

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 614 744**

51 Int. Cl.:

**C12N 15/55** (2006.01)

**C12N 9/16** (2006.01)

**A23K 20/189** (2006.01)

**A01H 5/00** (2006.01)

**A01K 67/027** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **04.10.2005 PCT/DK2005/000631**

87 Fecha y número de publicación internacional: **13.04.2006 WO06037327**

96 Fecha de presentación y número de la solicitud europea: **04.10.2005 E 05788896 (8)**

97 Fecha y número de publicación de la concesión europea: **14.12.2016 EP 1809747**

54 Título: **Polipéptidos con actividad de fitasa y polinucleótidos que codifican los mismos**

30 Prioridad:

**04.10.2004 DK 200401515**

**04.10.2004 DK 200401514**

**23.12.2004 DK 200401998**

**23.12.2004 DK 200401999**

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

**01.06.2017**

73 Titular/es:

**NOVOZYMES A/S (100.0%)**

**Krogshøjvej 36**

**2880 Bagsvaerd, DK**

72 Inventor/es:

**TAKAMIYA, MONICA y**

**SJOEHOLM, CARSTEN**

74 Agente/Representante:

**TOMAS GIL, Tesifonte Enrique**

ES 2 614 744 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

Polipéptidos con actividad de fitasa y polinucleótidos que codifican los mismos

5 Antecedentes de la invención

Campo de la invención

10 [0001] La presente invención se refiere a polipéptidos aislados que tienen actividad de fitasa y polinucleótidos aislados que codifican los polipéptidos.

Los polipéptidos se relacionan con fitasas Citrobacter derivadas de Citrobacter amalonaticus, la secuencia de aminoácidos de los cuales se muestra en el listado de secuencias anexo como SEQ ID n.º: 6.

15 La invención también se refiere a constructos de ácidos nucleicos, vectores y células huésped que comprenden los polinucleótidos al igual que métodos para producir y utilizar los polipéptidos, en particular, en el alimento para animales.

Descripción de la técnica relacionada

20 [0002] Las fitasas son enzimas bien conocidas como lo son las ventajas de añadirlas a productos alimenticios para animales, incluidos seres humanos.

Las fitasas han sido aisladas de muchas fuentes, incluido un número de cepas bacterianas y fúngicas.

25 [0003] La fosfatasa ácida de histidina appA de Escherichia coli al igual que otras fitasas bacterianas gram-negativas se conocen por tener una alta actividad específica.

[0004] La producción por Citrobacter braakii YH-15 de una fitasa intracelular se proporciona por Kim et al en Biotechnology Letters 25: 1231-1234, 2003. KR-2004-A-045267 y WO-2004/085638 revelan, como SEQ ID n.º: 7, la secuencia de aminoácidos de una fitasa de Citrobacter braakii YH-15, depositada como KCCM 10427.

30 La WO-2004/085638 se publicó en 07.10.2004, es decir, después de la primera fecha de prioridad de la presente solicitud.

[0005] Es un objeto de la presente invención proporcionar polipéptidos alternativos con actividad de fitasa y polinucleótidos que codifican los polipéptidos.

35 Los polipéptidos de la invención son preferiblemente de propiedades enmendadas, más preferiblemente, mejoradas, por ejemplo, de una especificidad de sustrato diferente, de una actividad específica más alta, de una estabilidad aumentada (tal como estabilidad al ácido, estabilidad térmica y/o estabilidad de proteasa, en particular, estabilidad de pepsina), de un pH enmendado óptimo (tal como un ipH óptimo mayor o menor), y/o de un rendimiento mejorado en el alimento para animales (tal como una liberación mejorada y/o degradación de fitato).

40 Resumen

[0006] La presente invención se refiere a polipéptidos aislados que tienen actividad de fitasa y una estabilidad al ácido de al menos un 60 % de la actividad residual tras la incubación durante la noche a 37 °C en el tampón de ácido clorhídrico/glicina pH 2.2, en relación con la actividad residual tras la incubación durante la noche a 37 °C en el tampón HEPES pH 7.0, seleccionados del grupo que consiste en:

(a) un polipéptido con una secuencia de aminoácidos que tiene al menos un 80 % de identidad con

(i) aminoácidos 1 a 414 de SEQ ID n.º: 6, y/o

(ii) la parte de polipéptido maduro de SEQ ID n.º: 6;

50 donde el grado de identidad se determina por el programa "align" utilizando la matriz de puntuación por defecto BLOSUM50 con penalización para el primer residuo en un espacio de -12 mientras las penalizaciones para residuos adicionales en un espacio son -2.

[0007] La invención también se refiere a polinucleótidos aislados que comprenden una secuencia de nucleótidos que codifica el polipéptido indicado arriba o, específicamente, polinucleótidos aislados que codifican un polipéptido con actividad de fitasa y una estabilidad al ácido de al menos un 60 % de actividad residual tras la incubación durante la noche a 37 °C en el tampón de ácido clorhídrico/glicina pH 2.2, en relación a la actividad residual tras la incubación durante la noche a 37 °C en el tampón HEPES pH 7.0, seleccionados del grupo que consiste en:

(a) un polinucleótido que codifica un polipéptido que tiene una secuencia de aminoácidos con al menos un 80 % de identidad con aminoácidos 1 a 414 de SEQ ID n.º: 6; y

60 (b) un polinucleótido con al menos un 80 % de identidad con nucleótidos 67 a 1308 de SEQ ID n.º: 5, donde el grado de identidad se determina por el programa "align" que es un alineamiento de Needleman-Wunsch (es decir alineamiento global), utilizando la matriz de identidad predeterminada, la penalización para el primer residuo en un espacio es -16 mientras las penalizaciones para residuos adicionales en un espacio son -4.

65 La invención también se refiere a constructos de ácidos nucleicos, vectores de expresión recombinantes y células huésped recombinantes que comprenden los polinucleótidos.

[0008] La invención también se refiere a métodos para producir tales polipéptidos con actividad de fitasa que comprenden (a) cultivo de una célula huésped recombinante que comprende un constructo de ácido nucleico que comprende un polinucleótido que codifica el polipéptido bajo condiciones conductivas para la producción del polipéptido; y (b) recuperación del polipéptido.

[0009] La invención también se refiere a métodos de uso de los polipéptidos de la invención en el alimento para animales, al igual que alimento para animales y composiciones de aditivo de alimento que contienen los polipéptidos.

## Definiciones

[0010] Actividad de fitasa: en el presente contexto un polipéptido que tiene actividad de fitasa (una fitasa) es una enzima que cataliza la hidrólisis de fitato (mioinositol hexaquis-fosfato) a (1) mioinositol y/o (2) mono-, di-, tri-, tetra- y/o penta-fosfatos de los mismos y (3) fosfato inorgánico.

[0011] El sitio enzimático en internet (<http://www.expasi.ch/enzyme/>) es un repositorio de información relativo a la nomenclatura de enzimas.

Principalmente, se basa en las recomendaciones del Comité de Nomenclatura de la Unión Internacional de Bioquímica y Biología Molecular (IUB-MB) y este describe cada tipo de enzima caracterizada para la que se ha proporcionado un número EC (comisión enzimática) (Bairoch A. The ENZYME database, 2000, Nucleic Acids Res 28:304-305).

Ver también el manual de nomenclatura enzimática de NC-IUBBM, 1992).

[0012] Según el sitio ENZYME, se conocen tres tipos diferentes de fitasas: una 3-fitasa (mioinositol hexafosfato 3-fosfohidrolasa, EC 3.1.3.8), una 6-fitasa (mioinositol hexafosfato 6-fosfohidrolasa, EC 3.1.3.26) y una 5-fitasa (EC 3.1.3.72).

Para los fines de la presente invención, todos los tipos se incluyen en la definición de fitasa.

[0013] En una forma de realización particular, las fitasas de la invención pertenecen a la familia de fosfatasa ácida de histidina, que incluye la fosfatasa ácida *Escherichia coli* pH 2,5 (gene *appA*) al igual que fitasas fúngicas tales como fitasas *Aspergillus awamorii* A y B (EC: 3.1.3.8) (gene *phyA* y *phyB*).

La fosfatasa de ácido de histidina comparten dos regiones de similitud de secuencia, cada una centrada alrededor de un residuo de histidina conservado.

Estas dos histidinas parecen estar implicadas en el mecanismo catalítico enzimático.

La primera histidina se localiza en la sección N-terminal y forma un fosfo-histidina intermedio mientras la segunda se localiza en la sección C-terminal y actúa posiblemente como donador de protón.

[0014] En otra divulgación, las fitasas tienen un motivo de sitio activo conservado, es decir, R-H-G-X-R-X-P, donde X designa cualquier aminoácido (ver aminoácidos 16 a 22 de SEQ ID n.º: 2, aminoácidos 16 a 22 de SEQ ID n.º: 4 y aminoácidos 16 a 22 de SEQ ID n.º: 6).

[0015] Para los fines de la presente invención, la actividad de fitasa se determina en la unidad de FYT, un FYT es la cantidad de enzimas que libera 1 micro-mol ortofosfato inorgánico por min. bajo las condiciones siguientes: pH 5,5; temperatura 37 °C; sustrato: fitato de sodio ( $C_6H_6O_{24}P_6Na_{12}$ ) en una concentración de 0,0050 mol/l.

Los ensayos de fitasa adecuados son el FYT y FTU descritos en ejemplo 1 de la WO 00/20569.

FTU es para determinar la actividad de fitasa en el alimento y premezcla.

La actividad de fitasa también se puede determinar utilizando el ensayo de fitasa del ejemplo 4 de aquí.

[0016] El pH óptimo de un polipéptido de la invención se determina por la incubación de fitasa a varios valores de pH, utilizando un sustrato en una concentración predeterminada y una temperatura de incubación fija.

Luego el pH óptimo se determina a partir de una representación gráfica de actividad de fitasa frente al pH.

En una forma de realización particular, se usa el ensayo FYT, es decir, el sustrato es 5 mM de fitato de sodio, la temperatura de reacción 37 °C y la actividad se determina en unidades FYT a varios valores de pH, por ejemplo pH 2-12, usando tampones adecuados, tales como: 100 mM de ácido succínico, 100 mM HEPES, 100 mM CHES, 100 mM CABS, 1 mM  $CaCl_2$ , 150 mM KCl, 0,01 % Tritón X-100 ajustado a valores de pH 2,0; 2,5; 3,0; 3,5; 4,0; 5,0; 6,0; 7,0; 8,0; 9,0; 10,0; 11,0 y 12,0 con HCl o NaOH.

En otra forma de realización particular, se usa el ensayo de fitasa del ejemplo 4, es decir, el sustrato es 0,5 mM, preferiblemente 5 mM, na-fitato, que se disuelve en un tampón con el pH deseado (tal como los mencionados anteriormente) y el fosfato soluble se determina por complexación con molibdato/hierro y medición de densidad óptica a 750 nm.

Ensayo de ciego: se mezclan 20ul de muestra, 100ul de sustrato y 120ul de reactivo de color, se incuban 5 min a 37 °C y el  $OD_{Blind}$  se mide a 750 nm.

Muestra: se mezclan 20ul de muestra, 100ul de sustrato, se incuban 30 min a 37 °C, se añade 120ul de reactivo de color, se incuban 5 min a 37 °C y se mide la  $OD_{sample}$  a 750nm.

La actividad de fitasa se mide como  $OD = OD_{sample} - OD_{Blind}$ .

Un pH óptimo relativamente bajo implica un pH óptimo por debajo de pH 5,0, por ejemplo debajo pH 4,5; 4,0; 3,5;

3,0; 2,5 o incluso por debajo de 2,0.

Un medio de pH óptimo relativamente alto implica un pH óptimo por encima de pH 5,0, por ejemplo, por encima de pH 5,5; 6,0; 6,5; 7,0; 7,5; 8,0; 8,5 o incluso por encima de 9,0.

5 [0017] Polipéptido aislado: el término "polipéptido aislado" como se usa aquí se refiere a un polipéptido que es al menos 20 % puro, preferiblemente, al menos 40 % puro, más preferiblemente, al menos 60 % puro, aún más preferiblemente al menos 80 % puro, de la forma más preferible al menos 90 % puro e incluso de la forma más preferible al menos 95 % puro, como se determina por SDS-PAGE.

10 [0018] Polipéptido sustancialmente puro: el término "polipéptido sustancialmente puro" denota aquí una preparación de polipéptido que contiene como mucho 10 %, preferiblemente, como mucho 8 %, más preferiblemente, como mucho 6 %, más preferiblemente, como mucho 5 %, más preferiblemente, como mucho 4 %, como mucho 3 %, aún más preferiblemente, como mucho 2 %, de la forma más preferible, como mucho 1 % e incluso de la forma más preferible como mucho 0,5 % en peso de otro material de polipéptido con el cual se asocia originalmente.

15 Por lo tanto, se prefiere que el polipéptido sustancialmente puro sea al menos 92 % puro, preferiblemente, al menos 94 % puro, más preferiblemente, al menos 95 % puro, más preferiblemente, al menos 96 % puro, más preferiblemente, al menos 96 % puro, más preferiblemente, al menos 97 % puro, más preferiblemente, al menos 98 % puro, aún más preferiblemente, al menos 99 %, de la forma más preferible al menos 99,5 % puro e incluso de la forma más preferible 100 % en peso puro del material de polipéptido total presente en la preparación.

20 [0019] Los polipéptidos de la presente invención son preferiblemente en una forma sustancialmente pura. En particular, se prefiere que los polipéptidos estén en "forma esencialmente pura", es decir, que la preparación de polipéptido esté esencialmente libre de otro material de polipéptido con el cual está asociado originalmente. Esto se puede realizar, por ejemplo, preparando el polipéptido mediante métodos recombinantes bien conocidos o por métodos de purificación tradicionales.

25 [0020] Aquí, el término "polipéptido sustancialmente puro" es sinónimo de los términos "polipéptido aislado" y "polipéptido en forma aislada"

30 [0021] Identidad: la relación entre dos secuencias de aminoácidos o entre dos secuencias de nucleótidos se describe por la "identidad" de parámetro.

[0022] Para fines de la presente invención, el grado de identidad entre dos secuencias de aminoácidos, al igual que el grado de identidad entre dos secuencias de nucleótidos, se determina por el programa "align" que es un alineamiento de Needleman-Wunsch (es decir, alineamiento global), útil para los alineamientos de proteína y ADN.

35 La matriz de puntuación por defecto BLOSUM50 y la matriz de identidad predeterminada se usan para alineamientos de proteína y ADN respectivamente.

La penalización para el primer residuo en un espacio es -12 para proteínas y -16 para ADN.

40 Mientras las penalizaciones para residuos adicionales en un espacio son -2 para proteínas y -4 para ADN.

[0023] "Align" es parte del FASTA package version v20u6 (ver W.R. Pearson y D. J. Lipman (1988), "Improved Tools for Biological Sequence Analysis", PNAS 85:2444-2448 y W.R. Pearson (1990) "Rapid and Sensitive Sequence Comparison with FASTP and FASTA," Methods in Enzymology 183:63-98).

45 Los alineamientos de proteína FASTA usan el algoritmo de Smith-Waterman sin limitación en tamaño del espacio (ver "Smith-Waterman algorithm", T.F. Smith y M.S. Waterman (1981) J.Mol.Biol. 147:195-197).

[0024] El algoritmo de Needleman-Wunsch se describe en Needleman, S.B. y Wunsch, C.D., (1970), Journal of Molecular Biology, 48: 443-453, y el programa de alineamiento por Myers y W. Miller in "Optimal Alignments in Linear Space" CABIOS (computer applications in the biosciences) (1988) 4:11-17.

50 [0025] El grado de identidad entre la secuencia meta (o muestra, o prueba) y una secuencia específica (por ejemplo aminoácidos 1 a 409 de SEQ ID n.º: 2) también se puede determinar de la siguiente manera: las secuencias se alinean utilizando el programa "align." El número de coincidencias perfectas ("N-coincidencia perfecta") en el alineamiento está determinado (una coincidencia perfecta significa el mismo residuo de aminoácido en misma posición del alineamiento, normalmente designada con un "I" en el alineamiento).

55 La longitud de la secuencia específica (el número de residuos de aminoácidos) está determinada ("N-específico", en el ejemplo mencionado anteriormente = 411).

El grado de identidad se calcula como la proporción entre "N-coincidencia perfecta" y "N-específico" (para conversión a porcentaje identidad, multiplicado por 100).

60 [0026] Alternativamente, el grado de identidad entre la secuencia diana (o muestra, o prueba) y la secuencia específica (por ejemplo, aminoácidos 1 a 409 de SEQ ID n.º: 2) se determina de la siguiente manera: las dos secuencias se alinean utilizando el programa "align." Se determina el número de coincidencias perfectas ("N-coincidencia perfecta") en el alineamiento (una correspondencia perfecta significa el mismo residuo de aminoácido en la misma posición del alineamiento, normalmente designado con "I" en el alineamiento).

65 También se determina la longitud común de las dos secuencias alineadas, es decir, el número total de aminoácidos

en la parte de superposición del alineamiento ("N-superposición").

El grado de identidad se calcula como la proporción entre "N-coincidencia perfecta" y "N-superposición" (para conversión a porcentaje de identidad, multiplicado por 100).

En una manera, n-solapamiento incluye espacios de cabeza y cola creados por el alineamiento, si hay.

5 En otra manera, n-solapamiento excluye espacios de cabeza y cola creados por el alineamiento, si hay.

[0027] El grado de identidad entre una secuencia diana (o muestra, o prueba) y una secuencia específica (por ejemplo, aminoácidos 1 a 409 de SEQ ID n.º: 2) también se puede determinar de la siguiente manera: las secuencias se alinean utilizando el programa "align." Se determina el número de coincidencias perfectas ("N-coincidencias perfectas") en el alineamiento (una correspondencia perfecta significa el mismo residuo de aminoácido en la misma posición del alineamiento, normalmente designado con "I" en el alineamiento).

Se determina la longitud de la secuencia diana (el número de residuos de aminoácidos) ("N-diana").

El grado de identidad se calcula como la proporción entre "N-coincidencia perfecta" y "N-diana" (para la conversión a porcentaje de identidad, multiplicado por 100).

15 [0028] Preferiblemente, el recubrimiento es al menos un 20 % de la secuencia específica ("N-superposición" tal como se ha definido anteriormente, dividido por el número de los aminoácidos en la secuencia específica ("N-específico") y multiplicado por 100), más preferiblemente, al menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 % o al menos 95 %.

20 Esto significa que al menos un 20 % (preferiblemente un 25-95 %) de los aminoácidos de la secuencia específica acaban siendo incluidos en el recubrimiento, cuando la secuencia de muestra se alinea a la secuencia específica.

[0029] En la alternativa, el recubrimiento es al menos 20 % de la secuencia diana (o muestra, o prueba) ("N-superposición" tal como se ha definido anteriormente, dividido por "N-diana" tal como se ha definido anteriormente y multiplicado por 100), más preferiblemente, al menos 25 %, 30 %, 35 %, 40 %, 45 %, 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 85 %, 90 % o al menos 95 %.

25 Esto significa que al menos el 20 % (preferiblemente 25-95 %) de los aminoácidos de la secuencia diana acaban siendo incluidos en el recubrimiento, cuando se alinea frente a la secuencia específica.

[0030] Fragmento de polipéptido: el término "fragmento polipéptido" se define aquí como un polipéptido con uno o más aminoácidos eliminados de amino y/o término de carboxilo de la parte de péptido maduro de SEQ ID n.º: 2, SEQ ID n.º: 4, SEQ ID n.º: 6 o una secuencia homóloga de cualquiera de estos, donde el fragmento tiene actividad de fitasa.

30 En formas de realización particulares, el fragmento contiene al menos 350, 360, 370, 380, 390, 400, 405 o al menos 35 410 de residuos de aminoácidos.

[0031] Subsecuencia: el término "subsecuencia" se define aquí como una secuencia de nucleótidos con uno o más nucleótidos eliminados del extremo 5' y/o 3' de la parte de codificación de péptido maduro de SEQ ID n.º: 1, SEQ ID n.º: 3, SEQ ID n.º: 5 o una secuencia homóloga de cualquiera de estas, donde la subsecuencia codifica un fragmento de polipéptido que tiene actividad de fitasa.

40 En formas de realización particulares, la subsecuencia contiene al menos 1050, 1080, 1110, 1140, 1170, 1200, 1215 o al menos 1230 nucleótidos.

[0032] Variante alélica: el término "variante alélica" denota aquí cualquiera de dos o más formas alternativas de un gen que ocupa el mismo locus cromosómico.

La variación alélica surge naturalmente por mutación y puede resultar en el polimorfismo dentro de poblaciones.

Las mutaciones de gen pueden ser silenciosas (ningún cambio en el polipéptido codificado) o pueden codificar polipéptidos con secuencias de aminoácidos alteradas.

Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

50 [0033] Polinucleótido sustancialmente puro: el término "polinucleótido sustancialmente puro" como se usa aquí se refiere a una preparación de polinucleótido libre de otros nucleótidos extraños o no deseados y en una forma adecuada para su uso dentro de sistemas de producción de proteína genéticamente modificada.

55 Así, un polinucleótido sustancialmente puro contiene como mucho 10 %, preferiblemente, como mucho 8 %, más preferiblemente, como mucho 6 %, más preferiblemente, como mucho 5 %, más preferiblemente, como mucho 4 %, más preferiblemente, como mucho 3 %, aún más preferiblemente, como mucho 2 %, de la forma más preferible, como mucho 1 % e incluso de la forma más preferible como mucho 0,5 % en peso de otro material de polinucleótido con el cual se asocia originalmente.

60 Sin embargo, un polinucleótido sustancialmente puro puede incluir regiones no traducidas de origen natural 5' y 3', tales como promotores y terminadores.

Se prefiere que el polinucleótido sustancialmente puro es al menos 90 % puro, preferiblemente, al menos 92 % puro, más preferiblemente, al menos 94 % puro, más preferiblemente, al menos 95 % puro, más preferiblemente, al menos 96 % puro, más preferiblemente, al menos 97 % puro, aún más preferiblemente, al menos 98 % puro, de la forma más preferible, al menos 99 % e incluso de la forma más preferible, al menos 99,5 % en peso puro.

65 Los polinucleótidos de la presente invención están preferiblemente en una forma sustancialmente pura.

En particular, se prefiere que los polinucleótidos que aquí se describen estén en "forma esencialmente pura", es

decir, que la preparación de polinucleótido esté esencialmente libre de otro material de polinucleótido con el cual se asocia originalmente.

Aquí, el término "polinucleótico sustancialmente puro" es sinónimo de los términos "polinucleótido aislado" y "polinucleótido en forma aislada". Los polinucleótidos pueden ser de origen genómico, de ADNc, de ARN, semisintéticos, sintéticos o cualquier combinación de los mismos.

[0034] Constructo de ácido nucleico: el término "constructo de ácido nucleico" como se utiliza en este caso se refiere a una molécula de ácido nucleico, bien uni- o bicatenaria, que se aísla de un gen de origen natural o que se modifica para contener segmentos de ácidos nucleicos de manera que de otro modo no existirían en la naturaleza.

El término constructo de ácido nucleico es sinónimo del término "casete de expresión" cuando el constructo de ácido nucleico contiene las secuencias de control requeridas para la expresión de una secuencia codificante de la presente invención.

[0035] Secuencia de control: el término "secuencias de control" se define aquí para incluir todos los componentes, que son necesarios o ventajosos para la expresión de un polinucleótido que codifica un polipéptido de la presente invención.

Cada secuencia de control puede ser nativa o extranjera de la secuencia de nucleótidos que codifica el polipéptido. Tales secuencias de control incluyen, pero de forma no limitativa, un líder, una secuencia de poliadenilación, secuencia de propéptido, promotor, secuencia de péptido señal y terminador de transcripción.

Como mínimo, las secuencias de control incluyen un promotor y señales de parada transcripcionales y traduccionales.

Las secuencias de control se pueden proporcionar con enlaces con motivo de introducir sitios de restricción específicos que faciliten la unión de las secuencias de control con la región de codificación de la secuencia de nucleótidos que codifican un polipéptido.

[0036] Operativamente enlazado: el término "operativamente enlazado" denota aquí una configuración en la que una secuencia de control se coloca en una posición apropiada relativa a la secuencia de codificación de la secuencia de polinucleótidos de tal manera que la secuencia de control dirija la expresión de la secuencia codificadora de un polipéptido.

[0037] Secuencia codificante: cuando se usa aquí, el término "secuencia codificante" significa una secuencia de nucleótidos, que especifica directamente la secuencia de aminoácidos de su producto proteínico.

Los límites de la secuencia codificante se determinan generalmente por un marco de lectura abierto, que empieza normalmente con el codón de inicio ATG o codones de inicio alternativos tales como GTG y TTG.

La secuencia codificante puede ser un ADN, ADNc o secuencia de nucleótidos recombinante.

[0038] Parte de polipéptido maduro: cuando se usan aquí, los términos "parte de polipéptido maduro" o "parte de péptido maduro" se refieren a la parte del polipéptido segregada por una célula que contiene, como parte de su equipo genético, un polinucleótido que codifica el polipéptido.

En otras palabras, la parte de polipéptido maduro se refiere a la parte del polipéptido que permanece después de que la parte de péptido señal se divida una vez que haya conseguido su función de dirigir el polipéptido codificado a la vía secretora de la célula.

La parte de péptido señal predicho de SEQ ID n.º: 2, SEQ ID n.º: 4 y SEQ ID n.º: 6 es los aminoácidos -22 a -1 del mismo, que significa que la parte de polipéptido maduro predicho de SEC n.º 2, 4 y 6, respectivamente, corresponde a aminoácidos 1 a 409, 1 a 410 y 1 a 414 del mismo, respectivamente.

Sin embargo, una ligera variación puede ocurrir de célula huésped a célula huésped y, por lo tanto, se prefiere la parte de polipéptido maduro de expresión.

[0039] Parte que codifica el polipéptido maduro: cuando se usa aquí, el término "parte que codifica el péptido maduro" o "secuencia que codifica el péptido maduro" se refiere a la parte del polinucleótido que codifica el polipéptido que codifica la parte de polipéptido maduro.

Por ejemplo, para SEC ID n.º: 1, 3 y 5, las partes de codificación de polipéptido maduro predicho corresponden a los nucleótidos 67 a 1293, 67 a 1296 y 67 a 1308 de los mismos, respectivamente (aminoácidos de codificación 1 a 409 de SEQ ID n.º: 2, aminoácidos 1 a 410 de SEQ ID n.º: 4 y aminoácidos 1 a 414 de SEQ ID n.º: 6, respectivamente).

[0040] Expresión: el término "expresión" incluye cualquier paso implicado en la producción del polipéptido incluyendo, pero de forma no limitativa, transcripción, modificación postranscripcional, traducción, modificación postraduccional y secreción.

[0041] Vector de expresión: el término "vector de expresión" se define aquí como una molécula de ADN lineal o circular que comprende un polinucleótido que codifica un polipéptido de la invención y que está operativamente enlazado a nucleótidos adicionales que proveen para su expresión.

[0042] Célula huésped: el término "célula huésped", como se utiliza en este caso, incluye cualquier tipo celular que sea susceptible de transformación, transfección, transducción y similar con un constructo de ácido nucleico que comprende un polinucleótido de la presente invención.

[0043] Modificación: el término "modificación" implica aquí cualquier modificación química de los polipéptidos que consiste en aminoácidos 1 a 409 de la SEQ ID n.º: 2, aminoácidos 1 a 410 de la SEQ ID n.º: 4 o aminoácidos 1 a 414 de la SEQ ID n.º: 6, al igual que la manipulación genética del ADN que codifica estos polipéptidos.

5 La(s) modificación(es) puede(n) ser sustitución(es), deleción(es) y/o inserció(nes) del aminoácido(s) al igual que sustitución(es) de cadena(s) lateral(es) de aminoácido.

[0044] Variante artificial: cuando se usa aquí, el término "variante artificial" significa un polipéptido con actividad de fitasa producida por un organismo que expresa una secuencia de nucleótidos modificada de SEQ ID n.º: 1, SEQ ID n.º: 3 o SEQ ID n.º: 5.

10 La secuencia de nucleótidos modificada se obtiene por intervención humana por modificación de la secuencia de nucleótidos descrita en la SEC ID n.º: 1, 3 y 5.

Descripción detallada

15 Polipéptidos con actividad de fitasa

[0045] En un primer aspecto, la presente invención se refiere a polipéptidos aislados

20 [0046] Con la actividad de fitasa y una estabilidad al ácido de al menos 60 % de actividad residual tras la incubación durante la noche a 37 °C en el tampón de ácido clorhídrico/glicina pH 2,2, con relación a la actividad residual tras la incubación durante la noche a 37 °C en el tampón HEPES pH 7,0, seleccionado del grupo que consiste en:

(a) un polipéptido con una secuencia de aminoácidos que tiene al menos un 80 % de identidad con

25 (i) aminoácidos 1 a 414 de SEQ ID n.º: 6 y/o

(ii) la parte de polipéptido maduro de SEQ ID n.º: 6;

donde el grado de identidad se determina por el programa "align" utilizando la matriz de puntuación por defecto BLOSUM50 con penalización para el primer residuo en un espacio de -12 mientras las penalizaciones para residuos adicionales en un espacio son -2.

30 [0047] En formas de realización particulares, el grado de identidad es al menos 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 o al menos 99 %.

[0048] En otras descripciones particulares, los polipéptidos tienen una secuencia de aminoácidos que difiere en veinte, dieciocho, dieciséis, catorce, doce, diez, ocho, seis, cinco, cuatro, tres, dos o en un aminoácido de cualquiera de los aminoácidos 1 a 414 de la SEQ ID n.º: 6.

35 [0049] La secuencia de nucleótidos de la SEQ ID n.º:5 o una subsecuencia de la misma, al igual que la secuencia de aminoácidos de SEQ ID n.º:6 se puede utilizar para diseñar una sonda de ácidos nucleicos para identificar y clonar polipéptidos de codificación de ADN con actividad de fitasa de cepas de diferentes géneros o especies, según métodos bien conocidos en la técnica. En particular, tales sondas se pueden usar para la hibridación con el genómico o ADNc del género o especies de interés, siguiendo los procedimientos estándar de hibridación Southern, para identificar y aislar el gen correspondiente del mismo.

40 Tales sondas pueden ser considerablemente más cortas que la secuencia entera, pero deberían ser al menos 14, preferiblemente, al menos 25, más preferiblemente, al menos 35 y de la forma más preferible al menos 70 nucleótidos de longitud.

Sin embargo, se prefiere que la sonda de ácido nucleico sea al menos 100 nucleótidos de longitud.

45 Por ejemplo, la sonda de ácidos nucleicos puede ser al menos 200 nucleótidos, preferiblemente, al menos 300 nucleótidos, más preferiblemente, al menos 400 nucleótidos o de la forma más preferible al menos 500 nucleótidos de longitud.

50 Incluso se pueden utilizar sondas más largas, por ejemplo, sondas de ácido nucleico que son al menos 600 nucleótidos, al menos preferiblemente, al menos 700 nucleótidos, más preferiblemente, al menos 800 nucleótidos o de la forma más preferible al menos 900 nucleótidos de longitud.

Se pueden usar sondas ADN y ARN.

55 Las sondas están típicamente marcadas para la detección del gen correspondiente (por ejemplo, con <sup>32</sup>P, <sup>3</sup>H, <sup>35</sup>S, biotina o avidina).

Tales sondas se incluyen en la presente descripción.

[0050] Una biblioteca de ADN genómico obtenida a partir de tales otros organismos se puede, por lo tanto, cribar de ADN que hibridiza con las sondas anteriormente descritas y que codifica un polipéptido con actividad de fitasa.

60 El ADN genómico u otro de tales otros organismos se puede separar por electroforesis en gel de poliacrilamida o agarosa, u otras técnicas de separación.

El ADN de las bibliotecas o el ADN separado se puede transferir e inmovilizar en la nitrocelulosa u otro material de portador adecuado.

65 Para identificar un clon o ADN que es homólogos de la SEC ID n.º:5 o una subsecuencia de la misma, el material portador se usa en una transferencia de Southern.

[0051] Para fines de la presente invención, la hibridación indica que la secuencia de nucleótidos hibridiza a una sonda de ácidos nucleicos marcada que corresponde con las secuencias de nucleótidos mostradas en la SEQ ID n.º: 5, la cadena complementaria o subsecuencias de las mismas, bajo condiciones de astringencia muy bajas a muy altas.

5 Las moléculas a las que la sonda de ácidos nucleicos hibridiza bajo estas condiciones se pueden detectar usando una película de rayos X.

[0052] En una divulgación particular, la sonda de ácidos nucleicos es cualquiera de las SEC ID n.º: 1, 3, 5 o 7-18.

10 En otro método particular, la sonda de ácidos nucleicos es la cadena complementaria de nucleótidos 67 a 450, nucleótidos 450 a 900 o nucleótidos 900 a 1293 de SEC ID n.º 1, 3 o 5.

En otra divulgación particular, la sonda de ácidos nucleicos es una secuencia de polinucleótidos que codifica el polipéptido de SEC ID n.º 2, 4 o 6, o una subsecuencia de cualquiera de estos.

Todavía en otra divulgación particular, la sonda de ácidos nucleicos es la SEQ ID n.º: 1, SEQ ID n.º: 3 o SEQ ID n.º: 5, en particular, las regiones de codificación de polipéptido maduro de las mismas.

15 [0053] Para sondas largas de al menos 100 nucleótidos de longitud, se definen condiciones de astringencia muy bajas a muy altas como prehibridación e hibridación a 42 °C en 5X SSPE, 0,3 % SDS, 200 µg/ml de ADN de esperma de salmón cortado y desnaturalizado, y bien 25 % de formamida para astringencias muy bajas y bajas, 35 % de formamida para astringencias medias y medio altas o 50 % de formamida para astringencias altas y muy altas, siguiendo los procedimientos estándar de transeferencia Southern durante 12 a 24 horas óptimamente.

20 [0054] Para sondas largas de al menos 100 nucleótidos de longitud, el material portador finalmente se lava tres veces, cada una durante 15 minutos utilizando 2X SSC, 0,2 % SDS preferiblemente al menos a 45 °C (astringencia muy baja), más preferiblemente, al menos a 50 °C (astringencia baja), más preferiblemente, al menos a 55 °C (astringencia media), más preferiblemente, al menos a 60 °C (astringencia medio alta), aún más preferiblemente, al menos a 65 °C (astringencia alta) y de la forma más preferible, al menos a 70 °C (astringencia muy alta).

25 [0055] Para sondas cortas que son aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, las condiciones de astringencia se definen como prehibridación, hibridación y post-hibridación de lavado en alrededor de 5 °C a aproximadamente 10 °C por debajo del T<sub>m</sub> calculado utilizando el cálculo según Bolton y McCarthy (1962, Proceeding of the National Academy of Sciences USA 48:1390) en 0,9 M NaCl, 0,09 M pH Tris-HCl 7.6,6 mM EDTA, 0,5 % NP-40, 1X solución de Denhardt, 1 mM pirofosfato de sodio, 1 mM fosfato monobásico de sodio, 0,1 mM ATP y 0,2 mg de levadura ARN por ml siguiendo los procedimientos estándar de transferencia Southern.

30 [0056] Para sondas cortas que son aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, el material portador se lava una vez en 6X SCC más 0,1 % SDS durante 15 minutos y dos veces, cada una durante 15 minutos utilizando 6X SSC de 5 °C a 10 °C por debajo del T<sub>m</sub> calculado.

35 [0057] Bajo condiciones de hibridación con contenido en sal, la T<sub>m</sub> efectiva es la que controla el grado de identidad requerido entre la sonda y el ADN ligado al filtro para que la hibridación tenga éxito. La T<sub>m</sub> efectiva se puede determinar utilizando la fórmula de abajo para determinar el grado de identidad requerida para dos ADNs para hibridar bajo varias condiciones de astringencia.

40 
$$T_m \text{ efectiva} = 81.5 + 16.6(\log M[\text{Na}^+]) + 0.41(\%G+C) - 0.72(\% \text{formamida})$$

45 (ver [www.ndsu.nodak.edu/instruct/mcclean/plsc731/dna/dna6.htm](http://www.ndsu.nodak.edu/instruct/mcclean/plsc731/dna/dna6.htm))

[0058] "G+C" designa el contenido de nucleótidos G y T en la sonda.

50 Para una astringencia media, por ejemplo, la formamida es 35 % y la concentración Na<sup>+</sup> para 5X SSPE es 0,75 M.

[0059] Como se indica, la presente invención se refiere a polipéptidos aislados que tienen actividad de fitasa y las siguientes propiedades fisicoquímicas (como se ha analizado en los polipéptidos sustancialmente puros):

55 (ii) estabilidad al ácido; tal como  
 (a) al menos 60 % de la actividad residual tras la incubación durante la noche a 37 °C en el tampón de ácido clorhídrico/glicina pH 2,2, en relación a la actividad residual tras la incubación durante la noche a 37 °C en el tampón HEPES pH 7,0;

[0060] La presente descripción se refiere también a variantes artificiales que comprenden una substitución conservadora, deleción y/o inserción de uno o más aminoácidos de SEC ID n.º 2, 4 o 6, o los polipéptidos maduros de cualquiera de estos.

Una inserción puede estar dentro de la molécula y/o en el extremo N- y/o C-terminal de la molécula en cuyo caso también se llama extensión.

60 Preferiblemente, los cambios de aminoácidos son de una naturaleza menor, que es sustituciones de aminoácido conservador; deleciones pequeñas, típicamente de uno a aproximadamente 30 aminoácidos; pequeñas extensiones



de amino- o carboxilo-terminales, tal como un residuo de metionina aminoterminal; un pequeño péptido enlazador de hasta aproximadamente 20-25 residuos; o una pequeña extensión que facilita la purificación mediante el cambio de la carga neta u otra función, tal como un tracto de polihistidina, un epítipo antigénico o un dominio de dominio de unión, en otras palabras: cambios que no afectan significativamente al plegado y/o la actividad de la proteína.

[0061] Ejemplos de sustituciones conservadoras se incluyen en el grupo de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrofóbicos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano y tirosina), y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina).

[0062] Otros ejemplos de sustituciones conservadoras son las sustituciones de 20 aminoácidos estándar con aminoácidos no estándar (tal como 4-hidroxiprolina, 6-N-metilo lisina, ácido 2-aminoisobutírico, isovalina y serina de alfa-metilo).

Las sustituciones conservadoras también pueden incluir una sustitución en aminoácidos que no se codifican por el código genético y amino innatural. Los "aminoácidos innaturales" han sido modificados después de la síntesis de proteína y/o tienen una estructura química en su(s) cadena(s) lateral(es) diferente de la de los aminoácidos estándar.

Los aminoácidos no naturales pueden ser químicamente sintetizados y, preferiblemente, están disponibles comercialmente e incluyen ácido pipecólico, ácido carboxílico de tiazolidina, dehidroprolina, 3- y 4-metilprolina, y 3,3-dimetilprolina.

[0063] Alternativamente, los cambios aminoácidos son de tal naturaleza que las propiedades físico químicas de los polipéptidos se alteran.

Por ejemplo, los cambios aminoácidos pueden mejorar la termoestabilidad del polipéptido, alterar la especificidad del sustrato, cambiar el pH óptimo y similar.

[0064] Los aminoácidos esenciales en el polipéptido progenitor se pueden identificar según procedimientos conocidos en la técnica, tales como mutagénesis dirigida al sitio o mutagénesis de escaneado de alanina (Cunningham and Wells, 1989, Science 244: 1081-1085).

En la técnica anterior, las mutaciones de alanina únicas se introducen en cada residuo en la molécula y las moléculas mutantes resultantes se evalúan para la actividad biológica (es decir, actividad de fitasa) para identificar residuos de aminoácidos que son críticos para la actividad de la molécula.

Ver también, Hilton et al., 1996, J. Biol. Chem. 271: 4699-4708.

El sitio activo de la enzima u otra interacción biológica también se puede determinar por análisis físico de estructura, como se ha determinado por tales técnicas como resonancia magnética nuclear, cristalografía, difracción electrónica o marcaje por fotoafinidad, conjuntamente con la mutación de aminoácidos de sitio de contacto putativo.

Ver, por ejemplo, de Vos et al., 1992, Science 255: 306-312; Smith et al., 1992, J. Mol. Biol. 224: 899-904; Wlodaver et al., 1992, FEBS Lett. 309:59-64.

Las identidades de aminoácidos esenciales también pueden ser inferidas de análisis de identidades con polipéptidos que se relacionan con un polipéptido según la invención.

[0065] Las sustituciones de aminoácidos únicas o múltiples se pueden hacer y evaluar usando métodos conocidos de mutagénesis, recombinación y/o redistribución, seguidos de un procedimiento de selección pertinente, tal como los descritos por Reidhaar-Olson y Sauer, 1988, Science 241: 53-57; Bowie y Sauer, 1989, Proc. Natl. Acad. Sci. USA 86: 2152-2156; WO 95/17413; o WO 95/22625.

Otros métodos que se pueden usar incluyen PCR con tendencia al error, presentación en fagos (por ejemplo, Lowman et al., 1991, Biochem. 30:10832-10837; U.S. Patente n.º 5,223,409; WO 92/06204) y mutagénesis dirigida por región (Derbyshire et al., 1986, Gene 46:145; Ner et al., 1988, ADN 7:127).

[0066] Los métodos de mutagénesis/redistribución se pueden combinar con alto rendimiento, métodos de selección automatizados para detectar la actividad de clonados, polipéptidos mutagenizados expresados por células huésped.

Las moléculas de ADN mutagenizadas que codifican polipéptidos activos se pueden recuperar de las células huésped y secuenciar rápidamente usando métodos estándar en la técnica. Estos métodos permiten la determinación rápida de la importancia de residuos de aminoácidos individuales en un polipéptido de interés y se pueden aplicar a polipéptidos de estructura desconocida.

[0067] El número total de sustituciones de aminoácidos (preferiblemente, sustituciones conservadoras), deleciones y/o inserciones en la secuencia de aminoácidos 1 a 414 de SEQ ID n.º: 6 es como mucho 10, preferiblemente, como mucho 9, más preferiblemente, como mucho 8, más preferiblemente, como mucho 7, más preferiblemente, como mucho 6, más preferiblemente, como mucho 5, más preferiblemente, como mucho 4, aún más preferiblemente, como mucho 3, de la forma más preferible, como mucho 2 e incluso de la forma más preferible 1.

[0068] El número total de sustituciones de aminoácidos, deleciones y/o inserciones de aminoácidos 1 a 414 de SEQ ID n.º: 6 es 10, preferiblemente 9, más preferiblemente 8, más preferiblemente 7, más preferiblemente como mucho 6, más preferiblemente como mucho 5, más preferiblemente 4, aún más preferiblemente 3, de la forma más preferible 2 e incluso de la forma más preferible 1.

En la alternativa, el número total de sustituciones de aminoácidos (sustituciones preferiblemente conservadoras), deleciones y/o inserciones en la secuencia de aminoácidos 1 a 414 de SEQ ID n.º: 6, está a más 50, 45, 40, 35, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12 o como mucho 11.

5 [0069] El polipéptido de la invención puede ser una variante hipoalérgica, diseñada para recurrir una respuesta inmunológica reducida cuando se expone a animales, incluyendo el hombre.  
El término respuesta inmunológica debe ser entendido como cualquier reacción por el sistema inmunológico de un animal expuesto al polipéptido.

10 Un tipo de respuesta inmunológica es una respuesta alérgica que conduce a niveles aumentados de IgE en el animal expuesto.

Se pueden preparar variantes hipoalérgicas utilizando técnicas conocidas en la técnica. Por ejemplo, el polipéptido se puede conjugar con partes de protección de fracciones poliméricas o epítomos del polipéptido implicado en una respuesta inmunológica.

15 La conjugación con polímeros puede implicar el acoplamiento químico in vitro de polímero al polipéptido, por ejemplo, como se describe en la WO 96/17929; la WO98/30682; la WO98/35026 y/o la WO99/00489.

La conjugación puede implicar además o alternativamente a ella el acoplamiento in vivo de polímeros al polipéptido. Tal conjugación se puede conseguir por ingeniería genética de la secuencia de nucleótidos que codifica el polipéptido, secuencias de consenso de inserción que codifican sitios de glicosilación adicionales en el polipéptido y que expresan el polipéptido en un huésped capaz de la glicosilación del polipéptido, ver por ejemplo la WO00/26354.

20 Otra vía de la provisión de variantes hipoalérgicas es ingeniería genética de la secuencia de nucleótidos que codifica el polipéptido para provocar que el polipéptido se auto-oligomerice, que provoca que los monómeros de polipéptido puedan proteger los epítomos de otros monómeros de polipéptido y así reduce la antigenicidad de los oligómeros.

Tales productos y su preparación se describe por ejemplo en la WO96/16177.

25 Los epítomos implicados en una respuesta inmunológica se pueden identificar por varios métodos tales como el método de presentación en fagos descrito en la WO 00/26230 y la WO 01/83559, o el método aleatorio descrito en la EP 561907.

30 Una vez un epítomo haya sido identificado, su secuencia de aminoácidos se puede alterar para producir propiedades inmunológicas alteradas del polipéptido por técnicas de manipulación génica conocidas tales como mutagénesis dirigida de sitio (ver por ejemplo la WO 00/26230, la WO 00/26354 y/o la WO00/22103) y/o la conjugación de un polímero puede realizarse en suficiente proximidad al epítomo para que el polímero proteja el epítomo.

#### Fuentes de polipéptidos con actividad de fitasa

35 [0070] Un polipéptido de la presente invención se puede obtener de microorganismos de cualquier género. Para fines de la presente invención, el término "obtenido de" como se utiliza en este caso en relación con una fuente dada debe significar que el polipéptido codificado por una secuencia de nucleótidos se produce por la fuente o por una cepa donde la secuencia de nucleótidos ha sido insertada de la fuente.

40 En un aspecto preferido, el polipéptido obtenido de una fuente dada se secreta extracelularmente.

[0071] Un polipéptido de la presente invención puede ser un polipéptido bacteriano.

Por ejemplo, el polipéptido puede ser un polipéptido bacteriano gram positivo, tal como un polipéptido de *Bacillus* o un polipéptido de *Streptomyces*; o un polipéptido bacteriano gram negativo, por ejemplo, un *Escherichia coli*, *Yersinia*, *Klebsiella*, *Citrobacter* o un polipéptido de *Pseudomonas*.

45 En una forma de realización particular, el polipéptido se deriva de Proteobacterias, tales como Gammaproteobacteria, por ejemplo, Enterobacteriales, tales como Enterobacteriaceae.

[0072] En un aspecto particular, el polipéptido derivado de Enterobacteriaceae es un polipéptido de *Citrobacter*, tal como un *Citrobacter amalonaticus*, *Citrobacter braakii*, *Citrobacter farmeri*, *Citrobacter freundii*, *Citrobacter gillennii*, *Citrobacter intermedius*, *Citrobacter koseri*, *Citrobacter murliniae*, *Citrobacter rodentium*, *Citrobacter sedlakii*, *Citrobacter werkmanii*, *Citrobacter youngae* o polipéptido de especies de *Citrobacter*.

[0073] En una divulgación más preferida, el polipéptido es un polipéptido *Citrobacter gillennii* y, de la forma más preferible, un polipéptido *Citrobacter gillennii* DSM 13694, por ejemplo, el polipéptido de SEQ ID n.º: 4.

55 La cepa específica está disponible públicamente de la colección alemana Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM).

[0074] En un aspecto más preferido, el polipéptido es un *Citrobacter amalonaticus* polipéptido y, de la forma más preferible, un *Citrobacter amalonaticus* ATCC 25405 o un *Citrobacter amalonaticus* polipéptido de ATCC 25407, por ejemplo, el polipéptido de SEQ ID n.º: 6.

60 Estas cepas específicas están disponibles públicamente de la colección americana American Type Culture Collection (ATCC).

[0075] Un polipéptido de la presente invención también puede ser un polipéptido fúngico, tal como un polipéptido de levadura o un polipéptido fúngico filamentoso.

65

[0076] Las cepas de los microorganismos anteriores son accesibles fácilmente al público en un número de colecciones de cultivo, tal como la American Type Culture Collection (ATCC), la Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM), Centraalbureau Voor Schimmelcultures (CBS) y la Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

[0077] Además, tales polipéptidos se pueden identificar y obtener de otras fuentes incluyendo microorganismos de aislados de la naturaleza (por ejemplo, tierra, abonos, agua, etc.) utilizando las sondas anteriormente mencionadas. Las técnicas para el aislamiento de microorganismos de hábitats naturales se conocen en la técnica. Luego, el polinucleótido puede ser obtenido seleccionando de forma similar de una biblioteca genómica o genoteca de otro microorganismo.

Una vez se haya detectado una secuencia de polinucleótidos que codifica un polipéptido con la sonda(s), el polinucleótido se puede aislar o clonar por la utilización de técnicas que son bien conocidas por aquellos técnicos en la materia (ver, por ejemplo, Sambrook et al., 1989, supra).

[0078] Los polipéptidos de la presente invención también incluyen polipéptidos fusionados o polipéptidos divisibles de fusión donde otro polipéptido se fusiona en el N-terminal o el C-terminal del polipéptido o fragmento del mismo.

Un polipéptido fusionado se produce por la fusión de una secuencia de nucleótidos (o una porción de la misma) que codifica otro polipéptido a una secuencia de nucleótidos (o una porción de la misma) de la presente invención.

En la técnica se conocen técnicas para producir los polipéptidos de fusión e incluyen el ligamiento de las secuencias de codificación que codifican los polipéptidos, de modo que estas están en el marco y que la expresión del polipéptido fusionado está bajo control del mismo promotor(es) y terminador.

#### Polinucleótidos

[0079] La presente invención también se refiere a polinucleótidos aislados con una secuencia de nucleótidos que codifica un polipéptido que tiene actividad de fitasa y una estabilidad al ácido de al menos un 60 % de la actividad residual tras la incubación durante la noche a 37 °C en el tampón de ácido clorhídrico/glicina pH 2,2, en relación a la actividad residual tras la incubación durante la noche a 37 °C en el tampón HEPES pH 7,0" seleccionado del grupo que consiste en:

(a) un polinucleótido que codifica un polipéptido con una secuencia de aminoácidos que tiene al menos un 80 % de identidad con aminoácidos 1 a 414 de SEQ ID n.º: 6; y

(b) un polinucleótido con al menos un 80 % de identidad con nucleótidos 67 a 1308 de SEQ ID n.º: 5, donde el grado de identidad se determina por el programa "align" que es un alineamiento de Needleman-Wunsch (es decir, alineamiento global), utilizando la matriz de identidad por defecto, la penalización para el primer residuo en un espacio es -16 mientras las penalizaciones para residuos adicionales en un espacio son -4.

La presente invención también se refiere a polinucleótidos mutantes que comprenden al menos una mutación en la secuencia de codificación del polipéptido maduro de SEC ID n.º:5, donde la secuencia de nucleótidos mutante codifica un polipéptido que consiste en aminoácidos de 1 a 414 de SEQ ID n.º: 6.

[0080] Las técnicas usadas para aislar o clonar un polinucleótido que codifica un polipéptido se conocen en la técnica e incluyen el aislamiento de ADN genómico, la preparación de ADNc o una combinación de los mismos.

La clonación de los polinucleótidos de la presente invención de tal ADN genómico pueden ser efectuadas, por ejemplo, usando la reacción en cadena de polimerasa bien conocida (PCR) o selección de anticuerpo de bibliotecas de expresión para detectar fragmentos de ADN clonados con características estructurales compartidas.

Ver, por ejemplo, Innis et al., 1990, PCR: A Guide to Methods and Application, Academic Press, New York.

Se pueden utilizar otros procedimientos de amplificación de ácido nucleico, tales como reacción en cadena de ligasa (LCR), transcripción activada ligada (LAT) y amplificación basada en secuencias de nucleótidos (NASBA).

Los polinucleótidos se pueden clonar a partir de una cepa de Citrobacter u otro organismo o relacionado y así, por ejemplo, puede ser una variante alélica o de especies de la región que codifica el polipéptido de la secuencia de nucleótidos.

[0081] La presente descripción también se refiere a polinucleótidos con secuencias de nucleótidos que tienen un grado de identidad con la secuencia que codifica el polipéptido maduro de SEQ ID n.º: 5 (es decir, 67 a 1308) de al menos 75 % y que codifica un polipéptido con actividad de fitasa.

En formas de realización particulares, el grado de identidad es al menos 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 % o al menos 99 %.

[0082] La modificación de una secuencia de nucleótidos que codifica un polipéptido de la presente invención puede ser necesaria para la síntesis de los polipéptidos sustancialmente similar al polipéptido.

El término "sustancialmente similar" al polipéptido se refiere a formas del polipéptido que se producen de forma no natural.

Estos polipéptidos pueden diferir en alguna vía modificada genéticamente del polipéptido aislado de su fuente nativa, por ejemplo, variantes artificiales que difieren en la actividad específica, termoestabilidad, pH óptimo o similar.

La secuencia variante se puede construir basándose en la secuencia de nucleótidos presentada como la región que codifica el polipéptido de SEC ID n.º:5, por ejemplo, una subsecuencia de la misma y/o por introducción de

sustituciones de nucleótidos que no dan lugar a otra secuencia de aminoácidos del polipéptido codificado por la secuencia de nucleótidos, pero que corresponden al uso de codón del organismo huésped destinado a la producción del polipéptido o por introducción de sustituciones de nucleótidos que pueden dar lugar a una secuencia de aminoácidos diferente.

5 Para una descripción general de la sustitución de nucleótido, ver, por ejemplo, Ford et al., 1991, *Protein Expression and Purification* 2: 95-107.

[0083] Será aparente para los expertos en la técnica que tales sustituciones se pueden hacer fuera de las regiones críticas a la función de la molécula y todavía suponen un polipéptido activo.

10 Los residuos de aminoácidos esenciales para la actividad del polipéptido codificado por un polinucleótido aislado de la invención y, por lo tanto, preferiblemente no sometidos a sustitución, se puede identificar según procedimientos conocidos en la técnica, tales como mutagénesis dirigida al sitio o mutagénesis de escaneado de alanina (ver, por ejemplo, Cunningham and Wells, 1989, *Science* 244: 1081-1085).

15 En la técnica anterior, las mutaciones se introducen en cada residuo positivamente cargado en la molécula y las moléculas mutantes resultantes se evalúan de actividad de fitasa para identificar residuos de aminoácidos que son críticos a la actividad de la molécula.

20 Los sitios de interacción de sustrato-polipéptido también se pueden determinar por análisis de la estructura tridimensional como se ha determinado por tales técnicas como análisis de resonancia magnética nuclear, cristalografía o etiquetado de fotoafinidad (ver, por ejemplo, de Vos et al., 1992, *Science* 255: 306-312; Smith et al., 1992, *Journal of Molecular Biology* 224: 899-904; Wlodaver et al., 1992, *FEBS Letters* 309: 59-64).

#### Constructos de ácidos nucleicos

25 [0084] La presente invención también se refiere a constructos de ácidos nucleicos que incluyen un polinucleótido aislado de la presente invención operativamente enlazado a una o más secuencias de control que dirigen la expresión de la secuencia de codificación en una célula huésped adecuada bajo condiciones compatibles con las secuencias de control.

30 [0085] Un polinucleótido aislado que codifica un polipéptido de la presente invención se puede manipular en una variedad de vías de proveer la expresión del polipéptido.

La manipulación de la secuencia del polinucleótido antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión.

35 Las técnicas de modificación de las secuencias polinucleótidas utilizando métodos de ADN recombinantes se conocen bien en la técnica.

[0086] La secuencia de control puede ser una secuencia promotora apropiada, una secuencia de nucleótidos que se reconoce por una célula huésped para la expresión de un polinucleótido que codifica un polipéptido de la presente invención.

40 La secuencia promotora contiene secuencias de control transcripcionales que median la expresión del polipéptido.

El promotor puede ser cualquier secuencia de nucleótidos que muestra una actividad transcripcional en la célula huésped de elección incluyendo promotores mutantes, truncados e híbridos, y se puede obtener de genes que codifican los polipéptidos extracelulares o intracelulares bien homólogos o heterólogos a la célula huésped.

45 [0087] Ejemplos de promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente invención, especialmente, en una célula huésped bacteriana son los promotores obtenidos del operón lac de *E. Coli*, gen de agarasa de *Streptomyces coelicolor* (dagA), gen de levansucrasa de *Bacillus subtilis* (sacB), gen de alfa-amilasa de *Bacillus licheniformis* (amilo), el gen de amilasa maltogénica de *Bacillus stearothermophilus* (amyM), gen de *Bacillus amyloliquefaciens* alfa-amilasa (amyQ), gen de penicilinas de *Bacillus licheniformis* (penP), genes xylA y xylB de *Bacillus subtilis* y gen de beta-lactamasa procariótica (Villa-Kamaroff et al., 1978, *Proceedings of the National Academy of Sciences USA* 75: 3727-3731), al igual que el promotor tac (DeBoer et al., 1983, *Proceedings of the National Academy of Sciences USA* 80: 21-25).

50 Otros promotores se describen en "Useful proteins from recombinant bacteria" en *Scientific American*, 1980, 242: 74-94; y en Sambrook et al., 1989, *supra*.

55 [0088] Los ejemplos de promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente invención en una célula huésped fúngica filamentosa son promotores obtenidos de los genes para TAKA amilasa de *Aspergillus oryzae*, proteinasa aspártica de *Rhizomucor miehei*, alfa-amilasa neutra de *Aspergillus niger*, alfa-amilasa estable en ácido de *Aspergillus niger*, glucoamilasa de *Aspergillus niger* o *Aspergillus awamori* (glaA), lipasa de *Rhizomucor miehei*, proteasa alcalina de *Aspergillus oryzae*, triosafosfato isomerasa de *Aspergillus oryzae*, acetamidasa de *Aspergillus nidulans*, amiloglucosidasa de *Fusarium venenatum* (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), proteasa de tipo tripsina de *Fusarium oxysporum* (WO 96/00787), beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, endoglucanasa de *Trichoderma reesei* I, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa de *Trichoderma reesei* IV, endoglucanasa de *Trichoderma reesei* V, xilanasa de *Trichoderma reesei* I, xilanasa de *Trichoderma reesei* II, beta-xilosidasa de *Trichoderma reesei*, al igual que el promotor NA2-tpi (un híbrido de los promotores de los genes para alfa-amilasa neutra de *Aspergillus niger* y triosafosfato isomerasa de

*Aspergillus oryzae*); y promotores mutantes, truncados e híbridos de los mismos.

5 [0089] En un huésped de levadura, los promotores útiles se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), galactocinasa de *Saccharomyces cerevisiae* (GAL1), deshidrogenasa de dehidrogenasa/gliceraldehído-3-fosfato de alcohol de *Saccharomyces cerevisiae* (ADH1, aDH2/GAP) triosafosfato isomerasa de *Saccharomyces cerevisiae*, (TPI) metalotionina de *Saccharomyces cerevisiae* (CUP1), quinasa de 3-fosfoglicerato de *Saccharomyces cerevisiae* y alcohol-oxidasa (AOX1) de *Pichia pastoris*. Otros promotores útiles para células huésped de levadura se describen en Romanos et al., 1992, levadura 8: 423-488.

10 [0090] La secuencia de control también puede ser una secuencia terminadora de transcripción adecuada, una secuencia reconocida por una célula huésped para terminar la transcripción. La secuencia terminadora está operativamente enlazada al 3' terminal de la secuencia de nucleótidos que codifica el polipéptido.  
15 Cualquier terminador que es funcional en la célula huésped de elección se puede utilizar en la presente invención.

[0091] Los terminadores preferidos para células huésped fúngicas filamentosas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae*, glucoamilasa *Aspergillus niger*, antranilato sintasa de *Aspergillus nidulans*, alfa-glucosidasa de *Aspergillus niger* y proteasa de tipo tripsina de *Fusarium oxysporum*.

20 [0092] Los terminadores preferidos para células huésped de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae*, citocromo de *Saccharomyces cerevisiae* C (CYC1) y gliceraldehído-3-fosfato-deshidrogenasa de *Saccharomyces cerevisiae*. Otros terminadores útiles para células huésped de levadura se describen por Romanos et al., 1992, supra.

25 [0093] La secuencia de control también puede ser una secuencia líder adecuada, una región no traducida de un ARNm que es importante para la traducción por parte de la célula huésped. La secuencia líder está operativamente enlazada al 5' terminal de la secuencia de nucleótidos que codifica el polipéptido.  
30 Cualquier secuencia líder que sea funcional en la célula huésped de elección se puede utilizar en la presente invención.

[0094] La secuencia de control también puede ser una secuencia líder adecuada, una región no traducida de un ARNm que es importante para la traducción por parte de la célula huésped.  
35 La secuencia líder está operativamente enlazada al 5' terminal de la secuencia de nucleótidos que codifica el polipéptido. Cualquier secuencia líder que sea funcional en la célula huésped de elección se puede utilizar en la presente invención.

40 [0095] Los líderes preferidos para células huésped fúngicas filamentosas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae* y triosafosfato isomerasa de *Aspergillus nidulans*.

45 [0096] Los líderes adecuados para células huésped de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), quinasa de 3-fosfoglicerato de *Saccharomyces cerevisiae*, alfa-factor de *Saccharomyces cerevisiae* y deshidrogenasa de dehidrogenasa/gliceraldehído-3-fosfato de alcohol de *Saccharomyces cerevisiae* (ADH2/GAP).

[0097] La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia operativamente enlazada al 3' terminal de la secuencia de nucleótidos y que, cuando se transcribe, se reconoce por la célula huésped como una señal para añadir residuos de poliadenosina al ARNm transcrito.  
50 Cualquier secuencia de poliadenilación que sea funcional en la célula huésped de elección se puede utilizar en la presente invención.

55 [0098] Las secuencias de poliadenilación preferidas para células huésped fúngicas filamentosas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae*, glucoamilasa de *Aspergillus niger*, antranilato sintasa de *Aspergillus nidulans*, proteasa de tipo tripsina de *Fusarium oxysporum* y alfa-glucosidasa de *Aspergillus niger*.

[0099] Secuencias de poliadenilación útiles para células huésped de levadura se describen en Guo y Sherman, 1995, *Molecular Cellular Biology* 15: 5983-5990.

60 [0100] La secuencia de control también puede ser una región codificante del péptido señal que codifica para una secuencia de aminoácidos enlazada al amino terminal de un polipéptido y dirige el polipéptido codificado en la vía secretora de la célula.  
65 El 5' extremo de la secuencia codificante de la secuencia de nucleótidos puede contener intrínsecamente una región codificante del péptido señal naturalmente enlazado en el marco de lectura de traducción con el segmento de la región de codificación que codifica el polipéptido segregado.

Alternativamente, el 5' extremo de la secuencia codificante puede contener una región codificante del péptido señal que sea extranjera a la secuencia codificante.

La región codificante del péptido señal foráneo se puede requerir donde la secuencia codificante naturalmente no contiene una región codificante del péptido señal.

5 Alternativamente, la región codificante del péptido señal foráneo sencillamente puede reemplazar la región que codifica el péptido señal natural para mejorar la secreción del polipéptido.

Sin embargo, en la presente invención se puede utilizar cualquier región que codifique el péptido señal que dirige el polipéptido expresado en la vía secretora de una célula huésped de elección.

10 [0101] Las regiones codificadas del péptido señal eficaces para células huésped bacterianas son las regiones codificadoras del péptido señal obtenidas de los genes para amilasa maltogénica de *Bacillus* NCIB 11837, alfa-amilasa de *Bacillus stearothermophilus*, subtilisina de *Bacillus licheniformis*, beta-lactamasa de *Bacillus licheniformis*, proteasas neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM) y *Bacillus subtilis* prsA.

Otros péptidos señal se describen en Simonen y Palva, 1993, *Microbiological Reviews* 57: 109-137.

15 [0102] Las regiones codificadoras del péptido señal eficaces para células huésped fúngicas filamentosas son las regiones codificadoras del péptido señal obtenidas de los genes para TAKA amilasa de *Aspergillus oryzae*, amilasa neutra de *Aspergillus niger*, glucoamilasa *Aspergillus niger*, proteinasa aspártica de *Rhizomucor miehei*, celulasa de *Humicola insolens* y lipasa *Humicola lanuginosa*.

20 [0103] Los péptidos señal útiles para células huésped de levadura se obtienen de los genes para alfa-factor de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*.

Otras regiones codificadoras del péptido señal útiles se describen en Romanos et al., 1992, supra y Xiong et al in *Journal of Applied Microbiology* 2005, 98, 418-428.

25 [0104] En un aspecto preferido, la región que codifica el péptido señal es nucleótidos 1 a 66 de SEC ID n.º 1, 3 o 5, que codifican aminoácidos 1 a 22 de SEC ID n.º 2, 4 y 6, respectivamente.

30 [0105] La secuencia de control también puede ser una región codificante de propéptido que codifica una secuencia de aminoácidos posicionada en el amino terminal de un polipéptido.

El polipéptido resultante se conoce como un propolipéptido o propolipéptido (o un zimógeno en algunos casos).

Un propolipéptido es inactivo generalmente y se puede convertir en un polipéptido activo maduro por escisión catalítica o autocatalítica del propéptido del propolipéptido.

35 La región codificante del propéptido se puede obtener de los genes para proteasa alcalina de *Bacillus subtilis* (aprE), proteasa neutra de *Bacillus subtilis* (nprT), alfa-factor de *Saccharomyces cerevisiae*, proteinasa aspártica de *Rhizomucor miehei* y lacasa *myceliophthora thermophila* (WO 95/33836).

40 [0106] Donde ambos péptidos señal y regiones de propéptido están presentes en el amino terminal de un polipéptido, la región de propéptido está situada junto al amino terminal de un polipéptido y la región de péptido señal está situada junto al amino terminal de la región de propéptido.

[0107] También puede ser deseable añadir secuencias reguladoras que permitan la regulación de la expresión del polipéptido en relación al crecimiento de la célula huésped.

45 Los ejemplos de sistemas reguladores son los que causan que la expresión del gen se active o desactive en respuesta a una sustancia química o estímulo físico con la presencia de un compuesto regulador.

Los sistemas reguladores en sistemas procarióticos incluyen los sistemas operadores lac, tac y trp.

En la levadura, se puede utilizar el sistema ADH2 o sistema GAL1.

En hongos filamentosos, el promotor de TAKA alfa-amilasa, promotor de glucoamilasa *Aspergillus niger* y promotor de glucoamilasa de *Aspergillus oryzae* se pueden utilizar como secuencias reguladoras.

50 Otros ejemplos de secuencias reguladoras son aquellos que permiten la amplificación génica.

#### Vectores de expresión

55 [0108] La presente invención también se refiere a vectores de expresión recombinantes que comprenden un polinucleótido de la presente invención, un promotor y señales de parada transcripcionales y traduccionales.

Los varios ácidos nucleicos y secuencias de control anteriormente descritas se pueden unir entre sí para producir un vector de expresión recombinante que pueda incluir uno o más sitios de restricción convenientes para permitir la inserción o sustitución de la secuencia de nucleótidos que codifican el polipéptido a tales sitios.

60 Alternativamente, una secuencia de nucleótidos de la presente invención se puede expresar por la inserción de la secuencia de nucleótidos o un constructo de ácidos nucleicos que comprende la secuencia en un vector apropiado para la expresión.

En la creación del vector de expresión, la secuencia codificante se localiza en el vector, de modo que la secuencia codificadora se enlazada operativamente con las secuencias de control apropiadas para la expresión.

65 [0109] El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido o virus) que puede ser convenientemente sometido a procedimientos de ADN recombinante y puede provocar la expresión de la

secuencia de nucleótidos.

La elección del vector típicamente dependerá de la compatibilidad del vector con la célula huésped en la que el vector será introducido.

Los vectores puede ser plásmidos circulares cerrados o lineales.

5 [0110] El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, la replicación de la cual es independiente de replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un minicromosoma o un cromosoma artificial.

El vector puede contener cualquier medio para asegurar la autorreplicación.

10 Alternativamente, el vector puede ser uno que, cuando se introduce en la célula huésped, se integra en el genoma y se replica con el cromosoma(s) en que se ha integrado.

Además, se puede utilizar un vector único o plásmido o dos o más vectores o plásmidos que juntos contienen el ADN total que se va a introducir en el genoma de la célula huésped o un transposón.

15 [0111] Un gen condicionalmente esencial puede funcionar como una etiqueta seleccionable no antibiótica.

Los ejemplos no limitativos de marcadores seleccionables no antibióticos bacterianos condicionalmente esenciales son los genes *dal* de *Bacillus subtilis*, *Bacillus licheniformis* u otros bacilos, que son solo esenciales cuando la bacteria se cultiva en ausencia de alanina-D.

20 También los genes que codifican enzimas implicadas en la producción de UDP-galactosa pueden funcionar como marcadores condicionalmente esenciales en una célula cuando la célula se cultiva en presencia de galactosa o crece en un medio que da lugar a la altura de la presencia de galactosa.

Los ejemplos no limitativos de tales genes son aquellos de *B. subtilis* o *B. licheniformis* que codifican fosforilasa dependiente de UTP (EC 2.7.7.10), uridililtransferasa dependiente de glucosa-UDP (EC 2.7.7.12) o epimerasa UDP-galactosa (EC 5.1.3.2).

25 También un gen de xilosa isomerasa tal como *xylA*, de *Bacilli* se puede usar como marcadores seleccionables en células que crecen en el medio mínimo con xilosa como única fuente de carbono.

Los genes necesarios para la utilización de gluconato, *gntK* y *gntP* también pueden usarse como marcadores seleccionables en células que crecen en el medio mínimo con gluconato como única fuente de carbono.

30 Otros ejemplos de genes condicionalmente esenciales se conocen en la técnica. Los marcadores seleccionables antibióticos confieren resistencia antibiótica a tales antibióticos como ampicilina, canamicina, cloranfenicol, eritromicina, tetraciclina, neomicina, higromicina o metotrexato.

[0112] Los marcadores adecuados para células huésped de levadura son *ADE2*, *HIS3*, *LEU2*, *LYS2*, *MET3*, *TRP1* y *URA3*.

35 Los marcadores seleccionables para usar en una célula huésped fúngica filamentosa incluyen, pero de forma no limitativa, *amdS* (acetamidasa), *argB* (ornitina-carbamoiltransferasa), *barra* (fosfonitrilina acetiltransferasa), *hph* (higromicina fosfotransferasa), *niaD* (nitrato-reductasa), *pyrG* (orotidina-5'-fosfato Descarboxilasa), *sC* (adeniltransferasa de sulfato) y *trpC* (antranilato sintasa), al igual que equivalentes de los mismos.

40 Para su uso en una célula de *Aspergillus* son los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces hygroscopicus*.

[0113] Los vectores de la presente invención preferiblemente contienen uno o más marcadores seleccionables que permiten una selección fácil de células transformadas.

45 Una etiqueta seleccionable es un gen el producto del cual proporciona resistencia vírica o biocida, resistencia a metales pesados, prototrofia a auxótrofos y similar.

[0114] Los ejemplos de marcadores seleccionables bacterianos son los genes *dal* de *Bacillus subtilis* o *Bacillus licheniformis*, o marcadores que confieren resistencia antibiótica, tal como ampicilina, canamicina, cloranfenicol o resistencia de tetraciclina.

50 [0115] Los vectores de la presente invención preferiblemente contienen un elemento(s) que permite la integración del vector en el genoma de la célula huésped o replicación autónoma del vector en la célula independiente del genoma.

55 [0116] Para la integración en el genoma de la célula huésped, el vector puede depender de la secuencia de polinucleótido que codifica el polipéptido o cualquier otro elemento del vector para la integración en el genoma por recombinación homóloga o no homóloga.

Alternativamente, el vector puede contener secuencias de nucleótidos adicionales para dirigir la integración por recombinación homóloga en el genoma de la célula huésped a una ubicación(es) precisa en el cromosoma(s).

60 Para aumentar la probabilidad de integración a una ubicación precisa, los elementos integracionales deberían contener preferiblemente un número suficiente de ácidos nucleicos, tal como 100 a 10,000 pares de bases, preferiblemente, 400 a 10,000 pares de bases y de la forma más preferible 800 a 10,000 pares de bases, que tienen un alto grado de identidad con la secuencia diana correspondiente para mejorar la probabilidad de recombinación homóloga.

65 Los elementos integracionales pueden ser cualquier secuencia que es homóloga a la secuencia diana en el genoma de la célula huésped.

Además, los elementos integracionales pueden ser secuencias de nucleótidos de codificación o no codificación. Por otro lado, el vector se puede integrar en el genoma de la célula huésped por recombinación no-homóloga.

5 [0117] Para replicación autónoma, el vector puede comprender además un origen de replicación que permita replicar al vector autónomamente en la célula huésped en cuestión.  
El origen de replicación puede ser cualquier replicador de plásmido que medie la replicación autónoma que funciona en la célula.  
El término "origen de la replicación" o "replicador plásmido" se define aquí como una secuencia de nucleótidos que permite que un plásmido o vector se replique in vivo.

10 [0118] Los ejemplos de orígenes bacterianos de replicación son los orígenes de replicación de plásmidos pBR322; pUC19; pACYC177 y pACYC184 que permiten la replicación en el E. Coli y pUB110; pE194; pTA1060 y pAMβ1 que permite la replicación en el Bacillus.

15 [0119] Los ejemplos de orígenes de replicación para su uso en una célula huésped de levadura son los orígenes de replicación de 2 micrones, ARS1, ARS4, la combinación de ARS1 y CEN3 y la combinación de ARS4 y CEN6.

[0120] Los ejemplos de orígenes de replicación útiles en una célula fúngica filamentosa son AMA1 y ANS1 (Gems et al., 1991, Gene 98:61-67; Cullen et al., 1987, Nucleic Acids Research 15: 9163-9175; WO 00/24883).  
20 El aislamiento del gen AMA1 y construcción de plásmidos o vectores que comprende el gen, se pueden realizar según los métodos descritos en la WO 00/24883.

[0121] Más de una copia de un polinucleótido de la presente invención se puede insertar en la célula huésped para aumentar la producción del producto génico.

25 Se puede obtener un aumento en el número de copias del polinucleótido por la integración de al menos una copia adicional de la secuencia en el genoma de la célula huésped o incluyendo un gen marcador seleccionable amplificable con el polinucleótido donde las células que contienen copias amplificadas del gen marcador seleccionable y, por lo tanto, las copias adicionales del polinucleótido, se pueden seleccionar por el cultivo de las células en presencia del agente seleccionable apropiado.

30 [0122] Los procedimientos usados para enlazar los elementos anteriormente descritos para construir los vectores de expresión recombinantes de la presente invención son conocidos para el experto en la materia (ver, por ejemplo, Sambrook et al., 1989, supra).

35 Células huésped

[0123] La presente invención también se refiere a células huésped recombinantes, que comprenden un polinucleótido de la presente invención, que son ventajosamente usadas en la producción recombinante de los polipéptidos.

40 Un vector que comprende un polinucleótido de la presente invención se introduce en una célula huésped, de modo que el vector se mantiene como un integrante cromosómico o como un vector extracromosomal que se autoreplica como se ha descrito anteriormente.

El término "célula huésped" abarca cualquier progenie de una célula madre que no es idéntica a la célula madre debido a mutaciones que se producen durante la replicación.

45 La elección de una célula huésped dependerá, en gran parte, del gen que codifica el polipéptido y su fuente.

[0124] La célula huésped puede ser un microorganismo unicelular, por ejemplo, un procarionta o un microorganismo no unicelular, por ejemplo, un eucariota.

50 [0125] Los microorganismos unicelulares útiles son células bacterianas tales como bacterias gram positivas que incluyen, pero de forma no limitativa, una célula de Bacillus, por ejemplo, Bacillus alkalophilus, Bacillus amyloliquefaciens, Bacillus brevis, Bacillus circulans, Bacillus clausii, Bacillus coagulans, Bacillus lautus, Bacillus lentus, Bacillus licheniformis, Bacillus megaterium, Bacillus stearothermophilus, Bacillus subtilis y Bacillus thuringiensis; o una célula de Streptomyces, por ejemplo, Streptomyces lividans y Streptomyces murinus o bacterias gram negativas, tales como E. Coli y Pseudomonas sp.

55 En un aspecto preferido, la célula huésped bacteriana es un Bacillus lentus, Bacillus licheniformis, Bacillus stearothermophilus o célula de Bacillus subtilis.

En otro aspecto preferido, la célula de Bacillus es un Bacillus alcalófilico.

60 [0126] La introducción de un vector en una célula huésped bacteriana, por ejemplo, se puede efectuar por transformación de protoplasto (ver, por ejemplo, Chang and Cohen, 1979, Molecular General Genetics 168: 111-115), que usa células competentes (ver, por ejemplo, Young and Spizizin, 1961, Journal of Bacteriology 81: 823-829 o Dubnau y Davidoff-Abelson, 1971, Journal of Molecular Biology 56: 209-221), electroporación (ver, por ejemplo, Shigekawa y Dower, 1988 Biotechniques 6: 742-751) o conjugación (ver, por ejemplo, Koehler and Thorne, 1987, Journal of Bacteriology 169: 5771-5278).

65



[0127] La célula huésped también puede ser un eucariota, tal como un mamífero, insecto, planta o célula fúngica.

[0128] En un aspecto preferido, la célula huésped es una célula fúngica. "Hongos" como se utiliza en este caso incluye el filo Ascomycota, Basidiomycota, Chytridiomycota y Zygomycota (como definido por Hawkswort et al., en, Ainswort and Bisby's Dictionary of The Fungi, 8th edición, 1995, CAB internacional. University Press, Cambridge, UK), al igual que la Oomycota (como se cita en el Hawkswort et al., 1995, supra, page 171) y todos los hongos mitospóricos (Hawkswort et al., 1995, supra).

[0129] En un aspecto más preferido, la célula huésped fúngica es una célula de levadura. "Levadura" como se utiliza en este caso, incluye levadura ascoesporógena (Endomicetales), levadura basidioesporogénea y levadura de los hongos imperfectos (Blastomicetos).

Ya que la clasificación de levadura puede cambiar en el futuro, para los fines de esta invención, la levadura se debe definir como se describe en Biology and Activities of Yeast (Skinner, F.A., Passmore, S.M., y Davenport, R.R., eds, Soc. App. Bacteriol. Symposium Series No. 9,1980).

[0130] En un aspecto aún más preferido, la célula huésped de levadura es una Candida, Hansenula, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces o célula de Yarrowia.

[0131] En un aspecto más preferido, la célula huésped de levadura es una Pichia pastoris, Pichia methanolica, Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, célula de Saccharomyces norbensis o Saccharomyces oviformis.

En otro aspecto más preferido, la célula huésped de levadura es una célula de Kluyveromyces lactis.

En otro aspecto más preferido, la célula huésped de levadura es una célula de Yarrowia lipolytica.

[0132] En otro aspecto más preferido, la célula huésped fúngica es una célula fúngica filamentosa. Los "hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión Eumycota y Oomycota (como se define por Hawkswort et al., 1995, supra).

Los hongos filamentosos están generalmente caracterizados por el hecho de que una pared micelial compuesta por quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos.

El crecimiento vegetativo es por alargamiento hifal y el catabolismo de carbono es estrictamente aeróbico.

En cambio, el crecimiento vegetativo por levaduras, tal como Saccharomyces cerevisiae es por el injerto de un talo unicelular y el catabolismo de carbono puede ser fermentativo.

[0133] En un aspecto aún más preferido, la célula huésped fúngica filamentosa es un Acremonium, Aspergillus, Aureobasidium, Bjerkandera, Ceriporiopsis, Coprinus, Coriolus, Cryptococcus, Filobasidium, Fusarium, Humicola, Magnaporthe, Mucor, Myceliophthora, Neocallimastix, Neurospora, Paecilomices, Penicillium, Phanerochaete, Phlebia, Piromices, Pleurotus, Schizophium, Talaromyces, Termoascus, Thielavia, Tolipocladium, Trametes o célula de Trichoderma.

[0134] En un aspecto más preferido, la célula huésped fúngica filamentosa es un Aspergillus awamori, Aspergillus fumigatus, Aspergillus foetidus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger o célula de Aspergillus oryzae.

En otro aspecto más preferido, la célula huésped fúngica filamentosa es una Fusarium bactridioides, Fusarium cerealis, Fusarium crookwellense, Fusarium culmorum, Fusarium graminearum, Fusarium graminum, Fusarium heterosporum, Fusarium negundi, Fusarium oxysporum, Fusarium reticulatum, Fusarium roseum, Fusarium sambucinum, Fusarium sarcochromum, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides o célula de Fusarium venenatum.

En otro aspecto más preferido, la célula huésped fúngica filamentosa es una Bjerkandera adusta, Ceriporiopsis aneirina, Ceriporiopsis aneirina, Ceriporiopsis caregiea, Ceriporiopsis gilvescens, Ceriporiopsis pannocinta, Ceriporiopsis rivulosa, Ceriporiopsis subrufa o Ceriporiopsis subvermisporea, Coprinus cinereus, Coriolus hirsutus, Humicola insolens, Humicola lanuginosa, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Penicillium purpurogenum, Phanerochaete chrysosporium, Phlebia radiata, Pleurotus eryngii, Thielavia terrestris, Trametes villosa, Trametes versicolor, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei o célula de cepa de Trichoderma viride.

[0135] Las células fúngicas se pueden transformar por un proceso que implica la formación de protoplasto, la transformación de los protoplastos y la regeneración de la pared celular en cierto modo conocida per se.

Los procedimientos adecuados para la transformación de Aspergillus y células huésped de trichoderma se describen en EP 238 023 y Yelton et al., 1984, Proceedings of the National Academy of Sciences USA 81: 1470-1474.

Los métodos adecuados para las especies de fusarium de transformación se describen en Malardier et al., 1989, Gene 78: 147-156 y la WO 96/00787.

La levadura se puede transformar utilizando los procedimientos descritos en Becker and Guarente, In Abelson, J.N. and Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York; Ito et al., 1983, Journal of Bacteriology 153: 163; and Hinnen et al., 1978, Proceedings of the National Academy of Sciences USA 75: 1920.

## Métodos de producción

[0136] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención, que comprende (a) el cultivo de una célula, que en su forma tipo salvaje es capaz de producir el polipéptido, bajo condiciones conductivas para la producción del polipéptido; y (b) recuperación del polipéptido. Preferiblemente, la célula es del género *Citrobacter* y, más preferiblemente, *Citrobacter gillenii* o *Citrobacter amalonaticus*.

[0137] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención, que comprende (a) el cultivo de una célula huésped recombinante que comprende un constructo de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica el polipéptido bajo condiciones propicias para la producción del polipéptido; y (b) la recuperación del polipéptido.

[0138] La presente descripción también se refiere a métodos para producir un polipéptido de la presente invención, que comprende (a) el cultivo de una célula huésped bajo condiciones propicias para la producción del polipéptido, donde la célula huésped comprende una secuencia de nucleótidos mutante que tiene al menos una mutación en la región de codificación de polipéptido maduro de SEC ID n.º 1, 3 o 5, donde la secuencia de nucleótidos mutante codifica un polipéptido que consiste en aminoácidos 1 a 409 de SEQ ID n.º: 2 o 4, o consiste en los aminoácidos 1 a 414 de SEQ ID n.º: 6, respectivamente y (b) la recuperación del polipéptido.

[0139] En los métodos de producción de la presente invención, las células se cultivan en un medio nutritivo adecuado para la producción del polipéptido usando métodos bien conocidos en la técnica. Por ejemplo, la célula se puede cultivar por cultivo en matraz de agitación y fermentación a pequeña escala o a gran escala (incluyendo, fermentaciones continuas, de lote, de lote alimentado o de estado sólido) en el laboratorio o fermentadores industriales realizadas en un medio adecuado y bajo condiciones que permiten al polipéptido que sea expresado y/o aislado.

El cultivo tiene lugar en un medio nutritivo adecuado que comprende fuentes de nitrógeno y carbono, y sales inorgánicas usando procedimientos conocidos en la técnica. Los medios adecuados están disponibles de proveedores comerciales o se pueden preparar según composiciones publicadas (por ejemplo, en catálogos de la American Type Culture Collection).

Si el polipéptido se segrega en el medio nutritivo, el polipéptido se puede recuperar directamente del medio. Si el polipéptido no se segrega, este se puede recuperar de lisatos de célula.

[0140] Los polipéptidos se pueden detectar utilizando métodos conocidos en la técnica que son específicos para los polipéptidos.

Estos métodos de detección pueden incluir el uso de anticuerpos específicos, formación de un producto de polipéptido o desaparición de un sustrato de polipéptido.

Por ejemplo, un ensayo de polipéptido se puede utilizar para determinar la actividad del polipéptido como se describe en este caso.

[0141] El polipéptido resultante se puede recuperar utilizando métodos conocidos en la técnica. Por ejemplo, el polipéptido se puede recuperar del medio nutritivo por procedimientos convencionales incluyendo, pero de forma no limitativa, centrifugación, filtración, extracción, secado por pulverización, evaporación o precipitación.

[0142] Los polipéptidos de la presente invención se pueden purificar por una variedad de procedimientos conocidos en la técnica incluyendo, pero de forma no limitativa, cromatografía (por ejemplo, intercambio iónico, afinidad, hidrofóbico cromatoenfoco y exclusión de tamaño), procedimientos electroforéticos (por ejemplo, isoelectroenfoco preparativo), solubilidad diferencial (por ejemplo, precipitación de sulfato amónico), SDS-PAGE o extracción (ver, por ejemplo, *Protein Purification*, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989).

Un esquema de purificación típica puede incluir la centrifugación, filtración de germen, precipitación de sulfato de amonio (usando por ejemplo un 80 % de saturación de sulfato de amonio), centrifugación, resuspensión de gránulos en el tampón A (por ejemplo 50 mM decacetato sódico, 1,5 M de sulfato de amonio pH 4,5), filtración, cromatografía de interacción hidrofóbica (Toyopearl de fenilo, cargando con tampón A, elución con tampón B (50 mM de acetato sódico pH 4,5)) y cromatografía de intercambio de cationes (SP-sefarosa, que carga con 10 mM de citrato sódico pH 4,0, elución con un gradiente de sal lineal (10 mM de citrato sódico pH 4,0 + 1 M NaCl)).

## Plantas transgénicas

[0143] La presente invención también se refiere a una planta transgénica, parte de planta o célula vegetal que ha sido transformada con una secuencia de nucleótidos que codifica un polipéptido que tiene actividad de fitasa de la presente invención, para expresar y producir el polipéptido en cantidades recuperables.

El polipéptido se puede recuperar de la planta o parte de planta.

Alternativamente, la planta o parte de planta con el polipéptido recombinante se puede utilizar como tal para mejorar la calidad de un alimento o pienso, por ejemplo, mejora del valor nutricional, palatabilidad y propiedades reológicas o destruir un factor antinutritivo.

[0144] En una forma de realización particular, el polipéptido está dirigido a las vacuolas de almacenamiento de endospermo en semillas.

Este se puede obtener sintetizándolo como un precursor con un péptido señal adecuado, ver Horvath et al en PNAS, Feb. 15, 2000, vol. 97, n.º 4, p. 1914-1919.

[0145] La planta transgénica puede ser dicotiledónea (una dicotiledónea) o monocotiledónea (una monocotiledónea) o variantes modificadas genéticamente de la misma.

Ejemplos de plantas monocotiledóneas son hierbas, tales como grama de prados (pasto azul, *Poa*), hierba forrajera tal como *Festuca*, *Lolium*, césped templado, tal como *Agrostis* y cereales, por ejemplo, trigo, avena, centeno, cebada, arroz, sorgo, triticale (híbrido estabilizado de trigo (*Triticum*) y centeno (*Secale*), y maíz (mazorca).

Ejemplos de plantas dicotiledóneas son tabaco, leguminosas, tales como girasol (helianto), algodón (*Gossypium*), altramuces, patata, remolacha azucarera, guisante, judía y soja, y plantas crucíferas (de familia *Brassicaceae*), tales como coliflor, semilla de colza y el organismo de modelo próximo relacionado *Arabidopsis thaliana*.

Plantas de bajo fitato como se describe por ejemplo en la patente US n.º 5,689,054 y patente US n.º 6,111,168 son ejemplos de plantas modificadas genéticamente.

[0146] Ejemplos de partes de planta son tallo, callo, hojas, raíz, frutas, semillas y tubérculos, al igual que los tejidos individuales que comprenden estas partes, por ejemplo epidermis, mesofilo, parénquima, tejidos vasculares, meristemos.

También los compartimentos de célula vegetal específica, tal como cloroplasto, apoplasto, mitocondria, vacuola, peroxisomas y citoplasma se consideran que son una parte de planta.

Además, cualquier célula vegetal, sea cual sea el origen del tejido, se considera que es una parte de planta.

Asimismo, partes de planta tales como tejidos específicos y células aisladas para facilitar la utilización de la invención están también consideradas partes de planta, por ejemplo, embriones, endospermas, aleurona y revestimientos de semilla.

[0147] También incluidas dentro del campo de la presente invención están la progenie de tales plantas, partes de planta y células vegetales.

[0148] La planta transgénica o célula vegetal que expresa un polipéptido de la presente invención se puede construir conforme a métodos conocidos en la técnica. Brevemente, la planta o célula vegetal se construye por la incorporación de uno o más constructos de expresión que codifican un polipéptido de la presente invención en el genoma huésped de planta y propagando la planta modificada resultante o célula vegetal en una planta transgénica o célula vegetal.

[0149] Convenientemente, el constructo de expresión es un constructo de ácido nucleico que comprende una secuencia de ácidos nucleicos que codifica un polipéptido de la presente invención operativamente vinculado a secuencias reguladoras apropiadas requeridas para la expresión de la secuencia de ácidos nucleicos en la planta o parte de planta elegida.

Además, el constructo de expresión puede comprender una etiqueta seleccionable útil para la identificación de células huésped en las que el constructo de expresión ha sido integrado y las secuencias de ADN necesarias para introducción del constructo en la planta en cuestión (el último depende del método de introducción de ADN para ser usado).

[0150] La elección de secuencias reguladoras, tales como promotor y secuencias terminadoras y, opcionalmente, señal o secuencias de tránsito se determinan, por ejemplo, basándose en cuándo, dónde y cómo se desea que se exprese el polipéptido.

Por ejemplo, la expresión del gen que codifica un polipéptido de la presente invención se puede constituir o inducir, o puede ser de desarrollo, fase o tejido específico, y el producto génico puede estar dirigido a un compartimento de célula específica, tejido o parte de planta, tal como semillas u hojas.

Las secuencias reguladoras están, por ejemplo, descritas en Tague et al., 1988, *Plant Physiology* 86: 506.

[0151] Para la expresión constitutiva, se pueden utilizar los promotores siguientes: el promotor 35S-CaMV (Franck et al., 1980, *Cell* 21: 285-294), la ubiquitina de maíz 1 (Christensen AH, Sharrock RA y Quail 1992. Maize polyubiquitin genes: structure, thermal perturbation of expression and transcript splicing, and promoter activity following transfer to protoplasts by electroporation) o el promotor de actina de arroz 1 (*Plant Mo. Biol.* 18,675-689.; Zhang W, McElroy D. and Wu R 1991, Analysis of rice Act1 5' region activity in transgenic rice plants. *Plant Cell* 3, 1155-1165).

Los promotores específicos de un órgano pueden ser, por ejemplo, un promotor de tejidos sumidero de almacenamiento tal como semillas, tubérculos de patata y frutas (Edwards & Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303) o de tejidos sumidero metabólicos tales como meristemos (Ito et al., 1994, *Plant Mol. Biol.* 24: 863-878), un promotor específico de semilla como la glutelina, prolamina, globulina o promotor de albúmina de arroz (Wu et al., 1998, *Plant and Cell Physiology* 39: 885-889), un promotor de Vicia faba de la legúmina B4 y el gen de proteína de semilla desconocida de Vicia faba (Conrad et al., 1998, *Journal of Plant Physiology* 152: 708-711), un promotor de una proteína de cuerpo oleoso de semilla (Chen et al., 1998, *Plant and Cell Physiology* 39: 935-941), el promotor de la proteína de almacenamiento napA de *Brassica napus* o cualquier otro promotor específico de semilla conocido en la técnica, por ejemplo, como se describe en la WO 91/14772.

- Además, el promotor puede ser un promotor específico de hoja tal como el promotor rbcS de arroz o tomate (Kozuka et al., 1993, Plant Physiology 102: 991-1000, el promotor de gen de adenina metiltransferasa del virus chlorella (Mitra and Higgins, 1994, Plant Molecular Biology 26: 85-93) o el promotor aldP de gen de arroz (Kagaya et al., 1995, Molecular and General Genetics 248: 668-674) o un promotor inducible por lesiones como el promotor de patata pin2 (Xu et al., 1993, Plant molecular Biology 22: 573-588).
- Asimismo, el promotor puede ser inducible por tratamientos abióticos tales como temperatura, sequía o alteraciones en la salinidad o inducible por sustancias aplicadas exógenamente que activan el promotor, por ejemplo etanol, estrógenos, hormonas de planta como etileno, ácido abscísico, ácido giberélico y/o metales pesados.
- [0152] También se puede usar un elemento intensificador del promotor para conseguir una expresión más alta del polipéptido en la planta.  
Por ejemplo, el elemento intensificador del promotor puede ser un intrón que se coloca entre el promotor y la secuencia de nucleótidos que codifica un polipéptido de la presente invención.  
Por ejemplo, Xu et al., 1993, supra revelan el uso del primer intrón del gen de actina de arroz 1 para mejorar la expresión.
- [0153] Además, todavía el uso de codón se puede optimizar para las especies de planta en cuestión para mejorar la expresión (ver Horvath et al referido arriba).
- [0154] El gen marcador seleccionable y cualquiera de las otras partes del constructo de expresión se pueden elegir de aquellas disponibles en la técnica.
- [0155] El constructo de ácido nucleico se incorpora en el genoma de planta según técnicas convencionales conocidas en la técnica, incluyendo transformación mediada por Agrobacterium, transformación mediada de virus, microinyección, bombardeo de partícula, transformación biolística y electroporación (Gasser et al., 1990, Science 244: 1293; Potrykus, 1990, Bio/Tecnology 8: 535; Shimamoto et al., 1989, Nature 338: 274).
- [0156] Actualmente, la transferencia de gen mediado por tumefaciens de Agrobacterium es el método de elección para generar dicotiledóneas transgénicas (para una revisión, ver Hooykas and Schilperoort, 1992, Plant Molecular biology 19: 15-38) y esto puede también usarse para la transformación de monocotiledóneas, aunque se usan más frecuentemente otros métodos de transformación para estas plantas.  
Actualmente, el método de elección para generar monocotiledóneas transgénicas, que complementa el método de Agrobacterium, es el bombardeo de partículas (partículas de oro o tungsteno microscópicas recubiertas con ADN transformante) de callos embriogénicos o embriones de desarrollo (Christou, 1992, Plant journal 2: 275-281; Shimamoto, 1994, Current Biotechnology 5: 158-162; Vasil et al., 1992, Bio/Tecnology 10: 667-674).  
Un método alternativo para la transformación de monocotiledóneas se basa en la transformación de protoplasto como se describe por Omirulleh et al., 1993, Plant Molecular Biology 21: 415-428.
- [0157] La transformación siguiente, los transformantes con el constructo de expresión incorporado en estos se seleccionan y regeneran en plantas enteras según métodos bien conocidos en la técnica. Frecuentemente, el procedimiento de transformación se diseña para la eliminación selectiva de genes de selección bien durante la regeneración o en las siguientes generaciones usando, por ejemplo, la cotransformación con dos constructos de T-ADN separados o escisión específica de sitio del gen de selección por una recombinasa específica.
- [0158] La presente descripción también se refiere a métodos para producir un polipéptido de la presente invención que comprende (a) el cultivo de una planta transgénica o una célula vegetal que comprende una secuencia de ácidos nucleicos que codifica un polipéptido con actividad de fitasa de la presente invención bajo condiciones propicias para la producción del polipéptido; y (b) la recuperación del polipéptido.
- 50 Composiciones y usos
- [0159] Todavía en otros aspectos, la presente invención se refiere a composiciones que comprenden un polipéptido de la presente invención, al igual que métodos de uso de estos.
- [0160] Las composiciones de polipéptido se pueden preparar conforme a métodos conocidos en la técnica y pueden ser en forma de un líquido o una composición seca.  
Por ejemplo, la composición de polipéptido puede ser en forma de granulados o microgranulados.  
El polipéptido que se va a incluir en la composición se puede estabilizar conforme a métodos conocidos en la técnica.
- [0161] La fitasa de la invención se puede usar para la degradación, en cualquier contexto industrial de, por ejemplo, fitato, ácido fítico y/o el mono-, di-, tri-, tetra- y/o pentafosfato de mioinositol.  
Es bien conocido que las fracciones de fosfato de estos compuestos quelan cationes bivalentes y trivalentes, tales como iones metálicos, i.a. los iones esenciales nutricionalmente de calcio, hierro, zinc y magnesio al igual que el manganeso de oligoelementos, cobre y molibdeno.  
Además, el ácido fítico también une hasta un cierto nivel proteínas por interacción electrostática.

[0162] Por consiguiente, usos preferidos de los polipéptidos de la invención están en preparaciones de alimentos para animales (incluyendo alimentos humanos) o en aditivos para este tipo de preparaciones.

5 [0163] En una forma de realización particular, el polipéptido de la invención se puede usar para mejorar el valor nutricional de un alimento para animales.

Ejemplos no limitativos de la mejora del valor nutricional de alimentos para animales (incluyendo alimentos humanos) son: mejora de la digestibilidad del alimento; crecimiento de promoción del animal; mejora de la utilización del alimento; mejora de biodisponibilidad de proteínas; aumento del nivel de fosfato digerible; mejora de la liberación y/o degradación de fitato; mejora de biodisponibilidad de oligoelementos; mejora de biodisponibilidad de macrominerales; eliminación de la necesidad de añadir fosfato adicional, oligoelementos y/o macrominerales; y/o mejora de la calidad de la capa del óvulo.

10 Por lo tanto, se aumenta el valor nutricional del alimento y el índice de crecimiento y/o aumento de peso y/o conversión de alimento (es decir, el peso del alimento ingerido con respecto a un aumento de peso) del animal se puede mejorar.

15 [0164] Además, el polipéptido de la invención se puede usar para reducir el nivel de fitato de estiércol.

Animales, alimento para animales y aditivos de alimento para animales

20 [0165] El término animal incluye todos los animales, incluyendo seres humanos.  
Ejemplos de animales son no rumiantes y rumiantes.

Los animales rumiantes incluyen, por ejemplo, animales tales como oveja, cabras, caballos y ganado bovino, por ejemplo, ganado bovino para carne, vacas y terneros jóvenes.

25 En una forma de realización particular, el animal es un animal no rumiante.

Los animales no rumiantes incluyen animales monogástricos, por ejemplo, cerdos o puercos (incluyendo, pero no de forma limitativa, lechones, cerdos en crecimiento y cerdas); aves tales como pavos, patos y pollo (incluyendo pero no de forma limitativa pollos de engorde y gallinas ponedoras); terneros jóvenes; y pescado (incluyendo pero no de forma limitativa salmón, trucha, tilapia, siluro y carpas; y crustáceos (incluyendo pero no de forma limitativa gambas y langostinos).

[0166] El término alimento o composición de alimento significa cualquier compuesto, preparación, mezcla o composición adecuada para o destinada a la ingesta de un animal.

35 [0167] En el uso, según la invención, el animal se puede alimentar del polipéptido antes, después o simultáneamente con la dieta.  
Se prefiere lo último.

[0168] En una forma de realización particular, el polipéptido, en la forma donde se añade al alimento o cuando se incluye en un aditivo de alimento es sustancialmente puro.

En una forma de realización particular está bien definido.

El término "bien definido" significa que la preparación de fitasa es al menos 50 % pura como se ha determinado por cromatografía de exclusión de tamaño (véase ejemplo 12 de la WO 01/58275).

45 En otras formas de realización particulares, la preparación de fitasa es al menos 60, 70, 80, 85, 88, 90, 92, 94 o al menos 95 % pura como se ha determinado por este método.

[0169] Una preparación de polipéptido sustancialmente pura y/o bien definida es ventajosa.

Por ejemplo, es mucho más fácil dosificar correctamente al alimento un polipéptido que está esencialmente libre de la interferencia o contaminación de otros polipéptidos.

50 El término dosificar correctamente se refiere, en particular, al objetivo de obtener resultados consistentes y constantes, y a la capacidad de optimizar la dosificación basada en el efecto deseado.

[0170] Sin embargo, para su uso en el alimento para animales, el polipéptido de fitasa de la invención no necesita ser puro; este puede incluir, por ejemplo, otros polipéptidos, en cuyo caso este podría denominarse una preparación de fitasa.

[0171] La preparación de fitasa se puede (a) añadir directamente al alimento (o usar directamente en un proceso de tratamiento de proteínas) o (b) esta se puede usar en la producción de una o más composiciones intermedias, tales como aditivos alimenticios o premezclas que sean posteriormente añadidas al alimento (o usada en un proceso de tratamiento).

60 El grado de pureza anteriormente descrito se refiere a la pureza de la preparación de polipéptido original, si se usa según (a) o (b) de arriba.

[0172] Las preparaciones de polipéptido con purezas de este orden de magnitud, en particular, son obtenibles usando métodos recombinantes de producción, mientras que estos no se obtienen tan fácilmente y también están sujetos a una variación de lote a lote mucho más alta cuando el polipéptido se produce por métodos de fermentación

tradicionales.

[0173] Tal preparación de polipéptido, por supuesto puede estar mezclada con otros polipéptidos.

5 [0174] El polipéptido se puede añadir al alimento en cualquier forma, ser este un polipéptido relativamente puro o en mezcla con otros componentes destinados a la adición a alimentos de animales, es decir, en forma de aditivos de alimentos para animales, tal como las denominadas premezclas para alimentos para animales.

10 [0175] En otro aspecto, la presente invención se refiere a composiciones para usar en el alimento para animales, tal como alimento para animales y aditivos de alimentos, por ejemplo premezclas.

[0176] A parte del polipéptido de la invención, los aditivos de alimentos de la invención contienen al menos una vitamina liposoluble y/o al menos una vitamina soluble en agua, y/o al menos un oligoelemento. El aditivo de alimentos también puede contener al menos un macromineral.

15 [0177] Además, ingredientes de aditivos de alimentos opcionales son agentes colorantes, por ejemplo, carotenoides tales como beta-caroteno, astaxantina y luteína; compuestos de aroma; estabilizadores; péptidos antimicrobianos; ácidos grasos poliinsaturados; especies generadoras de oxígeno reactivo; y/o al menos otro polipéptido seleccionado de entre fitasa (EC 3.1.3.8 o 3.1.3.26), xilanas (EC 3.2.1.8), galactanas (EC 3.2.1.89), alfa-galactosidas (EC 3.2.1.22), proteasa (EC 3.4.-.-), fosfolipasa A1 (EC 3.1.1.32); fosfolipasa A2 (EC 3.1.1.4); lisofosfolipasa (EC 3.1.1.5); fosfolipasa C (3.1.4.3); fosfolipasa D (EC 3.1.4.4); amilasa tal como, por ejemplo, alfa-amilasa (EC 3.2.1.1) y/o beta-glucanasa (EC 3.2.1.4 o EC 3.2.1.6).

25 [0178] En una forma de realización particular, estos otros polipéptidos están bien definidos (como se define arriba para preparaciones de fitasa).

[0179] En una forma de realización preferida, particularmente, la fitasa de la invención con un pH óptimo relativamente bajo se combina con al menos una fitasa con un pH óptimo más alto.

30 Ejemplos preferidos de fitasas de pH óptimo más alto son fitasas de Bacillus, tales como las fitasas de Bacillus licheniformis y Bacillus subtilis, al igual que derivados, variantes o fragmentos de la misma con actividad de fitasa.

[0180] La fitasa de la invención también se puede combinar con otras fitasas, por ejemplo, fitasas de ascomiceto tales como fitasas de Aspergillus, por ejemplo, derivadas de Aspergillus ficuum, Aspergillus niger o Aspergillus awamori; o fitasas de basidiomiceto, por ejemplo, derivadas de Peniophora lycii, Agrocybe pediades, Trametes pubescens o Paxillus involutus; o derivados, fragmentos o variantes de las mismas que tienen actividad de fitasa.

35 [0181] Así, en formas de realización preferidas del uso en el alimento para animales de la invención y en formas de realización preferidas del aditivo de alimento y el alimento para animales de la invención, la fitasa de la invención se combina con tales fitasas.

40 [0182] El ascomiceto anteriormente mencionado y fitasas de basidiomiceto, en particular, la fitasa RONOZYME P derivada de Peniophora lycii al igual que derivados, variantes y fragmentos del mismo, también se pueden combinar con fitasas de Bacillus, en particular, la fitasa B. licheniformis al igual que con un derivado, fragmento o variante de la misma, en particular para fines de alimento para animales.

45 [0183] Ejemplos de péptidos antimicrobianos (AMP's) son CAP18, Leucocina A, Tritrpticina, Protegrina-1, Tanatina, Defensina, Lactoferrina, Lactoferricina y Ovispirina tal como Novispirina (Robert Lehrer, 2000), Plectasinas y Estatinas, con los compuestos y los polipéptidos descritos en la WO 03/044049 y la WO 03/048148, al igual que variantes o fragmentos del anterior que retienen la actividad antimicrobiana.

50 [0184] Ejemplos de los polipéptidos antifúngicos (AFP's) son el Aspergillus giganteus y péptidos de Aspergillus niger, al igual que variantes y fragmentos de los mismos que retienen actividad antifúngica, como se describe en la WO 94/01459 y la WO 02/090384.

55 [0185] Ejemplos de ácidos grasos poliinsaturados son C18; C20 y C22 ácidos grasos poliinsaturados, tales como ácido araquidónico, ácido docosahexaenoico, ácido eicosapentanoico y ácido gamma-linolénico.

[0186] Ejemplos de especies generadoras de oxígeno reactivo son productos químicos tales como perborato, persulfato o percarbonato; y los polipéptidos tales como una oxidasa, una oxigenasa o una sintetasa.

60 [0187] Por lo general, la grasa y vitaminas hidrosolubles, al igual que oligoelementos forman parte de una denominada premezcla destinada a la adición al alimento, mientras que los macrominerales se añaden normalmente por separado al alimento.

65 Cualquiera de estos tipos de composición, cuando se ha enriquecido con un polipéptido de la invención, es un aditivo de alimento de animales de la invención.

[0188] En una forma de realización particular, el aditivo de alimento para animales de la invención está destinado a ser incluido (o prescrito como que se debe incluir) en dietas para animales o alimentos a niveles de 0,01 a 10,0 %; más particularmente 0,05 a 5,0 %; o 0,2 a 1,0 % (% significa g de aditivo por 100 g de alimento). Esto es así, en particular, para premezclas.

5 [0189] Las siguientes son listas no exclusivas de ejemplos de estos componentes:

Ejemplos de vitaminas liposolubles son vitamina A, vitamina D3, vitamina E y vitamina K, por ejemplo vitamina K3.

10 [0190] Ejemplos de vitaminas hidrosolubles son B12 vitamínico, biotina y colina, vitamina B1, vitamina B2, vitamina B6, niacina, ácido fólico y pantotenato, por ejemplo, ca-D-pantotenato.

[0191] Ejemplos de oligoelementos son manganeso, zinc, hierro, cobre, yodo, selenio y cobalto.

15 [0192] Ejemplos de macrominerales son calcio, fósforo y sodio.

[0193] Los requisitos nutricionales de estos componentes (ejemplificados con aves y lechones/cerdos) se enumeran en la tabla A de la WO 01/58275.

20 Requisito nutricional significa que estos componentes deberían ser proporcionados a la dieta en las concentraciones indicadas.

[0194] En la alternativa, el aditivo de alimento de animales de la invención comprende al menos uno de los componentes individuales en la tabla A específicos de la WO 01/58275.

25 Al menos uno significa cualquiera de uno o más de, uno, o dos, o tres, o cuatro etcétera hasta los trece o hasta los quince componentes individuales.

Más específicamente, al menos un componente individual se incluye en el aditivo de la invención en tal cantidad para proporcionar una concentración en el alimento en el rango indicado en la columna cuatro o columna cinco, o columna seis de la tabla A.

30 [0195] La presente invención también se refiere a composiciones de alimento para animales.

Las composiciones de alimento para animales o dietas tienen un contenido relativamente alto en proteína.

Las dietas de aves y cerdos se pueden caracterizar como se indica en la tabla B de la WO 01/58275, columnas 2-3.

Las dietas de pescado se pueden caracterizar como se indica en la columna 4 de esta tabla B. Además, tales dietas de pescado normalmente tienen un contenido de grasa cruda de 200-310 g/kg.

35 [0196] La WO 01/58275 corresponde a la US 09/779334 que está por la presente incorporada por referencia.

[0197] Una composición de alimento para animales según la invención tiene un contenido bruto en proteína de 50-800 g/kg y comprende, además, al menos un polipéptido como se reivindica aquí.

40 [0198] Además o en alternativa (al contenido bruto de proteína indicado arriba), la composición de alimento para animales de la invención tiene un contenido en energía metabolizable de 10-30 MJ/kg; y/o un contenido en calcio de 0,1-200 g/kg; y/o un contenido en fósforo disponible de 0,1-200 g/kg; y/o un contenido en metionina de 0,1-100 g/kg; y/o un contenido en metionina más cisteína de 0,1-150 g/kg; y/o un contenido en lisina de 0,5-50 g/kg.

45 [0199] En formas de realización particulares, el contenido de energía metabolizable, proteína cruda, calcio, fósforo, metionina, metionina más cisteína y/o lisina está dentro de cualquiera de los rangos 2, 3, 4 o 5 en la tabla B de la WO 01/58275 (R. 2-5).

50 [0200] La proteína cruda se calcula como nitrógeno (N) multiplicada por un factor 6,25, es decir, proteína cruda (g/kg) = N (g/kg) x 6,25.

El contenido de nitrógeno se determina por el método de Kjeldahl (A.O.A.C., 1984, Official Methods of Analysis 14th ed., Association of Official Chemists, Washington DC).

55 [0201] La energía metabolizable se puede calcular basándose en la NRC publication Nutrient requirements in swine, ninth revised edition 1988, subcommittee on swine nutrition, committee on animal nutrition, board of agriculture, national research council. National Academy Press, Washington, D.C., pp. 2-6, and the European Table of Energy Values for Poultry Feed-stuffs, Spelderholt centre for poultry research and extension, 7361 DA Beekbergen, The Netherlands.

60 Grafisch bedrijf Ponsen & looijen bv, Wageningen. ISBN 90-71463-12-5.

[0202] El contenido dietético de calcio, fósforo disponible y aminoácidos en dietas para animales completas se calcula basándose en tablas de alimento, tales como Veevoedertabel 1997, gegevens over chemische samenstelling, verteerbaarheid en voederwaarde van voedermiddelen, Central Veevoederbureau, Runderweg 6, 8219 pk Lelystad. ISBN 90-72839-13-7.

65

- [0203] En una forma de realización particular, la composición de alimento para animales de la invención contiene al menos una proteína.  
 La proteína puede ser una proteína animal, tal como harina de carne y hueso y/o harina de pescado; o esta puede ser una proteína vegetal.  
 El término proteínas vegetales, como se utiliza en este caso se refiere a cualquier compuesto, composición, preparación o mezcla que incluye al menos una proteína derivada de u originada de un vegetal, que incluye proteínas modificadas y derivados de proteína.  
 En formas de realización particulares, el contenido de proteína de las proteínas vegetales es al menos 10, 20, 30, 40, 50 o 60 % (p/p).
- [0204] Las proteínas vegetales se pueden derivar de fuentes de proteína vegetal, tales como leguminosas y cereales, por ejemplo, materiales de plantas de la familia Fabaceae (leguminosae), Cruciferaeae, Chenopodiaceae y Poaceae, tal como harina de soja, harina de lupino y alimento de harina de colza.
- [0205] En una forma de realización particular, la fuente de proteína vegetal es material de una o más plantas de la familia Fabaceae, por ejemplo, soja, altramuz, guisante o judía.
- [0206] En otra forma de realización particular, la fuente de proteína vegetal es material de una o más plantas de la familia Chenopodiaceae, por ejemplo, remolacha, remolacha azucarera, espinaca o quinoa.
- [0207] Otros ejemplos de fuentes de proteína vegetal son semilla de colza, semilla de girasol, semilla de algodón y repollo.
- [0208] La soja es una fuente de proteína vegetal preferida.
- [0209] Otros ejemplos de fuentes de proteína vegetal son cereales tales como cebada, trigo, centeno, avena, maíz (mazorca), arroz, triticale y sorgo.
- [0210] Todavía en otras formas de realización particulares, la composición de alimento para animales de la invención contiene 0-80 % de maíz; y/o 0-80 % de sorgo; y/o 0-70 % de trigo; y/o 0-70 % de cebada; y/o 0-30 % de avena; y/o 0-40 % de harina de soja; y/o 0-25 % de harina de pescado; y/o 0-25 % de harina de carne y hueso; y/o 0-20 % de suero de leche.
- [0211] Las dietas para animales se pueden fabricar, por ejemplo, como alimento triturado (no granulado) o alimento granulado.  
 Típicamente, los materiales de alimento molidos se mezclan y se añaden cantidades suficientes de vitaminas esenciales y minerales, según las especificaciones para las especies en cuestión.  
 Los polipéptidos se pueden adicionar como formulaciones de polipéptido sólidas o líquidas.  
 Por ejemplo, una formulación de polipéptido sólida se agrega típicamente antes o durante la etapa de mezcla; y una preparación de polipéptido líquida se añade típicamente después del paso de granulación.  
 El polipéptido también se puede incorporar a un aditivo de alimento o premezcla.
- [0212] La concentración de polipéptido final en la dieta está en el rango de 0,01-200 mg de proteína de polipéptido por kg de dieta, por ejemplo, en el rango de 5-30 mg de proteína de polipéptido por kg de dieta animal.
- [0213] La fitasa de la invención, por supuesto, debería aplicarse en una cantidad eficaz, es decir, en una cantidad adecuada para mejorar la solubilización y/o mejorar el valor nutricional del alimento.  
 Actualmente, se contempla que el polipéptido se administra en una o varias de las siguientes cantidades (rangos de dosificación): 0,01-200; 0,01-100; 0,5-100; 1-5; 5-100; 10-100; 0,05-50 o 0,10-10, todos estos rangos son proteína de polipéptido de fitasa en mg por kg (ppm) de alimento.
- [0214] Para la determinación de proteína de polipéptido de fitasa en mg por kg de alimento, la fitasa se purifica de la composición de alimento y la actividad específica de la fitasa purificada se determina usando un ensayo pertinente.  
 La actividad de fitasa de la composición de alimento como tal se determina también usando el mismo ensayo y, basándose en estas dos determinaciones, se calcula la dosificación de proteína de fitasa en mg por kg de alimento.
- [0215] Se pueden aplicar los mismos principios para determinar los mg de proteína de polipéptido de fitasa en aditivos alimenticios.  
 Por supuesto, si está disponible una muestra de la fitasa usada para la preparación del aditivo de alimento o el alimento, la actividad específica se determina de esta muestra (no es necesario purificar la fitasa de la composición de alimento o el aditivo).
- Péptido señal
- [0216] La presente descripción también se refiere a constructos de ácidos nucleicos que comprenden un gen que



codifica una proteína operativamente vinculada a una primera secuencia de nucleótidos que consiste en nucleótidos 1 a 66 de SEC ID n.º:5, que codifica un péptido señal que consiste en aminoácidos 1 a 22 de SEC ID n.º:6, respectivamente, donde el gen es extranjero para las primeras secuencias de nucleótidos.

5 [0217] La presente invención también se refiere a vectores de expresión recombinantes y células huésped recombinantes que comprenden tales constructos de ácido nucleico.

[0218] La presente invención también se refiere a métodos para producir una proteína que comprende (a) el cultivo de tal célula huésped recombinante bajo condiciones adecuadas para la producción de la proteína; y (b) la recuperación de la proteína.

[0219] Las primeras secuencias de nucleótidos pueden estar operativamente vinculadas a genes extranjeros individualmente con otras secuencias de control o en combinación con otras secuencias de control. Tales otras secuencias de control se describen supra.

15 [0220] La proteína puede ser nativa o heteróloga a una célula huésped. El término "proteína" aquí no se refiere a una longitud específica del producto codificado y, por lo tanto, abarca péptidos, oligopéptidos y proteínas.

El término "proteína" también abarca dos o más polipéptidos combinados para formar el producto codificado.

20 Las proteínas también incluyen los polipéptidos híbridos que comprenden una combinación de secuencias polipéptidas parciales o completas obtenidas de al menos dos proteínas diferentes donde una o más puede ser heteróloga o nativa a la célula huésped.

Las proteínas incluyen además variaciones alélicas de origen natural y modificadas de las proteínas anteriormente mencionadas y proteínas híbridas.

25 [0221] Preferiblemente, la proteína es una hormona o variante de la misma, polipéptido, receptor o porción del mismo, anticuerpo o porción del mismo, o indicador.

En un aspecto más preferido, la proteína es una oxidoreductasa, transferasa, hidrolasa, liasa, isomerasa o ligasa.

30 En un aspecto aún más preferido, la proteína es una aminopeptidasa, amilasa, carbohidrasa, carboxipeptidasa, catalasa, celulasa, quitinasa, cutinasa, ciclodextrina glicosiltransferasa, desoxiribonucleasa, esterasa, alfa-galactosidasa, beta-galactosidasa, glucoamilasa, alfa-glucosidasa, beta-glucosidasa, invertasa, lacasa, lipasa, manosidasa, mutanasa, oxidasa, polipéptido pectinólítico, peroxidasa, fitasa, polifenoloxidasa, polipéptido proteolítico, ribonucleasa, transglutaminasa o xilanasas.

35 [0222] El gen se puede obtener de cualquier procarionte, eucariote u otra fuente.

## Ejemplos

### Ejemplo 1: clonación de una fitasa *Citrobacter gillenii*

40 [0223] Se realizó un alineamiento múltiple de las siguientes fosfatasa de histidina ácidas: appA *Escherichia coli* (SPTREMBL:Q8GN88), phyk *Klebsiella terrigena* (SPTREMBL:Q7WSY1) y ypo1648 *Yersinia pestis* CO92 (SPTREMBL:Q8ZFP6).

Dos cebadores oligonucleótidos degenerados fueron diseñados basándose en secuencias de consenso:

45 5'- CGC GTG GTG ATT GTG TCC MGN CAY GGN - 3' (SEQ ID n.º: 7, cebador directo)  
5'- C CAG GTT GGT ATC ATG GCC NGC DAT RAA - 3' (SEQ ID n.º: 8, cebador inverso).

[0224] Los cebadores fueron usados para la selección de PCR de un número de especies bacterianas a temperaturas de recocimiento de 45, 48 y 50 °C.

50 [0225] Se identificó un gen de fitasa parcial en forma de un fragmento 900 bp ADN en el *Citrobacter gillenii* DSM 13694.

[0226] El fragmento de PCR fue clonado en la pEZSeq blunt cloning kit (catálogo n.º 40501-1 de Lucigen Corporation, 2120 West Greenview Dr., Ste 9, Middleton, WI 53562, US).

55 Primero, el fragmento PCR se trató con PCRTerminator End Repair Kit (parte del pEZSeq blunt cloning kit), que contiene una mezcla de actividades enzimáticas que ha sido optimizada para crear extremidades desafiladas 5'-fosforiladas en cualquier tipo de producto PCR.

Después de la clonación en el vector pEZSeq, el clon fue secuenciado utilizando dos cebadores de vector específico.

60 Mediante traducción de las secuencias de nucleótidos se confirmó que el fragmento de ADN clonado fue parte de un gen de fitasa.

[0227] Para obtener la secuencia de nucleótidos de longitud total del gen, se usó el ADN Walking SpeedUp Kit (DWSK-V102 de Seegene, Inc., 2nd Fl., Myungji Bldg., 142-21, Samsung-dong, Kangnam-gu, Seoul, 135-090, Corea), que está diseñado para capturar sitios objetivo desconocidos.

## ES 2 614 744 T3

Para este propósito, fueron diseñados y usados 4 oligonucleótidos específicos con el equipo.

TSP1N: 5'- AGACTTCCGCCAGCCCG-3' (SEQ ID n.º: 9)

TSP1C: 5'- AAGCAGCTGGGCAGTCTGC-3' (SEQ ID n.º: 10)

TSP2N: 5'-AAGCGCGTGAACCTTTGTCGG-3' (SEQ ID n.º: 11)

5 TSP2C:5'-ATGGGGACTGGCTTCAACCCTG-3' (SEQ ID n.º: 12)

[0228] La secuencia de nucleótidos de longitud total que codifica la fitasa de *Citrobacter gillenii* DSM 13694 se muestra en el listado de secuencias como SEQ ID n.º: 3 y la secuencia de aminoácidos codificada correspondiente tiene el SEQ ID n.º: 4.

10 Está previsto que los primeros 22 aminoácidos de SEC ID n.º: 2 y 4 sean un péptido señal (predicho por señal P V3.0).

SEQ ID n.º: 1 es una variante de SEQ ID n.º: 3 que comprende las siguientes 13 sustituciones: T99G, C102G, C105T, C108G, A109T, G110C, T111C, T117C, C975T, A989C, T991 (-); C992(-), A1005(-).

15 SEQ ID n.º: 4 es una variante de la SEQ ID n.º: 2 que comprende las siguientes 5 sustituciones: N330T, I331K, S332R, A334(-), L335I.

[0229] El gen de fitasa *Citrobacter gillenii* DSM 13694 fue clonado en el vector de expresión pET-30a(+) *E.coli* sin etiquetas de fusión (catálogo n.º 69909 de Novagen, disponible comercialmente de Bie & Berntsen A/S, 7 Sandbaekvej, DK-2610 Roedovre, Dinamarca).

20 En este sistema, la expresión del gen se induce mediante una fuente de T7 RNA-polimerasa en la cepa huésped *E. coli* BL21star(DE)pLisS (catálogo n.º 69388 de Novagen, disponible comercialmente de Bie & Berntsen) que contiene una copia cromosómica del T7 gen de RNA-polimerasa bajo el control del promotor lacUV5.

La inducción del gen diana se realizó añadiendo lactosa a los medios.

25 La lactosa enlazará con el represor e inducirá su disociación desde el operador, permitiendo así la transcripción del promotor.

[0230] Para la expresión del gen de fitasa, una colonia individual de la cepa de *E. coli* transformada se transfirió a un cultivo de inóculo en los medios de no inducción (que contiene glucosa como la fuente única de carbono) que no permite la expresión de la T7 RNA-polimerasa.

30 Como un control negativo se usó *E.coli* (BL21star(DE)pLisS) que contiene un vector vacío pET-30(+).

Una parte alícuota pequeña (aproximadamente 150 micro litro) del cultivo de inóculo fue transferida a matraces que contienen lactosa como la fuente única de carbono.

El cultivo de inducción creció durante toda la noche con la agitación de 300 r.p.m. a 37 °C.

35 [0231] Las células se cosecharon por centrifugación y 15 micro partes alícuotas de litro del sobrenadante se analizaron por SDS-PAGE.

Como marcador de peso molecular (MW) se usaron 10 micro litros de la proteína estándar Precision Plus (catálogo n.º.161-0363, disponible comercialmente de Bio-Rad Laboratories Headquarters, 1000 Alfred Nobel Drive, Hercules, CA 94547, US).

40 Una banda diferente de PM de aproximadamente 50 kDa se identificó en el sobrenadante desde la cepa de *E. coli* recombinante, pero no en el control negativo.

[0232] La célula granulada cosechada se alisó y la fracción intracelular soluble también se analizó por SDS-PAGE como se ha descrito anteriormente.

45 También aquí apareció una banda de MW 50 kDa.

[0233] Es evidente que la proteína de fitasa recombinante se segrega parcialmente a los medios.

Sin embargo, una estanque del enzimático todavía permanece en la fracción intracelular.

50 [0234] La actividad de fitasa del sobrenadante y la fracción intracelular se confirmó usando el ensayo de Ejemplo 4.

### **Ejemplo 2:** preparación de preparación de fitasa *Citrobacter gillenii*

55 [0235] Se cultivó *Citrobacter gillenii* DSM 13694 durante toda la noche con la agitación (225 r.p.m.) a 30 °C en el medio LB (25 g de caldo LB, Merck 0285, agua intercambiada de iones hasta 1000 ml) con adición de 0,1 % (p/p) de fitato de sodio.

Las células se cosecharon por centrifugación (4000 r.p.m., 60 min) y el sobrenadante se descartó.

El granulado celular fue resuspendido en dos volúmenes de agua destilada con 100 mg/ml de lisozima y lisado por incubación durante toda la noche a 37 °C.

60 Las células lisadas fueron centrifugadas (4000 r.p.m., 2 h) y el sobrenadante salvado y usado para el análisis de estabilidad al ácido.

### **Ejemplo 3:** estabilidad al ácido de la fitasa *Citrobacter gillenii*

65 [0236] 50 micro litros del lisado obtenido en el ejemplo 2 se mezcló con 50 micro litros de tampones de 100 mM con valores de pH de 2,2; 3,0 (glicina/ácido clorhídrico) y 7,0 (HEPES) respectivamente.

## ES 2 614 744 T3

Las muestras fueron incubadas durante la noche a 37 °C y analizadas para actividad de fitasa residual utilizando el procedimiento analítico descrito en el Ejemplo 4.

La actividad de fitasa residual, expresada como la densidad óptica, se muestra en la tabla 1 de abajo.

Además, la actividad se calcula en por ciento relativamente a la actividad residual a pH 7.

5

Tabla 1

Actividad Residual [OD] tras la incubación a pH:	PH 2,2	PH 2,2 con respecto a pH 7,0	PH 3,0	PH 3,0 con respecto a pH 7,0	PH 7,0	PH 7,0 con respecto a pH 7,0
Cepa:						
Citrobacter gillenii DSM13694	0,07	81	0,10	116	0,09	100

### Ejemplo 4: ensayo de fitasa

10 [0237] El ensayo se basa en la determinación de fosfato soluble por complexación con molibdato/hierro y medición fotométrica del color azul en placas de microtitulación.

[0238] El sustrato es 0,5 mM de na-fitato (Sigma; P-8810) disuelto en 0,1 M de tampón de acetato, pH=5,5. En una forma de realización particular, la concentración de sustrato es 5 mM.

15

[0239] El reactivo de color se prepara de la siguiente manera: 1 % de amonio molibdato (Merck 1181, (NH<sub>4</sub>)<sub>6</sub>Mo<sub>7</sub>O<sub>24</sub>, 4H<sub>2</sub>O) se disuelve en 3,2 % de ácido sulfúrico (Merck 731). 1,1 g de ferrosulfato (Merck 3965) se disuelve en 15 ml del reactivo de molibdato anterior y se añade 10 ml de 0,5 M de ácido sulfúrico.

Se prepara cada día reciente y se almacena a oscuras.

20

[0240] Análisis de ciego: se mezcla 20ul de muestra, 100ul de sustrato y 120ul de reactivo de color, se incuba 5 min a 37 °C y se mide OD<sub>Blind</sub> a 750 nm.

25

[0241] Muestra: se mezcla 20ul de muestra, 100ul de sustrato, se incuba 30 min a 37 °C, se añade 120ul de reactivo de color, se incuba 5 min a 37 °C y se mide OD<sub>sample</sub> a 750 nm.

$$OD = OD_{sample} - OD_{Blind}$$

### Ejemplo 5: clonación de fitasas Citrobacter amalonaticus

30

[0242] Un alineamiento múltiple se realizó de las siguientes fosfatasas ácidas de histidina: appA Escherichia coli (SPTREMBL:Q8GN88), phyk Klebsiella terrigena (SPTREMBL:Q7WSY1) y ypo1648 Yersinia pestis CO92 (SPTREMBL:Q8ZFP6).

Dos cebadores oligonucleótidos degenerados fueron diseñados basándose en secuencias de consenso:

35

5'- CGC GTG GTG ATT GTG TCC MGN CAY GGN GT - 3' (SEQ ID n.º: 13, cebador directo)

5'- C CAG GTT GGT ATC ATG GCC NGC DAT RAA - 3' (SEQ ID n.º: 14, cebador inverso).

[0243] Los cebadores se usaron para la selección de PCR de un número de especies bacterianas a temperaturas de recocimiento de 45, 48 y 50 °C.

40

[0244] Genes de fitasa parciales en forma de 900 fragmentos de ADN bp fueron identificados en el Citrobacter amalonaticus ATCC 25405 y Citrobacter amalonaticus ATCC 25407.

[0245] Los fragmentos PCR fueron clonados en el pEZSeq blunt cloning kit (catálogo n.º 40501-1 de Lucigen Corporation, 2120 West Greenview Dr., Ste 9, Middleton, WI 53562, US).

45

Primero, los fragmentos PCR se trataron con el PCR Terminator End Repair Kit (parte del pEZSeq blunt cloning kit), que contiene una mezcla de actividades de polipéptido que ha sido optimizada para crear extremidades afiladas 5'-fosforiladas en cualquier tipo de producto PCR.

Después de la clonación en el vector pEZSeq, los clones fueron secuenciados utilizando dos cebadores de vector específicos.

50

Por traducción de las secuencias de nucleótidos, se confirmó que los fragmentos de ADN clonados fueron parte de genes de fitasa.

[0246] Para obtener la secuencia de nucleótidos de longitud total de los genes se usó el ADN Walking SpeedUp Kit (DWSK-V102 de Seegene, Inc., 2nd Fl., Myungji Bldg., 142-21, Samsung-dong, Kangnam-gu, Seoul, 135-090, Corea), que se diseña para capturar sitios objetivo desconocidos.

55

Para este propósito, se diseñaron 4 oligonucleótidos específicos y se usaron con el equipo.

[0247] Por ejemplo, para la obtención de las secuencias N-terminal y C-terminal del gen de fitasa de *Citrobacter amalonaticus* ATCC 25407, oligonucleótidos TSP1N/TSP2N y TSP1C/TSP2C fueron usados:

TSP1 N: 5'- TTT ACG AGT GCG CTG ATC GG - 3' (SEQ ID n.º: 15)  
 TSP2N: 5'- ATA ACT GCC ACC TGT TCT GGT GAC G - 3' (SEQ ID N.º: 16)  
 TSP1 C: 5'- ACT CAG GGT TTC TGC GGA TAC C- 3' (SEQ ID N.º: 17)  
 TSP2C: 5'- AAT GCC AGA GGT TGC GTG GGG - 3' (SEQ ID n.º: 18)

[0248] La secuencia de nucleótidos de longitud total que codifica la fitasa de *Citrobacter amalonaticus* ATCC 25407 se muestra en el listado de secuencias como SEQ ID Nn.º: 5 y la secuencia de aminoácidos codificada correspondiente tiene SEQ ID n.º: 6.

Se prevé que los primeros 22 aminoácidos de SEQ ID n.º: 6 que son un péptido señal (predicho por señal P V3,0). La fitasa de *Citrobacter amalonaticus* ATCC 25405 tiene precisamente la misma secuencia de aminoácidos como se muestra en la SEQ ID n.º: 6 y la secuencia de nucleótidos correspondiente difiere por solo tres nucleótidos de la SEQ ID n.º: 5, es decir, T78C, C118T y T120G, todo los cuales son silenciosos en el sentido de que estos no dan lugar a cualquier aminoácido diferente.

[0249] El gen de fitasa *Citrobacter amalonaticus* ATCC 25407 fue clonado en el vector de expresión pET-30a(+)E.coli sin etiquetas de fusión (catálogo n.º 69909 de Novagen, disponible comercialmente de Bie & Berntsen A/S, 7 Sandbaekvej, DK-2610 Roedovre, Dinamarca).

En este sistema, la expresión del gen se induce mediante una fuente de T7 RNA-polimerasa en la cepa huésped E. Coli BL21star(DE)pLisS (catálogo n.º 69388 de Novagen, disponible comercialmente de Bie & Berntsen) que contiene una copia cromosómica del gen de T7 RNA-polimerasa bajo el control del promotor lacUV5.

La inducción del gen diana se realizó añadiendo lactosa a los medios. La lactosa enlazará con el represor e inducirá a su disociación desde el operador, permitiendo así la transcripción del promotor.

[0250] Para la expresión del gen de fitasa, una colonia individual de la cepa transformada de E. Coli fue transferida a un cultivo de inóculo en los medios de no inducción (que contienen glucosa como única fuente de carbono) que no permite la expresión de la T7 RNA-polimerasa.

Como un control negativo se usó E.coli (BL21star(DE)pLisS) con un vector vacío pET-30(+). Una parte alícuota pequeña (aproximadamente 150 micro litros) del cultivo de inóculo fue transferida a matraces que contienen lactosa como única fuente de carbono. El cultivo de inducción creció durante toda la noche con la agitación de 300 r.p.m. a 37 °C.

[0251] Las células se cosecharon por centrifugación y 15 micro litros de partes alícuotas del sobrenadante fueron analizados por SDS-PAGE.

Como un marcador de peso molecular (PM) de 10 micro litros del estándar de proteína Precision Plus (catálogo n.º 161-0363, disponible comercialmente de Bio-Rad Laboratories Headquarters, 1000 Alfred Nobel Drive, Hercules, CA 94547, US).

Una banda diferente de PM de aproximadamente 50 kDa se identificó en el sobrenadante de la cepa de E. coli recombinante, pero no en el control negativo.

[0252] El granulado de célula cosechada se lisó y la fracción intracelular soluble también se analizó por SDS-PAGE como se ha descrito anteriormente.

También aquí apareció una banda de MW 50 kDa.

[0253] Está demostrado que la proteína de fitasa recombinante se segrega parcialmente a los medios.

Sin embargo, una agrupación del polipéptido todavía permanece en la fracción intracelular.

[0254] La actividad de fitasa del sobrenadante y la fracción intracelular se confirmó usando el ensayo del ejemplo 4.

#### **Ejemplo 6:** preparación de una preparación de fitasa *Citrobacter amalonaticus*

[0255] *Citrobacter amalonaticus* ATCC 25405 creció durante toda la noche con la agitación (225 r.p.m.) a 30 °C en el medio LB (25 g de caldo LB, Merck 0285, agua intercambiada de iones hasta 1000 ml) con adición del 0,1 % (p/p) de fitato de sodio.

Las células se cosecharon por centrifugación (4000 r.p.m., 60 min) y se descartó el sobrenadante.

El granulado celular fue resuspendido en dos volúmenes de agua destilada con 100 mg/ml de lisozima y lisado por incubación durante toda la noche a 37 °C.

Las células lisadas fueron centrifugadas (4000 r.p.m., 2 h) y el sobrenadante salvado y usado para análisis de estabilidad ácida.

#### **Ejemplo 7:** estabilidad al ácido de la fitasa *Citrobacter amalonaticus*

[0256] 50 micro litros del lisado obtenido en el ejemplo 2 se mezcló con tampones 100 mM de 50 micro litros con

## ES 2 614 744 T3

valores de pH de 2,2; 3,0 (glicina/ácido clorhídrico) y 7,0 (HEPES) respectivamente.

Las muestras fueron incubadas durante la noche a 37 °C y analizadas para actividad de fitasa residual utilizando el procedimiento analítico descrito en el ejemplo 4.

La actividad de fitasa residual, expresada como la densidad óptica, se muestra en la tabla 2 de abajo.

5 Además, la actividad se calcula en por ciento con respecto a la actividad residual pH 7.

Tabla 2

Actividad Residual [OD] tras la incubación a pH:	PH 2,2	PH 2,2 con respecto a pH 7,0	PH 3,0	PH 3,0 con respecto a pH 7,0	PH 7,0	PH 7,0 con respecto a pH 7,0
Cepa:						
Citrobacter amalonaticus ATCC 25405	0,12	78	0,14	100	0,14	100

10 Listado de secuencias

[0257]

<110> Novozymes A/S

<120> Polypeptides having phytase activity and polynucleotides encoding same

<130> 10676.504-WO

<160> 18

<170> PatentIn version 3.3

<210> 1

<211> 1296

<212> DNA

<213> Artificial

<220>

<223> variant of SEQ ID NO: 3  
(T99G+C102G+C105T+C108G+A109T+G110C+T111C+T117C+C975T+A989C+T991(-)+C992(-)+A1005(-))

<220>

<221> CDS

<222> (1)..(1293)

<220>

<221> mat\_peptide

<222> (67)..(1293)

<400> 1

15

ES 2 614 744 T3

atg	agt	aca	ctg	atc	att	cgt	tta	ttg	ttc	tta	acg	att	ata	ttg	gcc	48
Met	Ser	Thr	Leu	Ile	Ile	Arg	Leu	Leu	Phe	Leu	Thr	Ile	Ile	Leu	Ala	
		-20					-15					-10				
cct	gtt	tca	tta	cgc	gcc	gat	gaa	cag	agc	gga	atg	cag	ctt	gag	cgt	96
Pro	Val	Ser	Leu	Arg	Ala	Asp	Glu	Gln	Ser	Gly	Met	Gln	Leu	Glu	Arg	
	-5				-1	1				5					10	
gtg	gtg	att	gtg	tcc	cgt	cac	ggc	gtc	agg	gca	ccg	aca	aag	ttc	acg	144
Val	Val	Ile	Val	Ser	Arg	His	Gly	Val	Arg	Ala	Pro	Thr	Lys	Phe	Thr	
				15					20					25		
ccg	ctt	atg	cag	caa	gtc	act	ccc	gac	cgc	tgg	ccg	caa	tgg	gac	gtt	192
Pro	Leu	Met	Gln	Gln	Val	Thr	Pro	Asp	Arg	Trp	Pro	Gln	Trp	Asp	Val	
			30					35					40			
cct	ctg	ggg	tgg	ttg	act	cct	cgc	ggc	ggg	gca	ctc	att	act	gag	tta	240
Pro	Leu	Gly	Trp	Leu	Thr	Pro	Arg	Gly	Gly	Ala	Leu	Ile	Thr	Glu	Leu	
		45					50					55				
gga	cgg	tat	caa	cgt	tta	cgc	ctg	gcg	gac	aaa	ggg	ctg	ctg	gat	aat	288
Gly	Arg	Tyr	Gln	Arg	Leu	Arg	Leu	Ala	Asp	Lys	Gly	Leu	Leu	Asp	Asn	
	60					65					70					
aaa	acg	tgt	cca	acg	gca	ggg	cag	gtc	gcg	gtc	att	gcc	gat	agc	gat	336
Lys	Thr	Cys	Pro	Thr	Ala	Gly	Gln	Val	Ala	Val	Ile	Ala	Asp	Ser	Asp	
					80					85					90	
caa	cgt	acc	cgt	aaa	acg	ggg	gaa	gca	ttc	ctg	gca	gga	ctg	gct	ccg	384
Gln	Arg	Thr	Arg	Lys	Thr	Gly	Glu	Ala	Phe	Leu	Ala	Gly	Leu	Ala	Pro	
				95					100					105		
gaa	tgt	aaa	gta	cag	gtt	tat	tat	caa	caa	gat	aag	tca	aaa	tct	gat	432
Glu	Cys	Lys	Val	Gln	Val	Tyr	Tyr	Gln	Gln	Asp	Lys	Ser	Lys	Ser	Asp	
			110					115					120			

ES 2 614 744 T3

ccc Pro	ctt Leu	ttt Phe 125	aat Asn	ccc Pro	atc Ile	aag Lys	gcg Ala 130	ggg Gly	cgg Arg	tgt Cys	tcg Ser	ctg Leu 135	aac Asn	aca Thr	tcg Ser	480
cag Gln	gtg Val 140	aaa Lys	gag Glu	gcc Ala	atc Ile	ctg Leu 145	acc Thr	cgg Arg	gct Ala	ggc Gly	gga Gly 150	agt Ser	ctt Leu	gat Asp	gag Glu	528
tac Tyr 155	acg Thr	cgc Arg	cac His	tac Tyr	caa Gln 160	ccc Pro	gca Ala	ttt Phe	caa Gln	gcc Ala 165	ctg Leu	gaa Glu	cgg Arg	gtg Val	tta Leu 170	576
aat Asn	ttc Phe	tcc Ser	cag Gln	tca Ser 175	gaa Glu	aag Lys	tgt Cys	caa Gln	gca Ala 180	gct Ala	ggg Gly	cag Gln	tct Ser	gca Ala 185	cag Gln	624
tgt Cys	acg Thr	cta Leu	acc Thr 190	gac Asp	gtc Val	tta Leu	cct Pro	gct Ala 195	gaa Glu	ctc Leu	aag Lys	gtc Val	tct Ser 200	cca Pro	gaa Glu	672
aat Asn	ata Ile	tcg Ser 205	ttg Leu	tca Ser	ggc Gly	tca Ser	tgg Trp 210	gga Gly	ctg Leu	gct Ala	tca Ser	acc Thr 215	ctg Leu	acg Thr	gaa Glu	720
atc Ile	ttc Phe 220	ctg Leu	ctg Leu	caa Gln	caa Gln	gca Ala 225	caa Gln	ggg Gly	atg Met	tcg Ser	cag Gln 230	gtg Val	gcc Ala	tgg Trp	ggg Gly	768
cgt Arg 235	att Ile	cat His	ggc Gly	gat Asp	aaa Lys 240	gaa Glu	tgg Trp	cgt Arg	aca Thr	tta Leu 245	tta Leu	agt Ser	ctg Leu	cac His	aat Asn 250	816
gcg Ala	cag Gln	ttt Phe	gac Asp	ctt Leu 255	ctg Leu	cag Gln	aaa Lys	acc Thr	ccg Pro 260	gag Glu	gtt Val	gcc Ala	cgt Arg	agc Ser 265	agg Arg	864
gcc Ala	aca Thr	ccg Pro	tta Leu 270	ctt Leu	gat Asp	ttg Leu	ata Ile	cgt Arg 275	aca Thr	gca Ala	ctc Leu	gta Val	aca Thr 280	cag Gln	ggg Gly	912
gca Ala	aca Thr 285	gaa Glu	aat Asn	aaa Lys	tac Tyr	gca Ala	att Ile 290	cag Gln	ttg Leu	ccc Pro	gtc Val	tct Ser 295	ttg Leu	ttg Leu	ttt Phe	960
att Ile	gcg Ala 300	ggg Gly	cat His	gat Asp	acc Thr	aat Asn 305	ctt Leu	gcc Ala	act Thr	aag Lys	cgg Arg 310	ggc Gly	att Ile	ggc Gly	ctt Leu	1008
aac Asn 315	gtg Val	ttt Phe	ctg Leu	ccc Pro	ggt Gly 320	cag Gln	cca Pro	gat Asp	aat Asn	acg Thr 325	ccg Pro	ccg Pro	ggt Gly	gga Gly	gag Glu 330	1056
ttt Phe	gtt Val	ttc Phe	gaa Glu	agg Arg 335	tgg Trp	aaa Lys	cgg Arg	gtc Val	agc Ser 340	gat Asp	cat His	tct Ser	gat Asp	tgg Trp 345	gtg Val	1104
cag Gln	gtt Val	tct Ser	ttt Phe 350	atg Met	tat Tyr	cag Gln	aca Thr 355	ttg Leu	cag Gln	gaa Glu	atg Met	cgt Arg	gat Asp 360	atg Met	caa Gln	1152
cct Pro	ttg Leu	tcg Ser 365	ttg Leu	caa Gln	tcg Ser	cct Pro	ccc Pro 370	gga Gly	aaa Lys	att Ile	gtg Val	ctg Leu 375	ccc Pro	tta Leu	gcg Ala	1200
gcc Ala	tgc Cys 380	gat Asp	gag Glu	aaa Lys	aat Asn	acg Thr 385	cag Gln	gga Gly	atg Met	tgc Cys	tca Ser 390	tta Leu	aaa Lys	aat Asn	ttt Phe	1248

ES 2 614 744 T3

tct gca ctg att gat tcc gtt cgc gtg tcc gaa tgt gct gag aaa taa 1296  
 Ser Ala Leu Ile Asp Ser Val Arg Val Ser Glu Cys Ala Glu Lys  
 395 400 405

<210> 2  
 <211> 431  
 <212> PRT  
 <213> Artificial

<220>  
 <223> synthetic Construct

<400> 2

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



ES 2 614 744 T3

Cys Thr Leu Thr Asp Val Leu Pro Ala Glu Leu Lys Val Ser Pro Glu  
 190 195 200  
 Asn Ile Ser Leu Ser Gly Ser Trp Gly Leu Ala Ser Thr Leu Thr Glu  
 205 210 215  
 Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Ser Gln Val Ala Trp Gly  
 220 225 230  
 Arg Ile His Gly Asp Lys Glu Trp Arg Thr Leu Leu Ser Leu His Asn  
 235 240 245 250  
 Ala Gln Phe Asp Leu Leu Gln Lys Thr Pro Glu Val Ala Arg Ser Arg  
 255 260 265  
 Ala Thr Pro Leu Leu Asp Leu Ile Arg Thr Ala Leu Val Thr Gln Gly  
 270 275 280  
 Ala Thr Glu Asn Lys Tyr Ala Ile Gln Leu Pro Val Ser Leu Leu Phe  
 285 290 295  
 Ile Ala Gly His Asp Thr Asn Leu Ala Thr Lys Arg Gly Ile Gly Leu  
 300 305 310  
 Asn Val Phe Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly Gly Glu  
 315 320 325 330  
 Phe Val Phe Glu Arg Trp Lys Arg Val Ser Asp His Ser Asp Trp Val  
 335 340 345  
 Gln Val Ser Phe Met Tyr Gln Thr Leu Gln Glu Met Arg Asp Met Gln  
 350 355 360  
 Pro Leu Ser Leu Gln Ser Pro Pro Gly Lys Ile Val Leu Pro Leu Ala  
 365 370 375  
 Ala Cys Asp Glu Lys Asn Thr Gln Gly Met Cys Ser Leu Lys Asn Phe  
 380 385 390  
 Ser Ala Leu Ile Asp Ser Val Arg Val Ser Glu Cys Ala Glu Lys  
 395 400 405

<210> 3  
 <211> 1296  
 <212> DNA  
 <213> Citrobacter gillenii DSM 13694

<220>  
 <221> CDS  
 <222> (1)..(1296)

<220>  
 <221> mat\_peptide

<222> (67)..(1296)

<400> 3

ES 2 614 744 T3

atg Met	agt Ser	aca Thr -20	ctg Leu	atc Ile	att Ile	cgt Arg	tta Leu -15	ttg Leu	ttc Phe	tta Leu	acg Thr	att Ile -10	ata Ile	ttg Leu	gcc Ala	48
cct Pro	gtt Val -5	tca Ser	tta Leu	cgc Arg	gcc Ala -1	gat Asp 1	gaa Glu	cag Gln	agc Ser	gga Gly 5	atg Met	cag Gln	ctt Leu	gag Glu	cgt Arg 10	96
gtt Val	gtc Val	atc Ile	gtc Val	agt Ser 15	cgt Arg	cat His	ggc Gly	gtc Val	agg Arg 20	gca Ala	ccg Pro	aca Thr	aag Lys	ttc Phe 25	acg Thr	144
ccg Pro	ctt Leu	atg Met	cag Gln 30	caa Gln	gtc Val	act Thr	ccc Pro	gac Asp 35	cgc Arg	tgg Trp	ccg Pro	caa Gln	tgg Trp 40	gac Asp	gtt Val	192
cct Pro	ctg Leu	ggg Gly 45	tgg Trp	ttg Leu	act Thr	cct Pro	cgc Arg 50	ggc Gly	ggg Gly	gca Ala	ctc Leu	att Ile 55	act Thr	gag Glu	tta Leu	240
gga Gly	cgg Arg 60	tat Tyr	caa Gln	cgt Arg	tta Leu 65	cgc Arg	ctg Leu	gcg Ala	gac Asp	aaa Lys	ggt Gly 70	ctg Leu	ctg Leu	gat Asp	aat Asn	288
aaa Lys 75	acg Thr	tgt Cys	cca Pro	acg Thr	gca Ala 80	ggg Gly	cag Gln	gtc Val	gcg Ala	gtc Val 85	att Ile	gcc Ala	gat Asp	agc Ser	gat Asp 90	336
caa Gln	cgt Arg	acc Thr	cgt Arg	aaa Lys 95	acg Thr	ggt Gly	gaa Glu	gca Ala	ttc Phe 100	ctg Leu	gca Ala	gga Gly	ctg Leu	gct Ala 105	ccg Pro	384
gaa Glu	tgt Cys	aaa Lys	gta Val 110	cag Gln	gtt Val	tat Tyr	tat Tyr	caa Gln 115	caa Gln	gat Asp	aag Lys	tca Ser	aaa Lys 120	tct Ser	gat Asp	432
ccc Pro	ctt Leu	ttt Phe 125	aat Asn	ccc Pro	atc Ile	aag Lys	gcg Ala 130	ggg Gly	cgg Arg	tgt Cys	tcg Ser	ctg Leu 135	aac Asn	aca Thr	tcg Ser	480
cag Gln	gtg Val 140	aaa Lys	gag Glu	gcc Ala	atc Ile	ctg Leu 145	acc Thr	cgg Arg	gct Ala	ggc Gly	gga Gly 150	agt Ser	ctt Leu	gat Asp	gag Glu	528
tac Tyr 155	acg Thr	cgc Arg	cac His	tac Tyr	caa Gln 160	ccc Pro	gca Ala	ttt Phe	caa Gln	gcc Ala 165	ctg Leu	gaa Glu	cgg Arg	gtg Val	tta Leu 170	576
aat Asn	ttc Phe	tcc Ser	cag Gln	tca Ser 175	gaa Glu	aag Lys	tgt Cys	caa Gln	gca Ala 180	gct Ala	ggg Gly	cag Gln	tct Ser	gca Ala 185	cag Gln	624
tgt Cys	acg Thr	cta Leu	acc Thr 190	gac Asp	gtc Val	tta Leu	cct Pro	gct Ala 195	gaa Glu	ctc Leu	aag Lys	gtc Val	tct Ser 200	cca Pro	gaa Glu	672
aat Asn	ata Ile	tcg Ser 205	ttg Leu	tca Ser	ggc Gly	tca Ser	tgg Trp 210	gga Gly	ctg Leu	gct Ala	tca Ser	acc Thr 215	ctg Leu	acg Thr	gaa Glu	720
atc Ile	ttc Phe 220	ctg Leu	ctg Leu	caa Gln	caa Gln	gca Ala 225	caa Gln	ggg Gly	atg Met	tcg Ser	cag Gln 230	gtg Val	gcc Ala	tgg Trp	ggg Gly	768
cgt	att	cat	ggc	gat	aaa	gaa	tgg	cgt	aca	tta	tta	agt	ctg	cac	aat	816

ES 2 614 744 T3

Arg 235	Ile	His	Gly	Asp	Lys 240	Glu	Trp	Arg	Thr	Leu 245	Leu	Ser	Leu	His	Asn 250		
gcg Ala	cag Gln	ttt Phe	gac Asp	ctt Leu 255	ctg Leu	cag Gln	aaa Lys	acc Thr	ccg Pro 260	gag Glu	ggt Val	gcc Ala	cgf Arg	agc Ser 265	agg Arg	864	
gcc Ala	aca Thr	ccg Pro	tta Leu 270	ctt Leu	gat Asp	ttg Leu	ata Ile	cgf Arg 275	aca Thr	gca Ala	ctc Leu	gta Val	aca Thr 280	cag Gln	ggg Gly	912	
gca Ala	aca Thr	gaa Glu 285	aat Asn	aaa Lys	tac Tyr	gca Ala	att Ile 290	cag Gln	ttg Leu	ccc Pro	gtc Val	tct Ser 295	ttg Leu	ttg Leu	ttt Phe	960	
att Ile	gcg Ala 300	ggg Gly	cat His	gac Asp	acc Thr	aat Asn 305	ctt Leu	gcc Ala	aat Asn	atc Ile	agc Ser 310	ggg Gly	gca Ala	tta Leu	ggc Gly	1008	
ctt Leu 315	aac Asn	gtg Val	ttt Phe	ctg Leu	ccc Pro 320	ggt Gly	cag Gln	cca Pro	gat Asp	aat Asn 325	acg Thr	ccg Pro	ccg Pro	ggt Gly	gga Gly 330	1056	
gag Glu	ttt Phe	ggt Val	ttc Phe	gaa Glu 335	agg Arg	tgg Trp	aaa Lys	cgg Arg	gtc Val 340	agc Ser	gat Asp	cat His	tct Ser	gat Asp 345	tgg Trp	1104	
gtg Val	cag Gln	ggt Val	tct Ser 350	ttt Phe	atg Met	tat Tyr	cag Gln	aca Thr 355	ttg Leu	cag Gln	gaa Glu	atg Met	cgf Arg 360	gat Asp	atg Met	1152	
caa Gln	cct Pro	ttg Leu 365	tcg Ser	ttg Leu	caa Gln	tcg Ser	cct Pro 370	ccc Pro	gga Gly	aaa Lys	att Ile	gtg Val 375	ctg Leu	ccc Pro	tta Leu	1200	
gcg Ala	gcc Ala 380	tgc Cys	gat Asp	gag Glu	aaa Lys	aat Asn 385	acg Thr	cag Gln	gga Gly	atg Met	tgc Cys 390	tca Ser	tta Leu	aaa Lys	aat Asn	1248	
ttt Phe 395	tct Ser	gca Ala	ctg Leu	att Ile	gat Asp 400	tcc Ser	ggt Val	cgc Arg	gtg Val	tcc Ser 405	gaa Glu	tgt Cys	gct Ala	gag Glu	aaa Lys 410	1296	

<210> 4

<211> 432

<212> PRT

<213> citrobacter gilleni DSM 13694

<400> 4

Met Ser Thr Leu Ile Ile Arg Leu Leu Phe Leu Thr Ile Ile Leu Ala  
 -20 -15 -10

Pro Val Ser Leu Arg Ala Asp Glu Gln Ser Gly Met Gln Leu Glu Arg  
 -5 -1 1 5 10

Val Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Phe Thr  
 15 20 25

Pro Leu Met Gln Gln Val Thr Pro Asp Arg Trp Pro Gln Trp Asp Val  
 30 35 40

Pro Leu Gly Trp Leu Thr Pro Arg Gly Gly Ala Leu Ile Thr Glu Leu

ES 2 614 744 T3

45 50 ----- 55

Gly Arg Tyr Gln Arg Leu Arg Leu Ala Asp Lys Gly Leu Leu Asp Asn  
60 65 70

Lys Thr Cys Pro Thr Ala Gly Gln Val Ala Val Ile Ala Asp Ser Asp  
75 80 85 90

Gln Arg Thr Arg Lys Thr Gly Glu Ala Phe Leu Ala Gly Leu Ala Pro  
95 100 105

Glu Cys Lys Val Gln Val Tyr Tyr Gln Gln Asp Lys Ser Lys Ser Asp  
110 115 120

Pro Leu Phe Asn Pro Ile Lys Ala Gly Arg Cys Ser Leu Asn Thr Ser  
125 130 135

Gln Val Lys Glu Ala Ile Leu Thr Arg Ala Gly Gly Ser Leu Asp Glu  
140 145 150

Tyr Thr Arg His Tyr Gln Pro Ala Phe Gln Ala Leu Glu Arg Val Leu  
155 160 165 170

Asn Phe Ser Gln Ser Glu Lys Cys Gln Ala Ala Gly Gln Ser Ala Gln  
175 180 185

Cys Thr Leu Thr Asp Val Leu Pro Ala Glu Leu Lys Val Ser Pro Glu  
190 195 200

Asn Ile Ser Leu Ser Gly Ser Trp Gly Leu Ala Ser Thr Leu Thr Glu  
205 210 215

Ile Phe Leu Leu Gln Gln Ala Gln Gly Met Ser Gln Val Ala Trp Gly  
220 225 230

Arg Ile His Gly Asp Lys Glu Trp Arg Thr Leu Leu Ser Leu His Asn  
235 240 245 250

Ala Gln Phe Asp Leu Leu Gln Lys Thr Pro Glu Val Ala Arg Ser Arg  
255 260 265

Ala Thr Pro Leu Leu Asp Leu Ile Arg Thr Ala Leu Val Thr Gln Gly  
270 275 280

Ala Thr Glu Asn Lys Tyr Ala Ile Gln Leu Pro Val Ser Leu Leu Phe  
285 290 295

Ile Ala Gly His Asp Thr Asn Leu Ala Asn Ile Ser Gly Ala Leu Gly  
300 305 310

Leu Asn Val Phe Leu Pro Gly Gln Pro Asp Asn Thr Pro Pro Gly Gly  
Page 7

ES 2 614 744 T3

```

315                               320          325          330
Glu Phe Val Phe Glu Arg Trp Lys Arg Val Ser Asp His Ser Asp Trp
      335
Val Gln Val Ser Phe Met Tyr Gln Thr Leu Gln Glu Met Arg Asp Met
      350      355      360
Gln Pro Leu Ser Leu Gln Ser Pro Pro Gly Lys Ile Val Leu Pro Leu
      365      370      375
Ala Ala Cys Asp Glu Lys Asn Thr Gln Gly Met Cys Ser Leu Lys Asn
      380      385      390
Phe Ser Ala Leu Ile Asp Ser Val Arg Val Ser Glu Cys Ala Glu Lys
      395      400      405      410

```

<210> 5  
 <211> 1311  
 <212> DNA  
 <213> citrobacter amalonaticus ATCC 25407

<220>  
 <221> CDS  
 <222> (1)..(1308)

<220>  
 <221> mat\_peptide  
 <222> (67)..(1308)

<400> 5

ES 2 614 744 T3

atg	aat	acg	cta	ctt	ttt	cga	tta	ata	atg	ttt	ata	ttc	atg	ttt	ggt	48
Met	Asn	Thr	Leu	Leu	Phe	Arg	Leu	Ile	Met	Phe	Ile	Phe	Met	Phe	Gly	
		-20					-15					-10				
tct	ttc	cca	tta	cag	gcg	gaa	gtg	cca	gat	gac	atg	aag	ctt	gaa	cga	96
Ser	Phe	Pro	Leu	Gln	Ala	Glu	Val	Pro	Asp	Asp	Met	Lys	Leu	Glu	Arg	
	-5				-1	1				5					10	
gtt	gtg	ata	gta	agt	cgc	cac	ggt	gta	aga	gca	cca	aca	aag	ttc	acc	144
Val	Val	Ile	Val	Ser	Arg	His	Gly	Val	Arg	Ala	Pro	Thr	Lys	Phe	Thr	
				15					20					25		
cca	ttg	atg	cag	gaa	atc	aca	cct	tac	cat	tgg	ccg	caa	tgg	gat	gtt	192
Pro	Leu	Met	Gln	Glu	Ile	Thr	Pro	Tyr	His	Trp	Pro	Gln	Trp	Asp	Val	
			30					35					40			
ccc	ctg	ggc	tgg	ttg	acg	gct	cgg	ggt	ggt	gag	ctc	gtc	acc	gaa	atg	240
Pro	Leu	Gly	Trp	Leu	Thr	Ala	Arg	Gly	Gly	Glu	Leu	Val	Thr	Glu	Met	
		45					50					55				
gga	cga	tat	caa	caa	aaa	gta	tta	atc	gat	aac	ggc	gtt	ctg	gaa	agt	288
Gly	Arg	Tyr	Gln	Gln	Lys	Val	Leu	Ile	Asp	Asn	Gly	Val	Leu	Glu	Ser	
	60					65					70					
aat	gta	tgt	ccg	tca	cca	gaa	cag	gtg	gca	gtt	att	gcc	gat	acc	gat	336
Asn	Val	Cys	Pro	Ser	Pro	Glu	Gln	Val	Ala	Val	Ile	Ala	Asp	Thr	Asp	
					80					85					90	
cag	cgc	act	cgT	aaa	acc	ggt	gag	gca	ttt	ctg	gct	gga	ttt	gcg	ccg	384
Gln	Arg	Thr	Arg	Lys	Thr	Gly	Glu	Ala	Phe	Leu	Ala	Gly	Phe	Ala	Pro	



ES 2 614 744 T3

95										100					105					
gga	tgt	aaa	aat	aag	gtt	cat	tat	caa	aaa	gat	cac	gat	aaa	aaa	gat	432				
Gly	Cys	Lys	Asn	Lys	Val	His	Tyr	Gln	Lys	Asp	His	Asp	Lys	Lys	Asp					
			110					115					120							
cct	ctt	ttt	aat	cca	gta	aaa	atg	ggg	gtg	tgc	gct	ttt	aat	gta	caa	480				
Pro	Leu	Phe	Asn	Pro	Val	Lys	Met	Gly	Val	Cys	Ala	Phe	Asn	Val	Gln					
		125					130					135								
aaa	act	cag	gaa	gcg	att	ctg	aca	cgt	gcg	gaa	gga	aac	att	gaa	cgg	528				
Lys	Thr	Gln	Glu	Ala	Ile	Leu	Thr	Arg	Ala	Glu	Gly	Asn	Ile	Glu	Arg					
	140					145				150										
tac	act	cag	cgt	tat	gac	tct	gca	ttc	cgt	act	ctg	gaa	cag	gtt	ctc	576				
Tyr	Thr	Gln	Arg	Tyr	Asp	Ser	Ala	Phe	Arg	Thr	Leu	Glu	Gln	Val	Leu					
155					160					165					170					
aat	ttc	tcc	cgg	tca	gca	gca	tgc	cga	tca	gca	agc	cag	tct	ggt	tgc	624				
Asn	Phe	Ser	Arg	Ser	Ala	Ala	Cys	Arg	Ser	Ala	Ser	Gln	Ser	Gly	Cys					
				175					180					185						
acg	cta	cca	gga	acc	tta	cct	tca	gaa	ctc	agg	gtt	tct	gcg	gat	acc	672				
Thr	Leu	Pro	Gly	Thr	Leu	Pro	Ser	Glu	Leu	Arg	Val	Ser	Ala	Asp	Thr					
			190					195					200							
gtt	tcc	tta	tct	ggc	gcg	tgg	agt	ctt	tct	tcc	atg	ctg	acg	gaa	ata	720				
Val	Ser	Leu	Ser	Gly	Ala	Trp	Ser	Leu	Ser	Ser	Met	Leu	Thr	Glu	Ile					
		205					210					215								
ttt	cta	ttg	caa	gag	gcg	cag	gga	atg	cca	gag	ggt	gcg	tgg	ggg	cga	768				
Phe	Leu	Leu	Gln	Glu	Ala	Gln	Gly	Met	Pro	Glu	Val	Ala	Trp	Gly	Arg					
	220					225					230									
att	cat	ggg	gag	aaa	gaa	tgg	aca	gcg	tta	tta	agt	ctg	cat	aat	gct	816				
Ile	His	Gly	Glu	Lys	Glu	Trp	Thr	Ala	Leu	Leu	Ser	Leu	His	Asn	Ala					
235					240					245					250					
cag	ttt	gac	ctt	ttg	caa	aga	act	ccc	gaa	ggt	gcc	cgc	agc	aga	gca	864				
Gln	Phe	Asp	Leu	Leu	Gln	Arg	Thr	Pro	Glu	Val	Ala	Arg	Ser	Arg	Ala					
				255					260					265						
aca	cca	tta	ctc	gat	ttg	atc	agc	gaa	gca	tta	gtg	agt	aat	ggg	tca	912				
Thr	Pro	Leu	Leu	Asp	Leu	Ile	Ser	Glu	Ala	Leu	Val	Ser	Asn	Gly	Ser					
			270					275					280							
aca	gaa	aat	cat	tac	gga	att	aaa	tta	ccc	gtc	tca	tta	ttg	ttt	att	960				
Thr	Glu	Asn	His	Tyr	Gly	Ile	Lys	Leu	Pro	Val	Ser	Leu	Leu	Phe	Ile					
		285					290					295								
gct	ggt	cat	gat	acc	aat	ctt	gca	aat	ctc	agt	ggg	gta	ttt	gat	ctt	1008				
Ala	Gly	His	Asp	Thr	Asn	Leu	Ala	Asn	Leu	Ser	Gly	Val	Phe	Asp	Leu					
	300					305					310									
aac	tgg	tct	cta	cct	ggg	cag	cca	gat	aat	aca	cct	cct	ggc	ggg	gag	1056				
Asn	Trp	Ser	Leu	Pro	Gly	Gln	Pro	Asp	Asn	Thr	Pro	Pro	Gly	Gly	Glu					
					320					325					330					
ctg	gtt	ttc	gaa	aga	tgg	acg	cga	gtg	agt	gat	aac	act	gac	tgg	att	1104				
Leu	Val	Phe	Glu	Arg	Trp	Thr	Arg	Val	Ser	Asp	Asn	Thr	Asp	Trp	Ile					
				335					340					345						
caa	att	tcg	ttt	gtt	tat	cag	act	ctt	caa	caa	atg	cgt	aag	ttt	aaa	1152				
Gln	Ile	Ser	Phe	Val	Tyr	Gln	Thr	Leu	Gln	Gln	Met	Arg	Lys	Phe	Lys					
			350					355					360							
cct	ttt	tca	tct	tcg	tct	ctc	cca	aac	aag	att	gtg	ctt	acg	ttg	ccc	1200				
Pro	Phe	Ser	Ser	Ser	Ser	Leu	Pro	Asn	Lys	Ile	Val	Leu	Thr	Leu	Pro					

ES 2 614 744 T3

```

          365          370 ----- 375
tct tgc cag gat aaa aat cct gag ggt atg tgt cca tta aag cat ttt 1248
Ser Cys Gln Asp Lys Asn Pro Glu Gly Met Cys Pro Leu Lys His Phe
    380          385          390

att gac att gtg cag aca gca cgt att cca caa tgt gca gtg atg gct 1296
Ile Asp Ile Val Gln Thr Ala Arg Ile Pro Gln Cys Ala Val Met Ala
    395          400          405          410

gat gta aac cgt taa 1311
Asp Val Asn Arg

```

<210> 6  
 <211> 436  
 <212> PRT  
 <213> *Citrobacter amalonaticus* ATCC 25407  
 <400> 6

```

Met Asn Thr Leu Leu Phe Arg Leu Ile Met Phe Ile Phe Met Phe Gly
      -20          -15          -10

Ser Phe Pro Leu Gln Ala Glu Val Pro Asp Asp Met Lys Leu Glu Arg
   -5          -1  1          5          10

Val Val Ile Val Ser Arg His Gly Val Arg Ala Pro Thr Lys Phe Thr
      15          20          25

Pro Leu Met Gln Glu Ile Thr Pro Tyr His Trp Pro Gln Trp Asp Val
      30          35          40

Pro Leu Gly Trp Leu Thr Ala Arg Gly Gly Glu Leu Val Thr Glu Met
      45          50          55

Gly Arg Tyr Gln Gln Lys Val Leu Ile Asp Asn Gly Val Leu Glu Ser
      60          65          70

Asn Val Cys Pro Ser Pro Glu Gln Val Ala Val Ile Ala Asp Thr Asp
      75          80          85          90

Gln Arg Thr Arg Lys Thr Gly Glu Ala Phe Leu Ala Gly Phe Ala Pro
      95          100          105

Gly Cys Lys Asn Lys Val His Tyr Gln Lys Asp His Asp Lys Lys Asp
      110          115          120

Pro Leu Phe Asn Pro Val Lys Met Gly Val Cys Ala Phe Asn Val Gln
      125          130          135

Lys Thr Gln Glu Ala Ile Leu Thr Arg Ala Glu Gly Asn Ile Glu Arg
      140          145          150

Tyr Thr Gln Arg Tyr Asp Ser Ala Phe Arg Thr Leu Glu Gln Val Leu
      155          160          165          170

```

ES 2 614 744 T3

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

<210> 7  
 <211> 29  
 <212> DNA  
 <213> Artificial

<220>  
<223> Primer

<220>  
<221> misc\_feature  
<222> (19)..(19)  
<223> m is A, or C

<220>  
<221> misc\_feature  
<222> (21)..(21)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (24)..(24)  
<223> y is C, or T

<220>  
<221> misc\_feature  
<222> (27)..(27)  
<223> n is a, c, g, or t

<400> 7  
cgcggtgga ttgtgccmg ncayggngt 29

<210> 8  
<211> 28  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<220>  
<221> misc\_feature  
<222> (20)..(20)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (23)..(23)  
<223> d is A, G, or T

<220>  
<221> misc\_feature  
<222> (26)..(26)  
<223> r is A, or G

<400> 8  
ccaggtggt atcatggccn godatraa 28

<210> 9  
<211> 17  
<212> DNA  
<213> Artificial

<220>

<223> Primer

<400> 9  
agacttccgc cagcccc 17

<210> 10  
<211> 19  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 10  
aagcagctgg gcagtctgc 19

<210> 11  
<211> 21  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 11  
aagcggcgtg aacttgctg g 21

<210> 12  
<211> 22  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 12  
atggggactg gctcaacc tg 22

<210> 13  
<211> 29  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<220>  
<221> misc\_feature  
<222> (19)..(19)  
<223> m is A, or C

<220>  
<221> misc\_feature  
<222> (21)..(21)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature

# ES 2 614 744 T3

<222> (24)..(24)  
<223> y is C, or T

<220>  
<221> misc\_feature  
<222> (27)..(27)  
<223> n is a, c, g, or t

<400> 13  
cgcgtggtga ttgtgtccmg ncayggngt 29

<210> 14  
<211> 28  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<220>  
<221> misc\_feature  
<222> (20)..(20)  
<223> n is a, c, g, or t

<220>  
<221> misc\_feature  
<222> (23)..(23)  
<223> d is A, G, or T

<220>  
<221> misc\_feature  
<222> (26)..(26)  
<223> r is A, or G

<400> 14  
ccaggttggt atcatggccn gcdatraa 28

<210> 15  
<211> 20  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 15  
tttacgagtg cgctgatcgg 20

<210> 16  
<211> 25  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 16  
ataactgcca cctgttctgg tgacg 25

# ES 2 614 744 T3

<210> 17  
<211> 22  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 17  
actcagggtt tctgcggata cc 22

<210> 18  
<211> 21  
<212> DNA  
<213> Artificial

<220>  
<223> Primer

<400> 18  
aatgccagag gttgcgtggg g 21

## REIVINDICACIONES

1. Polipéptido aislado con actividad de fitasa y una estabilidad al ácido de al menos 60 % de actividad residual tras la incubación durante la noche a 37 °C en el tampón de glicina/ácido clorhídrico pH 2,2, en relación con la actividad residual tras la incubación durante la noche a 37 °C en el tampón HEPES pH 7,0, seleccionado del grupo que consiste en:
- 5 (a) un polipéptido con una secuencia de aminoácidos que tiene al menos un 80 % de identidad con
- (i) aminoácidos 1 a 414 de SEQ ID n.º: 6 y/o
- 10 (ii) la parte del polipéptido maduro de SEQ ID n.º: 6;
- donde el grado de identidad se determina por el programa "align" utilizando la matriz de puntuación por defecto BLOSUM50 con penalización para el primer residuo en un espacio de -12 mientras las penalizaciones para residuos adicionales en un espacio son -2.
2. Polinucleótido aislado que comprende una secuencia de nucleótidos que codifica el polipéptido según la reivindicación 1.
- 15 3. Polinucleótido aislado que codifica un polipéptido con actividad de fitasa y una estabilidad al ácido de al menos 60 % de actividad residual tras la incubación durante la noche a 37 °C en el tampón de glicina/ácido clorhídrico pH 2,2, con relación a la actividad residual tras la incubación durante la noche a 37 °C en el tampón HEPES pH 7,0, seleccionado del grupo que consiste en:
- 20 (a) un polinucleótido que codifica un polipéptido con una secuencia de aminoácidos que presenta al menos un 80 % de identidad con aminoácidos 1 a 414 de SEQ ID n.º: 6; y
- (b) un polinucleótido con al menos un 80 % de identidad con nucleótidos 67 a 1308 de SEQ ID n.º: 5, donde el grado de identidad se determina por el programa "align" que es un alineamiento Needleman-Wunsch (es decir, alineamiento global), utilizando la matriz de identidad por defecto, la penalización para el primer residuo en un espacio es -16 mientras las penalizaciones para residuos adicionales en un espacio son -4.
- 25 4. Polinucleótido aislado de cualquiera de las reivindicaciones 2 y 3, con al menos una mutación en la secuencia que codifica el polipéptido maduro de SEQ ID n.º: 5, donde la secuencia de nucleótidos mutante codifica un polipéptido que comprende aminoácidos 1 a 414 de SEQ ID n.º: 6.
- 30 5. Constructo de ácidos nucleicos que comprende el polinucleótido según cualquiera de las reivindicaciones 2-4 operativamente enlazado a una o más secuencias de control que dirigen la producción del polipéptido en un huésped de expresión.
- 35 6. Vector de expresión recombinante que comprende el constructo de ácidos nucleicos según la reivindicación 5.
7. Célula huésped recombinante que comprende el constructo de ácidos nucleicos según la reivindicación 5.
- 40 8. Método para la producción del polipéptido, según la reivindicación 1, que comprende (a) el cultivo de una célula, que en su forma tipo salvaje es capaz de producir el polipéptido, bajo condiciones propicias para la producción del polipéptido; y (b) la recuperación del polipéptido.
- 45 9. Método para la producción del polipéptido según la reivindicación 1 que comprende (a) el cultivo de una célula huésped recombinante que comprende un constructo de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica el polipéptido bajo condiciones propicias para la producción del polipéptido; y (b) la recuperación del polipéptido.
- 50 10. Planta, parte de planta o célula vegetal transgénica que comprende el polipéptido, según la reivindicación 1, que ha sido transformada con un polinucleótido que codifica el polipéptido, según la reivindicación 1.
11. Uso de al menos un polipéptido, según la reivindicación 1, en el alimento para animales.
12. Uso de al menos un polipéptido, según la reivindicación 1, en la preparación de una composición para usar en el alimento para animales.
- 55 13. Método para mejorar el valor nutricional de un alimento para animales, donde al menos un polipéptido, según la reivindicación 1, se añade al alimento.
- 60 14. Aditivo de alimento para animales que comprende
- (a) al menos un polipéptido, según la reivindicación 1; y
- (b) al menos una vitamina soluble en grasa,
- (c) al menos una vitamina soluble en agua y/o
- (d) al menos un oligoelemento.
- 65 15. Aditivo de alimento para animales, según la reivindicación 14, que comprende además al menos una amilasa, al



menos una fitasa adicional, al menos una xilanasas, al menos una galactanasas, al menos una alfa-galactosidasas, al menos una proteasa, al menos una fosfolipasa y/o al menos una beta-glucanasas.

- 5 16. Composición de alimento para animales con un contenido en proteína bruto de 50 a 800 g/kg y que comprende al menos un polipéptido, según la reivindicación 1.