

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 649 217**

51 Int. Cl.:

**A23K 10/16** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.07.2014 PCT/EP2014/066078**

87 Fecha y número de publicación internacional: **29.01.2015 WO15011276**

96 Fecha de presentación y número de la solicitud europea: **25.07.2014 E 14744322 (0)**

97 Fecha y número de publicación de la concesión europea: **06.09.2017 EP 3030093**

54 Título: **Polipéptidos que tienen actividad de  $\alpha$ -L-galactosidasa y polinucleótidos que los codifican**

30 Prioridad:

**26.07.2013 EP 13178257**

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

**10.01.2018**

73 Titular/es:

**NOVOZYMES A/S (100.0%)  
Krogshoejvej 36  
2880 Bagsvaerd, DK**

72 Inventor/es:

**KROGH, KRISTIAN BERTEL RØMER M.;  
EKLÖF, JENS MAGNUS;  
JOHANSEN, KATJA SALOMON;  
ROGOWSKI, ARTUR;  
BOLAM, DAVID N. y  
GILBERT, HARRY J.**

74 Agente/Representante:

**TOMAS GIL, Tesifonte Enrique**

ES 2 649 217 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 que tienen actividad de  $\alpha$ -L-galactosidasa y polinucleótidos que los codifican.

5 Referencia a un listado de secuencias

[0001] Esta solicitud contiene un listado de secuencias en formato legible por ordenador.

10 Referencia a un depósito de material biológico

[0002] Esta solicitud contiene una referencia a un depósito público de material biológico, (ATCC 8483).

Antecedentes de la invención

15 Campo de la invención

[0003] La presente invención se refiere a polipéptidos que tienen actividad de  $\alpha$ -L-galactosidasa y un dominio catalítico que pertenece a la familia de las glucósido hidrolasas 95 (GH95, [www.cazy.org](http://www.cazy.org) PMID:18838391). También se describen construcciones de ácidos nucleicos, vectores, y células huésped que comprenden los polinucleótidos así como métodos para producir y utilizar los polipéptidos y el dominio catalítico.

20

Descripción de la técnica relacionada

[0004] La presencia de fracciones de azúcar  $\alpha$ -1-galactosil en xilano de maíz se ha descrito en la literatura (Allerdings et al. 2006, *Phytochemistry* 67: 1276-1286) pero una enzima responsable de la liberación de esta fracción de azúcar no se ha descrito anteriormente. La misma fracción de azúcar también se puede encontrar en xiloglucano de semillas de jojoba (Hantus et al. 1997, *Carbohydr. Res.* 304: 11-20) y en xiloglucano y ramnogalacturonano II de mutantes mur1 de Arabidopsis (Zablackis et al. 1996, *Science* 272: 1808-1810, Reuhs et al. 2004, *Planta* 219: 147-157) La presente invención proporciona polipéptidos de GH95 que tienen actividad de  $\alpha$ -L-galactosidasa y polinucleótidos que codifican los polipéptidos, especialmente la  $\alpha$ -L-galactosidasa GH95 designada como BACOVA 03438.

30

[0005] Se han publicado varias secuencias en la base de datos UNIPROT (referencia) que están indicadas como glicosil hidrolasas que pertenecen a la familia 31 (EC = 3.2.1.-;) o simplemente como "proteína". Esto también incluye las SEQ ID NO: 2 y 3 como se ejemplifica aquí. Las identidades de secuencia de estas secuencias respecto a la SEQ ID NO: 2 varían del 100% al 97,8%. Otras secuencias relacionadas (por identidad de secuencia) están por debajo del 77,2% de identidad.

35

[0006] De este modo, la SEQ ID NO: 2 está publicada como UNIPROT: A7M011 con la referencia Fulton, L Clifton, S Fulton, B Xu, J Minx, P Pepin, KH Johnson, M Thiruvilangam, P Bhonagiri, V Nash, WE Mardis, ER Wilson, RK. Introducida (MAR-2007) en las bases de datos EMBL/GenBank/DDBJ, y Sudarsanam, P Ley, R Guruge, J Turnbaugh, PJ Mahowald, M Liep, D Gordon, J. Borrador de secuencia de genoma de *Bacteroides ovatus* (ATCC 8483). Introducida (APR-2007) en las bases de datos EMBL/GenBank/DDBJ.

40

[0007] También se describen secuencias que tienen entre un 100 y un 92,6% de identidad con la SEQ ID NO: 2, por ejemplo UNIPROT: I8Z241 (100% *Bacteroides ovatus* CL03T12C18), UNIPROT: I8YAG0 (99.9% *Bacteroides ovatus* CL02T12C04), UNIPROT: 19US19 (98.9% *Bacteroides xylanisolvens* CL03T12C04), UNIPROT: D0TM59 (*Bacteroides* sp. Cepa 2\_1\_22), y algunas más.

45

50 Resumen de la invención

[0008] La presente invención se refiere al uso en pienso para animales de polipéptidos aislados que tienen actividad de  $\alpha$ -L-galactosidasa seleccionados del grupo que consiste en:

55 (a) un polipéptido que tiene al menos 80% de identidad de secuencia con el polipéptido maduro de la SEQ ID NO: 2;

(b) un polipéptido codificado por un polinucleótido que se hibrida bajo condiciones de astringencia alta, o muy alta con (i) la secuencia que codifica el polipéptido maduro de la SEQ ID NO: 1 o (ii) el complemento de longitud completa de (i);

60 (c) un polipéptido codificado por un polinucleótido que tiene al menos un 92% de identidad de secuencia con la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1

(d) una variante del polipéptido maduro de la SEQ ID NO: 2 que comprende una sustitución, deleción y/o inserción en una o más (por ejemplo, varias) posiciones; y

(e) un fragmento del polipéptido de (a), (b), (c) o (d) que tiene actividad de  $\alpha$ -L-galactosidasa.

65

[0009] La presente invención también se refiere a polipéptidos aislados que comprenden un dominio catalítico

seleccionado del grupo que consiste en:

- a) un dominio catalítico que tiene al menos un 80% de identidad de secuencia con los aminoácidos 20 a 811 de la SEQ ID NO: 2
- 5 b) un dominio catalítico codificado por un polinucleótido que se hibrida en condiciones de astringencia alta o muy alta con (i) los nucleótidos 60 a 2436 de la SEQ ID NO: 1, o (iii) el complemento de longitud completa de (i) o (ii);
- c) un dominio catalítico codificado por un polinucleótido que tiene al menos un 92% de identidad de secuencia con los nucleótidos 60 a 2436 de la SEQ ID NO: 1
- 10 d) una variante de los aminoácidos 20 a 811 de la SEQ ID NO: 2 que comprende una sustitución, deleción y/o inserción en una o más (por ejemplo, varias) posiciones; y
- e) un fragmento del dominio catalítico de (a), (b), (c) o (d) que tiene actividad de  $\alpha$ -L-galactosidasa.

15 [0010] La presente invención también se refiere a polinucleótidos aislados que codifican los polipéptidos de la presente invención; construcciones de ácidos nucleicos; vectores de expresión recombinantes; células huésped recombinantes que comprenden los polinucleótidos; y métodos para producir los polipéptidos.

20 [0011] La presente invención también se refiere a polipéptidos variantes que tienen actividad de  $\alpha$ -L-galactosidasa y que tienen al menos 80%, por ejemplo al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, o al menos 99% de identidad de secuencia con la SEQ ID NO: 3 que comprende al menos una sustitución, deleción y/o inserción de al menos uno o más (varios) aminoácidos de la SEQ ID NO: 3.

25 [0012] La presente invención se refiere además a composiciones que comprenden un polipéptido aislado que tiene actividad de  $\alpha$ -L-galactosidasa, seleccionado del grupo que consiste en:

- (a) un polipéptido que tiene al menos 80%, por ejemplo, al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, o 100% de identidad de secuencia con la SEQ ID NO: 3;
- (b) un polipéptido codificado por un polinucleótido que se hibrida bajo condiciones de astringencia alta o condiciones de astringencia muy alta con:

- 35 (i) la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1; y/o
- (ii) la cadena complementaria de longitud completa de (i);

- (c) un polipéptido codificado por un polinucleótido que tiene al menos 80%, por ejemplo, al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, o 100% de identidad de secuencia con la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1;
- 40 (d) una variante que comprende una sustitución, deleción y/o inserción de uno o más (varios) aminoácidos de la SEQ ID NO: 3; y
- (e) un fragmento de un polipéptido de (a), (b), (c) o (d), que tiene actividad de  $\alpha$ -L-galactosidasa.

45 [0013] Las composiciones pueden ser composiciones de pienso para animales. La presente invención también se refiere a polinucleótidos aislados que codifican los polipéptidos de la presente invención, construcciones de ácidos nucleicos, vectores de expresión recombinantes, células huésped recombinantes que comprenden los polinucleótidos, y a métodos de producción recombinante de los polipéptidos. La presente invención también se refiere a métodos para preparar una composición para uso en pienso para animales, a métodos para mejorar el valor nutricional de un pienso para animales, y a métodos para tratar proteínas para usar en composiciones para piensos para animales. En este documento también se describe un polinucleótido que codifica un péptido señal que comprende o que consiste en los aminoácidos 1 a 20 de la SEQ ID NO: 2, un polinucleótido que codifica un propéptido que comprende o consiste en los aminoácidos 1 a 20 de la SEQ ID NO: 2, o un polinucleótido que codifica un péptido señal y un propéptido que comprende o consiste en los aminoácidos 1 a 20 de la SEQ ID NO: 2, cada uno de los cuales está operativamente unido a un gen que codifica una proteína; construcciones de ácidos nucleicos, vectores de expresión y células huésped recombinantes que comprenden los polinucleótidos; y métodos para producir una proteína usada en la invención.

60 Breve descripción de las figuras

[0014]

65 La Fig. 1 muestra que la enzima de esta invención no hidroliza ninguno de los arabinosilanos o glucuronosilanos comercialmente disponibles testados. La figura muestra una cromatografía en capa fina que indica actividad de la enzima de la invención GH95 de *B.ovatus* (BACOVA\_03438) en xilano de salvado de maíz. GH95 de *B.ovatus* libera un solo producto del xilano de maíz (línea C). Línea (A) X1 a X5

mezcla de estándares de xilooligosacáridos. Línea (B) xilano de salvado de maíz sin tratar.

La Fig. 2 muestra una cromatografía líquida de alto rendimiento (HPLC) que muestra que la enzima de esta invención, GH95 (BACOVA 03438), libera L-galactosa y no fucosa de xilano de salvado de maíz.

5 La Fig. 3 muestra que la enzima GH95 de *B.ovatus* de la invención es -1000 veces más activa en xilano de maíz que GH95 de *B.bifidum*.

La Fig. 4 muestra la diferencia entre índices relativos de la GH95 (BACOVA 03438) y GH95 de *B.bifidum* (AAQ72464) contra xilano de salvado de maíz y 2'-fucosil lactosa. La  $\alpha$ -L-galactosidasa GH95 de *B. Ovatus* es al menos 200 veces más activa en xilano de maíz que la  $\alpha$ -fucosidasa GH95 de *B.bifidum* (panel A).

10 Índice relativo de las dos enzimas contra 2-fucosil lactosa (Panel B). La  $\alpha$ -L-galactosidasa GH95 de *B. ovatus* es al menos 5000 veces menos activa en 2'-fucosil lactosa que la  $\alpha$ -L-fucosidasa de *B.bifidum*.

La Fig. 5 muestra monosacáridos liberados de goma de fibra de maíz después de tres tratamientos enzimáticos diferentes. Las enzimas fueron  $\alpha$ -L-galactosidasa GH95, UirFlo L (de Novozymes A/S) y una combinación de las mismas. Los monosacáridos son los siguientes: arabinosa, galactosa, glucosa, xilosa.

15 La Fig. 6 muestra una combinación adicional con la  $\alpha$ -xilosidasa GH31 de *B. ovatus* (descrita en nuestra solicitud de patente presentada al mismo tiempo).

La Fig. 7 muestra el efecto de agregar la enzima GH95 de *B. ovatus* junto con una mezcla de hemicelulasas individuales en la liberación de monosacáridos de salvado de maíz desalmidonado. Los niveles de azúcar se expresan como el área bajo la curva multiplicada por un factor de dilución. La figura también muestra los efectos de la  $\alpha$ -L-galactosidasa GH95 y otras enzimas (GH11, GH43/51 y combinaciones).

20 Breve descripción del listado de secuencias

[0015] En el listado de secuencias

25 SEQ ID NO: 1 es la secuencia genómica/de ADNc de la  $\alpha$ -L-galactosidasa GH95 de *Bacteroides ovatus* ATCC 8483, que incluye del codón de inicio al codón de terminación sin intrones.

SEQ ID NO: 2 es la secuencia de proteínas sacada de la SEQ ID NO: 1 (igual que UNIPROT: A7M011).

SEQ ID NO: 3 es la secuencia de aminoácidos de *Bacteroides ovatus* CL03T12C18 (UNIPROT: I8Z241, igual que UNIPROT: A7M011)

30 SEQ ID NO: 4 es la secuencia de aminoácidos de *Bacteroides ovatus* CL02T12C04 (UNIPROT: I8YAG0)

SEQ ID NO: 5 es la secuencia de aminoácidos de *Bacteroides xylanisolvens* CL03T12C04 (UNIPROT: I9US19)

SEQ ID NO: 6 es la secuencia de aminoácidos de *Bacteroides* sp. cepa 2\_1\_22 (UNIPROT: D0TM59)

35 Matriz de identidad de secuencias:

SEQ ID NO:	2	3	4	5	6
2	100	100	99,9	98,9	92,2
3	100	100	98,2	99,0	99,2
4	99,9	99,9	100	98,6	92,1
5	98,9	98,9	98,8	100	92,6
6	92,2	92,2	92,1	92,6	100

#### Definiciones

40 [0016]  $\alpha$ -L-galactosidasa: el término " $\alpha$ -L-galactosidasa" significa una actividad enzimática que cataliza la hidrólisis de una unidad  $\alpha$ -L-galactosil a partir de un sustrato que libera L-galactosa (todavía no se ha establecido un número EC para esta actividad). Para los fines de la presente invención, la actividad de  $\alpha$ -L-galactosidasa se determina de acuerdo con el procedimiento descrito en los ejemplos. En un aspecto, los polipéptidos de la presente invención tienen al menos 20%, por ejemplo, al menos 40%, al menos 50%, al menos 60%, al menos 70%, al menos 80%, al menos 90%, al menos 95%, o al menos 100% de la actividad de  $\alpha$ -L-galactosidasa del polipéptido maduro de la SEQ ID NO: 2.

[0017] Variante alélica: el término "variante alélica" significa cualquiera de dos o más formas alternativas de un gen que ocupa el mismo locus cromosómico. La variación alélica se produce de manera natural a través de la mutación, y puede dar lugar a un polimorfismo dentro de las poblaciones. Las mutaciones génicas pueden ser silenciosas (sin cambios en el polipéptido codificado) o pueden codificar polipéptidos que tienen secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

55 [0018] Dominio catalítico: el término "dominio catalítico" significa la región de una enzima que contiene la maquinaria catalítica de la enzima.

[0019] ADNc: el término "ADNc" significa una molécula de ADN que se puede preparar mediante transcripción inversa a partir de una molécula de ARNm maduro y empalmado obtenida de una célula eucariota o procariota. El ADNc carece de secuencias de intrones que pueden estar presentes en el ADN genómico correspondiente. El

transcrito inicial primario de ARN es un precursor del ARNm que se procesa a través de una serie de pasos, que incluyen el corte y empalme, antes de aparecer como ARNm empalmado maduro.

5 [0020] Secuencia codificante: El término "secuencia codificante" significa un polinucleótido, que especifica directamente la secuencia de aminoácidos de un polipéptido. Los límites de la secuencia codificante generalmente están determinados por un marco de lectura abierto, que comienza con un codón de inicio tal como ATG, GTG o TTG y termina con un codón de terminación tal como TAA, TAG o TGA. La secuencia codificante puede ser un ADN genómico, ADNc, ADN sintético o una combinación de los mismos.

10 [0021] Secuencias de control: El término "secuencias de control" significa secuencias de ácidos nucleicos necesarias para la expresión de un polinucleótido que codifica un polipéptido maduro de la presente invención. Cada secuencia de control puede ser nativa (es decir, del mismo gen) o foránea (es decir, de un gen diferente) respecto del polinucleótido que codifica el polipéptido o nativas o foráneas entre sí. Dichas secuencias de control incluyen, pero no se limitan a, un líder, secuencia de poliadenilación, secuencia de propéptidos, promotor, 15 secuencia de péptidos señal y terminador de transcripción. Como mínimo, las secuencias de control incluyen un promotor y señales de terminación transcripcional y traslacional. Las secuencias de control pueden estar provistas de enlazadores con el fin de introducir sitios de restricción específicos que faciliten la ligación de las secuencias de control con la región codificante del polinucleótido que codifica un polipéptido.

20 [0022] Expresión: El término "expresión" incluye cualquier paso implicado en la producción de un polipéptido que incluye, pero no se limita a, transcripción, modificación postranscripcional, traducción, modificación postraduccional y secreción.

25 [0023] Vector de expresión: El término "vector de expresión" significa una molécula de ADN lineal o circular que comprende un polinucleótido que codifica un polipéptido y está operativamente enlazada a secuencias de control que permiten su expresión.

30 [0024] Fragmento: El término "fragmento" significa un polipéptido o un dominio catalítico que tiene uno o más (por ejemplo, varios) aminoácidos ausentes del extremo amino terminal y/o carboxi terminal de un polipéptido o dominio maduro; donde el fragmento tiene actividad de  $\alpha$ -L-galactosidasa. En un aspecto, un fragmento contiene al menos 365 residuos de aminoácidos (por ejemplo, los aminoácidos 321 a 686 de la SEQ ID NO: 2), al menos 666 residuos de aminoácidos (por ejemplo, los aminoácidos 20 a 686 de la SEQ ID NO: 2) , o al menos 760 residuos de aminoácidos (por ejemplo, los aminoácidos 20 a 780 de la SEQ ID NO: 2). En un aspecto, un 35 fragmento contiene al menos un 85%, al menos un 90% o al menos un 95% de los aminoácidos del péptido maduro.

40 [0025] Glucósido hidrolasas (EC 3.2.1.-): son un amplio grupo de enzimas que hidrolizan el enlace glicosídico entre dos o más carbohidratos o entre un carbohidrato y ua fracción no carbohidrato. La nomenclatura de enzimas IUBMB de las glucósido hidrolasas se basa en su especificidad de sustrato y ocasionalmente en su mecanismo molecular; dicha clasificación no refleja (y no pretendía hacerlo) las características estructurales de estas enzimas. Recientemente se ha propuesto una clasificación de glucósido hidrolasas en familias basada en las similitudes de las secuencias de aminoácidos (CAZy). Debido a que existe una relación directa entre las similitudes de plegamiento y las secuencias, esta clasificación:

- 45 (i) refleja las características estructurales de estas enzimas mejor que solo su especificidad de sustrato,  
 (ii) ayuda a revelar las relaciones evolutivas entre estas enzimas,  
 (iii) proporciona una herramienta conveniente para derivar información sobre el mecanismo, y  
 (iv) ilustra la dificultad de derivar relaciones entre la pertenencia a una familia y la especificidad del sustrato

50 [0026] La base de datos Carbohydrate-Active Enzymes (CAZy) proporciona una lista continuamente actualizada de las familias de glucósido hidrolasas. Debido a que el plegamiento de las proteínas se conserva mejor que sus secuencias, algunas de las familias se pueden agrupar en 'clanes' :

- 55 (i) cuando se descubren nuevas secuencias relacionadas con más de una familia,  
 (ii) cuando aumenta la sensibilidad de los métodos de comparación de secuencias o  
 (iii) cuando las determinaciones estructurales demuestran la semejanza entre miembros de diferentes familias [3]

60 [0027] Las enzimas GH95: se sabe que exhiben actividad de  $\alpha$ -1,2-L-fucosidasa (EC 3.2.1.63); y/o de  $\alpha$ -L-fucosidasa (EC 3.2.1.51).

#### Condiciones de astringencia

65 [0028] Condiciones de astringencia alta: El término "condiciones de astringencia alta" significa, para sondas de al menos 100 nucleótidos de longitud, prehibridación e hibridación a 42 ° C en SSPE 5X, SDS al 0,3%, 200 microgramos/ml de ADN de esperma de salmón fragmentado y desnaturalizado, y formamida al 50%, siguiendo

procedimientos Southern blot estándar durante 12 a 24 horas. El material portador finalmente se lava tres veces, cada una durante 15 minutos, usando 2X SSC, SDS al 0,2% a 65 ° C.

5 [0029] Condiciones de astringencia baja : el término "condiciones de astringencia baja" significa, para sondas de al menos 100 nucleótidos de longitud, prehibridación e hibridación a 42°C en 5X SSPE, SDS al 0,3% , 200 microgramos/ml de ADN de esperma de salmón fragmentado y desnaturalizado, y formamida al 25%, siguiendo procedimientos Southern blot estándar durante 12 a 24 horas. El material portador se lava tres veces, cada una durante 15 minutos, utilizando 2X SSC, SDS al 0,2% a 50°C.

10 [0030] Condiciones de astringencia media: el término "condiciones de astringencia media" significa, para sondas de al menos 100 nucleótidos de longitud, prehibridación e hibridación a 42°C en 5X SSPE, SDS al 0,3%, 200 microgramos/ml de ADN de esperma de salmón fragmentado y desnaturalizado y formamida al 35%, siguiendo procedimientos Southern blot estándar durante 12 a 24 horas. El material portador se lava finalmente tres veces, cada una durante 15 minutos, utilizando 2X SSC, SDS al 0,2% a 55°C.

15 [0031] Condiciones de astringencia media-alta: el término "condiciones de astringencia media-alta" significa, para sondas de al menos 100 nucleótidos de longitud, prehibridación e hibridación a 42°C en 5X SSPE, SDS al 3%, 200 microgramos/ml de ADN de esperma de salmón fragmentado y desnaturalizado, y formamida al 35%, siguiendo procedimientos Southern blot estándar durante 12 a 24 horas. El material portador se lava tres veces, cada una durante 15 minutos, utilizando 2X SSC, SDS al 0,2% a 60°C.

20 [0032] Célula huésped: el término "célula huésped" significa cualquier tipo de célula que es susceptible a una transformación, transfección, transducción o similar con una construcción de ácidos nucleicos o vector de expresión que comprende un polinucleótido de la presente invención. El término "célula huésped" abarca cualquier progenie de una célula progenitora que no es idéntica a la célula progenitora debido a mutaciones que ocurren durante la replicación.

25 [0033] Aislado: el término "aislado" significa una sustancia en una forma o entorno que no ocurre en la naturaleza.

30 Ejemplos no limitativos de sustancias aisladas incluyen (1) cualquier sustancia que no ocurra de forma natural, (2) cualquier sustancia incluyendo, pero sin limitarse a, cualquier enzima, variante, ácido nucleico, proteína, péptido o cofactor, que se elimina al menos parcialmente de uno o más o todos los componentes de origen natural con el cual/con los cuales está asociado en la naturaleza; (3) cualquier sustancia modificada por la mano del hombre en relación con esa sustancia que se encuentra en la naturaleza; o (4) cualquier sustancia modificada al aumentar la cantidad de la sustancia en relación con otros componentes con los que está naturalmente asociada (por ejemplo, producción recombinante en una célula huésped; copias múltiples de un gen que codifica la sustancia; y uso de un promotor más fuerte que el promotor naturalmente asociado al gen que codifica la sustancia).

35 [0034] Polipéptido maduro: el término "polipéptido maduro" significa un polipéptido en su forma final después de la traducción y cualquier modificación postraduccional, tal como procesamiento en el N-terminal, truncamiento en el C-terminal, glicosilación, fosforilación, etc. En un aspecto, el polipéptido maduro es los aminoácidos 20 a 811 de la SEQ ID NO: 2 basado en el programa SignalP (Nielsen et al., 1997, Protein Engineering 10: 1-6) que indica que los aminoácidos 1 a 20 de la SEQ ID NO: 2 son un péptido señal. Se sabe en la técnica que una célula huésped puede producir una mezcla de dos de más polipéptidos maduros diferentes (es decir, con un aminoácido diferente en el C-terminal y/o N-terminal) expresados por el mismo polinucleótido. También se sabe en la técnica que células huésped diferentes procesan los polipéptidos de manera diferente y, por lo tanto, una célula huésped que expresa un polinucleótido puede producir un polipéptido maduro diferente (por ejemplo, con un aminoácido diferente en el extremo C-terminal y/o N-terminal ) en comparación con otra célula huésped que expresa el mismo polinucleótido.

40 En un aspecto, un polipéptido maduro contiene hasta 792 residuos de aminoácidos (por ejemplo, los aminoácidos 1 a 792 de la SEQ ID NO: 2), hasta 780 residuos de aminoácidos (por ejemplo ,los aminoácidos 10 a 772 de la SEQ ID NO: 2),

45 [0035] Secuencia codificante de polipéptido maduro: el término "secuencia codificante de polipéptido maduro" significa un polinucleótido que codifica un polipéptido maduro que tiene actividad de  $\alpha$ -L-galactosidasa. En un aspecto, la secuencia codificante de polipéptido maduro es los nucleótidos 60 a 2433 de SEQ ID NO:1 según SignalP (Nielsen et al., 1997, supra), que indica que los nucleótidos 1 a 60 de la SEQ ID NO: 1 codifican un péptido señal.

50 [0036] Construcción de ácidos nucleicos: el término "construcción de ácidos nucleicos" significa una molécula de ácido nucleico, de una sola cadena o de doble cadena, que se aísla de un gen de origen natural o se modifica para que contenga segmentos de ácidos nucleicos de una manera que no existiría en la naturaleza de otro modo o que es sintético, que comprende una o más secuencias de control.

55 [0037] Operativamente enlazado: el término "operativamente enlazado" significa una configuración en la que una

secuencia de control se coloca en una posición apropiada respecto a la secuencia de codificación de un polinucleótido de manera que la secuencia de control dirige la expresión de la secuencia codificante.

5 [0038] Identidad de secuencia: la relación entre dos secuencias de aminoácidos o entre dos secuencias de nucleótidos es descrita por el parámetro "identidad de secuencia".

[0039] Para los fines de la presente invención, la identidad de secuencia entre dos secuencias de aminoácidos se determina usando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, J. Mol. Biol. 48:443-453) según está implementado en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends Genet. 16:276-277), preferiblemente la versión 5.0.0 o posterior. Los parámetros usados son penalización por apertura de hueco de 10, penalización por extensión de hueco de 0,5, y la matriz de sustitución EBLOSUM62 (versión de EMBOSS de BLOSUM62). El resultado de Needle etiquetado como "identidad más larga" (obtenido utilizando la opción -nobrief) se usa como identidad en porcentaje y se calcula de la siguiente manera:

15 
$$\text{(Residuos idénticos x 100) / (Longitud de alineamiento - Número total de huecos en alineamiento)}$$

[0040] Para los fines de la presente invención, la identidad de secuencia entre dos secuencias de desoxirribonucleótidos se determina usando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, supra) según está implementado en el programa Needle del paquete EMBOSS (EMBOSS: LThe European Molecular Biology Open Software Suite, Rice et al., 2000, supra), preferiblemente la versión 5.0.0 o posterior. Los parámetros utilizados son la penalización por apertura de hueco de 10, la penalización de extensión de hueco de 0,5 y la matriz de sustitución EADNFULL (versión de EMBOSS de NCBI NUC4.4). El resultado de Needle etiquetado como "identidad más larga" (obtenido mediante la opción -nobrief) se utiliza como identidad en porcentaje y se calcula de la siguiente manera:

25 
$$\text{(Desoxirribonucleótidos idénticos x 100) / (Lognitud de alineamento - Número total de huecos en alineamiento)}$$

30 [0041] Subsecuencia: El término "subsecuencia" significa un polinucleótido que tiene uno o más (por ejemplo, varios) nucleótidos ausentes del extremo 5 'y/o 3' terminal de una secuencia codificante de un polipéptido maduro; donde la subsecuencia codifica un fragmento que tiene actividad de  $\alpha$ -L-galactosidasa. En un aspecto, una subsecuencia contiene al menos 1095 nucleótidos (por ejemplo, los nucleótidos 963 a 2058 de la SEQ ID NO: 1), al menos 1998 nucleótidos (por ejemplo, los nucleótidos 60 a 2058 de la SEQ ID NO: 1), o al menos 2280 nucleótidos (por ejemplo, los nucleótidos 60 a 2340 de la SEQ ID NO: 1). En un aspecto, una subsecuencia contiene al menos 85%, 90% o 95% de los nucleótidos que codifican el polipéptido maduro de la SEQ ID NO:2

35 [0042] Variante: El término "variante" significa un polipéptido que tiene actividad de  $\alpha$ -L-galactosidasa que comprende una alteración, es decir, una sustitución, inserción y/o delección, en una o más (por ejemplo, varias) posiciones. Una sustitución significa la sustitución del aminoácido que ocupa una posición con un aminoácido diferente; una delección significa la eliminación de uno o más (por ejemplo, varios) aminoácidos, por ejemplo, 1-5 aminoácidos, adyacentes al aminoácido que ocupa una posición; y una inserción significa añadir uno o más (por ejemplo, varios) aminoácidos, por ejemplo, 1-5 aminoácidos, adyacentes al aminoácido que ocupa una posición.

45 Descripción detallada de la invención

Polipéptidos que tienen actividad de  $\alpha$ -L-galactosidasa

50 [0043] En Una forma de realización, la presente invención se refiere al uso en pienso para animales de polipéptidos aislados que tienen una identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 2 de al menos 80%, al menos 82%, al menos 84%, al menos 85 %, al menos 87%, al menos 89%, al menos 90%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98 %, al menos 99% o 100%, que tienen actividad de  $\alpha$ -L-galactosidasa. En un aspecto, los polipéptidos difieren en hasta 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11 o 10 aminoácidos, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10, del polipéptido maduro de la SEQ ID NO: 2.

60 [0044] Un polipéptido de la presente invención preferiblemente comprende o consiste en la secuencia de aminoácidos de la SEQ ID NO: 2 o una variante alélica de la misma; o es un fragmento de esta que tiene actividad de  $\alpha$ -L-galactosidasa. En otro aspecto, el polipéptido comprende o consiste en el polipéptido maduro de la SEQ ID NO: 2. En otro aspecto, el polipéptido comprende o consiste en los aminoácidos 25 a 750 de la SEQ ID NO: 2.

65 [0045] La presente invención también se refiere al uso en pienso para animales de polipéptidos aislados que tienen una identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 4 de al menos 80%, por ejemplo, al menos 85%, al menos 86%, al menos 87%. , al menos 88%, al menos 89%, al menos 90%, al menos

91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97% , al menos 98%, al menos 99% o 100%, que tienen actividad de  $\alpha$ -L-galactosidasa. En un aspecto, los polipéptidos difieren en no más de treinta y seis aminoácidos, por ejemplo, en treinta aminoácidos, en veinticinco aminoácidos, en veinte aminoácidos, en quince aminoácidos, en diez aminoácidos, en nueve aminoácidos, en ocho aminoácidos, en siete aminoácidos, en seis aminoácidos, en cinco aminoácidos, en cuatro aminoácidos, en tres aminoácidos, en dos aminoácidos, y en un aminoácido del polipéptido maduro de la SEQ ID NO: 4 .

[0046] La presente invención se refiere además al uso en pienso para animales de polipéptidos aislados que tienen una identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 5 de al menos 80%, por ejemplo, al menos 85%, al menos 86%, al menos 87% , al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97% , al menos 98%, al menos 99% o 100%, que tienen actividad de  $\alpha$ -L-galactosidasa. En un aspecto, los polipéptidos difieren en no más de treinta y seis aminoácidos, por ejemplo, en treinta aminoácidos, en veinticinco aminoácidos, en veinte aminoácidos, en quince aminoácidos, en diez aminoácidos, en nueve aminoácidos , en ocho aminoácidos, en siete aminoácidos, en seis aminoácidos, en cinco aminoácidos, en cuatro aminoácidos, en tres aminoácidos, en dos aminoácidos, y en un aminoácido del polipéptido maduro de la SEQ ID NO: 5.

[0047] La presente invención se refiere además al uso en pienso para animales de polipéptidos aislados que tienen una identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 6 de al menos 80%, por ejemplo, al menos 85%, al menos 86%, al menos 87% , al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97% , al menos 98%, al menos 99% o 100%, que tienen actividad de  $\alpha$ -L-galactosidasa. En un aspecto, los polipéptidos difieren en no más de treinta y seis aminoácidos, por ejemplo, en treinta aminoácidos, en veinticinco aminoácidos, en veinte aminoácidos, en quince aminoácidos, en diez aminoácidos, en nueve aminoácidos , en ocho aminoácidos, en siete aminoácidos, en seis aminoácidos, en cinco aminoácidos, en cuatro aminoácidos, en tres aminoácidos, en dos aminoácidos, y en un aminoácido del polipéptido maduro de la SEQ ID NO: 6.

[0048] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos 80% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.

[0049] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 82% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.

[0050] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos 84% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.

[0051] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 85% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.

[0052] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 86% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.

[0053] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 87% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.

[0054] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos 88% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.

[0055] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 89% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.

[0056] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 90% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.

[0057] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido para uso en pienso para animales o detergentes que tienen al menos un 91% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.



- 5 [0058] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 92% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.
- [0059] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 93% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.
- 10 [0060] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 94% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.
- 15 [0061] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 95% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.
- [0062] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 96% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.
- 20 [0063] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos 97% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.
- 25 [0064] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos 98% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.
- 30 [0065] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 99% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 2.
- 35 [0066] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene una identidad de secuencia del 100% con el polipéptido de la SEQ ID NO: 2.
- [0067] Cada una de las formas de realización anteriores también se refiere no solo a la SEQ ID NO: 2, sino también a cualquiera de las SEQ ID NO: 4, 5 y 6.
- 40 [0068] Por consiguiente, la invención también se refiere a:
- [0069] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 80% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 45 [0070] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 82% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 50 [0071] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos 84% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 55 [0072] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 85% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 60 [0073] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 86% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- [0074] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 87% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 65

- [0075] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos 88% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 5 [0076] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos 89% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 10 [0077] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 90% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 15 [0078] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido para usar en pienso para animales o detergentes que tienen al menos un 91% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 20 [0079] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 92% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 25 [0080] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 93% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 30 [0081] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 94% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 35 [0082] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 95% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 40 [0083] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 96% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 45 [0084] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 97% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 50 [0085] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos 98% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 55 [0086] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 99% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 4.
- 60 [0087] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene una identidad de secuencia del 100% con el polipéptido de la SEQ ID NO: 4.
- 65 [0088] Y  
Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 80% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.
- [0089] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 82% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.
- [0090] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 84% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.
- [0091] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un

polinucleótido en pienso para animales que tiene al menos un 85% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

5 [0092] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 86% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

10 [0093] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 87% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

15 [0094] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 88% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

[0095] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 89% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

20 [0096] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 90% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

25 [0097] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido para uso en pienso para animales o detergentes que tienen al menos un 91% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

30 [0098] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 92% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

35 [0099] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 93% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

[0100] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 94% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

40 [0101] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 95% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

45 [0102] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 96% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

50 [0103] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 97% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

55 [0104] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 98% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

[0105] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 99% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 5.

60 [0106] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene una identidad de secuencia del 100% con el polipéptido de la SEQ ID NO: 2.

65 [0107] Y  
Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 80% de identidad de secuencia respecto al polipéptido de la SEQ ID



[0124] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene al menos un 99% de identidad de secuencia respecto al polipéptido de la SEQ ID NO: 6.

5 [0125] Una forma de realización de la invención usa un polipéptido o un polipéptido codificado por un polinucleótido en pienso para animales que tiene una identidad de secuencia del 100% con el polipéptido de la SEQ ID NO: 6.

10 [0126] Un polipéptido para usar en la presente invención preferiblemente comprende o consiste en la secuencia de aminoácidos de la SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5 y/o SEQ ID NO: 6 o una variante alélica de la misma; o es un fragmento de la misma al que le faltan, por ejemplo, 30, 25, 20, 15, 10 o 5 aminoácidos del extremo N- y/o C-terminal y que tiene actividad de  $\alpha$ -L-galactosidasa. En otro aspecto, el polipéptido comprende o consiste en el polipéptido maduro de la SEQ ID NO: 2, SEQ ID NO: 4 SEQ ID NO: 5 y/o SEQ ID NO: 6. En otro aspecto, el polipéptido comprende o consiste en los aminoácidos 22 a 814 de la SEQ ID NO: 2, los aminoácidos 22 a 814 de la SEQ ID NO: 4, SEQ ID NO: 5 y/o los aminoácidos 22 a 814 de la SEQ ID NO: 6.

15 [0127] En otra forma de realización, la presente invención se refiere a un polipéptido aislado que tiene actividad de  $\alpha$ -L-galactosidasa codificado por un polinucleótido que se hibrida en condiciones de astringencia muy baja, condiciones de astringencia baja, condiciones de astringencia media, condiciones de astringencia media-alta, condiciones de astringencia alta o condiciones de astringencia muy alta con (i) la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1, o (ii) el complemento de longitud completa de (i) (Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2nd edition, Cold Spring Harbor, New York).

20 [0128] El polinucleótido de la SEQ ID NO: 1 o una subsecuencia de esta, así como el polipéptido de la SEQ ID NO: 2 o un fragmento de esta, pueden usarse para diseñar sondas de ácidos nucleicos para identificar y clonar ADN que codifica polipéptidos que tienen actividad de  $\alpha$ -L-galactosidasa procedentes de cepas de diferentes géneros o especies según métodos ampliamente conocidos en la técnica. En particular, tales sondas se pueden usar para hibridación con el ADN genómico o ADNc de una célula de interés, siguiendo procedimientos Southern Blot estándar, con el fin de identificar y aislar ahí el gen correspondiente. Tales sondas pueden ser considerablemente más cortas que la secuencia completa, pero deben ser de al menos 15, por ejemplo, al menos 25, al menos 35, o al menos 70 nucleótidos de longitud. Preferiblemente, la sonda de ácido nucleico tiene al menos 100 nucleótidos de longitud, por ejemplo, al menos 200 nucleótidos, al menos 300 nucleótidos, al menos 400 nucleótidos, al menos 500 nucleótidos, al menos 600 nucleótidos, al menos 700 nucleótidos, al menos 800 nucleótidos, o al menos 900 nucleótidos de longitud. Se pueden usar sondas de ADN y ARN. Las sondas suelen estar marcadas para detectar el gen correspondiente (por ejemplo, con  $^{32}\text{P}$ ,  $^3\text{H}$ ,  $^{35}\text{S}$ , biotina o avidina). Dichas sondas están abarcadas por la presente invención.

25 [0129] Una genoteca de ADN genómico o de ADNc preparada a partir de tales otras cepas puede rastrearse para detectar ADN que se hibride con las sondas descritas anteriormente y codifique un polipéptido que tiene actividad de  $\alpha$ -L-galactosidasa. El ADN genómico o de otro tipo de tales otras cepas se puede separar mediante electroforesis en gel de agarosa o poliacrilamida u otras técnicas de separación. El ADN de las genotecas o el ADN separado puede transferirse a nitrocelulosa u otro material portador adecuado e inmovilizarse en él. Con el fin de identificar un clon o ADN que se hibrida con la SEQ ID NO: 1 o una subsecuencia de la misma, el material de soporte se usa en un Southern blot.

30 [0130] Para los fines de la presente invención, la hibridación indica que el polinucleótido se hibrida con una sonda de ácido nucleico marcada correspondiente a (i) SEQ ID NO: 1; (ii) la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1; (iii) el complemento completo de esta; o (iv) una subsecuencia de esta; en condiciones de astringencia muy baja a muy alta. Las moléculas a las que se hibrida la sonda de ácido nucleico en estas condiciones se pueden detectar usando, por ejemplo, película de rayos X o cualquier otro medio de detección conocido en la técnica.

35 [0131] En un aspecto, la sonda de ácido nucleico es un polinucleótido que codifica el polipéptido de la SEQ ID NO: 2; el polipéptido maduro de la misma; o un fragmento de la misma. En otro aspecto, la sonda de ácido nucleico es la SEQ ID NO: 1. En otro aspecto, la sonda de ácido nucleico es la región codificante del polipéptido maduro contenida en *B. Ovatum* ATCC - 8483, donde el polinucleótido codifica un polipéptido que tiene actividad de  $\alpha$ -L-galactosidasa.

40 [0132] Para sondas largas de al menos 100 nucleótidos de longitud, las condiciones de astringencia alta a muy alta se definen como prehibridación e hibridación a 42 ° C en SSPE 5X, SDS al 0,3%, 200 microgramos/ml de ADN de esperma de salmón fragmentado y desnaturalizado, y formamida al 25% para astringencias muy bajas y bajas, formamida al 35% para astringencias medias y medias-altas, o formamida al 50% para astringencias altas y muy altas, siguiendo procedimientos Southern blot estándar durante 12 a 24 horas de manera óptima. El material portador finalmente se lava tres veces, cada una durante 15 minutos, usando SSC 2X, SDS al 0,2% a 65 ° C (astringencia alta), y a 70 ° C (astringencia muy alta).

- 5 [0133] Para sondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, las condiciones de astringencia se definen como prehibridación e hibridación a aproximadamente 5°C a aproximadamente 10°C por debajo de la  $T_m$  calculada usando el cálculo según Bolton y McCarthy (1962, Proc. Natl. Acad. Sci. USA 48: 1390) en 0,9 M NaCl, 0,09 M Tris-HCl, pH 7.6, 6 mM EDTA, NP-40 al 0,5%, solución de Denhardt 1x, 1 mM pirofosfato de sodio, 1 mM fosfato monobásico de sodio, 0,1 mM ATP y 0,2 mg de ARN de levadura por ml siguiendo procedimientos Southern Blot estándar durante 12 a 24 horas de forma óptima. El material portador se lava finalmente una vez en SCC 6X más SDS al 0,1% durante 15 minutos y dos veces, cada una durante 15 minutos, usando SSC6X a entre 5 ° C y 10 ° C por debajo de la  $T_m$  calculada.
- 10 [0134] En otra forma de realización, la presente invención se refiere a un polipéptido aislado que tiene actividad de  $\alpha$ -L-galactosidasa codificado por un polinucleótido que tiene una identidad de secuencia con la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1 de al menos 92%, al menos 93%, a al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100%.
- 15 [0135] En otra forma de realización, la presente invención se refiere a variantes del polipéptido maduro de la SEQ ID NO: 2 que comprende una sustitución, delección y/o inserción en una o más (por ejemplo, varias) posiciones. En una forma de realización, el número de sustituciones, delecciones y/o inserciones de aminoácidos introducidas en el polipéptido maduro de la SEQ ID NO: 2 es de hasta 10, por ejemplo, 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10. Los cambios en los aminoácidos pueden ser de naturaleza menor, es decir, sustituciones o inserciones conservadoras de aminoácidos que no afectan significativamente al plegamiento y/o actividad de la proteína; delecciones pequeñas, típicamente de 1-30 aminoácidos; pequeñas extensiones en los extremos amino o carboxilo terminales, tales como un residuo de metionina en el extremo amino terminal; un péptido enlazador pequeño de hasta 20-25 residuos; o una pequeña extensión que facilita la purificación al cambiar la carga neta u otra función, tal como un tracto de polihistidina, un epítipo antigénico o un dominio de unión.
- 20 [0136] Ejemplos de sustituciones conservadoras se encuentran dentro de los grupos de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrófobos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano y tirosina) y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina). Las sustituciones de aminoácidos que generalmente no alteran la actividad específica son conocidas en la técnica y se describen, por ejemplo, por H. Neurath y R.L. Hill, 1979, en *The Proteins*, Academic Press, New York. Las sustituciones comunes son Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, y Asp/Gly.
- 25 [0137] Alternativamente, los cambios en los aminoácidos son de tal naturaleza que las propiedades fisicoquímicas de los polipéptidos se ven alteradas. Por ejemplo, los cambios en los aminoácidos pueden mejorar la estabilidad térmica del polipéptido, alterar la especificidad del sustrato, cambiar el pH óptimo y similares.
- 30 [0138] Los aminoácidos esenciales en un polipéptido se pueden identificar de acuerdo con procedimientos conocidos en la técnica, tales como mutagénesis dirigida o mutagénesis de barrido con alanina (Cunningham y Wells, 1989, *Science* 244: 1081-1085) En esta última técnica, se introducen mutaciones únicas de alanina en cada residuo de la molécula, y las moléculas mutantes resultantes se analizan para determinar la actividad de  $\alpha$ -L-galactosidasa para identificar residuos de aminoácidos que son fundamentales para la actividad de la molécula. Véase también Hilton et al., 1996, *J. Biol. Chem.* 271:4699-4708. El sitio activo de la enzima u otra interacción biológica también puede determinarse por análisis físico de la estructura, según se determina mediante técnicas tales como resonancia magnética nuclear, cristalografía, difracción de electrones o marcaje por fotoafinidad, junto con la mutación de aminoácidos de sitio de contacto putativo. Véase, 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. La identidad de los aminoácidos esenciales también puede inferirse a partir de un alineamiento con un polipéptido relacionado.
- 35 [0139] Se considera que las enzimas GH95 emplean un mecanismo de reacción único, en el que Asn activada por Asp actúa como catalizador de base general que implica Glu566 y Asn423 (Nagae M, Tsuchiya A, Katayama T, Yamamoto K, Wakatsuki S y Kato R. Structural basis of the catalytic reaction mechanism of novel 1,2-alpha-L-fucosidase from *Bifidobacterium bifidum*. *J Biol Chem.* 2007 Jun 22;282(25):18497-509. Glu566 está unido a hidrógeno con Asn421. Asn423 es activada por la Asp766 cercana y, en consecuencia, activa una molécula de agua nucleofílica. Este modelo (activación mediada por ácido carboxílico del grupo amido) tiene alguna analogía con el mecanismo de participación del grupo cercano empleado por los miembros de GH18, GH20, GH25, GH56, GH84 y GH85. Estos cuatro residuos son invariables en los miembros de esta familia, y la sustitución con alanina o glicina disminuye las actividades de 1000 a 10 000 veces (Nagae M, Tsuchiya A, Katayama T, Yamamoto K, Wakatsuki S y Kato R. Structural basis of the catalytic reaction mechanism of novel 1,2-alpha-L-fucosidase from *Bifidobacterium bifidum*. *J Biol Chem.* 2007 Jun 22;282(25):18497-509.
- 40 [0140] Se pueden realizar y probar sustituciones, delecciones y/o inserciones de aminoácidos simples o múltiples usando métodos conocidos de mutagénesis, recombinación y/o mezcla, seguido de un procedimiento de selección relevante, tal como los descritos por Reidhaar-Olson y Sauer, 1988, *Science* 241: 53-57; Bowie y
- 45
- 50
- 55
- 60
- 65

Sauer, 1989, Proc. Natl. Acad. Sci. USA 86: 2152-2156; WO 95/17413; o WO 95/22625. Otros métodos que se pueden usar incluyen la PCR propensa a errores, el *phage display* (p. Lowman et al., 1991, Biochemistry 30: 10832-10837; U.S. Patent No. 5,223,409; WO 92/06204), y la mutagénesis dirigida (Derbyshire et al, 1986, Gene 46: 145; Ner et al., 1988, ADN 7: 127).

5

[0141] Los métodos de mutagénesis/barajado se pueden combinar con métodos de cribado automatizados de alto rendimiento para detectar la actividad de polipéptidos clonados y mutagenizados expresados por células huésped (Ness et al., 1999, Nature Biotechnology 17: 893-896) Las moléculas de ADN mutagenizadas que codifican polipéptidos activos pueden recuperarse de las células huésped y secuenciarse rápidamente usando métodos estándar en la técnica. Estos métodos permiten la determinación rápida de la importancia de residuos de aminoácidos individuales en un polipéptido.

10

[0142] El polipéptido puede ser un polipéptido híbrido en el que una región de un polipéptido está fusionado en el extremo N terminal o en el extremo C terminal de una región de otro polipéptido.

15

[0143] El polipéptido puede ser un polipéptido de fusión o un polipéptido de fusión escindible en el que otro polipéptido está fusionado en el extremo N terminal o en el extremo C terminal del polipéptido de la presente invención. Un polipéptido de fusión se produce fusionando un polinucleótido que codifica otro polipéptido con un polinucleótido de la presente invención. Las técnicas para producir polipéptidos de fusión son conocidas en la técnica e incluyen ligar las secuencias codificantes que codifican los polipéptidos para que estén dentro del marco y para que la expresión del polipéptido de fusión esté bajo el control del mismo promotor o promotores y terminador. Los polipéptidos de fusión también se pueden construir usando tecnología de inteína en la que los polipéptidos de fusión se crean post-traduccionalmente (Cooper et al., 1993, EMBO J. 12: 2575-2583; Dawson et al., 1994, Science 266: 776-779).

20

25

[0144] Un polipéptido de fusión puede comprender además un sitio de escisión entre los dos polipéptidos. Tras la secreción de la proteína de fusión, el sitio se escinde liberando los dos polipéptidos. Los ejemplos de sitios de escisión incluyen, pero no se limitan a, los sitios descritos en Martin et al, 2003, J. Ind. Microbiol. Biotechnol. 3:568-576; Svetina et al, 2000, J. Biotechnol. 76:245-251; Rasmussen-Wilson et al, 1997, Appl. Reinar. Microbiol. 63:3488-3493; Ward et al., 1995, Biotechnology 13: 498-503; y Contreras et al., 1991, Biotechnology 9: 378-381; Eaton et al., 1986, Biochemistry 25: 505-512; Collins-Racie et al., 1995, Biotechnology 13: 982-987; Carter et al., 1989, Proteins: Structure, Function and Genetics 6: 240-248; y Stevens, 2003, Drug Discovery World 4: 35-48.

30

35

Fuentes de polipéptidos que tienen actividad de  $\alpha$ -L-galactosidasa

[0145] Un polipéptido que tiene actividad de  $\alpha$ -L-galactosidasa de la presente invención se puede obtener a partir de microorganismos de cualquier género. Para los fines de la presente invención, el término "obtenido a partir de" como se usa en el presente documento en conexión con una fuente dada significará que el polipéptido codificado por un polinucleótido es producido por la fuente o por una cepa en la que se ha insertado el polinucleótido de la fuente. En un aspecto, el polipéptido obtenido de una fuente dada se secreta extracelularmente.

40

[0146] El polipéptido puede ser un polipéptido bacteriano. Por ejemplo, el polipéptido puede ser un polipéptido bacteriano Gram-negativo tal como un polipéptido de *Bacteroides*, *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, o *Ureaplasma* o un polipéptido bacteriano Gram-positivo tal como un polipéptido de *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, o *Streptomyces* que tiene actividad de alfa xilosidasa.

45

[0147] En un aspecto, el polipéptido es un polipéptido de *Bacteroides ovatus*, *Bacteroides acidifaciens*, *Bacteroides gracilis*, *Bacteroides oris*, *Bacteroides putredinis*, *Bacteroides pyogenes* o *Bacteroides vulgatus*.

50

[0148] En otro aspecto, el polipéptido es un polipéptido de *Bacteroides ovatus*, por ejemplo, un polipéptido obtenido de *Bacteroides ovatus* ATCC 8483.

55

[0149] En un aspecto, el polipéptido es un polipéptido de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, o *Bacillus thuringiensis*.

60

[0150] En otro aspecto, el polipéptido es un polipéptido de *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, o *Streptococcus equi* subsp. *Zooepidemicus*.

60

[0151] En otro aspecto, el polipéptido es un polipéptido de *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, o *Streptomyces lividans*.

65

[0152] El polipéptido puede ser un polipéptido fúngico. Por ejemplo, el polipéptido puede ser un polipéptido de levadura tal como un polipéptido de *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia*; o un polipéptido fúngico filamentoso tal como un polipéptido de *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Aureobasidium*, *Botryosphaeria*, *Ceriporiopsis*, *Chaetomidium*, *Chrysosporium*, *Claviceps*, *Cochliobolus*, *Coprinopsis*, *Coptotermes*, *Corynascus*, *Cryphonectria*, *Cryptococcus*, *Diplodia*, *Exidia*, *Filibasidium*, *Fusarium*, *Gibberella*, *Holomastigotoides*, *Humicola*, *Irpex*, *Lentinula*, *Leptosphaeria*, *Magnaporthe*, *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*, *Pseudoplectania*, *Pseudotriconympha*, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trichoderma*, *Trichophaea*, *Verticillium*, *Volvariella*, o *Xylaria*.

[0153] En otro aspecto, el polipéptido es un polipéptido de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis* o *Saccharomyces oviformis*.

[0154] En otro aspecto, el polipéptido es un polipéptido de *Acremonium cellulolyticus*, *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenlandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *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*, *Fusarium venenatum*, *Humicola grisea*, *Humicola insolens*, *Humicola lanuginosa*, *Irpex lacteus*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium funiculosum*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia setosa*, *Thielavia spededonium*, *Thielavia subthermophila*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.

[0155] Se entenderá que, para las especies mencionadas anteriormente, la invención abarca tanto el estado perfecto como el imperfecto, y otros equivalentes taxonómicos, por ejemplo anamorfos, independientemente del nombre de especie por el que se conocen. Los expertos en la técnica reconocerán fácilmente la identidad de los equivalentes adecuados.

[0156] Las cepas de estas especies son de fácil acceso para el público en una serie de colecciones de cultivos, como la American Type Culture Collection (ATCC), la Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), la Centraalbureau Voor Schimmelcultures (CBS) y el Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

[0157] El polipéptido puede identificarse y obtenerse a partir de otras fuentes, incluyendo microorganismos aislados de la naturaleza (por ejemplo tierra, compost, agua, etc.) o muestras de ADN obtenidas directamente de materiales naturales (por ejemplo tierra, compost, agua, etc.) utilizando las sondas anteriormente mencionadas. Las técnicas para aislar microorganismos y ADN directamente de hábitats naturales son ampliamente conocidas en la técnica. Se puede obtener un polinucleótido que codifica el polipéptido rastreando de forma similar una genoteca de ADN genómico o de ADNc de otro microorganismo o muestra de ADN mixto. Una vez que se ha detectado un polinucleótido que codifica un polipéptido con la(s) sonda(s), el polinucleótido puede aislarse o clonarse utilizando técnicas que son conocidas por los expertos en la técnica. (véase, por ejemplo, Sambrook *et al.*, 1989, *supra*).

#### Dominios catalíticos

[0158] En una forma de realización, la presente invención también se refiere a dominios catalíticos que tienen una identidad de secuencia respecto a los aminoácidos 20 a 811 de la SEQ ID NO: 2 de al menos el 80%, al menos el 85%, al menos el 86%, al menos el 87%, al menos el 88%, al menos el 89%, al menos el 90%, al menos el 91%, al menos el 92%, al menos el 93%, al menos el 94%, al menos el 95%, al menos el 96%, al menos el 97%, al menos el 98%, al menos el 99%, al menos el 92%, al menos el 95%, al menos el 96%, al menos el 97%, al menos el 98%, al menos el 99% o 100%. En un aspecto, los dominios catalíticos comprenden secuencias de aminoácidos que difieren en hasta 10 aminoácidos, por ejemplo 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10, de los aminoácidos 20 a 811 de la SEQ ID NO: 2.

[0159] El dominio catalítico comprende o consiste preferiblemente en los aminoácidos 20 a 811 de la SEQ ID NO: 2 o una variante alélica de la misma; o es un fragmento de la misma que tiene actividad de  $\alpha$ -L-galactosidasa.

[0160] En otra forma de realización, la presente invención también se refiere a dominios catalíticos codificados por polinucleótidos que se hibridan en condiciones de astringencia muy baja, condiciones de astringencia baja,



condiciones de astringencia media, condiciones de astringencia media-alta, condiciones de astringencia alta o condiciones de astringencia muy alta (como se ha definido anteriormente) con (i) los nucleótidos 60 a de la SEQ ID NO: 1, o (ii) el complemento de longitud completa de (i) (Sambrook *et al.*, 1989, *supra*).

5 [0161] En otra forma de realización, la presente invención también se refiere a dominios catalíticos codificados por polinucleótidos que tienen una identidad de secuencia respecto a los nucleótidos 60 a 2433 de la SEQ ID NO: 1 de al menos 80%, al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, a al menos 98%, al menos 99%, al menos 92%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99% o 100%.

[0162] El polinucleótido que codifica el dominio catalítico comprende preferiblemente o está constituido por los nucleótidos 60 a 2433 de la SEQ ID NO: 1 o es la secuencia contenida en ATCC 8483.

15 [0163] En otra forma de realización, la presente invención también se refiere a variantes de dominio catalítico de los aminoácidos 20 a 811 de la SEQ ID NO: 2 que comprenden una sustitución, delección y/o inserción en una o más (por ejemplo, varias) posiciones. En un aspecto, el número de sustituciones, delecciones y/o inserciones de aminoácidos introducidas en la secuencia de los aminoácidos 20 a 811 de la SEQ ID NO: 2 es de hasta 10, por ejemplo, 1, 2, 3, 4, 5, 6, 8, 9 o 10.

20 Polinucleótidos

[0164] La presente invención también se refiere a polinucleótidos aislados que codifican un polipéptido, un dominio catalítico de la presente invención, como se describe en este documento.

25 [0165] Las técnicas usadas para aislar o clonar un polinucleótido son conocidas en la técnica e incluyen el aislamiento de ADN genómico o ADNc, o una combinación de los mismos. La clonación de los polinucleótidos del ADN genómico se puede efectuar, por ejemplo, utilizando la ampliamente conocida reacción en cadena de la polimerasa (PCR) o el cribado de anticuerpos de bibliotecas de expresión para detectar fragmentos de ADN clonados con características estructurales compartidas. Véase, por ejemplo, Innis *et al.*, 1990, PCR: A Guide to Methods and Application, Academic Press, New York. Se pueden usar otros procedimientos de amplificación de ácidos nucleicos tales como la reacción en cadena de la ligasa (LCR), la transcripción activada por ligación (LAT) y la amplificación basada en polinucleótidos (NASBA). Los polinucleótidos pueden clonarse a partir de una cepa de *Bacteroides*, o un organismo relacionado y, por lo tanto, por ejemplo, puede ser una variante alélica o de especie de la región que codifica el polipéptido del polinucleótido.

40 [0166] La modificación de un polinucleótido que codifica un polipéptido de la presente invención puede ser necesaria para la síntesis de polipéptidos sustancialmente similares al polipéptido. El término "sustancialmente similares" al polipéptido se refiere a formas no naturales del polipéptido. Estos polipéptidos pueden diferir de alguna manera modificada del polipéptido aislado de su fuente nativa, por ejemplo, variantes que difieren en actividad específica, termoestabilidad, pH óptimo o similares. Las variantes se pueden construir sobre la base del polinucleótido presentado como la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1, por ejemplo una subsecuencia de la misma, y/o mediante la introducción de sustituciones de nucleótidos que no dan como resultado un cambio en la secuencia de aminoácidos del polipéptido, pero que corresponden al uso del codón del organismo huésped destinado a la producción de la enzima, o mediante la introducción de sustituciones de nucleótidos que pueden dar lugar a una secuencia de aminoácidos diferente. Para una descripción general de la sustitución de nucleótidos véase, por ejemplo, Ford *et al.*, 1991, Protein Expression and Purification 2: 95-107.

50 Construcciones de ácidos nucleicos

[0167] También se describen construcciones de ácidos nucleicos que comprenden un polinucleótido de la presente invención unido operativamente a una o más secuencias de control que dirigen la expresión de la secuencia codificante en una célula huésped adecuada en condiciones compatibles con las secuencias de control.

55 [0168] El polinucleótido puede manipularse de diversas maneras para proporcionar la expresión del polipéptido. La manipulación del polinucleótido antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar polinucleótidos utilizando métodos de ADN recombinante son ampliamente conocidas en la técnica.

60 [0169] La secuencia de control puede ser un promotor, un polinucleótido que es reconocido por una célula huésped para la expresión de un polinucleótido que codifica un polipéptido de la presente invención. El promotor contiene secuencias de control de la transcripción que median en la expresión del polipéptido. El promotor puede ser cualquier polinucleótido que muestre actividad transcripcional en la célula huésped incluyendo promotores mutantes, truncados e híbridos, y puede obtenerse a partir de genes que codifican polipéptidos extracelulares o intracelulares homólogos o heterólogos a la célula huésped.

65

[0170] Ejemplos de promotores adecuados para dirigir la transcripción de las construcciones de ácido nucleico aquí descritos en una célula huésped bacteriana son los promotores obtenidos del gen de alfa-amilasa (*amyQ*) de *Bacillus amyloliquefaciens*, gen de alfa-amilasa (*amyL*) de *Bacillus licheniformis*, gen de penicilinas (*PenP*) de *Bacillus licheniformis*, gen de amilasa maltogénica (*amyM*) de *Bacillus stearothermophilus*, gen de levansacarasa (*SacB*) de *Bacillus subtilis*, los genes *xylA* y *xylB* de *Bacillus subtilis*, el gen *cryIIIA* de *Bacillus thuringiensis* (Agaisse and Lereclus, 1994, Molecular Microbiology 13: 97-107), el operón *lac* de *E. coli*, el promotor *trc* de *E. coli* (Egon et al., 1988, Gene 69: 301-315), el gen de agarasa (*dagA*) de *Streptomyces coelicolor*, y gen de beta-lactamasa procariota (Villa-Kamaroff et al, 1978, Proc. Natl. Acad. Sci. USA 75: 3727-3731), así como el promotor *tac* (DeBoer et al, 1983, Proc. Natl. Acad. Sci. USA 80: 21-25) Otros promotores se describen en "Useful proteins from recombinant bacteria" en Gilbert et al., 1980, Scientific American 242: 74-94; y en Sambrook et al., 1989, *supra*. Algunos ejemplos de promotores en tándem se describen en WO 99/43835.

[0171] Ejemplos de promotores adecuados para dirigir la transcripción de las construcciones de ácidos nucleicos aquí descritos en una célula huésped fúngica filamentosa son promotores obtenidos a partir de los genes para acetamidasa de *Aspergillus nidulans*, alfa-amilasa neutra de *Aspergillus niger*, alfa-amilasa estable a los ácidos de *Aspergillus niger*, glucoamilasa (*glaA*) de *Aspergillus niger* o *Aspergillus awamori*, TAKA amilasa de *Aspergillus oryzae*, proteasa alcalina de *Aspergillus oryzae*, triosa fosfato isomerasa de *Aspergillus oryzae*, proteasa de tipo tripsina de *Fusarium oxysporum* (WO 96/00787), amiloglucosidasa de *Fusarium venenatum* (WO 00/56900), Daria de *Fusarium venenatum* (WO 00/56900), Quinn de *Fusarium venenatum* (WO 00/56900), lipasa de *Rhizomucor miehei*, proteinasa aspártica de *Rhizomucor miehei*, beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasa I de *Trichoderma reesei*, xilanasa II de *Trichoderma reesei*, xilanasa III de *Trichoderma reesei*, beta-xilosidasa de *Trichoderma reesei*, y factor de elongación de la traducción de *Trichoderma reesei*, así como el promotor NA2-tpi (un promotor modificado de un gen neutro de alfa-amilasa de *Aspergillus* en el que el líder no traducido ha sido reemplazado por un líder no traducido de un gen de triosa fosfato isomerasa de *Aspergillus*; ejemplos no limitativos incluyen promotores modificados de un gen neutro de alfa-amilasa de *Aspergillus niger* en el que el líder no traducido ha sido reemplazado por un líder no traducido de un gen de triosa fosfato isomerasa de *Aspergillus nidulans* o *Aspergillus oryzae*); y promotores mutantes, truncados e híbridos de los mismos. Otros promotores se describen en la patente de EE. UU. No. 6,011,147.

[0172] En un huésped de levadura, se obtienen promotores útiles de los genes para enolasa (ENO-1) de *Saccharomyces cerevisiae*, galactoquinasa (GAL1) de *Saccharomyces cerevisiae*, alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa (ADH1, ADH2/GAP) de *Saccharomyces cerevisiae*, triosa fosfato isomerasa (TPI) de *Saccharomyces cerevisiae*, metalotioneína (CUP1) de *Saccharomyces cerevisiae*, y 3-fosfoglicerato cinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para las células huésped de levadura son descritos por Romanos et al., 1992, Yeast 8: 423-488.

[0173] La secuencia de control también puede ser un terminador de la transcripción, que es reconocido por una célula huésped para terminar la transcripción. El terminador está unido operativamente al extremo 3' terminal del polinucleótido que codifica el polipéptido. Cualquier terminador que sea funcional en la célula huésped puede usarse en la presente invención.

[0174] Los terminadores preferidos para las células huésped bacterianas se obtienen de los genes para proteasa alcalina (*aprH*) de *Bacillus clausii*, alfa-amilasa (*amyL*) de *Bacillus licheniformis*, y ARN ribosómico (*rnmB*) de *Escherichia coli*.

[0175] Los terminadores preferidos para células huésped de hongos filamentosos se obtienen de los genes para acetamidasa de *Aspergillus nidulans*, antranilato sintasa de *Aspergillus nidulans*, glucoamilasa de *Aspergillus niger*, alfa-glucosidasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae*, proteasa de tipo tripsina de *Fusarium oxysporum*, beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasa I de *Trichoderma reesei*, xilanasa II de *Trichoderma reesei*, xilanasa III de *Trichoderma reesei*, beta-xilosidasa de *Trichoderma reesei*, y factor de elongación de la traducción de *Trichoderma reesei*.

[0176] Los terminadores preferidos para las células huésped de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae*, citocromo C (CYC1) de *Saccharomyces cerevisiae*, y gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae*. Otros terminadores útiles para las células huésped de levadura son descritos por Romanos et al., 1992, *supra*.

[0177] La secuencia de control también puede ser una región estabilizadora de ARNm en dirección 5' respecto de un promotor y en dirección 3' de la secuencia codificante de un gen que aumenta la expresión del gen.

[0178] Se obtienen ejemplos de regiones estabilizadoras de ARNm adecuadas a partir de un gen *cryIIIA* de

*Bacillus thuringiensis* (WO 94/25612) y un gen SP82 de *Bacillus subtilis* (Hue et al., 1995, Journal of Bacteriology 177: 3465-3471).

5 [0179] La secuencia de control también puede ser un líder, una región no traducida de un ARNm que es importante para la traducción por parte de la célula huésped. El líder está unido operativamente al extremo 5' terminal del polinucleótido que codifica el polipéptido. Se puede usar cualquier líder que sea funcional en la célula anfitriona.

10 [0180] Los líderes preferidos para las células huésped fúngicas filamentosas se obtienen de los genes de TAKA amilasa de *Aspergillus oryzae* y triosa fosfato isomerasa de *Aspergillus nidulans*.

15 [0181] Los líderes adecuados para las células huésped de levadura se obtienen de los genes de enolasa (ENO-1) de *Saccharomyces cerevisiae*, 3-fosfoglicerato cinasa de *Saccharomyces cerevisiae*, factor alfa de *Saccharomyces cerevisiae*, y alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa (ADH2/GAP) de *Saccharomyces cerevisiae*.

20 [0182] La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia operativamente unida al extremo 3' terminal del polinucleótido y, cuando se transcribe, es reconocida por la célula huésped como una señal para añadir residuos de poliadenosina al ARNm transcrito. Se puede usar cualquier secuencia de poliadenilación que sea funcional en la célula huésped.

25 [0183] Las secuencias de poliadenilación preferidas para las células huésped fúngicas filamentosas se obtienen de los genes para antranilato sintasa de *Aspergillus nidulans*, glucoamilasa de *Aspergillus niger*, alfa-glucosidasa de *Aspergillus niger*, TAKA amilasa de *Aspergillus oryzae*, y proteasa de tipo tripsina de *Fusarium oxysporum*.

[0184] En Guo y Sherman, 1995, Mol. Cellular Biol. 15:5983-5990 se describen secuencias de poliadenilación útiles para células huésped de levadura.

30 [0185] La secuencia de control también puede ser una región codificante de péptido señal que codifica un péptido señal unido al extremo N terminal de un polipéptido y dirige el polipéptido hasta la vía secretora de la célula. El extremo 5' terminal de la secuencia de codificación del polinucleótido puede contener inherentemente una secuencia de codificación de péptido señal unida naturalmente en el marco de lectura de traducción con el segmento de la secuencia de codificación que codifica el polipéptido. Alternativamente, el extremo 5' terminal de la secuencia de codificación puede contener una secuencia de codificación de péptido señal que es foránea a la secuencia de codificación. Puede requerirse una secuencia codificante de péptido señal foránea cuando la secuencia codificante no contiene naturalmente una secuencia codificante de péptido señal. Alternativamente, una secuencia codificante de péptido señal foránea puede simplemente reemplazar la secuencia codificante del péptido señal natural para potenciar la secreción del polipéptido. Sin embargo, puede usarse cualquier secuencia codificante de péptido señal que dirija el polipéptido expresado hasta la vía secretora de una célula huésped.

40 [0186] Secuencias codificantes de péptido señal eficaces para células huésped bacterianas son las secuencias codificantes de péptido señal obtenidas de los genes para amilasa maltogénica NCIB 11837 de *Bacillus subtilis*, subtilisina de *Bacillus licheniformis*, beta-lactamasa de *Bacillus licheniformis*, alfa-amilasa de *Bacillus stearothermophilus*, proteasas neutras (*nprT*, *nprS*, *nprM*) de *Bacillus stearothermophilus*, y *prsA* de *Bacillus subtilis*. Otros péptidos señal son descritos por Simonen y Palva, 1993, Microbiological Reviews 57: 109-137.

50 [0187] Secuencias codificantes de péptidos de señal eficaces para células huésped fúngicas filamentosas son las secuencias codificantes de péptido señal obtenidas de los genes para amilasa neutral de *Aspergillus niger*, nigerglucoamilasa de *Aspergillus*, TAKA amilasa de *Aspergillus oryzae*, celulasa de *Humicola insolens*, endoglucanasa V de *Humicola insolens*, lipasa de *Humicola lanuginosa*, y proteinasa aspártica de *Rhizomucor miehei*.

55 [0188] Péptidos señal útiles para células huésped de levadura se obtienen de los genes para factor alfa de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*. Otras secuencias codificantes de péptidos señal útiles son descritas por Romanos *et al.*, 1992, *supra*.

60 [0189] La secuencia de control también puede ser una secuencia codificante de propéptido que codifica un propéptido situado en el extremo N terminal de un polipéptido. El polipéptido resultante se conoce como una proenzima o propolipéptido (o un zimógeno en algunos casos). Un propolipéptido generalmente es inactivo y puede convertirse en un polipéptido activo por escisión catalítica o autocatalítica del propéptido respecto del propolipéptido. La secuencia de codificación del propéptido se puede obtener de los genes para proteasa alcalina (*aprE*) de *Bacillus subtilis*, proteasa neutra (*nprT*) de *Bacillus subtilis*, lacasa de *Myceliophthora thermophila* (WO 95/33836), proteinasa aspártica de *Rhizomucor miehei*, y factor alfa de *Saccharomyces cerevisiae*.

65 [0190] Cuando están presentes tanto las secuencias de péptido señal como de propéptido, la secuencia de propéptido se coloca junto al extremo N terminal de un polipéptido y la secuencia de péptido señal se posiciona

junto al extremo N terminal de la secuencia del propéptido.

[0191] También puede ser deseable añadir secuencias reguladoras que regulen la expresión del polipéptido en relación con el crecimiento de la célula huésped. Ejemplos de secuencias reguladoras son aquellas que hacen que la expresión del gen se active o desactive en respuesta a un estímulo químico o físico, incluyendo la presencia de un compuesto regulador. Las secuencias reguladoras en los sistemas procarióticos incluyen sistemas de operón *lac*, *tac*, y *trp*. En levaduras, se puede usar el sistema ADH2 o el sistema GAL1. En hongos filamentosos, se puede usar el promotor de glucoamilasa de *Aspergillus niger*, el promotor de TAKA alfa-amilasa de *Aspergillus oryzae*, y el promotor de glucoamilasa de *Aspergillus oryzae*, el promotor de celobiohidrolasa I de *Trichoderma reesei*, y el promotor de celobiohidrolasa II de *Trichoderma reesei*. Otros ejemplos de secuencias reguladoras son aquellas que permiten la amplificación génica. En los sistemas eucariotas, estas secuencias reguladoras incluyen el gen de dihidrofolato reductasa que se amplifica en presencia de metotrexato y los genes de metalotioneína que se amplifican con metales pesados. En estos casos, el polinucleótido que codifica el polipéptido estaría operativamente unido a la secuencia reguladora.

#### Vectores de expresión

[0192] También se describen vectores de expresión recombinantes que comprenden un polinucleótido de la presente invención, un promotor y señales de parada transcripcionales y traduccionales. Las diversas secuencias de nucleótidos y control se pueden unir para producir un vector de expresión recombinante que puede incluir uno o más sitios de restricción convenientes para permitir la inserción o sustitución del polinucleótido que codifica el polipéptido en dichos sitios. Alternativamente, el polinucleótido puede expresarse insertando el polinucleótido o una construcción de ácido nucleico que comprende el polinucleótido en un vector apropiado para la expresión. Al crear el vector de expresión, la secuencia de codificación se ubica en el vector de modo que la secuencia de codificación se une operativamente con las secuencias de control apropiadas para la expresión.

[0193] El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido o virus) que se puede someter convenientemente a procedimientos de ADN recombinante y puede provocar la expresión del polinucleótido. La elección del vector dependerá típicamente de la compatibilidad del vector con la célula huésped en la que se va a introducir el vector. El vector puede ser un plásmido circular lineal o cerrado.

[0194] El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la 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. Alternativamente, el vector puede ser uno que, cuando se introduce en la célula huésped, se integra en el genoma y se replica junto con el/los cromosoma (s) en el/los que se ha integrado. Además, se puede usar un único vector 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.

[0195] El vector preferiblemente contiene uno o más marcadores seleccionables que permiten la fácil selección de células transformadas, transfectadas, transducidas o similares. Un marcador seleccionable es un gen cuyo producto proporciona resistencia a biocidas o vírica, resistencia a metales pesados, prototrofia a auxótrofos y similares.

[0196] Ejemplos de marcadores seleccionables bacterianos son genes *dal* de *Bacillus licheniformis* o *Bacillus subtilis* o marcadores que confieren resistencia a antibióticos, tal como resistencia a ampicilina, cloranfenicol, kanamicina, neomicina, espectinomicina o tetraciclina. Los marcadores adecuados para células huésped de levadura incluyen, pero no se limitan a, ADE2, HIS3, LEU2, LYS2, MET3, TRP1 y URA3. Los marcadores seleccionables para uso en una célula huésped fúngica filamentosa incluyen, entre otros, *AdeA* (fosforribosilaminoimidazol-succinocarboxamida sintasa), *adeB* (fosforribosilaminoimidazol sintasa), *amdS* (acetamidasa), *argB* (ornitina carbamoiltransferasa), *bar* (fosfotricina acetiltransferasa), *hph* (higromicina fosfotransferasa), *niaD* (nitrato reductasa), *pyrG* (orotidina-5'-fosfato descarboxilasa), *sC* (sulfato de adeniltransferasa), y *trpC* (antranilato sintasa), así como sus equivalentes. Para usar en una célula de *Aspergillus* se prefieren los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y un gen *bar* *Streptomyces hygroscopicus*. Para usar en una célula de *Trichoderma* se prefieren los genes *adeA*, *adeB*, *amdS*, *hph*, y *pyrG*.

[0197] El marcador seleccionable puede ser un sistema de marcador seleccionable dual como se describe en WO 2010/039889. En un aspecto, el marcador seleccionable dual es un sistema de marcador dual seleccionable *hph-tk*.

[0198] El vector preferiblemente contiene un elemento o elementos que permite(n) la integración del vector en el genoma de la célula huésped o la replicación autónoma del vector en la célula independiente del genoma.

[0199] Para la integración en el genoma de la célula huésped, el vector puede basarse en la secuencia del polinucleótido que codifica el polipéptido o cualquier otro elemento del vector para su integración en el genoma mediante recombinación homóloga o no homóloga. Alternativamente, el vector puede contener polinucleótidos

adicionales para dirigir la integración mediante recombinación homóloga en el genoma de la célula huésped en una ubicación o ubicaciones precisas en el cromosoma o cromosomas. Para aumentar la probabilidad de integración en una ubicación precisa, los elementos de integración deben contener un número suficiente de ácidos nucleicos, tal como de 100 a 10 000 pares de bases, de 400 a 10 000 pares de bases y de 800 a 10 000 pares de bases, que tienen un alto grado de identidad de secuencia respecto a la secuencia diana correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos de integración pueden ser cualquier secuencia que sea homóloga con la secuencia diana en el genoma de la célula huésped. Además, los elementos de integración pueden ser polinucleótidos no codificantes o codificantes. Por otro lado, el vector puede integrarse en el genoma de la célula huésped mediante recombinación no homóloga.

[0200] Para la replicación autónoma, el vector puede comprender además un origen de replicación que permite que el vector se replique autónomamente en la célula huésped en cuestión. El origen de replicación puede ser cualquier replicador de plásmido que medie en la replicación autónoma que funciona en una célula. El término "origen de replicación" o "replicador de plásmido" significa un polinucleótido que permite la replicación de un plásmido o un vector *in vivo*.

[0201] Ejemplos de orígenes bacterianos de replicación son los orígenes de la replicación de los plásmidos pBR322, pUC19, pACYC177 y pACYC184 que permiten la replicación en *E coli*, y pUB110, pE194, pTA1060 y pAMβ1 que permiten la replicación en *Bacillus*.

[0202] Ejemplos de orígenes de replicación para usar en una célula huésped de levadura son el origen de replicación de 2 micras, ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6.

[0203] 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 Res. 15:9163-9175; WO 00/24883) El aislamiento del gen AMA1 y la construcción de plásmidos o vectores que comprenden el gen se pueden lograr según los métodos descritos en WO 00/24883.

[0204] Se puede insertar más de una copia de un polinucleótido de la presente invención en una célula huésped para aumentar la producción de un polipéptido. Puede obtenerse un aumento en el número de copias del polinucleótido integrando 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 en el que las células contienen copias amplificadas del gen marcador seleccionable, y por lo tanto, se pueden seleccionar copias adicionales del polinucleótido cultivando las células en presencia del agente seleccionable apropiado.

[0205] Los procedimientos usados para ligar los elementos descritos anteriormente para construir los vectores de expresión recombinantes de la presente invención son ampliamente conocidos por los expertos en la técnica (véase, por ejemplo, Sambrook *et al.*, 1989, *supra*).

#### Células huésped

[0206] En este documento se describen células huésped recombinantes, que comprenden un polinucleótido de la presente invención operativamente unido a una o más secuencias de control que dirigen la producción de un polipéptido de la presente invención. Se introduce una construcción o vector que comprende un polinucleótido en una célula huésped de modo que la construcción o vector se mantiene como un integrante cromosómico o como un vector extracromosómico autorreplicante como se ha descrito anteriormente. El término "célula huésped" abarca cualquier progenie de una célula progenitora que no es idéntica a la célula progenitora debido a mutaciones que se producen durante la replicación. La elección de una célula huésped dependerá en gran medida del gen que codifica el polipéptido y su fuente.

[0207] La célula huésped puede ser cualquier célula útil en la producción recombinante de un polipéptido de la presente invención, por ejemplo, una procariota o una eucariota.

[0208] La célula huésped procariota puede ser cualquier bacteria Gram-positiva o Gram-negativa. Las bacterias Gram-positivas incluyen, pero no están limitadas a, *Bacillus*, *Clostridium*, *Enterococcus*, *Geobacillus*, *Lactobacillus*, *Lactococcus*, *Oceanobacillus*, *Staphylococcus*, *Streptococcus*, y *Streptomyces*. Las bacterias Gram-negativas incluyen, entre otras, *Campylobacter*, *E. coli*, *Flavobacterium*, *Fusobacterium*, *Helicobacter*, *Ilyobacter*, *Neisseria*, *Pseudomonas*, *Salmonella*, y *Ureaplasma*.

[0209] La célula huésped bacteriana puede ser cualquier célula de *Bacillus* incluyendo, pero sin limitarse a, células de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, y *Bacillus turingiensis*.

[0210] La célula huésped bacteriana también puede ser cualquier célula de *Streptococcus*, incluyendo, pero sin limitarse a, células de *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, y *Streptococcus*

*equi* subsp. *Zooepidemicus*.

[0211] La célula huésped bacteriana también puede ser cualquier célula de *Streptomyces*, incluyendo, pero no sin limitarse a, células de *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, y *Streptomyces lividans*.

[0212] La introducción de ADN en una célula de *Bacillus* puede efectuarse por transformación de protoplastos (véase, por ejemplo, Chang y Cohen, 1979, Mol. Gen. Genet. 168:111-115), transformación de células competentes (véase, por ejemplo, Young y Spizizen, 1961, J. Bacteriol. 81:823-829, o Dubnau y Davidoff-Abelson, 1971, J. Mol. Biol. 56:209-221), electroporación (véase, por ejemplo, Shigekawa y Dower, 1988, Biotechniques 6: 742-751), o conjugación (véase, por ejemplo, Koehler y Thorne, 1987, J. Bacteriol. 169:5271-5278) La introducción de ADN en una célula de *E coli* puede efectuarse por transformación de protoplastos (véase, por ejemplo, Hanahan, 1983, J. Mol. Biol. 166:557-580) o electroporación (véase, por ejemplo, Dower et al., 1988, Nucleic Acids Res. 16:6127-6145) La introducción de ADN en una célula de *Streptomyces* puede efectuarse por transformación de protoplastos, electroporación (ver, por ejemplo, Gong et al, 2004, Folia Microbiol. (Praga) 49: 399-405), conjugación (ver, por ejemplo, Mazodier et al, 1989, J. Bacteriol. 171:3583-3585), o transducción (ver, por ejemplo, Burke y col., 2001, Proc. Natl. Acad. Sci. USA 98: 6289-6294) La introducción de ADN en una célula de *Pseudomonas* puede efectuarse por electroporación (ver, por ejemplo, Choi y col., 2006, J. Microbiol. Methods 64: 391-397) o conjugación (ver, por ejemplo, Pinedo y Smets, 2005, Appl. Environ. Microbiol. 71:51-57) La introducción del ADN en una célula de *Sstreptococcus* puede efectuarse por la competencia natural (ver, por ejemplo, Perry y Kuramitsu, 1981, Infect. Immun. 32:1295-1297), transformación de protoplastos (ver, por ejemplo, Catt y Jollick, 1991, Microbios 68: 189-207), electroporación (ver, por ejemplo, Buckley et al., 1999, Appl. Environ. Microbiol. 65:3800-3804), o conjugación (ver, por ejemplo, Clewell, 1981, Microbiol. Rdo. 45:409-436) Sin embargo, puede usarse cualquier método conocido en la técnica para introducir ADN en una célula huésped.

[0213] La célula huésped también puede ser eucariota, tal como una célula de mamífero, insecto, planta u hongo.

[0214] La célula huésped puede ser una célula fúngica. "Hongos" como se usa en este documento incluye los filos Ascomycota, Basidiomycota, Chytridiomycota y Zygomycota, así como los Oomycota y todos los hongos mitospóricos (tal como se define en Hawksworth et al., In, Ainsworth y Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, Reino Unido).

[0215] La célula huésped fúngica puede ser una célula de levadura. "Levadura" como se usa en el presente documento incluye levadura ascospórica (Endomycetales), levadura basidiospórica y levadura que pertenece a los Fungi Imperfecti (Blastomycetes). Dado que la clasificación de las levaduras puede cambiar en el futuro, para los fines de esta invención la levadura se definirá como se describe en Biology and Activities of Yeast (Skinner, Passmore, y Davenport, editors, Soc. Aplicación Bacteriol. Symposium Series No. 9, 1980).

[0216] La célula huésped de levadura puede ser una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia*, tal como una célula de *Kluyveromyces lactis*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis*, o *Yarrowia lipolytica*.

[0217] La célula huésped fúngica puede ser una célula fúngica filamentosa. Los "hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión Eumycota y Oomycota (según la definición de Hawksworth). *et al.*, 1995, *supra*) Los hongos filamentosos generalmente se caracterizan por una pared micelial compuesta de quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos. El crecimiento vegetativo es por crecimiento de la hifa y el catabolismo del carbono es obligatoriamente aeróbico. Por el contrario, el crecimiento vegetativo de levaduras como *Saccharomyces cerevisiae* es por gemación de un tallo unicelular y el catabolismo de carbono puede ser fermentativo.

[0218] La célula huésped fúngica filamentosa puede ser una célula de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolyocladium*, *Trametes*, o *Trichoderma*.

[0219] Por ejemplo, la célula huésped fúngica filamentosa puede ser una célula de *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus fumigatus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Chrysosporium inops*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium merdarium*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium tropicum*, *Chrysosporium zonatum*, *Coprinus cinereus*, *Coriolus hirsutus*, *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 sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *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 *Trichoderma viride*.

[0220] Las células fúngicas se pueden transformar mediante un proceso que implica la formación de protoplastos, la transformación de los protoplastos y la regeneración de la pared celular de una manera conocida *per se*. Ciertos procedimientos adecuados para la transformación de células huésped de *Aspergillus* y *Trichoderma* se describen en EP 238023, Yelton *et al*, 1984, Proc. Natl. Acad. Sci. USA 81: 1470-1474 y Christensen *et al*, 1988, Bio/Technology 6: 1419-1422. Métodos adecuados para la transformación de especies de *Fusarium* son descritos por Malardier *et al.*, 1989, Gene 78: 147-156y WO 96/00787. La levadura se puede transformar usando los procedimientos descritos por Becker y Guarente, en Abelson, J.N. y Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volumen 194, pp 182-187, Academic Press, Inc., New York; Ito *et al*, 1983, J. Bacteriol. 153:163; y Hinnen *et al*, 1978, Proc. Natl. Acad. Sci. Estados Unidos 75: 1920.

Métodos de producción

[0221] También se describen métodos para producir un polipéptido de la presente invención, que comprenden (a) cultivar una célula, que en su forma silvestre produce el polipéptido, en condiciones que conducen a la producción del polipéptido; y opcionalmente, (b) recuperar el polipéptido. En un aspecto, la célula es una célula de *Bacteroides*. En otro aspecto, la célula es una célula de *Bacteroides ovatus*. En otro aspecto, la célula es *Bacteroides ovatus* ATCC8483. Se divulgan métodos para producir un polipéptido de la presente invención, que comprenden (a) cultivar una célula huésped recombinante de la presente invención en condiciones que conducen a la producción del polipéptido; y opcionalmente, (b) recuperar el polipéptido.

[0222] Las células huésped se cultivan en un medio nutritivo adecuado para la producción del polipéptido usando métodos conocidos en la técnica. Por ejemplo, las células pueden cultivarse mediante cultivo en matraz de agitación, o fermentación a pequeña escala o a gran escala (que incluye fermentaciones continuas, discontinuas, de lote alimentado o en estado sólido) en fermentadores industriales o de laboratorio en un medio adecuado y en condiciones que permitan expresar y/o aislar el polipéptido. El cultivo tiene lugar en un medio nutriente adecuado que comprende fuentes de carbono y nitrógeno y sales inorgánicas, usando procedimientos conocidos en la técnica. Los medios adecuados están disponibles a través de proveedores comerciales o pueden prepararse según las composiciones publicadas (por ejemplo, en catálogos de la American Type Culture Collection). Si el polipéptido se secreta en el medio nutritivo, el polipéptido se puede recuperar directamente del medio. Si el polipéptido no se secreta, se puede recuperar de lisados celulares.

[0223] El polipéptido se puede detectar usando métodos conocidos en la técnica que son específicos para los polipéptidos. Estos métodos de detección incluyen, entre otros, el uso de anticuerpos específicos, la formación de un producto enzimático o la desaparición de un sustrato enzimático. Por ejemplo, un ensayo enzimático que mide la liberación de L-galactosa a partir de xilano de salvado de maíz se puede usar para determinar la actividad del polipéptido o para medir la liberación enzimática de para-nitrofenol a partir de un sustrato artificial de PNP  $\alpha$ -L-galactosa.

[0224] El polipéptido puede recuperarse usando métodos conocidos en la técnica. Por ejemplo, el polipéptido puede recuperarse del medio nutritivo mediante procedimientos convencionales que incluyen, pero no se limitan a, recolección, centrifugación, filtración, extracción, secado por pulverización, evaporación o precipitación. En un aspecto, se recupera un caldo de fermentación que comprende el polipéptido.

[0225] El polipéptido se puede purificar mediante una variedad de procedimientos conocidos en la técnica que incluyen, pero sin limitarse a, cromatografía (por ejemplo, de intercambio iónico, por afinidad, de interacción hidrofóbica, cromatografía de exclusión por tamaño), procedimientos de electroforesis (por ejemplo, enfoque isoelectrico preparativo), solubilidad diferencial (por ejemplo, precipitación con sulfato de amonio), SDS-PAGE o extracción (ver, por ejemplo, Protein Purification, Janson y Ryden, editors, VCH Publishers, New York, 1989) para obtener polipéptidos sustancialmente puros.

[0226] En un aspecto alternativo, el polipéptido no se recupera, sino que se usa una célula huésped de la presente invención que expresa el polipéptido como fuente del polipéptido.

Eliminación o reducción de la actividad de  $\alpha$  L-galactosidasa

[0227] Aquí también se describen métodos para producir una mutante de una célula progenitora, que comprenden alterar o delecionar un polinucleótido, o una porción del mismo, que codifica un polipéptido de la

presente invención, lo que da como resultado que la célula mutante produzca menos del polipéptido que la célula progenitora cuando se cultiva bajo las mismas condiciones.

[0228] La célula mutante puede construirse reduciendo o eliminando la expresión del polinucleótido usando métodos ampliamente conocidos en la técnica, por ejemplo, inserciones, interrupciones, reemplazos o deleciones. En un aspecto preferido, el polinucleótido es inactivado. El polinucleótido que se desea modificar o inactivar puede ser, por ejemplo, la región codificante o una parte de la misma esencial para la actividad, o un elemento regulador requerido para la expresión de la región codificante. Un ejemplo de dicha secuencia reguladora o de control puede ser una secuencia promotora o una parte funcional de la misma, es decir., una parte que es suficiente para afectar a la expresión del polinucleótido. Otras secuencias de control para una posible modificación incluyen, pero sin limitarse a, un líder, secuencia de poliadenilación, secuencia de propéptido, secuencia de péptido señal, terminador de la transcripción y activador de la transcripción.

[0229] La modificación o inactivación del polinucleótido se puede realizar sometiendo a la célula progenitora a mutagénesis y seleccionando células mutantes en las que se ha reducido o eliminado la expresión del polinucleótido. La mutagénesis, que puede ser específica o aleatoria, puede realizarse, por ejemplo, mediante el uso de un agente mutagenizante físico o químico adecuado, mediante el uso de un oligonucleótido adecuado, o sometiendo la secuencia de ADN a mutagénesis generada por PCR. Además, la mutagénesis puede realizarse mediante el uso de cualquier combinación de estos agentes mutagenizantes.

[0230] Los ejemplos de un agente mutagenizante físico o químico adecuado para el presente propósito incluyen radiación ultravioleta (UV), hidroxilamina, N-metil-N'-nitro-N-nitrosoguanidina (MNNG), O-metilhidroxilamina, ácido nitroso, metanosulfonato de etilo (EMS), bisulfito de sodio, ácido fórmico y análogos de nucleótidos.

[0231] Cuando se usan tales agentes, la mutagénesis se realiza típicamente incubando la célula progenitora que se desea mutagenizar en presencia del agente de mutagénesis de elección en condiciones adecuadas, y cribando y/o seleccionando células mutantes que muestren una expresión reducida o nula del gen.

[0232] La modificación o inactivación del polinucleótido puede realizarse por inserción, sustitución o deleción de uno o más nucleótidos en el gen o un elemento regulador requerido para la transcripción o traducción del mismo. Por ejemplo, se puede insertar o eliminar nucleótidos para dar como resultado la introducción de un codón de terminación, la eliminación del codón de inicio o un cambio en el marco de lectura abierto. Dicha modificación o inactivación se puede llevar a cabo mediante mutagénesis dirigida o mutagénesis generada por PCR de acuerdo con métodos conocidos en la técnica. Aunque, en principio, la modificación se puede realizar *in vivo*, es decir, directamente sobre la célula que expresa el polinucleótido que se desea modificar, se prefiere que la modificación se realice *in vitro* como se ejemplifica a continuación.

[0233] Un ejemplo de una forma conveniente de eliminar o reducir la expresión de un polinucleótido se basa en técnicas de sustitución génica, eliminación génica o interrupción génica. Por ejemplo, en el método de interrupción génica, una secuencia de ácido nucleico correspondiente al polinucleótido endógeno se mutageniza *in vitro* para producir una secuencia de ácido nucleico defectuosa que luego se transforma en la célula progenitora para producir un gen defectuoso. Mediante recombinación homóloga, la secuencia defectuosa de ácido nucleico reemplaza al polinucleótido endógeno. Puede ser deseable que el polinucleótido defectuoso también codifique un marcador que pueda usarse para la selección de transformantes en los que el polinucleótido se ha modificado o destruido. En un aspecto, el polinucleótido se rompe con un marcador seleccionable tal como los descritos en este documento. También se describen métodos para inhibir la expresión de un polipéptido que tiene actividad de  $\alpha$ -L-galactosidasa en una célula, que comprende administrar a la célula o expresar en la célula una molécula de ARN de doble cadena (ARNdc), en la que el ARNdc comprende una subsecuencia de un polinucleótido de la presente invención. En un aspecto preferido, el ARNdc tiene aproximadamente 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 o más nucleótidos dúplex de longitud.

[0234] El ARNdc es preferiblemente un ARN interferente pequeño (ARNip) o un micro ARN (miARN). En un aspecto preferido, el ARNdc es un ARN interferente pequeño para inhibir la transcripción. En otro aspecto preferido, el ARNdc es micro ARN para inhibir la traducción. También se describen moléculas de ARN de doble cadena (ARNdc), que comprenden una porción de la secuencia que codifica el polipéptido maduro de la SEQ ID NO: 1 para inhibir la expresión del polipéptido en una célula. Si bien la presente invención no está limitada por ningún mecanismo particular de acción, el ARNdc puede entrar en una célula y causar la degradación de un ARN monocatenario (ARNmc) de secuencias similares o idénticas, que incluyen ARNm endógenos. Cuando una célula está expuesta a ARNdc, el ARNm del gen homólogo se degrada de forma selectiva mediante un proceso denominado interferencia por ARN (iARN).

[0235] Los ARNdc pueden usarse en el silenciamiento génico. En un aspecto, se proporcionan métodos para degradar selectivamente ARN usando una iARNdc aquí descrita. El proceso puede ser practicado *in vitro*, *ex vivo* o *in vivo*. En un aspecto, las moléculas de ARNdc se pueden usar para generar una mutación de pérdida de función en una célula, un órgano o un animal. Los métodos para fabricar y usar moléculas de ARNdc para degradar RNA selectivamente son ampliamente conocidos en la técnica; véase, por ejemplo, las patentes de EE.



UU. nº 6,489,127; 6,506,559; 6,511,824; y 6,515,109. También se describe una célula mutante de una célula progenitora que comprende una interrupción o delección de un polinucleótido que codifica el polipéptido o una secuencia de control del mismo o un gen silenciado que codifica el polipéptido, que da como resultado que la célula mutante produzca menos polipéptidos o ningún polipéptido en comparación con la célula progenitora.

5 [0236] Las células mutantes deficientes en polipéptidos son particularmente útiles como células huésped para la expresión de polipéptidos nativos y heterólogos. Por lo tanto, la presente invención se refiere además a métodos para producir un polipéptido nativo o heterólogo, que comprende (a) cultivar la célula mutante en condiciones que conducen a la producción del polipéptido; y (b) recuperar el polipéptido. El término "polipéptidos heterólogos" significa polipéptidos que no son nativos de la célula huésped, por ejemplo, una variante de una proteína nativa. La célula huésped puede comprender más de una copia de un polinucleótido que codifica el polipéptido nativo o heterólogo.

10 [0237] Los métodos usados para el cultivo y la purificación del producto de interés se pueden realizar mediante métodos conocidos en la técnica.

15 [0238] Los métodos aquí descritos para producir un producto esencialmente libre de  $\alpha$ -L-galactosidasa son de particular interés en la producción de polipéptidos eucariotas, en particular proteínas fúngicas tales como enzimas. Las células deficientes en  $\alpha$ -L-galactosidasa también pueden usarse para expresar proteínas heterólogas de interés farmacéutico tales como hormonas, factores de crecimiento, receptores y similares. El término "polipéptidos eucariotas" incluye no solo polipéptidos nativos, sino también aquellos polipéptidos, por ejemplo, enzimas, que han sido modificados por sustituciones, deleciones o adiciones de aminoácidos, u otras modificaciones de este tipo para mejorar la actividad, la termoestabilidad, la tolerancia del pH y similares.

20 [0239] En otro aspecto, también se divulga un producto proteínico esencialmente libre de actividad de  $\alpha$ -L-galactosidasa que se produce mediante un método de la presente invención.

Formulaciones de caldo de fermentación o composiciones celulares

25 [0240] La presente invención también se refiere a una formulación de caldo de fermentación o una composición celular que comprende un polipéptido de la presente invención. El producto de caldo de fermentación comprende adicionalmente ingredientes adicionales utilizados en el proceso de fermentación, tales como, por ejemplo, células (incluyendo las células huésped que contienen el gen que codifica el polipéptido de la presente invención, que se usan para producir el polipéptido de interés), restos celulares, biomasa, medios de fermentación y/o productos de fermentación. En algunas formas de realización, la composición es un caldo entero de células inactivadas que contiene ácido(s) orgánico(s), células inactivadas y/o restos celulares, y medio de cultivo.

30 [0241] El término "caldo de fermentación", como se usa en el presente documento, se refiere a una preparación producida por fermentación celular que se somete a una recuperación y/o purificación nula o mínima. Por ejemplo, los caldos de fermentación se producen cuando los cultivos microbianos se cultivan hasta la saturación, se incuban en condiciones limitativas de carbono para permitir la síntesis de proteínas (por ejemplo, la expresión de enzimas por células huésped) y se secretan en un medio de cultivo celular. El caldo de fermentación puede contener contenidos no fraccionados o fraccionados de los materiales de fermentación derivados al final de la fermentación. Típicamente, el caldo de fermentación no está fraccionado y comprende el medio de cultivo agotado y los restos celulares presentes después de las células microbianas (por ejemplo, células de hongos filamentosos) se eliminan, por ejemplo, por centrifugación. En algunas realizaciones, el caldo de fermentación contiene medio de cultivo celular gastado, enzimas extracelulares y células microbianas viables y/o no viables.

35 [0242] En una forma de realización, la formulación de caldo de fermentación y las composiciones celulares comprenden un primer componente de ácido orgánico que comprende al menos un ácido orgánico de 1-5 carbonos y/o una sal del mismo y un segundo componente de ácido orgánico que comprende al menos un ácido orgánico de 6 o más carbonos y/o una sal del mismo. En una forma de realización específica, el primer componente de ácido orgánico es ácido acético, ácido fórmico, ácido propiónico, una sal del mismo, o una mezcla de dos o más de los anteriores y el segundo componente ácido orgánico es ácido benzoico, ácido ciclohexanocarboxílico, ácido 4-metilvalérico, ácido fenilacético, una sal del mismo, o una mezcla de dos o más de los anteriores.

40 [0243] En un aspecto, la composición contiene uno o varios ácido(s) orgánico(s) y, opcionalmente, además contiene células inactivadas y/o restos celulares. En una forma de realización, las células inactivadas y/o los restos celulares se eliminan de un caldo completo con inactivación celular para proporcionar una composición que está libre de estos componentes.

45 [0244] Las formulaciones de caldo de fermentación o las composiciones celulares pueden comprender además un agente conservante y/o antimicrobiano (por ejemplo, bacteriostático), que incluye, pero no se limita a, sorbitol, cloruro de sodio, sorbato de potasio y otros conocidos en la técnica.

[0245] El caldo entero o composición con inactivación celular puede contener los contenidos no fraccionados de los materiales de fermentación derivados al final de la fermentación. Típicamente, el caldo entero o composición con inactivación celular contiene el medio de cultivo gastado y los restos celulares presentes después de que las células microbianas (por ejemplo, células fúngicas filamentosas) se hayan cultivado hasta la saturación, incubadas en condiciones limitativas de carbono para permitir la síntesis de proteínas. En algunas formas de realización, el caldo entero o composición con inactivación celular contiene el medio de cultivo celular gastado, las enzimas extracelulares y las células fúngicas filamentosas inactivadas. En algunas formas de realización, las células microbianas presentes en el caldo entero o composición con inactivación celular se pueden permeabilizar y/o lisar usando métodos conocidos en la técnica.

[0246] Un caldo completo o composición celular como se describe en este documento es típicamente un líquido, pero puede contener componentes insolubles, tales como células inactivadas, restos celulares, componentes de medios de cultivo, y/o enzimas insolubles. En algunas formas de realización, los componentes insolubles se pueden eliminar para proporcionar una composición líquida clarificada.

[0247] Las formulaciones de caldo entero y las composiciones celulares de la presente invención se pueden producir mediante un método descrito en WO 90/15861 o WO 2010/096673.

#### Composiciones enzimáticas

[0248] La presente invención también se refiere a composiciones que comprenden un polipéptido de la presente invención. Preferiblemente, las composiciones están enriquecidas con dicho polipéptido. El término "enriquecido" indica que la actividad de  $\alpha$ -L-galactosidasa de la composición se ha incrementado, por ejemplo, con un factor de enriquecimiento de al menos 1,1.

[0249] En un aspecto, la composición comprende un polipéptido aislado que tiene actividad de  $\alpha$ -L-galactosidasa, seleccionado del grupo que consiste en:

(a) un polipéptido que tiene al menos 80%, por ejemplo, al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, o 100% de identidad de secuencia con la SEQ ID NO: 3 ( No. 1-792);

(b) un polipéptido codificado por un polinucleótido que se hibrida en condiciones de astringencia alta o en condiciones de astringencia muy alta con:

(i) la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1; y/o  
(iii) la cadena complementaria en toda su longitud de (i);

(c) un polipéptido codificado por un polinucleótido que tiene al menos 80%, por ejemplo, al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91 %, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, o 100% de identidad de secuencia con la secuencia codificante de polipéptido maduro de la SEQ ID NO: 1;

(d) una variante que comprende una sustitución, delección y/o inserción de uno o más (varios) aminoácidos del polipéptido maduro de la SEQ ID NO: 3; y

(e) un fragmento de un polipéptido de (a), (b), (c) o (d), que tiene actividad de  $\alpha$ -xilosidasa.

[0250] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 85% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.

[0251] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 86% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.

[0252] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 87% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.

[0253] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 88% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.

[0254] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 89% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.

[0255] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 90% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.

[0256] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 91% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.

- [0257] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 92% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.
- 5 [0258] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 93% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.
- [0259] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 94% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.
- 10 [0260] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 95% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.
- [0261] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 96% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.
- 15 [0262] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 97% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.
- [0263] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 98% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.
- 20 [0264] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene al menos un 99% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 3.
- 25 [0265] Una forma de realización de la invención es una composición que comprende un polipéptido aislado que tiene una identidad de secuencia del 100% con el polipéptido maduro de la SEQ ID NO: 3.
- [0266] En un aspecto, la composición comprende o consiste en la secuencia de aminoácidos madura de la SEQ ID NO: 3 o una variante alélica de la misma; o es un fragmento de la misma que tiene actividad de  $\alpha$ -L-galactosidasa. En otro aspecto, la composición comprende o consiste en el polipéptido maduro de la SEQ ID NO: 2. En un aspecto adicional, la composición comprende o consiste en el polipéptido maduro de la SEQ ID NO: 3. En otro aspecto, la composición comprende o consiste en los aminoácidos 1 a 160 de la SEQ ID NO: 2, los aminoácidos 5 a 154 de la SEQ ID NO: 2, o los aminoácidos 10 a 149 de la SEQ ID NO: 2. En otro aspecto, el polipéptido comprende o consiste en los aminoácidos 1 a 160 de la SEQ ID NO: 3, los aminoácidos 5 a 154 de la SEQ ID NO: 3, o los aminoácidos 10 a 149 de la SEQ ID NO: 3.
- 30 [0267] La presente invención también se refiere a composiciones que comprenden polipéptidos aislados que tienen actividad de  $\alpha$ -L-galactosidasa que son codificados por polinucleótidos que se hibridan en condiciones de astringencia alta o condiciones de astringencia muy alta con (i) la secuencia codificante de polipéptido maduro de la SEQ ID NO: 1, y/o (ii) la cadena complementaria de longitud completa de (i) (J. Sambrook, E.F. Fritsch, y T. Maniatis, 1989, Molecular Cloning, A Laboratory Manual, 2nd edition, Cold Spring Harbor, New York.).
- 40 [0268] La presente invención se refiere además a composiciones que comprenden polipéptidos aislados que difieren en no más de cien residuos de aminoácidos, por ejemplo, en noventa aminoácidos, en ochenta aminoácidos, en setenta aminoácidos, en sesenta aminoácidos, en cincuenta aminoácidos, en cuarenta aminoácidos, en treinta aminoácidos, en veinticinco aminoácidos, en veinte aminoácidos, en quince aminoácidos, en diez aminoácidos, en ocho aminoácidos, en siete aminoácidos, en seis aminoácidos, en cinco aminoácidos, en cuatro aminoácidos, en tres aminoácidos, en dos aminoácidos y en un aminoácido del polipéptido de la SEQ ID NO: 3.
- 45 [0269] La presente invención también se refiere a composiciones que comprenden variantes que comprenden una sustitución, delección y/o inserción de uno o más (o varios) aminoácidos de la SEQ ID NO: 3 o una secuencia homóloga de los mismos. El número total de posiciones que tienen sustituciones, delecciones y/o inserciones de aminoácidos en la SEQ ID NO: 3 no es más de 100, por ejemplo, 1,2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 40, 50, 60, 70, 80, 90 o 100. Preferiblemente, los cambios de aminoácidos son de naturaleza menor, es decir, sustituciones, inserciones o eliminaciones de aminoácidos conservadoras que no afectan significativamente al plegamiento y/o a la actividad de la proteína; delecciones pequeñas, típicamente de uno a aproximadamente 30 aminoácidos; extensiones pequeñas en los extremos amino o carboxilo terminales, tales como un residuo de metionina en el extremo amino terminal; un péptido enlazador pequeño de hasta aproximadamente 20-25 residuos; o una extensión pequeña que facilita la purificación al cambiar la carga neta u otra función, tal como un tracto de polihistidina, un epítipo antigénico o un dominio de unión.
- 60 [0270] Las formas de realización indicadas anteriormente con respecto a la SEQ ID NO: 3 se aplican correspondientemente a composiciones que comprenden o consisten en polipéptidos como se proporciona en la
- 65

SEQ ID NO: 2.

5 [0271] Las formas de realización indicadas anteriormente con respecto a la SEQ ID NO: 3 se aplican de manera correspondiente a las composiciones que comprenden o consisten en polipéptidos como se proporciona en la SEQ ID NO: 4.

10 [0272] Las formas de realización indicadas anteriormente con respecto a la SEQ ID NO: 3 se aplican de manera correspondiente a las composiciones que comprenden o consisten en polipéptidos como se proporciona en la SEQ ID NO: 5.

[0273] Las formas de realización indicadas anteriormente con respecto a la SEQ ID NO: 3 se aplican de manera correspondiente a las composiciones que comprenden o consisten en polipéptidos como se proporciona en la SEQ ID NO: 6.

15 [0274] Las composiciones pueden comprender  $\alpha$ -L-galactosidasa de la presente invención como el principal componente enzimático, por ejemplo, una composición monocomponente. Alternativamente, las composiciones pueden comprender múltiples actividades enzimáticas, tales como una o más (por ejemplo, varias) enzimas seleccionadas del grupo que consiste en hidrolasa, isomerasa, ligasa, liasa, oxidorreductasa o transferasa, por ejemplo, una alfa-galactosidasa, alfa-glucosidasa, aminopeptidasa, amilasa, beta-galactosidasa, beta-glucosidasa, beta-xilosidasa, carbohidrasa, carboxipeptidasa, catalasa, celobiidrolasa, celulasa, quitinasa, cutinasa, ciclodextrina glicosiltransferasa, desoxirribonucleasa, endoglucanasa, esterasa, glucoamilasa, invertasa, lacasa, lipasa, manosidasa, mutanasa, oxidasa, enzima pectinolítica, peroxidasa, fitasa, polifenoloxidasa, enzima proteolítica, ribonucleasa, transglutaminasa o xilanasa.

25 [0275] La(s) enzima(s) adicional(es) pueden ser producidas, por ejemplo, por un microorganismo tal como bacterias u hongos o por plantas o por animales. Las composiciones se pueden preparar de acuerdo con métodos conocidos en la técnica y pueden estar en forma de una composición líquida o seca. Por ejemplo, la composición puede estar en forma de un granulado o un microgranulado. La  $\alpha$ -L-galactosidasa se puede estabilizar de acuerdo con métodos conocidos en la técnica.

30 [0276] Las composiciones se pueden preparar de acuerdo con métodos conocidos en la técnica y pueden estar en forma de una composición líquida o seca. Las composiciones se pueden estabilizar de acuerdo con métodos conocidos en la técnica.

35 [0277] Las composiciones pueden comprender un polipéptido de la presente invención como el principal componente enzimático, por ejemplo, una composición monocomponente. Alternativamente, las composiciones pueden comprender múltiples actividades enzimáticas, tales como una o más (por ejemplo, varias) enzimas seleccionadas del grupo que consiste en hidrolasa, isomerasa, ligasa, liasa, oxidorreductasa o transferasa, por ejemplo, una alfa-galactosidasa, alfa-glucosidasa, aminopeptidasa, amilasa, beta-galactosidasa, beta-glucosidasa, beta-xilosidasa, carbohidrasa, carboxipeptidasa, catalasa, celobihidrolasa, celulasa, quitinasa, cutinasa, ciclodextrina glicosiltransferasa, desoxirribonucleasa, endoglucanasa, esterasa, glucoamilasa, invertasa, lacasa, lipasa, manosidasa, mutanasa, oxidasa, enzima pectinolítica, peroxidasa, fitasa, polifenoloxidasa, enzima proteolítica, ribonucleasa, transglutaminasa o xilanasa.

45 [0278] Las composiciones se pueden preparar de acuerdo con métodos conocidos en la técnica y pueden estar en forma de una composición líquida o seca. Las composiciones se pueden estabilizar de acuerdo con métodos conocidos en la técnica.

50 [0279] A continuación se dan ejemplos de usos preferidos de las composiciones de la presente invención. La dosificación de la composición y otras condiciones bajo las cuales se usa la composición se pueden determinar en función de métodos conocidos en la técnica.

Uso de  $\alpha$ -L-galactosidasas de la invención en pienso para animales

55 [0280] El término animal incluye todos los animales. Ejemplos de animales son los no rumiantes y los rumiantes. Los animales rumiantes incluyen, por ejemplo, animales tales como ovejas, cabras y ganado bovino, por ejemplo ganado para carne, vacas y terneros jóvenes. En una forma de realización particular, el animal es un animal no rumiante. Los animales no rumiantes incluyen animales monogástricos, por ejemplo cerdos o ganado porcino (incluyendo, entre otros, lechones, cerdos en crecimiento y cerdas); aves de corral tales como pavos, patos y pollo (incluidos, entre otros, pollos de engorde, gallinas ponedoras); caballos (incluidos, entre otros, de sangre caliente, sangre fría y sangre templada), terneros jóvenes; y peces (incluidos, entre otros, salmón, trucha, tilapia, bagre y carpas); y crustáceos (incluidos, entre otros, gambas y camarones).

65 [0281] El término pienso o composición de pienso significa cualquier compuesto, preparación, mezcla o composición adecuada para un animal o destinada a ser ingerida por un animal. En el uso según la invención, la  $\alpha$ -L-galactosidasa puede suministrarse al animal antes, después o simultáneamente con la dieta. Se prefiere este

último modo.

[0282] En una forma de realización particular, la  $\alpha$ -L-galactosidasa, en la forma en la que se agrega al pienso, o cuando se incluye en un aditivo de pienso, está bien definida. Bien definida significa que la preparación de  $\alpha$ -xilosidasa es al menos 50% pura según se determina mediante cromatografía de exclusión por tamaño (véase el Ejemplo 12 de WO 01/58275) En otras formas de realización particulares, la preparación de  $\alpha$ -L-galactosidasa tiene al menos 60, 70, 80, 85, 88, 90, 92, 94, o al menos un 95% de pureza según se determina por este método.

[0283] Una preparación de  $\alpha$ -L-galactosidasa bien definida es ventajosa. Por ejemplo, es mucho más fácil dosificar correctamente en el pienso una  $\alpha$ -xilosidasa que está esencialmente libre de interferir o contaminar otras proteasas u otras proteínas en general. El término dosificar correctamente se refiere en particular al objetivo de obtener resultados estables y constantes, y a la capacidad de optimizar la dosificación en función del efecto deseado.

[0284] Para el uso en pienso para animales, sin embargo, la  $\alpha$ -L-galactosidasa no necesita ser tan pura; puede, por ejemplo, incluir otras enzimas, en cuyo caso se podría denominar una preparación de  $\alpha$ -L-galactosidasa.

[0285] La preparación de  $\alpha$ -L-galactosidasa puede (a) agregarse directamente al pienso (o usarse directamente en un proceso de tratamiento con proteínas), o (b) puede usarse en la producción de una o más composiciones intermedias tales como aditivos para piensos o premezclas que posteriormente se agregan al pienso (o se usan en un proceso de tratamiento). El grado de pureza descrito anteriormente se refiere a la pureza de la preparación de  $\alpha$ -L-galactosidasa original, se use de acuerdo con (a) o con (b) anteriores.

[0286] Las preparaciones de  $\alpha$ -L-galactosidasa con purezas de este orden de magnitud se pueden obtener en particular usando métodos de producción recombinantes, mientras que no se obtienen tan fácilmente y también están sujetas a una variación mucho mayor entre lotes cuando la  $\alpha$ -L-galactosidasa se produce mediante métodos de fermentación tradicionales. Tal preparación de  $\alpha$ -L-galactosidasa puede, por supuesto, mezclarse con otras enzimas para obtener una preparación con dos o más enzimas purificadas con actividades diferentes o similares.

[0287] La proteína del sustrato puede ser una proteína animal, tal como harina de carne y hueso, harina de plumas y/o harina de pescado; o puede ser una proteína vegetal.

[0288] La fuente del sustrato es preferentemente de origen vegetal. El término vegetal o proteína vegetal como se usa en el presente documento se refiere a cualquier compuesto, composición, preparación o mezcla que incluya al menos una proteína derivada de un vegetal o que se origine a partir de un vegetal, incluyendo proteínas modificadas y derivados de proteínas. En formas de realización particulares, el contenido en proteínas de las proteínas vegetales es al menos 10, 20, 30, 40, 50 o 60% (p/p).

[0289] Las proteínas vegetales pueden derivarse de fuentes vegetales que contienen componentes de la pared celular de la galactosa, como las legumbres y los cereales (Bach Knudsen, K. E. (1997). Carbohydrate and lignin contents of plant materials used in animal feeding. *Animal Feed Science and Technology* 67: 319-338), por ejemplo materiales de plantas de las familias Fabaceae (Leguminosae), Cruciferae, Chenopodiaceae y Poaceae, tales como harina de soja, harina de altramuz y harina de semilla de colza. En una forma de realización particular, la fuente de proteína vegetal es material procedente de una o más plantas de la familia Fabaceae, por ejemplo soja, altramuz, guisante o alubia. En otra forma de realización particular, la fuente de proteína vegetal es material procedente de una o más plantas de la familia Chenopodiaceae, por ejemplo remolacha, remolacha azucarera, espinaca o quinoa. Otros ejemplos de fuentes de proteínas vegetales son colza, semillas de girasol, semillas de algodón y repollo. La soja es una fuente de proteína vegetal preferida. Otros ejemplos de fuentes de proteínas vegetales son los cereales como la cebada, el trigo, el centeno, la avena, el maíz, el arroz, el triticale, el sorgo, los granos secos de destilería con solubles (DDGS) y las microalgas.

[0290] En una forma de realización particular de un proceso de tratamiento, la(s)  $\alpha$ -L-galactosidasa(s) en cuestión tiene efecto en (o actúan sobre, o ejercen su influencia hidrolizante o degradante sobre) las proteínas, tales como proteínas vegetales o fuentes de proteínas. Para lograr esto, la proteína o fuente de proteína se suspende típicamente en un disolvente, por ejemplo un disolvente acuoso tal como agua, y los valores de pH y temperatura se ajustan teniendo en cuenta las características de la enzima en cuestión. Por ejemplo, el tratamiento puede tener lugar a un valor de pH en el cual la actividad de la  $\alpha$ -L-galactosidasa real es de al menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70 %, 80% o al menos 90%. Asimismo, por ejemplo, el tratamiento puede tener lugar a una temperatura a la cual la actividad de la  $\alpha$ -L-galactosidasa real es de al menos 5%, 10%, 20%, 30%, 40%, 50%, 60%, 70 %, 80% o al menos 90%. Las indicaciones de actividad porcentual anteriores son relativas a las actividades máximas. La reacción enzimática se continúa hasta que se alcanza el resultado deseado, después de lo cual puede o no detenerse inactivando la enzima, por ejemplo mediante un paso de tratamiento térmico.

[0291] En otra forma de realización particular de un proceso de tratamiento de la invención, se añade  $\alpha$ -L-

galactosidasa a la fuente de proteína vegetal, pero su influencia hidrolizante no se activa hasta más tarde cuando se desee, una vez que se establecen las condiciones de hidrolización adecuadas, o una vez que se inactivan los inhibidores de enzimas, o cualquier otro medio que se pueda haber aplicado para posponer la acción de la enzima.

5 [0292] En una forma de realización, el tratamiento es un pretratamiento de pienso para animales o proteínas para uso en pienso para animales, es decir, las fuentes de proteínas vegetales se hidrolizan antes de la ingesta.

10 [0293] El término mejora del valor nutricional de un pienso para animales significa la mejora de la disponibilidad de nutrientes en el pienso. En esta invención, la mejora de los valores nutricionales se refiere en particular a la mejora de la disponibilidad de la fracción proteica del pienso, lo que conduce a una mayor extracción de proteínas, a mayores rendimientos de proteínas y/o a una mejor utilización de las proteínas. Cuando se aumenta el valor nutricional del pienso, la digestibilidad de la proteína y/o aminoácidos aumenta y la velocidad de crecimiento y/o el aumento de peso y/o la transformación del pienso (es decir, el peso del pienso ingerido en relación con el aumento de peso) del animal pueden mejorar.

15 [0294] La  $\alpha$ -L-galactosidasa puede agregarse al pienso en cualquier forma, ya sea como una  $\alpha$ -xilosidasa relativamente pura, o en mezcla con otros componentes destinados a la adición a pienso para animales, es decir, en forma de aditivos para piensos, tales como las llamadas premezclas para pienso para animales.

20 [0295] En un aspecto adicional, la presente invención se refiere a composiciones para uso en pienso para animales, tales como pienso para animales y aditivos para piensos, por ejemplo premezclas.

25 [0296] Además de la  $\alpha$ -L-galactosidasa de la invención, los aditivos para piensos de la invención contienen al menos una vitamina liposoluble y/o al menos una vitamina soluble en agua y/o al menos un oligoelemento, y/o al menos un macromineral.

30 [0297] Además, otros ingredientes aditivos para el pienso opcionales son agentes colorantes, por ejemplo carotenoides tales como betacaroteno, astaxantina y luteína; estabilizantes; aditivos que mejoran el crecimiento y compuestos aromáticos/aromatizantes, por ejemplo creosol, anetol, deca-, undeca- y/o dodeca-lactonas, iononas, irona, gingerol, piperidina, propilideno-ftálico, butilideno-ftálico, capsaicina y/o tanino; péptidos antimicrobianos; ácidos grasos poliinsaturados (PUFA); especies generadoras de oxígeno reactivo; también se puede usar un soporte que puede contener, por ejemplo, 40-50% en peso de fibras de madera, 8-10% en peso de estearina, 4-5% en peso de polvo de cúrcuma, 4-58% en peso de polvo de romero, 22-28% en peso de piedra caliza, 1-3% en peso de una goma, tal como goma arábiga, 5-50% en peso de azúcar y/o almidón y 5-15% en peso de agua.

35 [0298] Un pienso o un aditivo para piensos de la invención también puede comprender al menos otra enzima seleccionada de entre fitasa (EC 3.1.3.8 o 3.1.3.26); xilanasas (EC 3.2.1.8); galactanasas (EC 3.2.1.89); alfa-galactosidasa (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).

40 [0299] En una forma de realización particular, el pienso o un aditivo para piensos de la invención también comprende una proteasa (EC 3.4).

[0300] En una forma de realización particular, el pienso o un aditivo para piensos de la invención también comprende una fitasa (EC 3.1.3.8 o 3.1.3.26).

45 [0301] En una forma de realización particular, el pienso o un aditivo para piensos de la invención también comprende una xilanasas (EC 3.2.1.8).

50 [0302] Un pienso o un aditivo para piensos de la invención también puede comprender al menos un microbio probiótico o de alimentación directa (DFM) opcionalmente junto con una o más enzimas diferentes seleccionadas de entre fitasa (EC 3.1.3.8 o 3.1.3.26); xilanasas (EC 3.2.1.8); galactanasas (EC 3.2.1.89); alfa-galactosidasa (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).

55 [0303] El DFM se puede agregar al pienso para animales de manera que la dosis diaria de DFM esté entre 1x10<sup>5</sup> UFC y 1x10<sup>13</sup> UFC, preferiblemente entre 1x10<sup>6</sup> UFC y 1x10<sup>12</sup> UFC, más preferiblemente entre 1x10<sup>7</sup> UFC y 1x10<sup>11</sup> UFC e incluso más preferiblemente entre 5x10<sup>7</sup> UFC y 1x10<sup>10</sup> UFC. Alternativamente, el DFM puede agregarse al pienso para animales de manera que la concentración de DFM en el pienso esté entre 1x10<sup>3</sup> UFC/g de pienso y 1x10<sup>8</sup> UFC/g de pienso, preferiblemente entre 5x10<sup>3</sup> UFC/g de pienso y 1x10<sup>7</sup> UFC/g de pienso, más preferiblemente entre 1x10<sup>4</sup> UFC/g de pienso y 5x10<sup>6</sup> UFC/g de pienso e incluso más preferiblemente entre 2.5x10<sup>4</sup> UFC/g de pienso y 1x10<sup>6</sup> UFC/g de pienso.

5 [0304] El microbiano alimentado directamente puede ser una bacteria de uno o más de los siguientes géneros: Lactobacillus, Lactococcus, Streptococcus, Bacillus, Pediococcus, Enterococcus, Leuconostoc, Carnobacterium, Propionibacterium, Bifidobacterium, Clostridium y Megasphaera o cualquier combinación de los mismos, preferiblemente de Bacillus subtilis, Bacillus licheniformis, Bacillus amyloliquefaciens, Enterococcus faecium, Enterococcus spp, y Pediococcus spp, Lactobacillus spp, Bifidobacterium spp, Lactobacillus acidophilus, 10 Pediococcus acidilactici, Lactococcus lactis, Bifidobacterium bifidum, Propionibacterium thoenii, Lactobacillus farciminus, Lactobacillus rhamnosus, Clostridium butyricum, Bifidobacterium animalis ssp. animalis, Lactobacillus reuteri, Bacillus cereus, Lactobacillus salivarius ssp. salivarius, Megasphaera elsdenii, Propionibacteria sp y más preferiblemente de las cepas de Bacillus subtilis 3A-P4 (PTA-6506); 15A-P4 (PTA-6507); 22C-P1 (PTA-6508); 2084 (NRRL B-500130); LSSA01 (NRRL-B-50104); BS27 (NRRL B-501 05); BS 18 (NRRL B-50633); y BS 278 (NRRL B-50634).

15 [0305] En una forma de realización particular, estas otras enzimas están bien definidas (como se ha definido anteriormente). En una forma de realización particular, el pienso o un aditivo para piensos de la invención también comprende una xilanasa (EC 3.2.1.8).

20 [0306] Algunos ejemplos de péptidos antimicrobianos (AMP) son CAP18, Leucocina A, Tritripina, Protegrina-1, Thanatina, Defensina, Lactoferrina, Lactoferrina y Ovispirina tales como Novispirina (Robert Lehrer, 2000), Plectasinas y Estatinas, que incluyen los compuestos y polipéptidos divulgados en WO 03/044049 y WO 03/048148, así como variantes o fragmentos de los anteriores que conservan la actividad antimicrobiana.

25 [0307] Algunos ejemplos de polipéptidos antifúngicos (AFP) son los péptidos Aspergillus giganteus y Aspergillus niger, así como variantes y fragmentos de los mismos que retienen actividad antifúngica, como se describe en WO 94/01459 y WO 02/090384.

[0308] Algunos ejemplos de ácidos grasos poliinsaturados son los ácidos grasos poliinsaturados C18, C20 y C22, tales como ácido araquidónico, ácido docosohexaenoico, ácido eicosapentaenoico y ácido gamma-linoleico.

30 [0309] Algunos ejemplos de especies generadoras de oxígeno reactivo son productos químicos tales como perborato, persulfato o percarbonato; y enzimas tales como una oxidasa, una oxigenasa o una sintetasa.

35 [0310] Por lo general, las vitaminas liposolubles y solubles en agua, así como los oligoelementos, forman parte de la denominada premezcla destinada a ser añadida al pienso, mientras que los macrominerales generalmente se añaden por separado al pienso. Cualquiera de estos tipos de composición, cuando está enriquecido con una  $\alpha$ -xylosidasa de la invención, es un aditivo para piensos de la invención.

40 [0311] En una forma de realización particular, el aditivo para piensos de la invención está destinado a ser incluido (o prescrito como que debe incluirse) en dietas para animales o pienso a niveles de 0,01 a 10,0%; más particularmente de 0,05 a 5,0%; o 0,2 a 1,0% (% significa g de aditivo por 100 g de pienso). Esto es así en particular para las premezclas.

[0312] Las siguientes son listas no exclusivas de ejemplos de estos componentes:

45 [0313] Ejemplos de vitaminas liposolubles son la vitamina A, vitamina D3, vitamina E y vitamina K, por ejemplo la vitamina K3.

50 [0314] Los ejemplos de vitaminas hidrosolubles son la vitamina B12, biotina y colina, vitamina B1, vitamina B2, vitamina B6, niacina, ácido fólico y pantotenato, por ejemplo Ca-D-pantotenato.

[0315] Ejemplos de oligoelementos son manganeso, zinc, hierro, cobre, yodo, selenio y cobalto.

[0316] Ejemplos de macrominerales son calcio, fósforo y sodio.

55 [0317] Los requisitos nutricionales de estos componentes (ejemplificados con aves de corral y lechones/cerdos) se enumeran en la Tabla A de WO 01/58275. Requisito nutricional significa que estos componentes se deben proporcionar en la dieta en las concentraciones indicadas.

60 [0318] Alternativamente, el aditivo para piensos de la invención comprende al menos uno de los componentes individuales especificados en la Tabla A de WO 01/58275. Al menos uno significa uno, uno o más de uno, dos, tres o cuatro, y así sucesivamente hasta los trece, o hasta los quince componentes individuales. Más específicamente, este al menos un componente individual está incluido en el aditivo de la invención en una cantidad tal que proporciona una concentración de pienso dentro del intervalo indicado en la columna cuatro, o columna cinco, o columna seis de la Tabla A.

65 [0319] En otra forma de realización más, el aditivo para piensos para animales de la invención comprende al

menos una de las siguientes vitaminas, preferiblemente para proporcionar una concentración en pienso dentro de los intervalos especificados en la Tabla 1 a continuación (para dietas de lechones y dietas de pollos de engorde, respectivamente).

5

Tabla 1: recomendaciones típicas de vitaminas

Vitamina	Dieta de lechón	Dieta de pollo de engorde
Vitamina A	10, 000-15,000 UI/kg de pienso	8-12,500 UI/kg de pienso
Vitamina D3	1800-2000 UI/kg de pienso	3000-5000 UI/kg de pienso
Vitamina E	60-100 mg/kg de pienso	150-240 mg/kg de pienso
Vitamina K3	2-4 mg/kg de pienso	2-4 mg/kg de pienso
Vitamina B1	2-4 mg/kg de pienso	2-3 mg/kg de pienso
Vitamina B2	6-10 mg/kg de pienso	7-9 mg/kg de pienso
Vitamina B6	4-8 mg/kg de pienso	3-6 mg/kg de pienso
Vitamina B12	0,03-0,05 mg/kg de pienso	0,015-0,04 mg/kg de pienso
Niacina (Vitamina B3)	30-50 mg/kg de pienso	50-80 mg/kg de pienso
Ácido pantoténico	20-40 mg/kg de pienso	10-18 mg/kg de pienso
Ácido fólico	1-2 mg/kg de pienso	1-2 mg/kg de pienso
Biotina	0,15-0,4 mg/kg de pienso	0,15-0,3 mg/kg de pienso
Cloruro de colina	200-400 mg/kg de pienso	300-600 mg/kg de pienso

10 [0320] La presente invención también se refiere a composiciones de pienso para animales. Las composiciones o las dietas de pienso para animales tienen un contenido relativamente alto de proteína. Las dietas para aves y cerdos se pueden caracterizar como se indica en la Tabla B de WO 01/58275, columnas 2-3. Las dietas para peces se pueden caracterizar como se indica en la columna 4 de esta Tabla B. Además, tales dietas de peces generalmente tienen un contenido de grasa bruta de 200-310 g/kg. WO 01/58275 corresponde a US 09/779334.

15 [0321] Una composición de pienso para animales de acuerdo con la invención tiene un contenido de proteína bruta de 50-800 g/kg, y además comprende al menos una  $\alpha$ -L-galactosidasa como se reivindica en este documento.

20 [0322] Además, o en alternativa (al contenido de proteína bruta indicado anteriormente), la composición de pienso para animales de la invención tiene un contenido de energía metabolizable de 10-30 MJ/kg; y/o un contenido de calcio de 0,1-200 g/kg; y/o un contenido de fósforo disponible de 0,1-200 g/kg; y/o un contenido de metionina de 0,1-100 g/kg; y/o un contenido de metionina más cisteína de 0,1-150 g/kg; y/o un contenido de lisina de 0,5-50 g/kg.

25 [0323] En formas de realización particulares, el contenido de energía metabolizable, proteína bruta, calcio, fósforo, metionina, metionina más cisteína, y/o lisina está dentro de cualquiera de los intervalos 2, 3, 4 o 5 en la Tabla B de WO 01/58275 (R. 2-5).

30 [0324] La proteína bruta se calcula como nitrógeno (N) multiplicado por un factor de 6,25, es decir proteína bruta (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 Analytical Chemists, Washington DC).

35 [0325] La energía metabolizable se puede calcular sobre la base de la publicación del NRC 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, y European Table of Energy Values for Poultry Feed-stuffs, Spelderholt centre for poultry research and extension, 7361 DA Beekbergen, The Netherlands. Grafisch bedrijf Ponsen & Iooijen bv, Wageningen. ISBN 90-71463-12-5.

40 [0326] El contenido nutricional de calcio, fósforo disponible y aminoácidos en dietas completas para animales se calcula sobre la base de tablas de alimentación 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.

45 [0327] En una forma de realización particular, la composición de pienso para animales de la invención contiene al menos una proteína vegetal como se ha definido anteriormente.

50 [0328] La composición de pienso para animales de la invención también puede contener proteína animal, tal como carne y harina de huesos, harina de plumas, y/o harina de pescado, típicamente en una cantidad de 0-25%. La composición de pienso para animales de la invención también puede comprender granos de destilería secos con solubles (DDGS), típicamente en cantidades de 0-30%.



[0329] En otras formas de realización particulares, la composición de pienso 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% harina de soja; y/o harina de pescado al 0-25%; y/o 0-25% de harina de carne y hueso; y/o 0-20% de suero de leche.

5

[0330] Las dietas para animales se pueden fabricar por ejemplo como pienso en harina (no granulado) o pienso granulado. Típicamente, los piensos molidos se mezclan y se añaden cantidades suficientes de vitaminas y minerales esenciales de acuerdo con las especificaciones para la especie en cuestión. Las enzimas se pueden agregar como formulaciones enzimáticas sólidas o líquidas. Por ejemplo, para un pienso en harina se puede

agregar una formulación enzimática sólida o líquida antes o durante el paso de mezcla de ingredientes. Para el pienso granulado, también se puede agregar la preparación de  $\alpha$ -xilosidasa/enzima (líquida o sólida) antes o durante el paso del ingrediente del pienso. Típicamente, después del paso de granulación se agrega una preparación líquida de  $\alpha$ -xilosidasa/enzima. La enzima también se puede incorporar en un aditivo o premezcla para pienso.

10

15

[0331] La concentración final de enzima en la dieta está dentro del rango de 0,01-200 mg de proteína enzimática por kg de dieta, por ejemplo en el rango de 0,5-25 mg de proteína enzimática por kg de dieta animal.

[0332] La  $\alpha$ -L-galactosidasa debería aplicarse, por supuesto, en una cantidad efectiva, es decir, en una cantidad adecuada para mejorar la hidrólisis de las proteínas, la digestibilidad de proteínas y aminoácidos, y/o para mejorar el valor nutricional del pienso. En la actualidad se contempla que la enzima se administre en una o más de las siguientes cantidades (rangos de dosificación): 0,01 - 200; 0,01-100; 0,5-100; 1-50; 5-100; 10-100; 0,05-50; o 0,10-10 - todos estos rangos están en mg de proteína  $\alpha$ -xilosidasa por kg de pienso (ppm).

20

25

[0333] Para determinar los mg de proteína  $\alpha$ -L-galactosidasa por kg de pienso, la  $\alpha$ -L-galactosidasa se purifica a partir de la composición de pienso, y la actividad específica de la  $\alpha$ -L-galactosidasa purificada se determina usando un ensayo relevante (véase actividad de  $\alpha$ -L-galactosidasa, sustratos y ensayos). La actividad de  $\alpha$ -L-galactosidasa de la composición de pienso como tal también se determina usando el mismo ensayo, y en función de estas dos determinaciones se calcula la dosificación en mg de proteína  $\alpha$ -L-galactosidasa por kg de pienso.

30

[0334] Los mismos principios se aplican para determinar los mg de proteína  $\alpha$ -L-galactosidasa en los aditivos para pienso. Por supuesto, si se dispone de una muestra de la  $\alpha$ -L-galactosidasa utilizada para preparar el aditivo de pienso o el pienso, la actividad específica se determina a partir de esta muestra (no es necesario purificar la  $\alpha$ -L-galactosidasa de la composición de pienso o el aditivo).

35

Péptido señal

[0335] Aquí se describe un polinucleótido aislado que codifica un péptido señal que comprende o consiste en los aminoácidos 1 a 19 de la SEQ ID NO: 2. Los polinucleótidos pueden comprender además un gen que codifica una proteína, que está operativamente unida al péptido señal. La proteína es preferiblemente foránea al péptido señal. En un aspecto, el polinucleótido que codifica el péptido señal son los nucleótidos 1 a 57 de la SEQ ID NO: 1.

40

Construcciones de ácidos nucleicos, vectores de expresión, células huésped recombinantes y métodos para la producción de  $\alpha$ -L-galactosidasa

45

[0336] Se describen construcciones de ácidos nucleicos, vectores de expresión y células huésped recombinantes que comprenden tales polinucleótidos. También se describen métodos para producir una proteína, que comprenden (a) cultivar una célula huésped recombinante que comprende dicho polinucleótido; y (b) recuperar la proteína.

50

[0337] La proteína puede ser nativa o heteróloga respecto de una célula huésped. El término "proteína" no se refiere aquí a una longitud específica del producto codificado y, por lo tanto, abarca péptidos, oligopéptidos y polipéptidos. El término "proteína" también abarca dos o más polipéptidos combinados para formar el producto codificado. Las proteínas también incluyen polipéptidos híbridos y polipéptidos fusionados.

55

[0338] Preferiblemente, la proteína es una hormona, enzima, receptor o parte del mismo, anticuerpo o parte del mismo, o reportero. Por ejemplo, la proteína puede ser una hidrolasa, isomerasa, ligasa, liasa, oxidoreductasa o transferasa, por ejemplo una alfa-galactosidasa, alfa-glucosidasa, aminopeptidasa, amilasa, beta-galactosidasa, beta-glucosidasa, beta-xilosidasa, carbohidrasa, carboxipeptidasa, catalasa, celobiohidrolasa, celulasa, quitinasa, cutinasa, ciclodextrina glicosiltransferasa, desoxirribonucleasa, endoglucanasa, esterasa, glucoamilasa, invertasa, lacasa, lipasa, manosidasa, mutanasa, oxidasa, enzima pectinolítica, peroxidasa, fitasa, polifenoloxidasa, enzima proteolítica, ribonucleasa, transglutaminasa o xilanasas.

60

[0339] Preferiblemente, la proteína es una  $\alpha$ -L-galactosidasa.

65

[0340] El gen puede obtenerse de cualquier fuente procariota, eucariota u otra fuente.

[0341] La presente invención se describe adicionalmente mediante los siguientes ejemplos que no deben interpretarse como limitativos del alcance de la invención.

5

## Ejemplos

### Cepas

10 [0342] Se ha hecho referencia al depósito disponible públicamente de la cepa microbiana *Bacteroides ovatus* (ATCC 8483) depositado en la ATCC en 1933 por Eggerth (Eggerth AH, Gagnon BH. The bacteroides of human feces. J. Bacteriol. 25:389-413, 1933).

### Ejemplo 1:

15

Construcción de plásmidos y producción de proteínas

[0343] Para producir construcciones de proteínas, secuencias de ADN que codifican GH31 (BACOVA 03422; depósito A7LZZ5) y GH95 (BACOVA\_03438; depósito A7M011) de *Bacteroides ovatus* se amplificaron por PCR de ADN genómico de *B. ovatus* ATCC8483 usando cebadores adecuados. El fragmento de ADN amplificado que codifica GH95 se clonó en pET21a (Novagen) restringido con NdeI y XhoI.

20

[0344] Se produjeron proteínas recombinantes solubles inoculando *E. coli* recombinante (BL21; Novagen) en medio LB (1 L) en matraces de 2 L suplementados con ampicilina (50 µg/mL) e incubando los cultivos a 37 ° C con agitación (200 rpm) hasta que la OD600 alcanzó 0,4. Se añadió isopropil β-D-tiogalactopiranosido a la concentración final de 1 mM para inducir la producción de proteínas, y el cultivo se incubó durante 5 h adicionales a 37°C. Las células se recogieron por centrifugación (ángulo fijo) a 5000 g/10 min/4°C, se resuspendieron en 10 ml de Tris-HCl 20 mM, pH 8,0, que contenía 300 mM de NaCl (tampón A) por litro de cultivo y se rompieron por sonicación. Tras la clarificación del lisado celular mediante centrifugación a 15000 g/30 min/4°C, el polipéptido se purificó mediante cromatografía de afinidad de metal inmovilizado usando resina Talon (Clontech). Las proteínas se eluyeron con tampón A que contenía 150 mM de imidazol. GH95 de *B. ovatus* se dializó contra tampón Na<sub>2</sub>PO<sub>4</sub> 50 mM, pH 7,5.

25

30

[0345] La secuencia de ADN clonada de BACOVA\_03438 GH95 se proporciona como la SEQ ID NO: 1 y la secuencia de aminoácidos deducida como SEQ ID NO: 2. La secuencia de aminoácidos madura se proporciona como SEQ ID NO: 3

35

### Ejemplo 2:

40 Caracterización bioquímica de BACOVA\_03438 GH95

Fuentes de carbohidratos utilizadas:

45 [0346] Los arabinoxilanos y xilooligosacáridos de trigo y centeno fueron de Megazyme (Megazyme International Ireland). Los xilanos de abedul y cascarilla de avena y todos los monosacáridos fueron de Sigma. El xilano de salvado de maíz fue una donación de los doctores Kevin Hicks y Madhav Yadav (USDA).

45

Cromatografía en Capa Fina (TLC):

50 [0347] La actividad de la enzima GH95 frente a una gama de xilanos decorados se evaluó inicialmente mediante TLC. Se incubaron xilanos (1% p/v final) de centeno, trigo, abedul, avena y salvado de maíz con 0,5 µM de cada enzima. Las reacciones se llevaron a cabo en K<sub>2</sub>PO<sub>4</sub> 50 mM, ácido cítrico 120 mM, pH 6,5. Los polisacáridos se incubaron con enzimas durante 5 horas a 37 ° C. A continuación, se cargaron muestras (de 6 µl cada una) a intervalos de 1 cm en la placa de TLC y se secaron usando un secador de pelo después de cada carga de 2 ul.

55 La placa de TLC se colocó luego en un tanque de vidrio de 1 l que contenía tampón de migración (butanol/ácido acético/agua a 2:2:1), hasta una profundidad de 0,5 cm y se selló con una placa de vidrio. Cuando el tampón de funcionamiento alcanzó ~ 1 cm desde la parte superior de la placa, la placa se extrajo, se secó con un secador de pelo y se sumergió durante unos segundos en un reactivo de ácido sulfúrico de orcinol (ácido sulfúrico/etanol/agua 3:70:20 v/v , orcinol 1%), y se secó a 120 ° C hasta que se revelaron los azúcares (~5-10min).

55

60

Cromatografía líquida de alta resolución (HPLC):

65 [0348] La HPLC se usó para cuantificar los productos de la hidrólisis de xilano de maíz mediante las enzimas GH31 y GH95, usando una columna analítica de intercambio aniónico CARBOPAC™ PA-100 (Dionex) equipada con una columna de protección CARBOPAC™ PA-100. El sistema totalmente automático tenía un tamaño de

bucle de 100 µl, caudal de 1,0 ml/min, presión de -2000 psi y azúcares detectados por detección amperométrica pulsada (PAD). La configuración de la PAD fue E1 = +0,05, E2 = +0,6 y E3 = - 0,6. Las condiciones de elución usadas fueron 0-5 minutos de 100 mM de NaOH, 5-15 minutos 100 mM de NaOH con gradiente de acetato sódico 0-75 mM y 15-25 minutos de 100 mM de NaOH que contenía 75 mM de acetato sódico. Antes y después de cada paso, la columna se lavó con 500 mM de acetato sódico durante 10 minutos y luego con 500 mM de hidróxido sódico durante 5 minutos y luego se equilibró con 100 mM de hidróxido sódico durante 10 minutos.

[0349] Las reacciones enzimáticas se realizaron en 50 mM de tampón de PC, pH 6,5 que contenía 1 mg/ml de BSA. El volumen de reacción final se calculó para permitir que se analizaran 4-8 alícuotas de cada reacción, que se inició mediante la adición de la enzima adecuadamente diluida (1/100 del volumen final). Las reacciones se terminaron llevando a ebullición alícuotas durante 10 min. Los datos fueron recogidos y analizados utilizando el software XChrom V.2.04 (LabSystems) mediante un VG Chromatography Server (Fisons Instruments).

### Ejemplo 3:

Actividad de GH95 de *B. ovatus*, BACOVA\_03438.

[0350] El análisis de TLC reveló que BACOVA\_03438 GH95 no era activa contra xilanos decorados de trigo, centeno, avena espelta o madera de abedul (datos no mostrados), sino que la enzima liberaba un solo producto de xilano de salvado de maíz que migraba conjuntamente con L-galactosa (Figura 1).

[0351] Se usó HPLC para investigar más a fondo la identidad del producto de GH95 de *B. ovatus* liberado a partir de xilano de salvado de maíz. Los datos revelaron un solo producto liberado por GH95 de *B. ovatus* a partir de xilano de salvado de maíz (Figura 2). Todas las enzimas caracterizadas dentro de GH95 son α-L-fucosidasas, pero el producto de GH95 no se coeluye con L-fucosa por HPLC (Figura 10). La L-fucosa es un azúcar desoxi hexosa que carece de un grupo hidroxilo en el carbono en la posición 6 (C-6) y es equivalente a 6-desoxi-L-galactosa. El producto de la actividad de GH95 en xilano de salvado de maíz se coeluye con L-galactosa en la HPLC, indicando que la enzima es una α-L-galactosidasa, una nueva actividad enzimática.

[0352] Luego comparamos la actividad de la enzima GH95 de *B. ovatus* (BACOVA\_03438) con una α-L-fucosidasa GH95 previamente caracterizada de *Bifidobacterium bifidum* (AAQ72464). Los datos (Figuras 3 y 4) mostraron que aunque la enzima de *B. bifidum* era capaz de liberar galactosa a partir de xilano de salvado de maíz, el índice de GH95 de *B. ovatus* en la misma cantidad de sustrato polimérico fue al menos 200 veces mayor. A partir de estos datos está claro que la actividad de la α-L-galactosidasa no es una propiedad general de las enzimas GH95.

### Ejemplo 5:

Desalmidonado de maíz y extracción de goma de fibra de maíz (CFG)

Maíz desalmidonado

[0353] Se mezclaron 107 kg de maíz pre-molido (<1,0 mm) con 253 kg de agua a 53°C. La mezcla se calentó a 95°C y el pH se ajustó a 6,2 con NaOH 1M. Se añadieron 1,12 kg, 120 L de Termamyl y la reacción se mantuvo a alrededor de 90 ° C durante 3 horas. Después de 3 horas, se añadió agua fría hasta un peso de reacción total de 600 kg. La separación del sólido y el líquido se realizó con un decantador Westfalia, CA-225-110, 4950 rpm, flujo 600 L/hora. La fracción de fibra sólida se volvió a suspender y se separó dos veces con agua hasta un peso total de 600 kg seguido de separación como se ha descrito anteriormente. La fracción de fibra final se dividió en porciones más pequeñas y se liofilizó durante 3-4 días.

Extracción de goma de fibra de maíz

[0354] Se añadieron 20 g de maíz desalmidonado a 200 ml de agua de MilliQ™ en ebullición junto con 0,8 g de NaOH y 0,8 g de Ca(OH)<sub>2</sub>. La mezcla se mantuvo durante 1 hora a 96°C y luego se centrifugó durante 20 minutos a 6000 g.

[0355] El sobrenadante tenía aspecto lechoso y graso y, por lo tanto, las grasas se extrajeron mediante agitación con hexano (1 parte de hexano por 4 partes de goma de fibra de maíz). Después de un tiempo de reposo, el hexano se eliminó mediante pipeteo.

[0356] Se realizó una segunda extracción (CFG2) disolviendo el sedimento en 200 ml de NaOH 1M. La mezcla se mantuvo durante 1 hora a 96°C y luego se centrifugó durante 20 min a 6000 g y el sobrenadante se ajustó a pH 6 con 4M de HCl.

[0357] Ambas fracciones de CFG se concentraron usando un rotavapor y se precipitaron en EtOH a una concentración final del 90%. El sedimento del NaOH 0,1 M se denominará a partir de ahora CFG1. Debido al alto

pH en el extracto de NaOH 1 M (CFG2), esta fracción se dializó en un cilindro de medición de 2 L con agua desionizada, con el grifo goteando durante la noche usando una membrana de diálisis de 3 kDa.

5 [0358] Antes de los ensayos de aplicación, tanto CFG1 como CFG2 se centrifugaron a 25000 g durante 25 min y se filtraron a través de un filtro de jeringa de 0,44 µm. La materia seca en las fracciones de CFG2 se determinó con un Mettler Toledo HR73.

Hidrólisis enzimática de la goma de fibra de maíz

10 [0359] La hidrólisis enzimática se realizó usando la fracción de CFG2 (véase arriba).

[0360] Los ensayos de 400 µL se llevaron a cabo a 40 ° C en un termomezclador Eppendorph Thermomixer con agitación, 1200 rpm durante 17 horas y las reacciones se detuvieron mediante calor a 97 ° C durante 10 min. La configuración del ensayo se puede encontrar en la Tabla 1. Se realizaron siete reacciones diferentes en paralelo y las enzimas en las respectivas muestras se pueden encontrar en la Tabla 2.

Tabla 1.

	µL
CFG2 2,3% MS	350
0,25 M Na citrato pH6	20
Enzima <sup>a</sup>	10
MilliQ	a 400
<sup>a</sup> 10 µL por enzima. Dosificación enzimática: GH31, 0.7 g/L; GH95, 0,7 g/l; UltraFlo L 8% (v/v)	

Tabla 2.

Muestra	Enzimas en muestras
1	GH31
2	GH95
3	GH31 + GH95
4	UltraFlo L + GH31
5	UltraFlo L + GH95
6	UltraFlo L + GH31 + GH95
7	UltraFlo L

20 Análisis por HPLC de los monosacáridos producidos

[0361] Los sobrenadantes se analizaron en un Dionex ICS-5000 mediante cromatografía de intercambio aniónico de alto rendimiento con detección amperométrica pulsada (HPAEC-PAD) usando una columna analítica y de protección CarboPac-20. El sistema estaba controlado por Chromeleon v. 6.8. Los monosacáridos liberados se cuantificaron frente a una curva estándar de 6 puntos de arabinosa, galactosa, glucosa y xilosa desde 0,0002-0,02 g/l. Se inyectaron 25 µL de cada muestra usando el programa de la Tabla 3. Los eluyentes se desgasificaron por burbujeo de helio durante 10 minutos y tenían la siguiente composición: A, agua de MilliQ; B, NaOH 0,5 M; C, NaOAc 0,5 M ; D, NaOAc 60 mM. Las muestras 1-3 se diluyeron a 1:50 y las muestras 4-7 se diluyeron a 1: 200, ambas en agua, antes de la inyección.

Tabla 3. Programa de eluyente PA-20 con un flujo de 0,5 ml/min.

min	Eluyente (%)				Curva
	A	B	C	D	
0-8,5	80	0	0	20	5
8,5-25	80-50	0-30	20	0	4
25-27	0-100	0	100-0	0	9
27-40	80	0	0	20	5

35 [0362] GH31 libera xilosa de goma de fibra de maíz y GH95 libera L-galactosa. Tanto GH31 como GH95 parecen contribuir a la hidrólisis lograda por UltraFlo L. Además, la actividad de GH95 falta claramente en UltraFlo L ya que no se libera galactosa en ausencia de GH95 (Figura 1).

[0363] Los datos de las Figuras 5 y 6 demuestran claramente que tanto GH31 como GH95 podrían usarse para ayudar en la degradación del xilano de maíz así como para liberar monosacáridos específicos de xilano de maíz.

**Ejemplo 6:**

Materiales y métodos



Blanco	1	0
GH95	6	0
GH11, 43, 51	9	23
GH11, 43, 51, 95	32	34

Material biológico

5 [0369] En este documento se ha hecho referencia al depósito de libre acceso de la cepa microbiana *Bacteroides ovatus* (ATCC 8483) depositado en la ATCC en 1933 por Eggerth (Eggerth AH, Gagnon BH. The bacteroides of human feces. J. Bacteriol. 25:389-413, 1933).

10 [0370] La invención descrita y reivindicada en este documento no debe verse limitada por los aspectos específicos aquí descritos, ya que estos aspectos pretenden ser ilustraciones de varios aspectos de la invención. De hecho, varias modificaciones de la invención además de las mostradas y descritas en la presente serán evidentes para los expertos en la técnica a partir de la descripción anterior. En el caso de conflicto, la presente divulgación, incluidas las definiciones, prevalecerá.

15 LISTADO DE SECUENCIAS

[0371]

<110> Novozymes A/S

20 <120> POLIPÉPTIDOS QUE TIENEN ACTIVIDAD DE ALFA-L-GALACTOSIDASA Y POLINUCLEÓTIDOS QUE LOS CODIFICAN

<130> 12570-WO-PCT

25 <150> EP13178257.5

<151> 2013-07-26

<160> 6

30 <170> versión de PatentIn 3.5

<210> 1

<211> 2436

<212> ADN

35 <213> *Bacteroides ovatus* ATCC 8483

<220>

<221> CDS

40 <222> (1)..(2433)

<220>

<221> sig\_peptide

<222> (1)..(57)

45

<220>

<221> mat\_peptide

<222> (58)..(2433)

50 <400> 1

atg aac aga aac act ttt ttt gta tta ttc cta tta ata agc aac ctt  
48

Met Asn Arg Asn Thr Phe Phe Val Leu Phe Leu Leu Ile Ser Asn Leu  
-15 -10 -5

55



ES 2 649 217 T3

Val Asn Val Gln Asn Asp Gln Leu Thr Val Thr Cys Gln Gly Lys Glu  
175 180 185

5 cag gaa ggt ttg aaa gcc gca cta cgg gca gaa tgt caa atc cag gta  
672  
Gln Glu Gly Leu Lys Ala Ala Leu Arg Ala Glu Cys Gln Ile Gln Val  
190 195 200 205

10 aaa acg aat ggc act ctc cgc ccg gca gga aat acg ctc caa ata aat  
720  
Lys Thr Asn Gly Thr Leu Arg Pro Ala Gly Asn Thr Leu Gln Ile Asn  
210 215 220

15 gaa gga aca gaa gcc acc ctt tat ata tcg gca gca acc aac tac gta  
768  
Glu Gly Thr Glu Ala Thr Leu Tyr Ile Ser Ala Ala Thr Asn Tyr Val  
225 230 235

20 aac tac cag gat gtc agc gcc gat gaa tcc cac cga acc agc gaa tat  
816  
Asn Tyr Gln Asp Val Ser Ala Asp Glu Ser His Arg Thr Ser Glu Tyr  
240 245 250

25 ctg aaa aga gca atg caa ata cct tac gaa aaa gct ctc aaa aac cat  
864  
Leu Lys Arg Ala Met Gln Ile Pro Tyr Glu Lys Ala Leu Lys Asn His  
255 260 265

30 atc gct tat tat aaa aaa caa ttt gac cgt gta cgc ctt aca tta ccg  
912  
Ile Ala Tyr Tyr Lys Lys Gln Phe Asp Arg Val Arg Leu Thr Leu Pro  
270 275 280 285

35 gca ggc aaa gct tcc caa ctg gag act ccc aaa cga att gaa aac ttc  
960  
Ala Gly Lys Ala Ser Gln Leu Glu Thr Pro Lys Arg Ile Glu Asn Phe  
290 295 300

40 gga aat ggg gaa gac atg gca atg gca gct cta ctt ttc cat tat gga  
1008  
Gly Asn Gly Glu Asp Met Ala Met Ala Ala Leu Leu Phe His Tyr Gly  
305 310 315

45 cgt tat ctg ctg att tct tct tcg caa ccg ggt ggg caa ccg gct aat  
1056  
Arg Tyr Leu Leu Ile Ser Ser Ser Gln Pro Gly Gly Gln Pro Ala Asn  
320 325 330

50 ctg caa ggt ata tgg aat aac agt act cat gca ccc tgg gat agc aaa  
1104  
Leu Gln Gly Ile Trp Asn Asn Ser Thr His Ala Pro Trp Asp Ser Lys  
335 340 345

55 tac acc atc aat atc aat aca gaa atg aac tat tgg ccg gca gaa gtc  
1152  
Tyr Thr Ile Asn Ile Asn Thr Glu Met Asn Tyr Trp Pro Ala Glu Val  
350 355 360 365



ES 2 649 217 T3

act aat tta agt gaa aca cat agt ccg tta ttc tct atg ttg aaa gat  
 1200  
 5 Thr Asn Leu Ser Glu Thr His Ser Pro Leu Phe Ser Met Leu Lys Asp  
 370 375 380

tta tcc gtc aca gga gca gaa aca gcc cga acc atg tat gat tgt cgg  
 1248  
 10 Leu Ser Val Thr Gly Ala Glu Thr Ala Arg Thr Met Tyr Asp Cys Arg  
 385 390 395

gga tgg gtg gca cat cat aac act gac tta tgg aga atc tgt ggt gta  
 1296  
 15 Gly Trp Val Ala His His Asn Thr Asp Leu Trp Arg Ile Cys Gly Val  
 400 405 410

gtc gat ttt gca gca gcc ggc atg tgg cct agc ggc ggc gca tgg tta  
 1344  
 20 Val Asp Phe Ala Ala Ala Gly Met Trp Pro Ser Gly Gly Ala Trp Leu  
 415 420 425

gct caa cat atc tgg caa cat tac ctc ttt aca gga aat aaa gag ttc  
 1392  
 25 Ala Gln His Ile Trp Gln His Tyr Leu Phe Thr Gly Asn Lys Glu Phe  
 430 435 440 445

ctg aaa gag tat tat ccc atc ctg aaa gga aca gcg cag ttc tac atg  
 1440  
 30 Leu Lys Glu Tyr Tyr Pro Ile Leu Lys Gly Thr Ala Gln Phe Tyr Met  
 450 455 460

gat ttc ttg gta gaa cat cca gtg tat aaa tgg ttg gta gtg tcg cct  
 1488  
 35 Asp Phe Leu Val Glu His Pro Val Tyr Lys Trp Leu Val Val Ser Pro  
 465 470 475

tct gtg tcg ccg gaa cac ggt ccc atc acg gcc ggt tgc aca atg gac  
 1536  
 40 Ser Val Ser Pro Glu His Gly Pro Ile Thr Ala Gly Cys Thr Met Asp  
 480 485 490

aat cag att gcc ttt gat gca ctg cat aat acc tta ttg gct tca tac  
 1584  
 45 Asn Gln Ile Ala Phe Asp Ala Leu His Asn Thr Leu Leu Ala Ser Tyr  
 495 500 505

atc gca ggt gaa gcc ccc tct ttc caa gac tcc ctc aaa caa aca cta  
 1632  
 50 Ile Ala Gly Glu Ala Pro Ser Phe Gln Asp Ser Leu Lys Gln Thr Leu  
 510 515 520 525

gag aag ttg cca cct atg caa ata gga aag cat aac caa cta caa gaa  
 1680  
 55 Glu Lys Leu Pro Pro Met Gln Ile Gly Lys His Asn Gln Leu Gln Glu  
 530 535 540

ES 2 649 217 T3

tgg ctt gaa gat att gac aat ccc aag gac gag cac cgt cat att tca  
1728  
Trp Leu Glu Asp Ile Asp Asn Pro Lys Asp Glu His Arg His Ile Ser  
545 550 555

5  
cat ttg tac ggg cta tat ccg agt aat caa att tct cca tac tca aat  
1776  
His Leu Tyr Gly Leu Tyr Pro Ser Asn Gln Ile Ser Pro Tyr Ser Asn  
560 565 570

10  
ccc gag tta ttc cag gca gca aga aac aca ttg ctt caa cgc ggt gat  
1824  
Pro Glu Leu Phe Gln Ala Ala Arg Asn Thr Leu Leu Gln Arg Gly Asp  
575 580 585

15  
aag gct acc ggt tgg agt atc ggc tgg aaa gtc aat ttc tgg gca cgt  
1872  
Lys Ala Thr Gly Trp Ser Ile Gly Trp Lys Val Asn Phe Trp Ala Arg  
590 595 600 605

20  
atg ctg gat gga aac cat gcc ttc caa atc att aaa aat atg att caa  
1920  
Met Leu Asp Gly Asn His Ala Phe Gln Ile Ile Lys Asn Met Ile Gln  
610 615 620

25  
cta ttg ccc aac gat cac ttg gct aaa gaa tat ccg aac ggg cgt acc  
1968  
Leu Leu Pro Asn Asp His Leu Ala Lys Glu Tyr Pro Asn Gly Arg Thr  
625 630 635

30  
tat ccc aat atg ctg gat gcc cat ccg cct ttc cag att gac ggt aat  
2016  
Tyr Pro Asn Met Leu Asp Ala His Pro Pro Phe Gln Ile Asp Gly Asn  
640 645 650

35  
ttc ggt tat aca gcc gga gta gcg gaa atg ttg ctc caa agt cat gac  
2064  
Phe Gly Tyr Thr Ala Gly Val Ala Glu Met Leu Leu Gln Ser His Asp  
655 660 665

40  
ggc gcc gta cat tta ttg ccg gca ttg cct gat gca tgg gaa gaa ggg  
2112  
Gly Ala Val His Leu Leu Pro Ala Leu Pro Asp Ala Trp Glu Glu Gly  
670 675 680 685

45  
agt gtg aag ggg ctg gtg gca cgt gga aac ttt act gtc gat atg gac  
2160  
Ser Val Lys Gly Leu Val Ala Arg Gly Asn Phe Thr Val Asp Met Asp  
690 695 700

50  
tgg aaa aac aat gtg tta aat aaa gcg atc atc cga tcg aat ata ggc  
2208  
Trp Lys Asn Asn Val Leu Asn Lys Ala Ile Ile Arg Ser Asn Ile Gly  
705 710 715

55  
agt acg ctc cgc atc cgg tct tat gtc cca cta aaa ggg aaa ggc tta  
2256

ES 2 649 217 T3

Ser Thr Leu Arg Ile Arg Ser Tyr Val Pro Leu Lys Gly Lys Gly Leu  
720 725 730

5 aaa caa gta aac ggg aaa gag tgc tca aat aga tta ttc gcc acc act  
2304  
Lys Gln Val Asn Gly Lys Glu Cys Ser Asn Arg Leu Phe Ala Thr Thr  
735 740 745

10 cct atc aag caa cct ctt gta gcc aaa gga gta agt gcg cag tct ccc  
2352  
Pro Ile Lys Gln Pro Leu Val Ala Lys Gly Val Ser Ala Gln Ser Pro  
750 755 760 765

15 aaa ctg cag aag gtt tac gaa tat gac atc gaa aca aag gcg ggt aaa  
2400  
Lys Leu Gln Lys Val Tyr Glu Tyr Asp Ile Glu Thr Lys Ala Gly Lys  
770 775 780

20 acc tat att gta aat aca atc gaa gga aaa caa taa  
2436  
Thr Tyr Ile Val Asn Thr Ile Glu Gly Lys Gln  
785 790

25 <210> 2  
<211> 811  
<212> PRT  
<213> Bacteroides ovatus ATCC 8483

30 <400> 2

Met Asn Arg Asn Thr Phe Phe Val Leu Phe Leu Leu Ile Ser Asn Leu  
-15 -10 -5

35 Val Ser Gly Gln Asp Leu Lys Leu Trp Tyr Ser Gln Pro Ala Gln Asn  
-1 1 5 10

40 Trp Ser Glu Ala Leu Pro Ile Gly Asn Ser Arg Leu Gly Ala Met Val  
15 20 25

45 Tyr Gly Gly Ile Glu Arg Glu Glu Leu Gln Leu Asn Glu Glu Thr Phe  
30 35 40 45

50 Trp Ala Gly Ser Pro Tyr Asn Asn Asn Asn Pro Asn Ala Val His Val  
50 55 60

55 Leu Pro Val Val Arg Lys Leu Ile Phe Glu Gly Arg Asn Lys Glu Ala  
65 70 75

Gln Arg Leu Ile Asp Ala Asn Phe Leu Thr Gln Gln His Gly Met Ser  
80 85 90

ES 2 649 217 T3

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

ES 2 649 217 T3

Arg Tyr Leu Leu Ile Ser Ser Ser Gln Pro Gly Gly Gln Pro Ala Asn  
 320 325 330

5

Leu Gln Gly Ile Trp Asn Asn Ser Thr His Ala Pro Trp Asp Ser Lys  
 335 340 345

10

Tyr Thr Ile Asn Ile Asn Thr Glu Met Asn Tyr Trp Pro Ala Glu Val  
 350 355 360 365

15

Thr Asn Leu Ser Glu Thr His Ser Pro Leu Phe Ser Met Leu Lys Asp  
 370 375 380

20

Leu Ser Val Thr Gly Ala Glu Thr Ala Arg Thr Met Tyr Asp Cys Arg  
 385 390 395

25

Gly Trp Val Ala His His Asn Thr Asp Leu Trp Arg Ile Cys Gly Val  
 400 405 410

30

Val Asp Phe Ala Ala Ala Gly Met Trp Pro Ser Gly Gly Ala Trp Leu  
 415 420 425

35

Ala Gln His Ile Trp Gln His Tyr Leu Phe Thr Gly Asn Lys Glu Phe  
 430 435 440 445

40

Leu Lys Glu Tyr Tyr Pro Ile Leu Lys Gly Thr Ala Gln Phe Tyr Met  
 450 455 460

45

Asp Phe Leu Val Glu His Pro Val Tyr Lys Trp Leu Val Val Ser Pro  
 465 470 475

50

Ser Val Ser Pro Glu His Gly Pro Ile Thr Ala Gly Cys Thr Met Asp  
 480 485 490

55

Asn Gln Ile Ala Phe Asp Ala Leu His Asn Thr Leu Leu Ala Ser Tyr  
 495 500 505

Ile Ala Gly Glu Ala Pro Ser Phe Gln Asp Ser Leu Lys Gln Thr Leu  
 510 515 520 525

Glu Lys Leu Pro Pro Met Gln Ile Gly Lys His Asn Gln Leu Gln Glu  
 530 535 540

ES 2 649 217 T3

Trp Leu Glu Asp Ile Asp Asn Pro Lys Asp Glu His Arg His Ile Ser  
545 550 555

5 His Leu Tyr Gly Leu Tyr Pro Ser Asn Gln Ile Ser Pro Tyr Ser Asn  
560 565 570

10 Pro Glu Leu Phe Gln Ala Ala Arg Asn Thr Leu Leu Gln Arg Gly Asp  
575 580 585

15 Lys Ala Thr Gly Trp Ser Ile Gly Trp Lys Val Asn Phe Trp Ala Arg  
590 595 600 605

Met Leu Asp Gly Asn His Ala Phe Gln Ile Ile Lys Asn Met Ile Gln  
610 615 620

20 Leu Leu Pro Asn Asp His Leu Ala Lys Glu Tyr Pro Asn Gly Arg Thr  
625 630 635

25 Tyr Pro Asn Met Leu Asp Ala His Pro Pro Phe Gln Ile Asp Gly Asn  
640 645 650

30 Phe Gly Tyr Thr Ala Gly Val Ala Glu Met Leu Leu Gln Ser His Asp  
655 660 665

35 Gly Ala Val His Leu Leu Pro Ala Leu Pro Asp Ala Trp Glu Glu Gly  
670 675 680 685

Ser Val Lys Gly Leu Val Ala Arg Gly Asn Phe Thr Val Asp Met Asp  
690 695 700

40 Trp Lys Asn Asn Val Leu Asn Lys Ala Ile Ile Arg Ser Asn Ile Gly  
705 710 715

45 Ser Thr Leu Arg Ile Arg Ser Tyr Val Pro Leu Lys Gly Lys Gly Leu  
720 725 730

50 Lys Gln Val Asn Gly Lys Glu Cys Ser Asn Arg Leu Phe Ala Thr Thr  
735 740 745

55 Pro Ile Lys Gln Pro Leu Val Ala Lys Gly Val Ser Ala Gln Ser Pro  
750 755 760 765

Lys Leu Gln Lys Val Tyr Glu Tyr Asp Ile Glu Thr Lys Ala Gly Lys



ES 2 649 217 T3

Thr Arg Tyr Gln Val Asp Asp Val Thr Tyr Thr Arg Thr Thr Phe Ala  
 130 135 140  
 5  
 Ser Phe Thr Asp Asn Val Ile Ile Met His Ile Lys Ala Ser Lys Ala  
 145 150 155  
 10  
 Asn Ala Leu Asn Phe Thr Ile Ala Tyr Asn Cys Pro Leu Val His Lys  
 160 165 170  
 15  
 Val Asn Val Gln Asn Asp Gln Leu Thr Val Thr Cys Gln Gly Lys Glu  
 175 180 185  
 20  
 Gln Glu Gly Leu Lys Ala Ala Leu Arg Ala Glu Cys Gln Ile Gln Val  
 190 195 200 205  
 25  
 Lys Thr Asn Gly Thr Leu Arg Pro Ala Gly Asn Thr Leu Gln Ile Asn  
 210 215 220  
 30  
 Glu Gly Thr Glu Ala Thr Leu Tyr Ile Ser Ala Ala Thr Asn Tyr Val  
 225 230 235  
 35  
 Asn Tyr Gln Asp Val Ser Ala Asp Glu Ser His Arg Thr Ser Glu Tyr  
 240 245 250  
 40  
 Leu Lys Arg Ala Met Gln Ile Pro Tyr Glu Lys Ala Leu Lys Asn His  
 255 260 265  
 45  
 Ile Ala Tyr Tyr Lys Lys Gln Phe Asp Arg Val Arg Leu Thr Leu Pro  
 270 275 280 285  
 50  
 Ala Gly Lys Ala Ser Gln Leu Glu Thr Pro Lys Arg Ile Glu Asn Phe  
 290 295 300  
 55  
 Gly Asn Gly Glu Asp Met Ala Met Ala Ala Leu Leu Phe His Tyr Gly  
 305 310 315  
 Arg Tyr Leu Leu Ile Ser Ser Ser Gln Pro Gly Gly Gln Pro Ala Asn  
 320 325 330  
 60  
 Leu Gln Gly Ile Trp Asn Asn Ser Thr His Ala Pro Trp Asp Ser Lys  
 335 340 345



ES 2 649 217 T3

Tyr Thr Ile Asn Ile Asn Thr Glu Met Asn Tyr Trp Pro Ala Glu Val  
 350 355 360 365  
 5  
 Thr Asn Leu Ser Glu Thr His Ser Pro Leu Phe Ser Met Leu Lys Asp  
 370 375 380  
 10  
 Leu Ser Val Thr Gly Ala Glu Thr Ala Arg Thr Met Tyr Asp Cys Arg  
 385 390 395  
 15  
 Gly Trp Val Ala His His Asn Thr Asp Leu Trp Arg Ile Cys Gly Val  
 400 405 410  
 20  
 Val Asp Phe Ala Ala Ala Gly Met Trp Pro Ser Gly Gly Ala Trp Leu  
 415 420 425  
 25  
 Ala Gln His Ile Trp Gln His Tyr Leu Phe Thr Gly Asn Lys Glu Phe  
 430 435 440 445  
 30  
 Leu Lys Glu Tyr Tyr Pro Ile Leu Lys Gly Thr Ala Gln Phe Tyr Met  
 450 455 460  
 35  
 Asp Phe Leu Val Glu His Pro Val Tyr Lys Trp Leu Val Val Ser Pro  
 465 470 475  
 40  
 Ser Val Ser Pro Glu His Gly Pro Ile Thr Ala Gly Cys Thr Met Asp  
 480 485 490  
 45  
 Asn Gln Ile Ala Phe Asp Ala Leu His Asn Thr Leu Leu Ala Ser Tyr  
 495 500 505  
 50  
 Ile Ala Gly Glu Ala Pro Ser Phe Gln Asp Ser Leu Lys Gln Thr Leu  
 510 515 520 525  
 55  
 Glu Lys Leu Pro Pro Met Gln Ile Gly Lys His Asn Gln Leu Gln Glu  
 530 535 540  
 60  
 Trp Leu Glu Asp Ile Asp Asn Pro Lys Asp Glu His Arg His Ile Ser  
 545 550 555  
 65  
 His Leu Tyr Gly Leu Tyr Pro Ser Asn Gln Ile Ser Pro Tyr Ser Asn  
 560 565 570

ES 2 649 217 T3

Pro Glu Leu Phe Gln Ala Ala Arg Asn Thr Leu Leu Gln Arg Gly Asp  
 575 580 585

5 Lys Ala Thr Gly Trp Ser Ile Gly Trp Lys Val Asn Phe Trp Ala Arg  
 590 595 600 605

10 Met Leu Asp Gly Asn His Ala Phe Gln Ile Ile Lys Asn Met Ile Gln  
 610 615 620

15 Leu Leu Pro Asn Asp His Leu Ala Lys Glu Tyr Pro Asn Gly Arg Thr  
 625 630 635

20 Tyr Pro Asn Met Leu Asp Ala His Pro Pro Phe Gln Ile Asp Gly Asn  
 640 645 650

Phe Gly Tyr Thr Ala Gly Val Ala Glu Met Leu Leu Gln Ser His Asp  
 655 660 665

25 Gly Ala Val His Leu Leu Pro Ala Leu Pro Asp Ala Trp Glu Glu Gly  
 670 675 680 685

30 Ser Val Lys Gly Leu Val Ala Arg Gly Asn Phe Thr Val Asp Met Asp  
 690 695 700

35 Trp Lys Asn Asn Val Leu Asn Lys Ala Ile Ile Arg Ser Asn Ile Gly  
 705 710 715

Ser Thr Leu Arg Ile Arg Ser Tyr Val Pro Leu Lys Gly Lys Gly Leu  
 720 725 730

40 Lys Gln Val Asn Gly Lys Glu Cys Ser Asn Arg Leu Phe Ala Thr Thr  
 735 740 745

45 Pro Ile Lys Gln Pro Leu Val Ala Lys Gly Val Ser Ala Gln Ser Pro  
 750 755 760 765

50 Lys Leu Gln Lys Val Tyr Glu Tyr Asp Ile Glu Thr Lys Ala Gly Lys  
 770 775 780

55 Thr Tyr Ile Val Asn Thr Ile Glu Gly Lys Gln  
 785 790

<210> 4

ES 2 649 217 T3

<211> 811  
 <212> PRT  
 <213> Bacteroides ovatus CL02T12C04

5

<220>  
 <221> SEÑAL  
 <222> (-19)..(-1)

10

<220>  
 <221> mat\_peptide  
 <222> (1)..(792)

<400> 4

15

Met Asn Arg Asn Thr Phe Phe Val Leu Phe Leu Leu Ile Ser Asn Leu  
 -15 -10 -5

20

Val Ser Gly Gln Asp Leu Lys Leu Trp Tyr Ser Gln Pro Ala Gln Asn  
 -1 1 5 10

25

Trp Ser Glu Ala Leu Pro Ile Gly Asn Ser Arg Leu Gly Ala Met Val  
 15 20 25

30

Tyr Gly Gly Ile Glu Arg Glu Glu Leu Gln Leu Asn Glu Glu Thr Phe  
 30 35 40 45

Trp Ala Gly Ser Pro Tyr Asn Asn Asn Asn Pro Asn Ala Val His Val  
 50 55 60

35

Leu Pro Val Val Arg Lys Leu Ile Phe Glu Gly Arg Asn Lys Glu Ala  
 65 70 75

40

Gln Arg Leu Ile Asp Ala Asn Phe Leu Thr Gln Gln His Gly Met Ser  
 80 85 90

45

Tyr Leu Thr Leu Gly Ser Leu Tyr Leu Glu Phe Pro Glu His Gln Asn  
 95 100 105

50

Gly Ser Gly Phe Tyr Arg Glu Leu Asn Leu Glu Asn Ala Thr Thr Thr  
 110 115 120 125

Thr Arg Tyr Gln Val Asp Asp Val Thr Tyr Thr Arg Thr Thr Phe Ala  
 130 135 140

55

Ser Phe Thr Asp Asn Val Ile Ile Met His Ile Lys Ala Ser Lys Ala  
 145 150 155

ES 2 649 217 T3

5 Asn Ala Leu Asn Phe Thr Ile Ala Tyr Asn Cys Pro Leu Val His Lys  
 160 165 170  
 Val Asn Val Gln Asn Asp Gln Leu Thr Val Thr Cys Gln Gly Lys Glu  
 175 180 185  
 10 Gln Glu Gly Leu Lys Ala Ala Leu Arg Ala Glu Cys Gln Ile Gln Val  
 190 195 200 205  
 15 Lys Thr Asn Gly Thr Leu Arg Pro Ala Gly Asn Thr Leu Gln Ile Asn  
 210 215 220  
 20 Glu Gly Thr Glu Ala Thr Leu Tyr Ile Ser Ala Ala Thr Asn Tyr Val  
 225 230 235  
 25 Asn Tyr Gln Asp Val Ser Ala Asp Glu Ser His Arg Thr Ser Glu Tyr  
 240 245 250  
 30 Leu Lys Arg Ala Met Gln Ile Pro Tyr Glu Lys Ala Leu Lys Asn His  
 255 260 265  
 Ile Ala Tyr Tyr Lys Lys Gln Phe Asp Arg Val Arg Leu Thr Leu Pro  
 270 275 280 285  
 35 Ala Gly Lys Ala Ser Gln Leu Glu Thr Pro Lys Arg Ile Glu Asn Phe  
 290 295 300  
 40 Gly Asn Gly Glu Asp Met Ala Met Ala Ala Leu Leu Phe His Tyr Gly  
 305 310 315  
 45 Arg Tyr Leu Leu Ile Ser Ser Ser Gln Pro Gly Gly Gln Pro Ala Asn  
 320 325 330  
 Leu Gln Gly Ile Trp Asn Asn Ser Thr His Ala Pro Trp Asp Ser Lys  
 335 340 345  
 50 Tyr Thr Ile Asn Ile Asn Thr Glu Met Asn Tyr Trp Pro Ala Glu Val  
 350 355 360 365  
 55 Thr Asn Leu Ser Glu Thr His Ser Pro Leu Phe Ser Met Leu Lys Asp  
 370 375 380

ES 2 649 217 T3

Leu Ser Val Thr Gly Ala Glu Thr Ala Arg Thr Met Tyr Asp Cys Arg  
 385 390 395

5

Gly Trp Val Ala His His Asn Thr Asp Leu Trp Arg Ile Cys Gly Val  
 400 405 410

10

Val Asp Phe Ala Ala Ala Gly Met Trp Pro Ser Gly Gly Ala Trp Leu  
 415 420 425

15

Ala Gln His Ile Trp Gln His Tyr Leu Phe Thr Gly Asn Lys Glu Phe  
 430 435 440 445

20

Leu Lys Glu Tyr Tyr Pro Ile Leu Lys Gly Thr Ala Gln Phe Tyr Met  
 450 455 460

25

Asp Phe Leu Val Glu His Pro Val Tyr Lys Trp Leu Val Val Ser Pro  
 465 470 475

30

Ser Val Ser Pro Glu His Gly Pro Ile Thr Ala Gly Cys Thr Met Asp  
 480 485 490

35

Asn Gln Ile Ala Phe Asp Ala Leu His Asn Thr Leu Leu Ala Ser Tyr  
 495 500 505

40

Ile Ala Gly Glu Ala Pro Ser Phe Gln Asp Ser Leu Lys Gln Thr Leu  
 510 515 520 525

45

Glu Lys Leu Pro Pro Met Gln Ile Gly Lys His Asn Gln Leu Gln Glu  
 530 535 540

50

Trp Leu Glu Asp Ile Asp Asn Pro Lys Asp Glu His Arg His Ile Ser  
 545 550 555

55

His Leu Tyr Gly Leu Tyr Pro Ser Asn Gln Ile Ser Pro Tyr Ser Asn  
 560 565 570

60

Pro Glu Leu Phe Gln Ala Ala Arg Asn Thr Leu Leu Gln Arg Gly Asp  
 575 580 585

65

Lys Ala Thr Gly Trp Ser Ile Gly Trp Lys Val Asn Phe Trp Ala Arg  
 590 595 600 605

# ES 2 649 217 T3

```

Met Leu Asp Gly Asn His Ala Phe Gln Ile Ile Lys Asn Met Ile Gln
      610                               615                               620

5  Leu Leu Pro Asn Asp His Leu Ala Lys Glu Tyr Pro Asn Gly Arg Thr
      625                               630                               635

10 Tyr Pro Asn Met Leu Asp Ala His Pro Pro Phe Gln Ile Asp Gly Asn
      640                               645                               650

15 Phe Gly Tyr Thr Ala Gly Val Ala Glu Met Leu Leu Gln Ser His Asp
      655                               660                               665

20 Gly Ala Val His Leu Leu Pro Ala Leu Pro Asp Ala Trp Glu Glu Gly
      670                               675                               680

25 Ser Val Lys Gly Leu Val Ala Arg Gly Asn Phe Thr Val Asp Met Asp
      690                               695                               700

30 Trp Lys Asn Asn Val Leu Asn Lys Ala Ile Ile Arg Ser Asn Ile Gly
      705                               710                               715

35 Ser Thr Leu Arg Ile Arg Ser Tyr Val Pro Leu Lys Gly Lys Gly Leu
      720                               725                               730

40 Lys Gln Val Asn Gly Lys Glu Cys Ser Asn Arg Leu Phe Ala Thr Thr
      735                               740                               745

45 Pro Ile Lys Gln Pro Leu Val Ala Lys Gly Val Ser Ala Gln Ser Pro
      750                               755                               760

50 Lys Leu Gln Lys Val Tyr Glu Tyr Asp Ile Glu Thr Lys Ala Gly Lys
      770                               775                               780

55 Thr Tyr Ile Val Asn Thr Ile Glu Gly Lys Gln
      785                               790

<210> 5
<211> 811
<212> PRT
<213> Bacteroides xylanisolvens CL03T12C04

<220>
<221> SEÑAL
<222> (-1)..(-1)

```

ES 2 649 217 T3

<220>  
 <221> mat\_peptide  
 <222> (1)..(792)

5

<400> 5

Met Asn Arg Asn Thr Phe Phe Val Leu Phe Leu Leu Ile Ser Asn Leu  
 -15 -10 -5

10

Val Ser Gly Gln Asp Leu Lys Leu Trp Tyr Ser Gln Pro Ala Gln Asn  
 -1 1 5 10

15

Trp Ser Glu Ala Leu Pro Ile Gly Asn Ser Arg Leu Gly Ala Met Val  
 15 20 25

20

Tyr Gly Gly Ile Glu Arg Glu Glu Leu Gln Leu Asn Glu Glu Thr Phe  
 30 35 40 45

25

Trp Ala Gly Ser Pro Tyr Asn Asn Asn Asn Pro Asn Ala Val His Val  
 50 55 60

30

Leu Pro Ile Val Arg Lys Leu Ile Phe Glu Gly Arg Asn Lys Glu Ala  
 65 70 75

35

Gln Arg Leu Ile Asp Ala Asn Phe Leu Thr Gln Gln His Gly Met Ser  
 80 85 90

40

Tyr Leu Thr Leu Gly Ser Leu Tyr Leu Glu Phe Pro Glu His Gln Asn  
 95 100 105

45

Ala Ser Gly Phe Tyr Arg Asp Leu Asn Leu Glu Asn Ala Thr Thr Thr  
 110 115 120 125

Thr Arg Tyr Gln Val Asp Asp Val Thr Tyr Thr Arg Thr Thr Phe Ala  
 130 135 140

50

Ser Phe Thr Asp Asn Val Ile Ile Met His Ile Lys Ala Ser Lys Ala  
 145 150 155

55

Asn Ala Leu Asn Phe Thr Ile Ala Tyr Asn Cys Pro Leu Val His Lys  
 160 165 170

Val Asn Val Gln Asn Asp Gln Leu Thr Val Thr Cys Gln Gly Lys Glu  
 175 180 185

ES 2 649 217 T3

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



ES 2 649 217 T3

Val Asp Phe Ala Ala Ala Gly Met Trp Pro Ser Gly Gly Ala Trp Leu  
415 420 425

5  
Ala Gln His Ile Trp Gln His Tyr Leu Phe Thr Gly Asn Lys Glu Phe  
430 435 440 445

10  
Leu Lys Glu Tyr Tyr Pro Ile Leu Lys Gly Thr Ala Gln Phe Tyr Met  
450 455 460

15  
Asp Phe Leu Val Glu His Pro Val Tyr Lys Trp Leu Val Val Ser Pro  
465 470 475

20  
Ser Val Ser Pro Glu His Gly Pro Ile Thr Ala Gly Cys Thr Met Asp  
480 485 490

25  
Asn Gln Ile Ala Phe Asp Ala Leu His Asn Thr Leu Leu Ala Ser Tyr  
495 500 505

30  
Ile Ala Gly Glu Ala Pro Ser Phe Gln Asp Ser Leu Lys Gln Thr Leu  
510 515 520 525

35  
Glu Lys Leu Pro Pro Met Gln Ile Gly Lys His Asn Gln Leu Gln Glu  
530 535 540

40  
Trp Leu Glu Asp Ile Asp Asn Pro Lys Asp Glu His Arg His Ile Ser  
545 550 555

45  
His Leu Tyr Gly Leu Tyr Pro Ser Asn Gln Ile Ser Pro Tyr Ser Asn  
560 565 570

50  
Pro Glu Leu Phe Gln Ala Ala Arg Asn Thr Leu Leu Gln Arg Gly Asp  
575 580 585

55  
Lys Ala Thr Gly Trp Ser Ile Gly Trp Lys Val Asn Phe Trp Ala Arg  
590 595 600 605

Met Leu Asp Gly Asn His Ala Phe Gln Ile Ile Lys Asn Met Ile Gln  
610 615 620

Leu Leu Pro Asn Asp His Leu Ala Lys Glu Tyr Pro Asn Gly Arg Thr  
625 630 635

ES 2 649 217 T3

Tyr Pro Asn Met Leu Asp Ala His Pro Pro Phe Gln Ile Asp Gly Asn  
 640 645 650

5 Phe Gly Tyr Thr Ala Gly Val Ala Glu Met Leu Leu Gln Ser His Asp  
 655 660 665

10 Gly Ala Val His Leu Leu Pro Ala Leu Pro Asp Ala Trp Glu Glu Gly  
 670 675 680 685

15 Ser Val Lys Gly Leu Val Ala Arg Gly Asn Phe Thr Val Asp Met Asp  
 690 695 700

20 Trp Lys Asn Asn Val Leu Asn Lys Ala Ile Ile Arg Ser Asn Ile Gly  
 705 710 715

25 Gly Thr Leu Arg Ile Arg Ser Tyr Val Pro Leu Lys Gly Lys Gly Leu  
 720 725 730

30 Lys Gln Val Asn Gly Lys Glu Cys Ser Asn Arg Leu Phe Ala Thr Thr  
 735 740 745

35 Pro Ile Lys Arg Pro Leu Val Ala Lys Gly Val Ser Ala Gln Ser Pro  
 750 755 760 765

40 Lys Leu Gln Lys Val Tyr Glu Tyr Asp Ile Glu Thr Lys Ala Gly Lys  
 770 775 780

45 Thr Tyr Ile Val Asn Thr Ile Glu Gly Lys Gln  
 785 790

<210> 6  
 <211> 811  
 <212> PRT  
 <213> Bacteroides sp. 2\_1\_22

50 <220>  
 <221> SEÑAL  
 <222> (-1)..(-1)

55 <220>  
 <221> mat\_peptide  
 <222> (1)..(792)

<400> 6  
 Met Asn Arg Asn Thr Leu Phe Leu Leu Leu Leu Leu Thr Ser Asn Leu



ES 2 649 217 T3

5 Gly Gly Thr Glu Ala Thr Leu Tyr Ile Ser Ala Ala Thr Asn Tyr Val  
225 230 235

10 Asn Tyr Gln Asn Val Ser Ala Asp Glu Ser Arg Arg Thr Thr Asp Tyr  
240 245 250

15 Leu Glu Glu Ala Ile Leu Ile Pro Tyr Glu Lys Ala Leu Lys Glu His  
255 260 265

20 Ile Ala Phe Tyr Lys Lys Gln Phe Asp Arg Val Gln Leu His Leu Pro  
270 275 280 285

25 Ser Ser Glu Ala Ser Gln Ile Glu Thr Pro Arg Arg Ile Glu Asn Phe  
290 295 300

30 Gly Gln Gly Asn Asp Met Ala Met Ala Ala Leu Leu Phe Gln Tyr Gly  
305 310 315

35 Arg Tyr Leu Leu Ile Ser Ser Ser Gln Pro Gly Gly Gln Pro Ala Asn  
320 325 330

40 Leu Gln Gly Ile Trp Asn Asn Ser Thr His Ala Pro Trp Asp Ser Lys  
335 340 345

45 Tyr Thr Ile Asn Ile Asn Thr Glu Met Asn Tyr Trp Pro Ala Glu Val  
350 355 360 365

50 Thr Asn Leu Ser Glu Thr His Ser Pro Leu Phe Ser Met Leu Lys Asp  
370 375 380

55 Leu Ser Val Thr Gly Ala Glu Thr Ala Arg Thr Met Tyr Asp Cys Trp  
385 390 395

60 Gly Trp Val Ala His His Asn Thr Asp Leu Trp Arg Ile Cys Gly Val  
400 405 410

65 Val Asp Phe Ala Ala Ala Gly Met Trp Pro Ser Gly Gly Ala Trp Leu  
415 420 425

70 Ala Gln His Ile Trp Gln His Tyr Leu Phe Thr Gly Asn Lys Glu Phe  
430 435 440 445

ES 2 649 217 T3

Leu Lys Glu Tyr Tyr Pro Ile Leu Lys Gly Thr Ala Gln Phe Tyr Met  
 450 455 460  
 5  
 Asp Phe Leu Val Glu His Pro Thr Tyr Lys Trp Leu Val Val Ser Pro  
 465 470 475  
 10 Ser Val Ser Pro Glu His Gly Pro Ile Thr Ala Gly Cys Thr Met Asp  
 480 485 490  
 15 Asn Gln Ile Ala Phe Asp Ala Leu His Asn Thr Leu Leu Ala Ser Tyr  
 495 500 505  
 20 Ile Ala Gly Glu Ala Pro Ser Phe Gln Asp Ser Leu Lys Gln Thr Leu  
 510 515 520 525  
 25 Glu Lys Leu Pro Pro Met Gln Ile Gly Lys His Asn Gln Leu Gln Glu  
 530 535 540  
 30 Trp Leu Glu Asp Ile Asp Asn Pro Lys Asp Glu His Arg His Ile Ser  
 545 550 555  
 35 His Leu Tyr Gly Leu Tyr Pro Ser Asn Gln Ile Ser Pro Tyr Ser Asn  
 560 565 570  
 40 Pro Glu Leu Phe Gln Ala Ala Arg Asn Thr Leu Leu Gln Arg Gly Asp  
 575 580 585  
 45 Lys Ala Thr Gly Trp Ser Ile Gly Trp Lys Val Asn Phe Trp Ala Arg  
 590 595 600 605  
 50 Met Leu Asp Gly Asn His Ala Phe Gln Ile Ile Lys Asn Met Ile Gln  
 610 615 620  
 55 Leu Leu Pro Asn Asp His Leu Ala Lys Glu Tyr Pro Asn Gly Arg Thr  
 625 630 635  
 60 Tyr Pro Asn Met Leu Asp Ala His Pro Pro Phe Gln Ile Asp Gly Asn  
 640 645 650  
 65 Phe Gly Tyr Thr Ala Gly Val Ala Glu Met Leu Leu Gln Ser His Asp  
 655 660 665

# ES 2 649 217 T3

	Gly	Ala	Val	His	Leu	Leu	Pro	Ala	Leu	Pro	Asp	Ala	Trp	Glu	Glu	Gly
	670					675					680					685
5	Ser	Val	Lys	Gly	Leu	Val	Ala	Arg	Gly	Asn	Phe	Thr	Val	Asp	Met	Asp
					690					695					700	
10	Trp	Lys	Asn	Asn	Val	Leu	Asn	Lys	Ala	Ile	Ile	Arg	Ser	Asn	Ile	Gly
			705						710					715		
15	Gly	Thr	Leu	Arg	Ile	Arg	Ser	Tyr	Val	Pro	Leu	Lys	Gly	Lys	Gly	Leu
			720					725					730			
20	Lys	Gln	Val	Asn	Gly	Lys	Glu	Cys	Ser	Asn	Arg	Leu	Phe	Ala	Thr	Thr
		735					740					745				
25	Pro	Ile	Lys	Arg	Pro	Leu	Val	Ala	Lys	Gly	Ile	Ser	Ala	Gln	Ser	Pro
	750					755					760					765
30	Lys	Leu	Gln	Lys	Val	Tyr	Glu	Tyr	Asp	Ile	Glu	Thr	Lys	Ala	Gly	Lys
					770					775					780	
35	Thr	Tyr	Ile	Val	Asn	Thr	Ile	Glu	Gly	Lys	Gln					
				785					790							

## REIVINDICACIONES

- 5 1. Uso de un polipéptido aislado que tiene actividad de  $\alpha$ -L-galactosidasa, seleccionado del grupo consistente en:
- (a) un polipéptido con al menos 80%, por ejemplo al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, o al menos 99% de identidad de secuencia respecto al polipéptido maduro de la SEQ ID NO: 2 o la SEQ ID NO:3;
- 10 (b) un polipéptido codificado por un polinucleótido que se hibridiza bajo condiciones de astringencia alta o condiciones de astringencia muy alta con (i) la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1, o (ii) el complemento en toda su longitud de (i);
- (c) un polipéptido codificado por un polinucleótido al menos 80%, por ejemplo al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, o al menos 99% de identidad de secuencia respecto a la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1
- 15 (d) una variante del polipéptido maduro de la SEQ ID NO: 2 o SEQ ID NO: 3 que comprende una sustitución, delección, y/o inserción en una o más posiciones, donde los polipéptidos difieren en no más de 10 aminoácidos; y
- 20 (e) un fragmento del polipéptido de (a); (b), (c), o (d) que tiene actividad de  $\alpha$ -L-galactosidasa donde el fragmento contiene al menos 90% de los aminoácidos del péptido maduro,
- en pienso para animales.
- 25 2. Uso según la reivindicación 1, donde el polipéptido comprende o consiste en la SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 o SEQ ID NO: 6 o el polipéptido maduro de la SEQ ID NO: 2, SEQ ID NO: 4 SEQ ID NO: 5 o SEQ ID NO: 6.
- 30 3. Uso según cualquiera de las reivindicaciones 1 a 2, donde el polipéptido maduro es los aminoácidos 20 a 811 de la SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 o SEQ ID NO: 6.
4. Composición que comprende un polipéptido aislado que tiene actividad de  $\alpha$ -L-galactosidasa, seleccionado del grupo que consiste en:
- 35 (a) un polipéptido que tiene al menos 80%, por ejemplo, al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91%, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, o 100% de identidad de secuencia con la SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 o SEQ ID NO: 6;
- 40 (b) un polipéptido codificado por un polinucleótido que se hibrida bajo condiciones de astringencia alta o condiciones de astringencia muy alta con:
- (i) la secuencia codificante del polipéptido maduro de la SEQ ID NO: 1; y/o
- (ii) la cadena complementaria de longitud completa de (i);
- 45 (c) un polipéptido codificado por un polinucleótido que tiene al menos 80%, por ejemplo, al menos 85%, al menos 86%, al menos 87%, al menos 88%, al menos 89%, al menos 90%, al menos 91 %, al menos 92%, al menos 93%, al menos 94%, al menos 95%, al menos 96%, al menos 97%, al menos 98%, al menos 99%, o 100% de identidad de secuencia con la secuencia codificante del polipéptido maduro de la SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 o SEQ ID NO: 6;
- 50 (d) una variante que comprende una sustitución, delección y/o inserción de uno o más (varios) aminoácidos de la SEQ ID NO: 3, donde los polipéptidos difieren en no más de 30 aminoácidos; y
- (e) un fragmento de un polipéptido de (a), (b), (c) o (d), que tiene actividad de  $\alpha$ -L-galactosidasa, donde el fragmento contiene al menos 90% de los aminoácidos del péptido maduro.
- 55 5. Composición de la reivindicación 4, donde el polipéptido comprende o consiste en la SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5 o SEQ ID NO: 6 o el polipéptido maduro de la SEQ ID NO: : 2, SEQ ID NO: 4 SEQ ID NO: 5 o SEQ ID NO: 6.
- 60 6. Composición de cualquiera de las reivindicaciones 4 a 5, que está en forma de un granulado o un microgranulado.
7. Formulación de caldo completo o composición de cultivo celular que comprende un polipéptido indicado en cualquiera de las reivindicaciones 1-5.
- 65 8. Uso de al menos un polipéptido indicado en una cualquiera de las reivindicaciones 1 a 5

- 5 en pienso para animales;  
en aditivos para pienso para animales;  
en la preparación de una composición para uso en pienso para animales;  
para mejorar el valor nutricional de un pienso para animales;  
para aumentar la proteína digerible y/o soluble en pienso para animales;  
para aumentar el grado de hidrólisis de las proteínas en las dietas para animales; y/o  
para el tratamiento de proteínas.
- 10 9. Método para mejorar el valor nutricional de un pienso para animales, donde al menos un polipéptido indicado en cualquiera de las reivindicaciones 1 a 5 se añade al pienso.
- 15 10. Aditivo para pienso que comprende  
al menos un polipéptido indicado en cualquiera de las reivindicaciones 1 a 5; y  
al menos una vitamina soluble en grasa, y/o  
al menos una vitamina soluble en agua, y/o  
al menos un oligoelemento.
- 20 11. Aditivo para pienso para animales de la reivindicación 10, donde el pienso para animales comprende una o más enzimas adicionales, donde las enzimas adicionales se seleccionan del grupo que comprende amilasas; fitasas; xilanasas; galactanasas; alfa-galactosidasas; proteasas, fosfolipasas; y beta-glucanasas, o cualquier mezcla de las mismas.
- 25 12. Aditivo para pienso para animales de cualquiera de las reivindicaciones 10 a 11, que comprende al menos un microbiano probiótico o de alimentación directa (DFM).
- 30 13. Pienso para animales que tiene un contenido de proteína bruta de 50 a 800 g/kg y que comprende al menos un polipéptido indicado en cualquiera de las reivindicaciones 1 a 5.
- 35 14. Método para el tratamiento de proteínas, que comprende el paso de añadir al menos un polipéptido indicado en cualquiera de las reivindicaciones 1 a 5 a al menos una proteína o fuente de proteína.
15. Método de la reivindicación 14, donde se incluye grano, maíz, soja o harina de soja en la al menos una fuente de proteína.



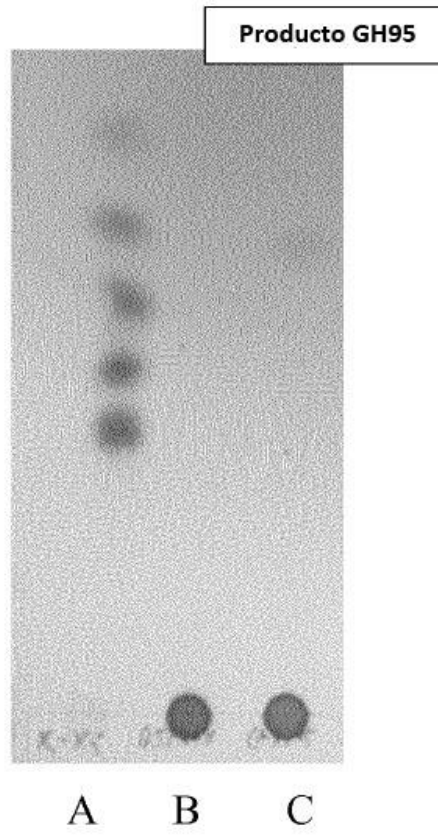


Fig. 1

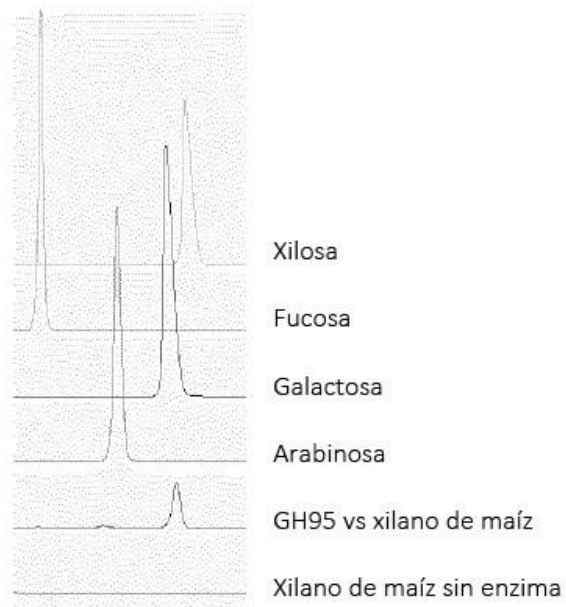


Fig. 2

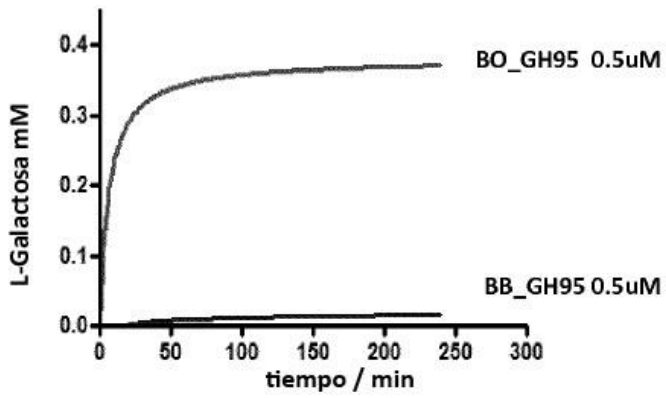


Fig. 3

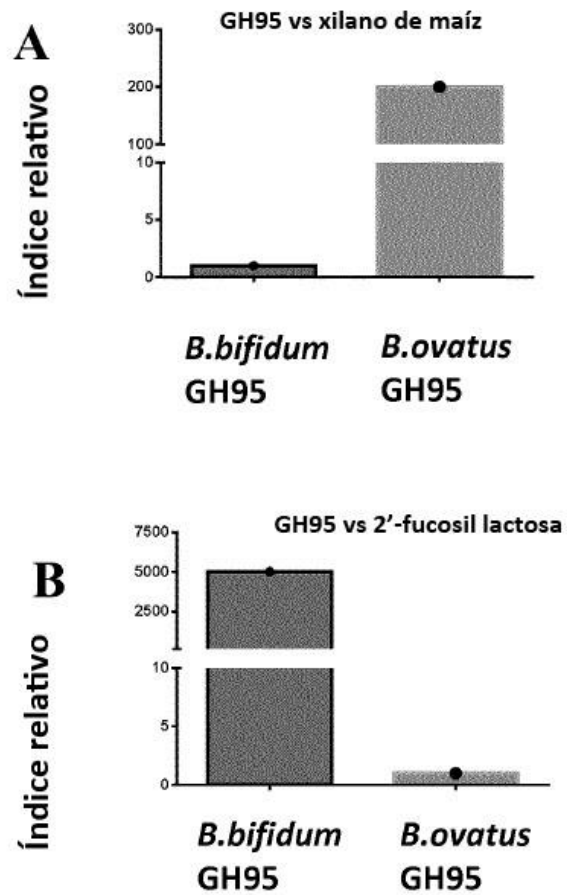


Fig. 4

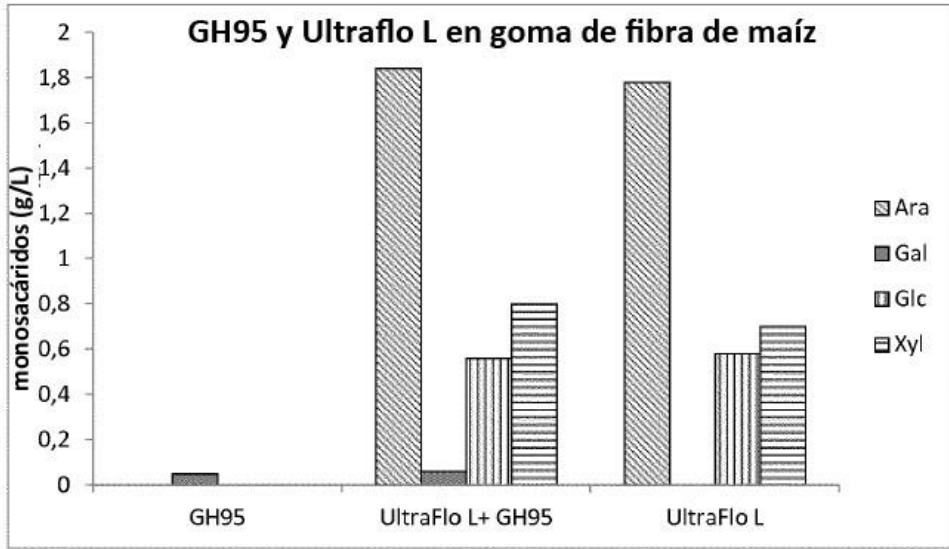


Fig. 5

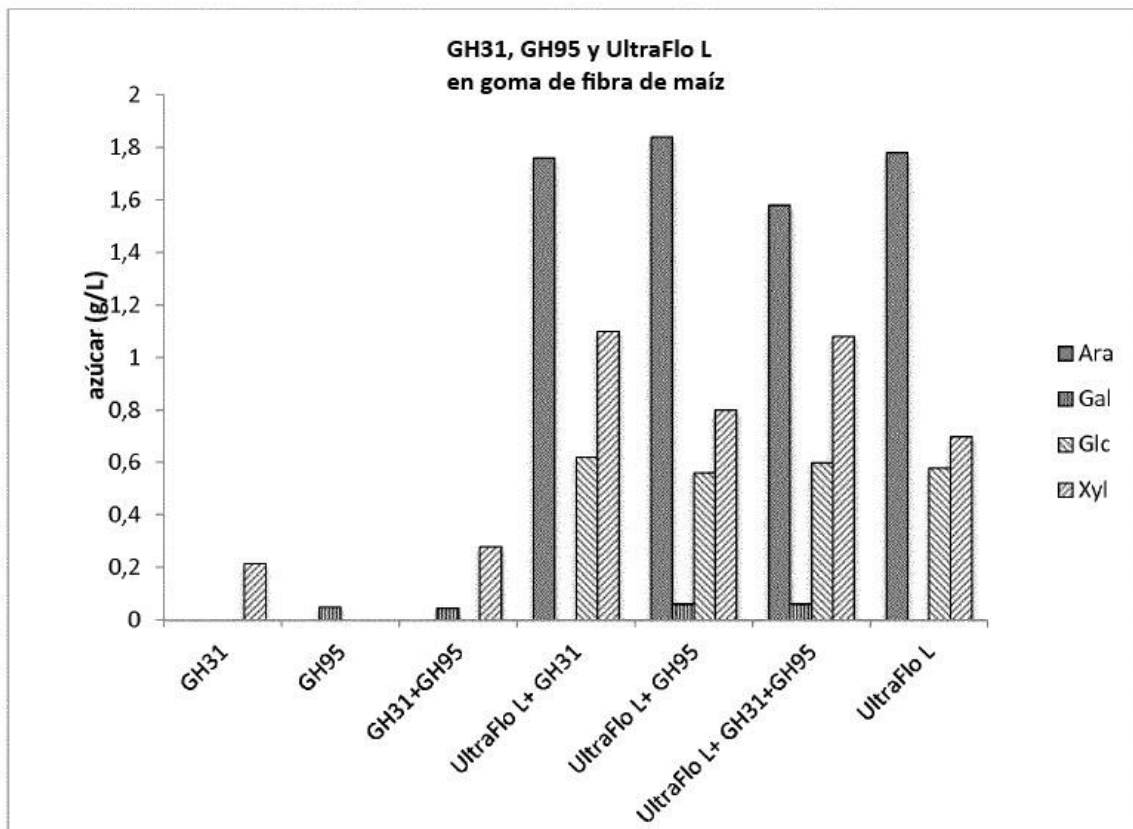


Fig. 6

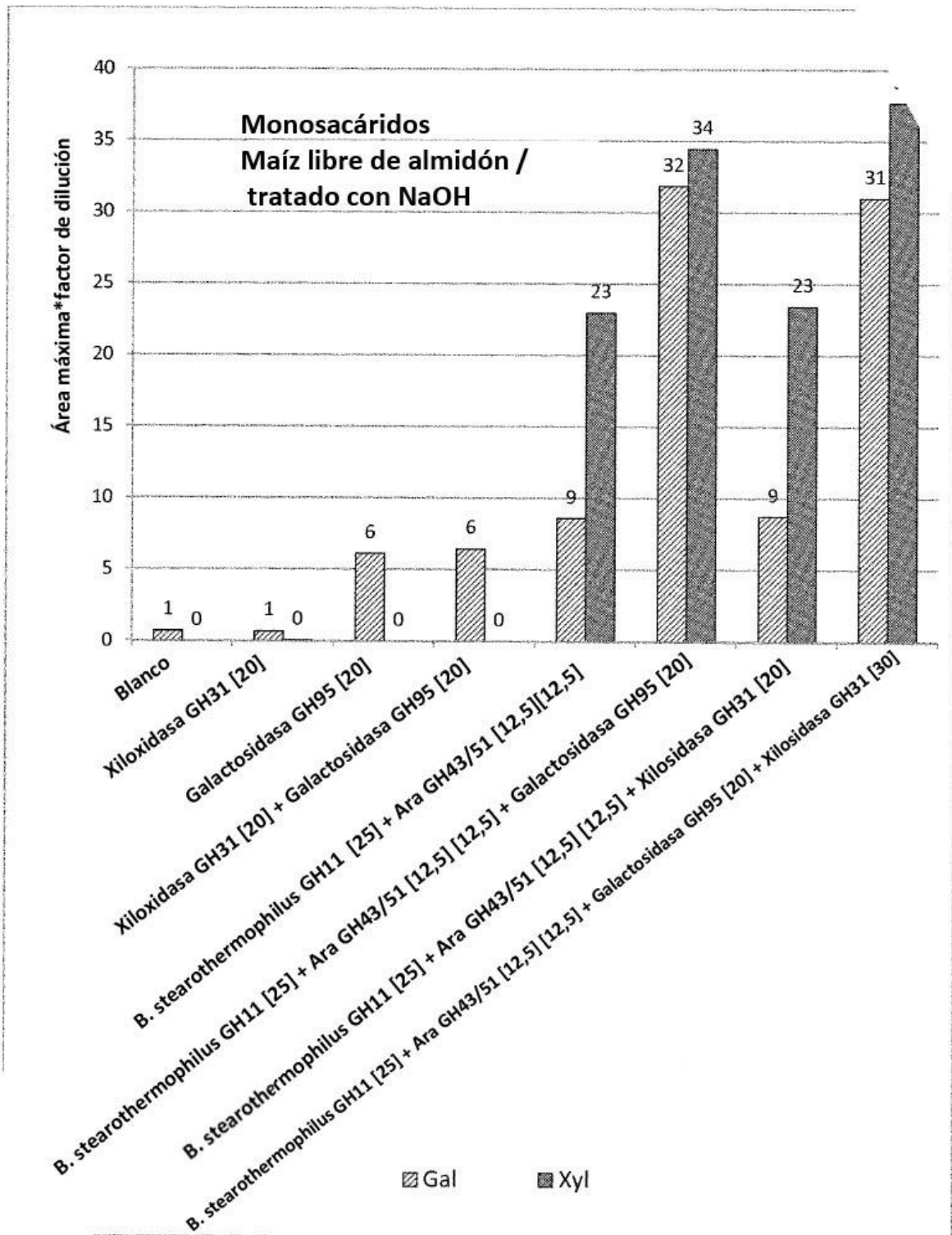


Fig. 7