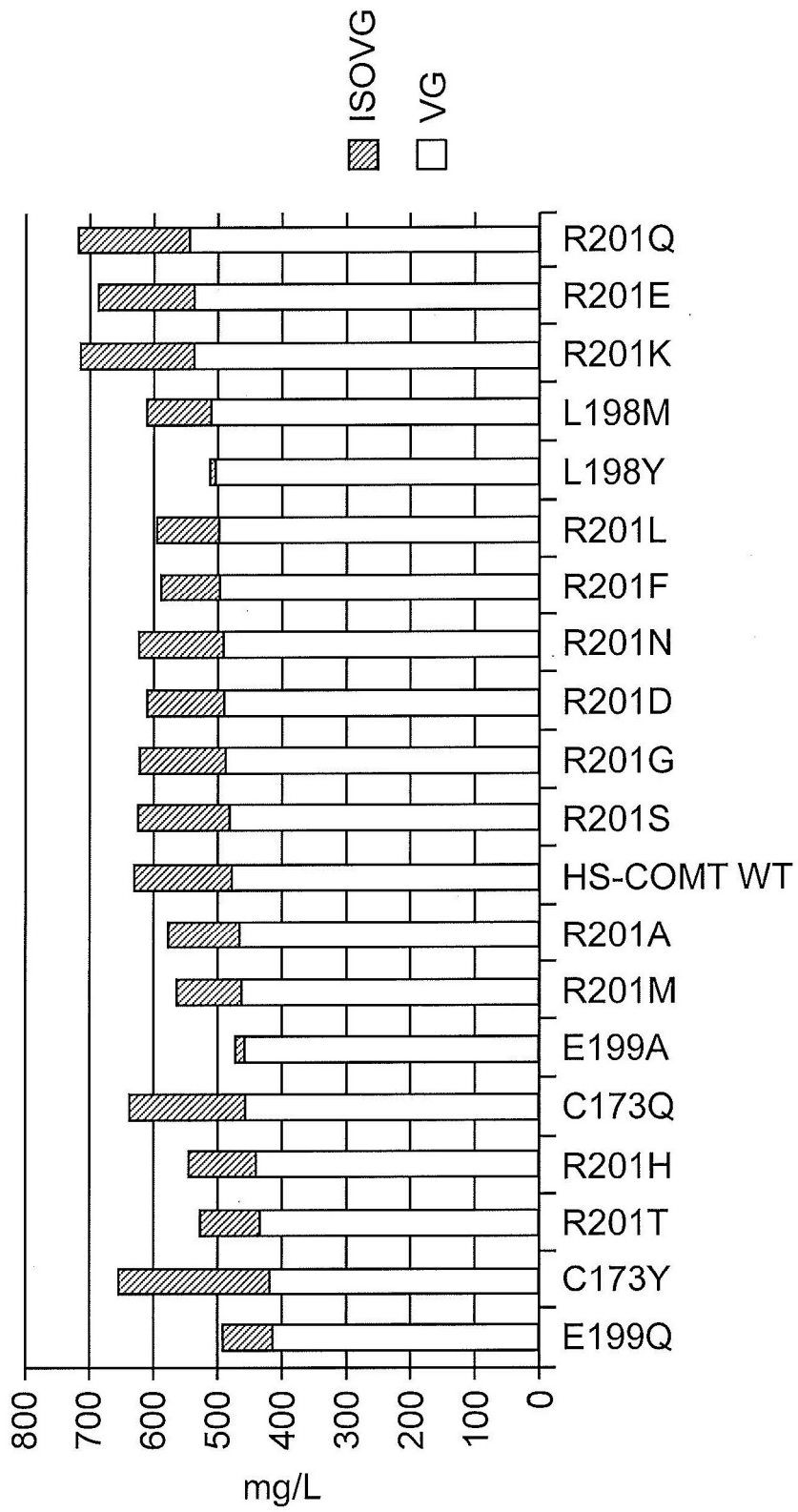


FIG. 3A

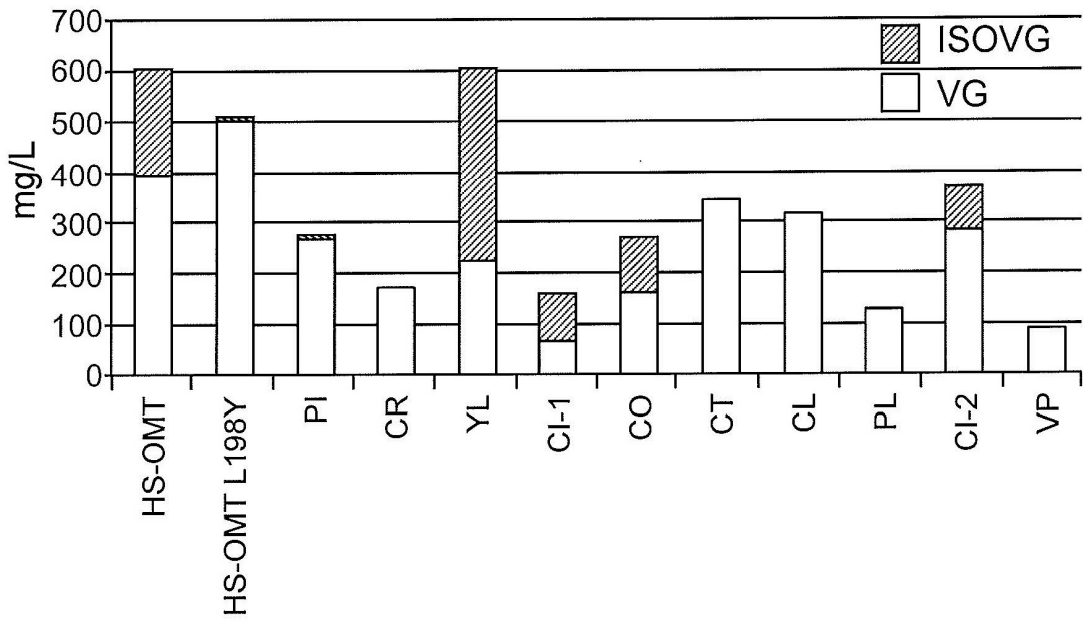


FIG. 3B

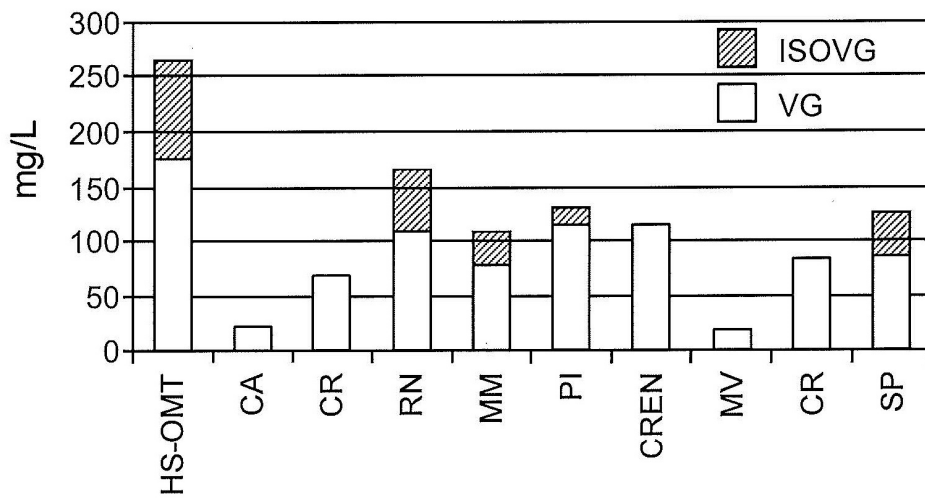


FIG. 3C

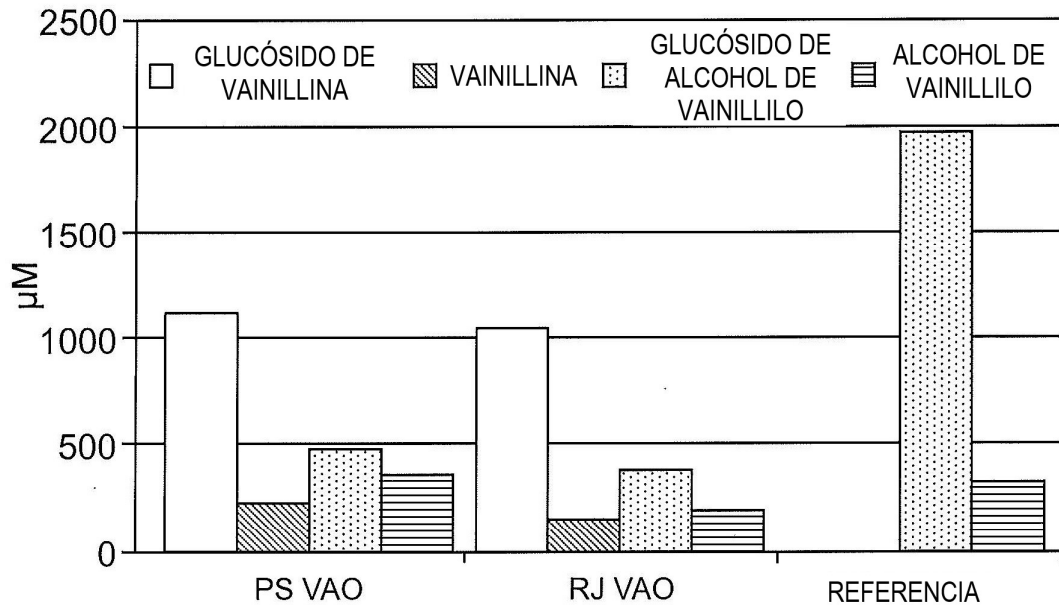


FIG. 4

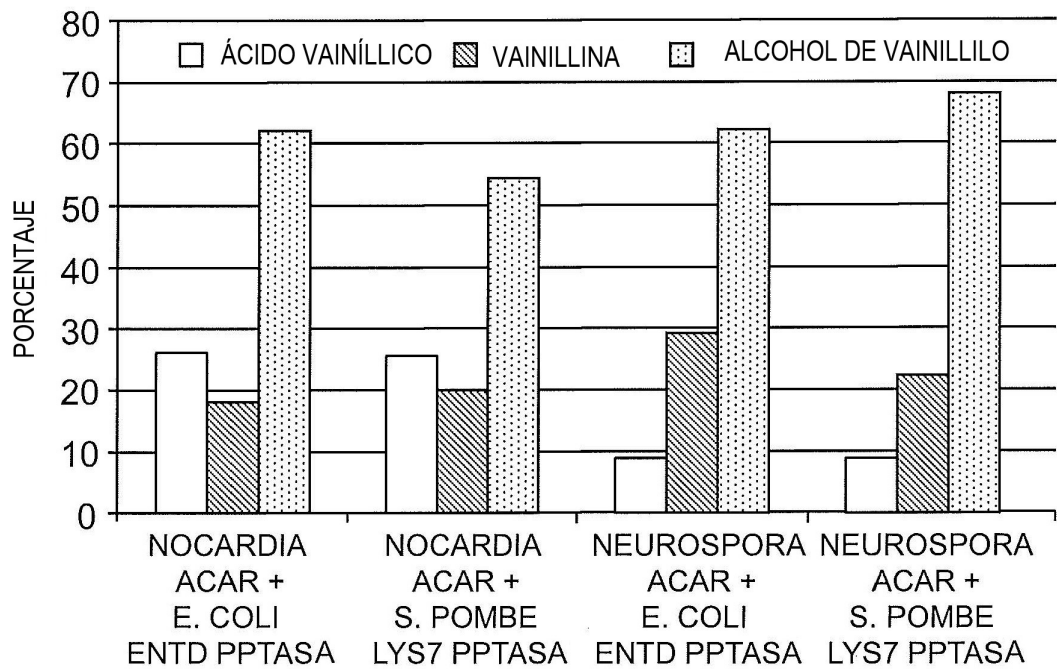


FIG. 5