

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 795 989**

51 Int. Cl.:

C12N 9/34 (2006.01)
C12N 15/56 (2006.01)
C12N 1/15 (2006.01)
C12N 15/80 (2006.01)
C12P 19/02 (2006.01)
C12P 19/14 (2006.01)
C12P 7/06 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **20.11.2008** **E 17205190 (6)**

97 Fecha y número de publicación de la concesión europea: **25.03.2020** **EP 3323889**

54 Título: **Variantes de glucoamilasa con propiedades modificadas**

30 Prioridad:

20.11.2007 US 989426 P

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

25.11.2020

73 Titular/es:

DANISCO US INC. (100.0%)
925 Page Mill Road
Palo Alto, California 94304 , US

72 Inventor/es:

AEHLE, WOLFGANG;
BOTT, RICHARD, R.;
SCHEFFERS, MARTIJIN;
VAN SOLINGEN, PIET y
VROEMEN, CASPER

74 Agente/Representante:

FLORES DREOSTI, Lucas

ES 2 795 989 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

Variantes de glucoamilasa con propiedades modificadas

5 CAMPO DE LA INVENCION

10 [0001] Las variantes de glucoamilasas presentan de forma beneficiosa propiedades modificadas (p. ej., termoestabilidad y/o actividad especifica mejoradas). Se presentan composiciones que comprenden las glucoamilasas variantes, construcciones de ADN que codifican las variantes y métodos para producir las variantes de glucoamilasa en células huésped.

ANTECEDENTES

15 [0002] Las enzimas de glucoamilasa (glucano 1,4- α -glucohidrolasas, EC 3.2.1.3) son carbohidrasas hidrolizantes de almidón que actúan en exo, que catalizan la eliminación de unidades de glucosa sucesivas de los extremos no reductores de moléculas relacionadas con oligosacáridos y polisacáridos o almidón. Las glucoamilasas pueden hidrolizar tanto los enlaces glucosídicos lineales como ramificados de almidón (p.ej., amilosa y amilopectina).

20 [0003] Las glucoamilasas se producen mediante numerosas cepas de bacterias, hongos, levaduras y plantas. Especialmente interesantes, y comercialmente importantes, las glucoamilasas son enzimas fúngicas que se producen de forma extracelular, por ejemplo, a partir de cepas de *Aspergillus* (Svensson *et al.* (1983) Carlsberg Res. Commun. 48:529-544; Boel *et al.*, (1984) EMBO J. 3:1097-1102; Hayashida *et al.*, (1989) Agric. Biol. Chem. 53:923-929; patente estadounidense n.º 5024941; patente estadounidense n.º 4794175 y WO 88/09795); *Talaromyces* (patente estadounidense n.º 4247637; patente estadounidense n.º 6255084 y patente estadounidense n.º 6620924); *Rhizopus* (Ashikari *et al.*, (1986) Agric. Biol. Chem. 50:957-964; Ashikari *et al.*, (1989) App. Microbiol. Biotech. 32:129-133 y patente estadounidense n.º 4863864); *Humicola* (WO 05/052148 y patente estadounidense n.º 4618579) y *Mucor* (Houghton-Larsen *et al.*, (2003) Appl. Microbiol. Biotechnol. 62:210-217). Muchos de los genes que codifican estas enzimas han sido clonados y expresados en células bacterianas, fúngicas y/o levadura.

30 [0004] Comercialmente, las glucoamilasas son enzimas muy importantes y se han utilizado en una amplia variedad de aplicaciones que requieren la hidrólisis de almidón (p. ej., para la producción de glucosa y otros monosacáridos a partir de almidón). Las glucoamilasas se utilizan para producir edulcorantes de maíz con alto contenido de fructosa, que comprenden más del 50 % del mercado de edulcorantes en Estados Unidos. En general, las glucoamilasas pueden utilizarse, y normalmente se utilizan, con alfa-amilasas en procesos hidrolizantes de almidón para hidrolizar almidón en dextrinas y posteriormente glucosa. La glucosa puede entonces convertirse en fructosa por parte de otras enzimas (p. ej., isomerasas de glucosa); cristalizarse; o utilizarse en fermentaciones con el fin de producir numerosos productos finales (p. ej., etanol, ácido cítrico, ácido láctico, succinato, intermediarios de ácido ascórbico, ácido glutámico, glicerol y 1,3-propanodiol). El etanol producido mediante el uso de glucoamilasas en la fermentación de material que contiene celulosa y/o almidón puede utilizarse como una fuente de combustible o para consumo alcohólico.

40 WO 2006/060062 expone la glucoamilasa *Trichoderma reesei* y homólogos de esta.

45 [0005] US 2003/0032163 describe una variante de una glucoamilasa fúngica precursora, cuya variante muestra una estabilidad térmica mejorada y/o una actividad específica aumentada mediante sustratos sacáridos.

[0006] WO 01/04273 describe una variante de una glucoamilasa fúngica, cuya variante muestra propiedades modificadas, en concreto una estabilidad térmica mejorada y/o una actividad específica aumentada.

50 [0007] Aunque las glucoamilasas se han utilizado con éxito en aplicaciones comerciales durante muchos años, todavía existe la necesidad de nuevas glucoamilasas con propiedades modificadas, tal como actividad específica mejorada o termoestabilidad aumentada.

55 [0008] Se han realizado diferentes mutaciones en glucoamilasas de *Aspergillus* que potencian la estabilidad térmica y la actividad específica. Se hace referencia a la patente estadounidense n.º 6537792; la patente estadounidense n.º 6352851; Chen *et al.* (1996) Prot. Eng. 9:499-505, Chen *et al.*, (1995) Prot Eng. 8:575-582; Fierobe *et al.* (1996) Biochem. 35:8698-8704; y Li *et al.*, (1997) Prot. Eng. 10:1199-1204. Todavía existe la necesidad de proporcionar variantes de glucoamilasa con propiedades modificadas en relación con su precursora.

SUMARIO

[0009] La presente exposición se refiere a variantes de glucoamilasa de una glucoamilasa precursora. En un primer aspecto, la presente invención proporciona una variante de glucoamilasa que comprende al menos:

- 5 (i) una sustitución de aminoácidos en la posición 431 en SEQ ID NO: 2 o una posición equivalente basada en la alineación de la secuencia de aminoácidos con la secuencia de SEQ ID NO: 2; y
- (ii) una o más sustituciones de aminoácidos en posiciones correspondientes a las posiciones 61, 73, 417, 430, 503, 511, 535, 539 o 563 de SEQ ID NO: 2 o posiciones equivalentes basadas en la alineación de la secuencia de aminoácidos con la secuencia de SEQ ID NO: 2;
- 10 en la que la variante de glucoamilasa presenta al menos un 93% de identidad de secuencia con SEQ ID NO: 2, y en la que la variante de glucoamilasa muestra termoestabilidad aumentada y/o actividad específica aumentada en comparación con la glucoamilasa de SEQ ID NO: 2.

[0010] En alguna forma de realización, la sustitución en la posición 431 en SEQ ID NO: 2 se elige de: A431L/Q.

- 15 **[0011]** En algunas partes de la exposición, la variante es una variante de una glucoamilasa precursora que presenta un dominio catalítico con al menos un 80 % de identidad de secuencia con SEQ ID NO: 1 o 2, o que presenta un dominio de unión a almidón con al menos un 80 % de identidad de secuencia con SEQ ID NO: 1 o 2.

- 20 **[0012]** En algunas partes de la exposición, la variante de glucoamilasa presenta al menos un 95 % o al menos un 99,5 % de identidad de secuencia con SEQ ID NO: 2, o la variante es una variante de una glucoamilasa precursora que es SEQ ID NO: 1 o 2.

- 25 **[0013]** En algunas formas de realización, la variante comprende sustituciones de aminoácidos correspondientes a A431L/Q y uno o más de N61I, G73F, L417R/V, T430A/M, E503A/V, Q511H, A535R, A539R o N563I/K en SEQ ID NO: 2.

- 30 **[0014]** En algunas partes de la exposición, la variante comprende además una o más de las siguientes sustituciones: D4L/E/R/S/C/A/Q/W, F5C/M/N/R/S/T/V/W, I12L/R, D24E/L/Y/T, F29L/I/D/C/S/V/W, I43F/R/D/Y/S/Q, D44E/H/K/S/N/Y/F/R/C, Y47W, Y49N, Q70R/K/M/P/G/L/F, Q75R/K/A, R76L/M/K/T/P, P94L, D100W/I/Q/M/P/A/N, N119P/T/Y/D/E, N146S/G/C/H/E/D/T/W/L/F/M, Q148V/Y/H/A/C/D/G/M/R/S/T, Y169D/F, Q172C/A/D/R/E/F/H/V/L/M/N/S/T/V, F175H/A/G/R/S/T/C/W/Y, W178A/C/D/E/F/G/H/K/N/R/S/T/V/Y, E180A/C/G/H/I/L/N/P/Q/R/S/T/V/Y, V181E/C/D/G/H/I/P/T/Y/S/L/K/F/A, Q208L/A/C/E/N/F/H/T, S211C/R/E/A/Y/W/M/H/L/I/R/Q/T, E243S/R/N/M/Y/A/L, R245A/E/M/I/P/V, I292D/H/P/R/T/N/V/F/L, G294C/D/E/T/Q/I/A, K297F/L/P/T/M/D/N/Q/A/Y/H/S/R/W, R309A/C/G/H/I/N/P/Q/S/T/W/Y/L,
- 35 Y310E/G/L/P/S/W/R/Q, D313Q, V314A/R/N/D/C/E/Q/G/H/I/L/K/M/F/P/S/T/W/Y, Y315F, Y316Q/R, N317T/H, K340D/T, K341F/D/P/V/G/S, T350S/E/A/N, Q356H/D/E, T363L/R/C/H/W, S368W/D/F/L, S369F, N376Q/T/H/S/V, Y395Q/R/S, A398S/I/T, S401C/V, R408S, N409W/T/K, T412A/H/K/G, R433H/Q, 1436A/T, o S451 M/T/H en SEQ ID NO: 2.

- 40 **[0015]** En algunas partes de la exposición, la variante de glucoamilasa comprende múltiples sustituciones en las posiciones:

- I43/N61/T430/**A431**/Q511/A539;
 I43/L417/T430/**A431**/Q511/A539;
 I43/T430/**A431**/E503/Q511;
 45 I43/T430/**A431**/Q511;
 I43/T430/**A431**/Q511/A539;
 I43/**A431**/Q511;
 G294/L417/**A431**;
 G294/L417/**A431**/Q511;
 50 L417/T430/**A431**/Q511/A535/A539/N563; y
 L417/**A431**/Q511.

[0016] En algunas formas de realización, la variante de glucoamilasa comprende uno de los conjuntos de sustitución:

- 55 I43Q/N61I/T430A/**A431L**/Q511H/A539R;

- I43Q/L417V/T430A/A431L/Q511H/A539R;
 I43Q/T430A/A431L/E503A/Q511H;
 I43Q/T430A/A431L/Q511H;
 I43Q/T430A/A431L/Q511H/A539R;
 5 I43Q/A431L/Q511H;
 G294C/L417R/A431L;
 G294C/L417R/A431L,Q/Q511H;
 G294C/L417V/A431Q;
 L417R,V/A431L,Q/Q511H; y
 10 L417V/T430A/A431L,Q/Q511H/A535R/A539R/N563I, de SEQ ID NO: 2.

[0017] En algunas formas de realización, la glucoamilasa precursora se elige de una glucoamilasa obtenida de una *Trichoderma* spp., una *Aspergillus* spp., una *Humicola* spp., una *Penicillium* spp., una *Talaromyces* spp., o una *Schizosaccharomyces* spp.

15

[0018] En un segundo aspecto, la presente invención proporciona un polinucleótido que codifica la variante de glucoamilasa del primer aspecto de la invención.

20

[0019] En un tercer aspecto, la presente invención proporciona un vector que comprende el polinucleótido del segundo aspecto de la invención.

[0020] En un cuarto aspecto, la presente invención proporciona una célula huésped que comprende el vector del tercer aspecto de la invención.

25

[0021] En un quinto aspecto, la presente invención proporciona una composición de enzima que comprende la variante de glucoamilasa de cualquiera del primer aspecto de la invención.

30

[0022] En un sexto aspecto, la presente invención proporciona un método para producir una variante de glucoamilasa en una célula huésped que comprende el cultivo de la célula huésped del cuarto aspecto de la invención con las condiciones adecuadas para la expresión y producción de la variante de glucoamilasa y la producción de la variante de glucoamilasa.

35

[0023] En algunas formas de realización, el método comprende además la recuperación de la variante de glucoamilasa del cultivo.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0024]

40

La FIG. 1A representa la glucoamilasa precursora *Trichoderma reesei* (TrGA) con 632 aminoácidos (SEQ ID NO: 1). El péptido señal se subraya, la región catalítica (SEQ ID NO: 3) que comienza con los residuos de aminoácidos SVDDFI (SEQ ID NO: 160) y que presenta 453 residuos de aminoácidos está en negrita; la región de enlace está en cursiva y el dominio de unión a almidón (SEQ ID NO: 161) está tanto en cursiva como subrayado. La proteína madura, que incluye el dominio catalítico (SEQ ID NO: 3), la región de enlace y el dominio de unión a almidón (SEQ ID NO: 161) se representa mediante SEQ ID NO: 2. La FIG. 1B representa el ADNc (SEQ ID NO: 4) que codifica la TrGA.

45

La FIG. 2 representa el plásmido pDONR-TrGA que incluye el ADNc (SEQ ID NO: 4) de la TrGA precursora.

50

La FIG. 3 representa los plásmidos pREP3Y-DEST (A) y pREP3Y-TrGA (B).

55

Las FIGS. 4A-4B representan una comparación de la alineación de los dominios catalíticos de glucoamilasas precursoras que incluye glucoamilasa derivada de *Aspergillus awamori* (AaGA) (SEQ ID NO: 5); *Aspergillus niger* (AnGA) (SEQ ID NO: 6); *Aspergillus oryzae* (AoGA) (SEQ ID NO: 7); *Trichoderma reesei* (TrGA) (SEQ ID NO: 3); *Humicola grisea* (HgGA) (SEQ ID NO: 8); e *Hypocrea vinosa* (HvGA) (SEQ ID NO: 9). Los aminoácidos idénticos se indican con un asterisco (*),

Las FIGS. 4C-4D representan una alineación que compara el dominio de unión a almidón (SBD) de glucoamilasas precursoras entre las que se incluyen *Trichoderma reesei* (SEQ ID NO: 161), *Humicola grisea* (HgGA) (SEQ ID NO: 162), *Thielavia terrestris* (TtGA) (SEQ ID NO: 163), *Thermomyces lanuginosus* (ThGA) (SEQ ID NO: 164), *Talaromyces emersonii* (TeGA) (SEQ ID NO: 165), *Aspergillus niger* (AnGA) (SEQ ID NO: 166), y *Aspergillus awamori* (AaGA) (SEQ ID NO: 167). Los aminoácidos idénticos se indican con un punto (.).

La FIG. 5A representa el plásmido pTrex3g-DEST. La FIG. 5B representa el plásmido pTrex3g-TrGA. Los plásmidos se utilizaron como vectores de expresión para la expresión y producción de glucoamilasas variantes en un huésped *Trichoderma reesei*.

La FIG. 6 representa la comparación V_{max} (μM glucosa/s) entre la TrGA precursora (de tipo silvestre) y las variantes, V314, S211R, Q172F y Q208N a 60 °C y 32 °C, como se detalla en el Ejemplo 8.

La FIG. 7 representa la actividad de variantes combinatorias en un sustrato de almidón. Entre las variantes combinatorias descritas en el presente documento se incluyen ET7-1 (D24Y/V181L/Q208C/G294A/T353R/N375N/N409W), LR8 (Q172F/Q208N), LR12 (Q172F/S211R), LR6 (Q172F/Q208N/V314H), ET8-1 (D24E/V181K/E243Y/I292V/G294Q/N409K) y ET7-2 (Q24L/V181L/Q208C/G294A/T353R/N375Q/N409W). La actividad se representa en unidades de absorción a 340 nm como función de ng de las variantes de glucoamilasa indicadas.

La FIG. 8 representa la actividad de variantes de sitio único en un sustrato de maicena. Entre las variantes de sitio único descritas en el presente documento se incluyen V314H, G294Q, S211R, Q208N, Q172F, G294I y P94N. La actividad se representa en unidades de absorción a 340 nm como función de ng de las variantes de glucoamilasa indicadas.

La FIG. 9 representa la actividad de glucoamilasa de TrGA y la variante de TrGA LR8 (Q172F/Q208N) sobre una muestra de sustrato licuado de pulpa de maíz (NE).

La FIG. 10 representa el perfil de actividad de TrGA y la variante de TrGA LR8 (Q172F/Q208N) sobre una muestra de sustrato licuado de pulpa de maíz (BSE).

La FIG. 11 representa el perfil de actividad de TrGA y la variante de TrGA LR8 (Q172F/Q208N) sobre una muestra de sustrato de maicena soluble.

La FIG. 12 representa una comparación de las estructuras tridimensionales de la glucoamilasa *Trichoderma reesei* (negra) (SEQ ID NO: 2) y la glucoamilasa *Aspergillus awamori* (gris), vistas de lado.

La FIG. 13 representa una comparación de las estructuras tridimensionales de la glucoamilasa *Trichoderma reesei* (negra) y la glucoamilasa *Aspergillus awamori* (gris) vistas desde la parte superior.

La FIG. 14A representa el plásmido pTTT-Dest. La FIG. 14B representa el plásmido pTTT-TrGA(B).

DESCRIPCIÓN DETALLADA

[0025] Las glucoamilasas son enzimas comercialmente importantes en una amplia variedad de aplicaciones que requieren la hidrólisis de almidón. Las variantes de glucoamilasa descritas en el presente documento contienen sustituciones de aminoácidos dentro del dominio catalítico o del dominio de unión a almidón. Las variantes pueden mostrar propiedades modificadas, tal como termoestabilidad y/o actividad específica mejoradas. Las variantes con termoestabilidad y/o actividad específica mejoradas pueden mejorar de forma significativa la eficacia de la producción de etanol carburante y glucosa a partir de maicena, por ejemplo.

1. Definiciones

[0026] A menos que se defina lo contrario, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado según lo entiende comúnmente alguien especializado en la técnica a la que esta invención pertenece. Singleton, *et al.*, DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 2D ED., John Wiley and Sons, New York (1994), y Hale & Markham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, N.Y. (1991) proporcionan al experto el significado general de muchos de los términos empleados en el presente documento. Determinados términos se definen a continuación para una mayor claridad y facilidad de referencia.

[0027] El término "glucoamilasa (EC 3.2.1.3)" se refiere en el presente documento a una enzima que cataliza la liberación de D-glucosa desde los extremos no reductores de almidón y polisacáridos y oligosacáridos relacionados.

5 **[0028]** El término "precursor" o "secuencia precursora" hace referencia a una secuencia que es natural o que se da de forma natural en una célula huésped. Entre las secuencias precursoras se incluyen, aunque sin carácter limitativo, las secuencias de glucoamilasa establecidas en SEQ ID NO: 1, 2, 3, 5, 6, 7, 8 y 9.

10 **[0029]** El término "posición equivalente" se refiere en el presente documento a una posición que es común a las dos secuencias precursoras que se basa en una alineación de la secuencia de aminoácidos de la glucoamilasa precursora en cuestión así como una alineación de la estructura tridimensional de la glucoamilasa precursora en cuestión con la secuencia de aminoácidos de la glucoamilasa de referencia TrGA (SEQ ID NO: 2) y la secuencia tridimensional.

15 **[0030]** El término "TrGA" se refiere a una secuencia de glucoamilasa *Trichoderma reesei* precursora que presenta la secuencia de proteína madura ilustrada en SEQ ID NO: 2 que incluye el dominio catalítico que presenta la secuencia ilustrada en SEQ ID NO: 3. El aislamiento, clonación y expresión de la TrGA se describen en la patente estadounidense n.º 7413887. En algunas formas de realización, la secuencia precursora hace referencia a una glucoamilasa que es el punto inicial para la ingeniería de proteínas. La numeración de los aminoácidos de la glucoamilasa en el presente documento se basa en la alineación de secuencia de una glucoamilasa con TrGA (SEQ ID NO: 2 y SEQ ID NO: 3).

20 **[0031]** La frase "forma madura de una proteína o polipéptido" hace referencia a la forma funcional final de la proteína o polipéptido. Con el fin de ejemplificarlo, una forma madura de la TrGA incluye el dominio catalítico, la región de enlace y el dominio de unión a almidón que presenta la secuencia de aminoácidos de SEQ ID NO: 2.

25 **[0032]** Los términos "variante de glucoamilasa" y "variante" se emplean en el presente documento en referencia a glucoamilasas que presentan un grado de identidad de secuencia de aminoácidos con una secuencia de glucoamilasa precursora y que pueden conservar características funcionales de una glucoamilasa. Una variante es similar a una secuencia precursora pero presenta al menos una sustitución, delección o inserción en su secuencia de aminoácidos que hace que sea diferente de una glucoamilasa precursora por la secuencia. En algunos casos, las variantes han sido manipuladas y/o diseñadas para que incluyan al menos una sustitución, delección o inserción en su secuencia de aminoácidos que las diferencie de una precursora por la secuencia.

30 **[0033]** Las "variantes" pueden presentar al menos un 99,5 %, al menos un 99 %, al menos un 98 %, al menos un 97 %, al menos un 96 %, al menos un 95 %, al menos un 94 %, al menos un 93 %, al menos un 92 %, al menos un 91 %, al menos un 90 %, al menos un 88 %, al menos un 85 %, al menos un 80 %, al menos un 75 %, al menos un 70 %, al menos un 65 %, al menos un 60 %, al menos un 55 %, al menos un 50 % o al menos un 45 % de identidad de secuencia con una secuencia de polipéptidos cuando se alinea de forma óptima para su comparación.

35 En algunas partes de la exposición, la variante de glucoamilasa puede presentar al menos un 99,5 %, al menos un 99 %, al menos un 98 %, al menos un 97 %, al menos un 96 %, al menos un 95 %, al menos un 94 %, al menos un 93 %, al menos un 92 %, al menos un 91 %, al menos un 90 %, al menos un 88 %, al menos un 85 %, al menos un 80 %, al menos un 75 %, al menos un 70 %, al menos un 65 %, al menos un 60 %, al menos un 55 %, al menos un 50 % o al menos un 45 % de identidad de secuencia con el dominio catalítico de una glucoamilasa precursora.

40 En algunas partes de la exposición, la variante de glucoamilasa puede presentar al menos un 99,5 %, al menos un 99 %, al menos un 98 %, al menos un 97 %, al menos un 96 %, al menos un 95 %, al menos un 94 %, al menos un 93 %, al menos un 92 %, al menos un 91 %, al menos un 90 %, al menos un 88 %, al menos un 85 %, al menos un 80 %, al menos un 75 %, al menos un 70 %, al menos un 65 %, al menos un 60 %, al menos un 55 %, al menos un 50 % o al menos un 45 % de identidad de secuencia con el dominio de unión a almidón de una glucoamilasa precursora.

45 La identidad de secuencia puede medirse sobre toda la longitud de la secuencia precursora o variante.

50 **[0034]** La identidad de secuencia se determina utilizando técnicas convencionales conocidas en la técnica (véase, p. ej., Smith and Waterman, Adv. Appl. Math. 2: 482 (1981); Needleman and Wunsch, J. Mol. Biol. 48: 443 (1970); Pearson and Lipman, Proc. Natl. Acad. Sci. EEUU 85: 2444 (1988); programas como GAP, BESTHT, FASTA, y TFASTA en Wisconsin Genetics Software Package (Genetics Computer Group, Madison, WI); y Devereux *et al.*, Nucleic Acid Res., 12: 387-395 (1984)).

55 **[0035]** El "porcentaje (%) de identidad de secuencia de ácido nucleico" o "porcentaje (%) de identidad de secuencia de aminoácidos" se define como el porcentaje de residuos de nucleótidos o residuos de aminoácidos en una secuencia candidata que son idénticos a los residuos de nucleótidos o residuos de aminoácidos de la secuencia inicial (p. ej., TrGA). La identidad de secuencia puede medirse sobre toda la longitud de la secuencia inicial (p. ej., SEQ ID NO: 2).

- 5 **[0036]** La identidad de secuencia se determina mediante métodos conocidos de alineación de secuencias. Un método de alineación comúnmente conocido es BLAST descrito por Altschul *et al.*, (Altschul *et al.*, J. Mol. Biol. 215: 403-410 (1990); y Karlin *et al.*, Proc. Natl. Acad. Sci. EEUU 90: 5873-5787 (1993)). Un programa BLAST especialmente útil es el programa WU-BLAST-2 (véase Altschul *et al.*, Meth. Enzymol. 266: 460-480 (1996)). WU-BLAST-2 utiliza diferentes parámetros de búsqueda, la mayoría de los cuales están establecidos a los valores por defecto. Los parámetros regulables están configurados con los siguientes valores: periodo de solapamiento = 1, fracción de solapamiento = 0,125, umbral de palabras (T) = 11. Los parámetros HSP S y HSP S2 son valores dinámicos y los establece el propio programa dependiendo de la composición de la secuencia específica y la composición de la base de datos específica contra la que se busca la secuencia de interés. Sin embargo, los valores pueden regularse con el fin de aumentar la sensibilidad. Un valor de % de identidad de secuencia de aminoácidos se determina por el número de residuos idénticos que coinciden dividido por el número total de residuos de la secuencia "más larga" en la región alineada. La secuencia "más larga" es la que presenta los residuos más reales en la región alineada (se ignoran los huecos introducidos por WU-Blast-2 con el fin de maximizar la puntuación de alineación).
- 10 **[0037]** Otros métodos pueden ser útiles para alinear secuencias. Un ejemplo de un algoritmo útil es PILEUP. PILEUP crea una alineación de secuencias múltiple de un grupo de secuencias relacionadas utilizando alineaciones por pares progresivas. También puede trazar un árbol que muestre las relaciones de agrupamiento utilizadas para crear la alineación. PILEUP utiliza una simplificación del método de alineación progresiva de Feng y Doolittle (Feng y Doolittle, J. Mol. Evol. 35: 351-360 (1987)). El método es similar al descrito por Higgins y Sharp (Higgins y Sharp, CABIOS 5: 151-153 (1989)). Parámetros útiles de PILEUP incluyen un peso del hueco por defecto de 3,00, un peso de la longitud del hueco por defecto de 0,10 y huecos de extremo ponderados.
- 15 **[0038]** El término "alineación óptima" hace referencia a la alineación que proporciona la mayor puntuación de porcentaje de identidad.
- 20 **[0039]** El término "dominio catalítico" hace referencia en el presente documento a una región estructural de un polipéptido que contiene el centro activo para la hidrólisis del sustrato.
- 30 **[0040]** El término "enlazador" se refiere a una secuencia de aminoácidos corta que presenta generalmente entre 3 y 40 residuos de aminoácidos que une de forma covalente una secuencia de aminoácidos que comprende un dominio de unión a almidón con una secuencia de aminoácidos que comprende un dominio catalítico.
- 35 **[0041]** El término "dominio de unión a almidón" se refiere a una secuencia de aminoácidos que se une preferentemente a un sustrato de almidón.
- 40 **[0042]** Los términos "secuencia mutante" y "gen mutante" se utilizan en el presente documento indistintamente y hacen referencia a una secuencia de polinucleótidos que presenta una alteración en al menos un codón que se da en la secuencia original de una célula huésped. El producto de expresión de la secuencia mutante es una proteína variante con una secuencia de aminoácidos modificada en relación con la precursora. El producto de expresión puede presentar una capacidad funcional modificada (p. ej., actividad enzimática mejorada).
- 45 **[0043]** El término "propiedad" o equivalentes gramaticales del mismo en el contexto de un polipéptido se refiere en el presente documento a cualquier característica o atributo de un polipéptido que puede elegirse o detectarse. Entre estas propiedades se incluyen, sin carácter limitativo, estabilidad oxidativa, especificidad de sustrato, actividad catalítica, estabilidad térmica, perfil de actividad de pH, resistencia a degradación proteolítica, K_M , K_{CAT} , relación K_M , K_{CAT} , plegamiento de proteínas, capacidad de unirse a un sustrato y capacidad de ser secretado.
- 50 **[0044]** El término "propiedad" o equivalentes gramaticales del mismo en el contexto de un ácido nucleico se refiere en el presente documento a cualquier característica o atributo de un ácido nucleico que pueda elegirse o detectarse. Entre estas propiedades se incluyen, sin carácter limitativo, una propiedad que afecta a la transcripción de genes (p. ej., fuerza del promotor o reconocimiento del promotor), una propiedad que afecta al procesamiento de ARN (p. ej., corte y empalme de ARN y estabilidad de ARN), una propiedad que afecta a la traducción (p. ej., regulación, unión de ARNm a proteínas ribosómicas).
- 55 **[0045]** Los términos "estable térmicamente" y "termoestable" hacen referencia a variantes de glucoamilasa de la presente invención que retienen una cantidad específica de actividad enzimática tras la exposición a temperaturas sobre un periodo de tiempo determinado en condiciones que se mantienen durante la hidrólisis de sustratos de almidón, por ejemplo, mientras se exponen a temperaturas modificadas.

[0046] El término "estabilidad mejorada" en el contexto de una propiedad tal como termoestabilidad se refiere a una mayor actividad hidrolítica de almidón retenido con el tiempo en comparación con otra glucoamilasas de referencia (es decir, precursoras).

5 **[0047]** El término "estabilidad disminuida" en el contexto de una propiedad tal como termoestabilidad se refiere a una menor actividad hidrolítica de almidón retenido con el tiempo en comparación con otra glucoamilasa de referencia.

10 **[0048]** El término "actividad específica" se define como la actividad por mg de proteína de glucoamilasa. En algunas formas de realización, la actividad de glucoamilasa se determina mediante el análisis de etanol descrito en el presente documento y se expresa como la cantidad de glucosa que se produce del sustrato de almidón. En algunas partes de la exposición, la concentración de proteína puede determinarse mediante el análisis Caliper descrito en el presente documento.

15 **[0049]** Los términos "activa" y "biológicamente activa" se refieren a una actividad biológica asociada con una proteína específica. Por consiguiente, la actividad biológica de una proteína determinada se refiere a cualquier actividad biológica típicamente atribuida a aquella proteína por los expertos en la técnica. Por ejemplo, una actividad enzimática asociada a una glucoamilasa es hidrolítica y, por tanto, una glucoamilasa activa presenta actividad hidrolítica.

20

[0050] Los términos "polinucleótidos" y "ácido nucleico", utilizados indistintamente en el presente documento, se refieren a una forma polimérica de nucleótidos de cualquier longitud, ya sean ribonucleótidos o desoxirribonucleótidos. Entre estos términos se incluyen, aunque sin carácter limitativo, ADN de cadena simple, doble o triple, ADN genómico, ADNc, ARN, híbrido ADN-ARN o un polímero que comprenda bases de pirimidina y purina u otras bases de nucleótidos naturales, modificadas de forma química, bioquímica, no naturales o derivatizadas.

25

[0051] Los términos "construcción de ADN, "ADN transformante" y "vector de expresión" se usan indistintamente para referirse al ADN utilizado para introducir secuencias en una célula huésped u organismo. El ADN puede generarse in vitro mediante PCR o cualquier otra técnica adecuada conocida por los expertos en la técnica. La construcción de ADN, el ADN transformante o el casete de expresión recombinante pueden incorporarse en un plásmido, cromosoma, ADN mitocondrial, ADN plastidial, virus o fragmento de ácido nucleico. Normalmente, la parte del casete de expresión recombinante de un vector de expresión, construcción de ADN o ADN transformante incluye, entre otras secuencias, una secuencia de ácido nucleico que ha de transcribirse y un promotor. En algunas partes de la exposición, los vectores de expresión tienen la capacidad de incorporar y expresar fragmentos de ADN heterólogos en una célula huésped.

30

35

[0052] El término "vector" se refiere en el presente documento a una construcción de polinucleótidos diseñada para introducir ácidos nucleicos en uno o más tipos de células. Entre los vectores se incluyen vectores de clonación, vectores de expresión, vectores transportadores, plásmidos, casetes y similares.

40

[0053] Tal y como se emplea en el presente documento en el contexto de introducción de una secuencia de ácido nucleico en una célula, el término "introducido" hace referencia a cualquier método adecuado para transferir la secuencia de ácido nucleico en la célula. Tales métodos para la introducción incluyen sin carácter limitativo fusión de protoplasto, transfección, transformación, conjugación y transducción.

45

[0054] Los términos "transformada" y "transformada de forma estable" hacen referencia en el presente documento a una célula que tiene una secuencia de polinucleótidos no natural (heteróloga) integrada en su genoma o como un plásmido episomal que se mantiene durante al menos dos generaciones.

50

[0055] Los términos "marcador seleccionable" y "marcador selectivo" hacen referencia en el presente documento a un ácido nucleico (p. ej., un gen) capaz de expresarse en células huésped que permite la facilidad de selección de aquellos huéspedes que contienen el vector. Normalmente, los marcadores seleccionables son genes que conceden resistencia antimicrobiana o una ventaja metabólica en la célula huésped con el fin de permitir que las células que contienen el ADN exógeno se distingan de las células que no han recibido ninguna secuencia exógena durante la transformación.

55

[0056] El término "promotor/a" hace referencia en el presente documento a una secuencia de ácido nucleico que funciona para dirigir la transcripción de un gen aguas abajo. El promotor, junto con otras secuencias de ácido nucleico reguladoras de traducción y transcripción (también denominadas "secuencias de control"), es necesario para expresar un gen determinado. En general, entre las secuencias reguladoras de traducción y transcripción se

60

incluyen, sin carácter limitativo, secuencias promotoras, sitios de unión ribosómicos, secuencias de transcripción de inicio y parada, secuencias de traducción de inicio y de parada y secuencias potenciadoras o activadoras.

5 **[0057]** Un ácido nucleico está "unido de forma operativa" cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, ADN que codifica un líder secretor (es decir, un péptido señal), está unido de forma operativa a ADN para un polipéptido si se expresa como una preproteína que participa en la secreción del polipéptido. Generalmente, "unido de forma operativa" indica que las secuencias de ADN que están unidas son contiguas y, en el caso de un líder secretor, contiguas y en fase de lectura.

10 **[0058]** El término "gen" se refiere en el presente documento a un polinucleótido (p. ej., un segmento de ADN) que codifica un polipéptido e incluye regiones que preceden y siguen a las regiones codificantes así como secuencias intermedias (intrones) entre segmentos codificantes individuales (exones).

15 **[0059]** Los términos "ortólogo" y "genes ortólogos" se refieren en el presente documento a genes en diferentes especies que han evolucionado de un gen ancestral común (es decir, un gen homólogo) mediante especiación. Normalmente, los ortólogos conservan la misma función durante el transcurso de la evolución. La identificación de ortólogos es útil a la hora de predecir de forma fiable la función del gen en genomas secuenciados recientemente.

20 **[0060]** Los términos "parálogo" y "genes parálogos" se refieren en el presente documento a genes que están relacionados por duplicación dentro de un genoma. Mientras que los ortólogos conservan la misma función durante el transcurso de la evolución, los parálogos desarrollan nuevas funciones, aunque algunas funciones están relacionadas normalmente con la original. Entre los ejemplos de genes parálogos se incluyen, sin carácter limitativo, genes que codifican tripsina, quimotripsina, elastasa y trombina, que son todas serina proteasas y se dan de forma conjunta en las mismas especies.

25 **[0061]** El término "hibridación" se refiere en el presente documento al proceso por el que una hebra de ácido nucleico se une con una hebra complementaria a través del apareamiento de bases, como se conoce en la técnica.

30 **[0062]** Se considera que una secuencia de ácido nucleico es "hibrizable de forma selectiva" en referencia a una secuencia de ácido nucleico si las dos secuencias se hibridan una a la otra de forma específica bajo condiciones de lavado e hibridación con una astringencia de moderada a alta. Las condiciones de hibridación se basan en la temperatura de fusión (T_m) de la sonda o complejo que une ácido nucleico. Por ejemplo, la "astringencia máxima" normalmente se da a aproximadamente $T_m - 5^\circ\text{C}$ (5° por debajo de la T_m de la sonda); "astringencia alta" a aproximadamente $5-10^\circ\text{C}$ por debajo de la T_m ; "astringencia intermedia" a aproximadamente $10-20^\circ\text{C}$ por debajo de la T_m de la sonda y "baja astringencia" a aproximadamente $20-25^\circ\text{C}$ por debajo de la T_m . De forma funcional, pueden utilizarse las condiciones de astringencia máxima para identificar secuencias que tengan identidad precisa o casi precisa con la sonda de hibridación; mientras que se puede utilizar una hibridación de astringencia intermedia o baja para identificar o detectar homólogos de secuencia de polinucleótidos.

40 **[0063]** Las condiciones de hibridación de astringencia moderada y alta son ya conocidas en la técnica. Un ejemplo de condiciones de astringencia alta incluye la hibridación a aproximadamente 42°C en formamida 50 %, 5X SSC, 5X solución Denhardt, SDS 0,5 % y 100 $\mu\text{g/ml}$ de ADN portador desnaturalizado seguido de dos lavados en 2X SSC y SDS 0,5 % a temperatura ambiente y dos veces adicionales en 0,1X SSC y SDS 0,5 % a 42°C . Un ejemplo de condiciones astringentes moderadas incluye una incubación durante la noche a 37°C en una solución que comprende formamida 20 %, 5 X SSC (150 mM NaCl, 15 mM citrato trisódico), 50 mM fosfato de sodio (pH 7,6), 5 X solución Denhardt, sulfato de dextrano 10 % y 20 mg/ml ADN de esperma de salmón cortado desnaturalizado seguido del lavado de los filtros en 1 X SSC a aproximadamente $37 - 50^\circ\text{C}$. Aquellos expertos en la técnica saben cómo regular la temperatura, la fuerza iónica, etc. según sea necesario para adaptarse a factores tales como la longitud de sonda y similares.

50 **[0064]** El término "recombinante" en el presente documento incluye referencia a una célula o vector que ha sido modificado mediante la introducción de una secuencia de ácido nucleico homóloga o heteróloga o que la célula se ha derivado de una célula modificada de ese modo. Por lo tanto, por ejemplo, las células recombinantes expresan genes que no se hallan en forma idéntica dentro de la forma natural (no recombinante) de la célula o expresan genes naturales que se han expresado de otro modo de forma anormal, subexpresado o no expresado en general como resultado de una intervención humana deliberada.

60 **[0065]** En una forma de realización de la presente exposición, las secuencias de ADN mutado se generan con mutagénesis por saturación del sitio en al menos un codón. En otra parte de la exposición, la mutagénesis por saturación del sitio se lleva a cabo en dos o más codones. En una parte adicional de la exposición, las secuencias de ADN presentan más del 50 %, más del 55 %, más del 60 %, más del 65 %, más del 70 %, más del 75 %, más del 80 %, más del 85 %, más del 90 %, más del 95 % o más del 98 % de homología con la secuencia precursora. En partes alternativas de la exposición, el ADN mutante se genera in vivo utilizando cualquier procedimiento

mutagénico conocido tal como, por ejemplo, radiación, nitrosoguanidina y similares. A continuación, la secuencia de ADN deseada se aísla y se utiliza en los métodos proporcionados en el presente documento.

5 **[0066]** El término "proteína heteróloga" se refiere en el presente documento a una proteína o polipéptido que no se da de forma natural en la célula huésped.

[0067] Una enzima está "sobrexpresada" en una célula huésped si la enzima se expresa en la célula con un nivel mayor que el nivel al que se expresa en una célula de tipo silvestre correspondiente.

10 **[0068]** Los términos "proteína" y "polipéptido" se utilizan indistintamente en el presente documento. En la presente exposición y reivindicaciones, se utilizan los códigos de una letra y tres letras convencionales para residuos de aminoácidos. Los códigos de 3 letras para aminoácidos se definen según la Comisión sobre nomenclatura bioquímica IUPAC-IUB (JCBN). También se entiende que un polipéptido puede codificarse mediante más de una secuencia de nucleótidos debido a la degeneración del código genético.

15 **[0069]** Se describen variantes de la presente exposición mediante la siguiente nomenclatura: [residuo aminoácido original/posición/residuo aminoácido sustituido]. Por ejemplo, la sustitución de leucina por arginina en la posición 76 se representa como R76L. Cuando se sustituye más de un aminoácido en una determinada posición, la sustitución se representa como 1) Q172C, Q172D o Q172R; 2) Q172C, D, o R o 3) Q172C/D/R. Cuando se
20 identifica en la misma una posición adecuada para la sustitución sin un aminoácido específico sugerido, se ha de entender que cualquier residuo de aminoácido puede sustituirse por el residuo de aminoácido presente en la posición. Cuando una glucoamilasa variante contiene una delección en comparación con otras glucoamilasas la delección se indica con "*". Por ejemplo, una delección en la posición R76 se representa como R76*. Una delección de dos o más aminoácidos consecutivos se indica por ejemplo como (76-78)*.

25 **[0070]** Una "prosecuencia" es una secuencia de aminoácidos entre la secuencia señal y la proteína madura que es necesaria para la secreción de la proteína. La escisión de la prosecuencia tendrá como resultado una proteína activa madura.

30 **[0071]** El término "secuencia señal" o "péptido señal" se refiere a cualquier secuencia de nucleótidos y/o aminoácidos que puedan participar en la secreción de las formas precursoras o maduras de la proteína. Esta definición de secuencia señal es una funcional, que pretende incluir todas las secuencias de aminoácidos codificadas por la parte N-terminal del gen de la proteína, que participan en la realización de la secreción de
35 proteína. Normalmente, aunque no de forma universal, están ligados a la parte N-terminal de una proteína o a la parte N-terminal de una proteína precursora. La secuencia señal puede ser endógena o exógena. La secuencia señal puede estar asociada normalmente a la proteína (p. ej., glucoamilasa) o puede ser de un gen que codifica otra proteína secretada.

40 **[0072]** El término forma "precursora" de una proteína o péptido se refiere a la forma madura de la proteína que tiene una prosecuencia unida de forma operativa al carbonilo o amino-terminal de la proteína. El precursor puede también presentar una secuencia "señal" unida de forma operativa al amino-terminal de la prosecuencia. El precursor puede también presentar polipéptidos adicionales que estén involucrados en la actividad traduccional posterior (p. ej., polinucleótidos escindidos de este para dejar la forma madura de una proteína o péptido).

45 **[0073]** Los términos "cepa huésped" o "célula huésped" se refieren a un huésped adecuado para un vector de expresión que comprende ADN según la presente exposición.

50 **[0074]** Los términos "derivado/a de" y "obtenido/a de" se refieren no solo a una glucoamilasa producida o que se puede producir mediante una cepa del organismo en cuestión, sino también a una glucoamilasa codificada por una secuencia de ADN aislada de tal cepa y producida en un organismo huésped que contiene dicha secuencia de ADN. Además, el término se refiere a una glucoamilasa que está codificada por una secuencia de ADN de origen sintético y/o ADNc y que presenta las características de identificación de la glucoamilasa en cuestión.

55 **[0075]** Un "derivado" dentro del alcance de esta definición generalmente conserva la característica actividad hidrolizante observada en la forma de tipo silvestre, natural o precursora hasta el punto de que el derivado es útil para fines similares como la forma precursora, natural o de tipo silvestre. Los derivados funcionales de glucoamilasas abarcan péptidos o fragmentos de péptidos producidos de forma natural, sintética o recombinante que presentan las características generales de las glucoamilasas de la presente exposición.

60 **[0076]** El término "aislado" se refiere a un material que se ha extraído de su medio natural si se da de forma natural.

5 **[0077]** Una proteína "purificada" hace referencia a una proteína que está al menos parcialmente purificada para su homogeneidad. En algunas partes de la exposición, una proteína purificada presenta más de un 10 % de pureza, preferiblemente más de un 20 % de pureza e incluso más preferiblemente más de un 30 % de pureza, según se determina mediante SDS-PAGE. Aspectos adicionales de la presente exposición incluyen la proteína en una forma altamente purificada (es decir, más de un 40 % de pureza, más de un 60 % de pureza, más de un 80 % de pureza, más de un 90 % de pureza, más de un 95 % de pureza, más de un 97 % de pureza e incluso más de un 99 % de pureza) según se determina mediante SDS-PAGE.

10 **[0078]** El término "mutagénesis combinatoria" hace referencia en el presente documento a métodos en los que se generan las bibliotecas de variantes de una secuencia inicial. En estas bibliotecas, las variantes contienen una o más mutaciones elegidas de un conjunto predefinido de mutaciones. Además, los métodos proporcionan medios para introducir mutaciones aleatorias que no eran elementos del conjunto predefinido de mutaciones. En algunas partes de la exposición, los métodos incluyen aquellos descritos en la patente estadounidense n.º 6582914. En partes alternativas de la exposición, los métodos de mutagénesis combinatoria incluyen kits disponibles comercialmente (p. ej., Quik-Change® Multisite, Stratagene, San Diego, CA).

20 **[0079]** El término "biblioteca de mutantes" hace referencia en el presente documento a una población de células que son idénticas en la mayor parte de su genoma pero incluyen diferentes homólogos de uno o más genes. Tales bibliotecas pueden utilizarse, por ejemplo, para identificar genes u operones con rasgos mejorados.

[0080] El término "contenido de sólidos secos (DS o ds por sus siglas en inglés)" hace referencia en el presente documento a los sólidos totales de una suspensión en % sobre una base de peso en seco.

25 **[0081]** El término "resultado inicial" hace referencia en el presente documento a una variante que se identificó mediante la detección de una biblioteca de mutagénesis combinatoria por consenso. En algunas partes de la exposición, los resultados iniciales presentan características de actuación mejoradas, comparadas con el gen inicial.

30 **[0082]** El término "resultado mejorado" hace referencia en el presente documento a una variante que se identificó mediante la detección de una biblioteca de mutagénesis combinatoria por consenso mejorada.

35 **[0083]** El término "propiedad diana" hace referencia en el presente documento a la propiedad del gen inicial que ha de modificarse. No se pretende que la presente exposición esté limitada por ninguna propiedad diana específica. Sin embargo, en algunas partes de la exposición, la propiedad diana es la estabilidad de un producto génico (p. ej., resistencia a la desnaturalización, proteólisis u otros factores de degradación), mientras que en otras partes de la exposición, el nivel de producción en un huésped de producción se modifica. De hecho, se contempla que cualquier propiedad de un gen inicial será útil en la presente exposición. Pueden aparecer otras definiciones de los términos en toda la memoria.

40 **[0084]** En los casos en los que se presente una gama de valores, ha de entenderse que también se indica de forma específica cada valor intermedio, hasta la décima de la unidad del límite inferior a menos que el contexto indique claramente lo contrario, entre los límites superior e inferior de cada gama. Cada gama menor entre cualquier valor indicado o valor intermedio en una gama indicada y cualquier otro valor indicado o intermedio en esa gama indicada se incluye dentro de la exposición. Los límites superior e inferior de estas gamas menores pueden incluirse o excluirse de forma independiente en la gama, y cada gama en la que cualquier límite, ambos o ninguno se incluyan en las gamas menores también se incluye dentro de la exposición, sujeta a cualquier límite excluido de forma específica en la gama indicada. En los casos en los que la gama indicada incluya uno o ambos límites, las gamas que excluyan alguno de esos límites incluidos o ambos también se incluyen en la invención.

50 **[0085]** Antes de describir las formas de realización de ejemplo con más detalle, ha de entenderse que la presente exposición no está limitada a las formas de realización específicas descritas, ya que tales pueden, obviamente, variar. Pese a que en la práctica o el estudio de la presente exposición pueden utilizarse métodos y materiales similares o equivalentes a aquellos descritos en el presente documento, los métodos y materiales adecuados se describen a continuación.

55 **[0086]** Tal y como se utilizan en el presente documento y en las reivindicaciones adjuntas, las formas singulares "un/a" y "el/la" incluyen referentes plurales a menos que el contexto indique claramente lo contrario. Por lo tanto, por ejemplo, la referencia a "un gen" incluye una pluralidad de tales agentes candidatos y la referencia a "la célula" incluye referencia a una o más células y equivalentes de estas conocidos por los expertos en la técnica, etc.

60 **[0087]** Las publicaciones detalladas en el presente documento se presentan únicamente para su exposición anterior a la fecha de presentación de la presente solicitud. Nada de lo que aparece en el presente documento ha

de entenderse como una admisión de que la presente invención no tiene derecho a preceder a tal publicación en virtud de la invención anterior.

2. Glucoamilasas precursoras:

5

[0088] En algunas formas de realización, la presente exposición proporciona una variante de glucoamilasa. La variante de glucoamilasa es una variante de una glucoamilasa precursora, que puede comprender tanto un dominio catalítico como un dominio de unión a almidón. En algunas partes de la exposición, la glucoamilasa precursora comprende un dominio catalítico que presenta una secuencia de aminoácidos como se ilustra en SEQ ID NO: 1, 2, 3, 5, 6, 7, 8 o 9 o que presenta una secuencia de aminoácidos que muestra al menos un 80 % de identidad de secuencia con una o más de las secuencias de aminoácidos ilustradas en SEQ ID NO: 1, 2, 3, 5, 6, 7, 8 o 9. En otras partes de la exposición, la glucoamilasa precursora comprende un dominio catalítico codificado por una secuencia de ADN que hibrida bajo condiciones astringentes, medias o altas con un ADN que codifica el dominio catalítico de una glucoamilasa que presenta una de las secuencias de aminoácidos de SEQ ID NO: 1, 2 o 3.

15

[0089] En algunas partes de la exposición, la glucoamilasa precursora comprende un dominio de unión a almidón que presenta una secuencia de aminoácidos según se ilustra en SEQ ID NO: 1, 2, 161, 162, 163, 164, 165, 166 o 167 o que presenta una secuencia de aminoácidos que muestra al menos un 80 % de identidad de secuencia con una o más de las secuencias de aminoácidos ilustradas en SEQ ID NO: 1, 2, 161, 162, 163, 164, 165, 166, o 167.

20

En otras partes de la exposición, la glucoamilasa precursora comprende un dominio de unión a almidón codificado mediante una secuencia de ADN que hibrida bajo condiciones astringentes, altas o medias con un ADN que codifica el dominio de unión a almidón de una glucoamilasa que presenta una de las secuencias de aminoácidos de SEQ ID NO: 1 o 2.

25

[0090] La estructura predicha y las secuencias conocidas de glucoamilasas se conservan entre especies fúngicas (Coutinho *et al.*, 1994, Protein Eng., 7:393-400 y Coutinho *et al.*, 1994, Protein Eng., 7: 749-760). En algunas partes de la exposición, la glucoamilasa precursora es una glucoamilasa fúngica filamentosa. En algunas formas de realización, la glucoamilasa precursora se obtiene de una cepa de *Trichoderma* (p. ej., *T. reesei*, *T. longibrachiatum*, *T. strictipilis*, *T. asperellum*, *T. konilangbra* y *T. hazianum*), una cepa de *Aspergillus* (p. ej., *A. niger*, *A. nidulans*, *A. kawachi*, *A. awamori* y *A. oryzae*), una cepa de *Talaromyces* (p. ej., *T. emersonii*, *T. thermophilus* y *T. dupontii*), una cepa de *Hypocrea* (p. ej., *H. gelatinosa*, *H. orientalis*, *H. vinosa* y *H. citrina*), una cepa de *Fusarium* (p. ej., *F. oxysporum*, *F. roseum*, y *F. venenatum*), una cepa de *Neurospora* (p. ej., *N. crassa*) y una cepa de *Humicola* (p. ej., *H. grisea*, *H. insolens* y *H. lanuginosa*), una cepa de *Penicillium* (p. ej., *P. notatum* o *P. chrysogenum*) o una cepa de *Saccharomycopsis* (p. ej., *S. fibuligera*).

35

[0091] En algunas partes de la exposición, la glucoamilasa precursora puede ser una glucoamilasa bacteriana. Por ejemplo, el polipéptido puede obtenerse de una cepa bacteriana gram positiva tal como *Bacillus* (p. ej., *B. alkalophilus*, *B. amyloliquefaciens*, *B. lentus*, *B. licheniformis*, *B. stearothermophilus*, *B. subtilis* y *B. thuringiensis*) o una cepa de *Streptomyces* (p. ej., *S. lividans*).

40

[0092] En algunas partes de la exposición, la glucoamilasa precursora comprenderá un dominio catalítico que presenta al menos un 80 % de identidad de secuencia, al menos un 85 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 97 % de identidad de secuencia y también al menos un 98 % de identidad de secuencia con el dominio catalítico de la secuencia de aminoácidos de TrGA de SEQ ID NO: 3.

45

[0093] En otras partes de la exposición, la glucoamilasa precursora comprenderá un dominio catalítico que presenta al menos un 90 % de identidad de secuencia, al menos un 93 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 96 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia y también al menos un 99 % de identidad de secuencia con el dominio catalítico de la glucoamilasa precursora *Aspergillus* de SEQ ID NO: 5 o SEQ ID NO: 6.

50

[0094] En otras partes de la exposición, la glucoamilasa precursora comprenderá un dominio catalítico que presenta al menos un 90 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 97 % de identidad de secuencia y también al menos un 99 % de identidad de secuencia con el dominio catalítico de la glucoamilasa precursora *Humicola grisea* (HgGA) de SEQ ID NO: 8.

55

[0095] En algunas formas de realización, la glucoamilasa precursora comprenderá un dominio de unión a almidón con al menos un 80 % de identidad de secuencia, al menos un 85 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 97 % de identidad de secuencia y también al menos un 98 % de identidad de secuencia con el dominio de unión a almidón de la secuencia de aminoácidos de TrGA de SEQ ID NO: 1, o 2.

60

[0096] En otras partes de la exposición, la glucoamilasa precursora comprenderá un dominio de unión a almidón con al menos un 90% de identidad de secuencia, al menos un 95% de identidad de secuencia, al menos un 97 % de identidad de secuencia y también al menos un 99 % de identidad de secuencia con el dominio catalítico de la glucoamilasa *Humicola grisea* (HgGA) de SEQ ID NO: 162.

5

[0097] En otras partes de la exposición, la glucoamilasa precursora comprenderá un dominio de unión a almidón que presenta al menos un 90% de identidad de secuencia, al menos un 95% de identidad de secuencia, al menos un 97 % de identidad de secuencia y también al menos un 99 % de identidad de secuencia con el dominio catalítico de la glucoamilasa *Thielavia terrestris* (TtGA) de SEQ ID NO: 163.

10

[0098] En otras partes de la exposición, la glucoamilasa precursora comprenderá un dominio de unión a almidón que presenta al menos un 90% de identidad de secuencia, al menos un 95% de identidad de secuencia, al menos un 97 % de identidad de secuencia y también al menos un 99 % de identidad de secuencia con el dominio catalítico de la glucoamilasa *Thermomyces lanuginosus* (ThGA) de SEQ ID NO: 164.

15

[0099] En otras partes de la exposición, la glucoamilasa precursora comprenderá un dominio de unión a almidón que presenta al menos un 90% de identidad de secuencia, al menos un 95% de identidad de secuencia, al menos un 97 % de identidad de secuencia y también al menos un 99 % de identidad de secuencia con el dominio catalítico de la glucoamilasa *Talaromyces emersonii* (TeGA) de SEQ ID NO: 165.

20

[0100] En otras partes de la exposición, la glucoamilasa precursora comprenderá un dominio de unión a almidón que presenta al menos un 90% de identidad de secuencia, al menos un 93 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 96 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia y también al menos un 99 % de identidad de secuencia con el dominio de unión a almidón de la glucoamilasa precursora *Aspergillus* de SEQ ID NO: 166 o 167.

25

[0101] En algunas formas de realización, la glucoamilasa precursora presentará al menos un 80 % de identidad de secuencia, al menos un 85 % de identidad de secuencia, al menos un 88 % de identidad de secuencia, al menos un 90 % de identidad de secuencia, al menos un 93 % de identidad de secuencia, al menos un 95 % de identidad de secuencia, al menos un 96 % de identidad de secuencia, al menos un 97 % de identidad de secuencia, al menos un 98 % de identidad de secuencia y también al menos un 99 % de identidad de secuencia con la secuencia de aminoácidos TrGA de SEQ ID NO: 2.

30

[0102] En partes adicionales de la exposición, se obtendrá un homólogo de glucoamilasa *Trichoderma* a partir de una cepa de *Trichoderma* o *Hypocrea*. Algunos homólogos típicos de glucoamilasa *Trichoderma* se describen en la patente estadounidense n.º 7413887 y se hace referencia de forma específica a las secuencias de aminoácidos establecidas en SEQ ID NO: 17-22 y 43-47 de dicha referencia.

35

[0103] En algunas formas de realización, la glucoamilasa precursora es TrGA que comprende la secuencia de aminoácidos de SEQ ID NO: 2 o un homólogo de la glucoamilasa *Trichoderma* con al menos un 80 %, al menos un 85 %, al menos un 88 %, al menos un 90 %, al menos un 93 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 %, al menos un 99 % de identidad de secuencia con la secuencia de TrGA (SEQ ID NO: 2).

40

[0104] Una glucoamilasa precursora puede aislarse y/o identificarse mediante técnicas de ADN recombinante convencionales. Se puede utilizar cualquier técnica convencional conocida por los expertos. Por ejemplo, se pueden utilizar sondas y/o cebadores específicos para áreas conservadas de la glucoamilasa con el fin de identificar homólogos en células bacterianas o fúngicas (el dominio catalítico, el centro activo, etc.). De forma alternativa, se puede utilizar una PCR con cebadores degenerados para identificar homólogos en células bacterianas o fúngicas. En algunos casos, se pueden analizar secuencias conocidas, tales como en una base de datos, para determinar la identidad estructural y/o de secuencia con una de las glucoamilasas conocidas, entre las que se incluyen SEQ ID NO: 2 o dominios de unión a almidón conocidos, entre los que se incluyen SEQ ID NO: 161. También se pueden utilizar análisis funcionales para identificar la actividad de glucoamilasa en una célula bacteriana o fúngica. Las proteínas con actividad de glucoamilasa se pueden aislar y cambiar el orden con el fin de aislar la secuencia de ADN correspondiente. El experto en la técnica conoce tales métodos.

45

50

55

3. Homología estructural de la glucoamilasa:

[0105] El dogma central de la biología molecular es que la secuencia de ADN que codifica un gen para una enzima específica determina la secuencia de aminoácidos de la proteína, secuencia que a su vez determina el plegamiento tridimensional de la enzima. plegamiento reúne los residuos dispares que crean un centro catalítico y una superficie de unión al sustrato y tiene como resultado una alta especificidad y actividad de las enzimas en cuestión.

60

[0106] Las glucoamilasas consisten en hasta tres dominios estructurales distintos, un dominio catalítico de aproximadamente 450 residuos que está conservado de forma estructural en todas las glucoamilasas, generalmente seguido de una región de enlace consistente en entre 30 y 80 residuos que están conectados a un dominio de unión a almidón de aproximadamente 100 residuos. La estructura de la glucoamilasa *Trichoderma reesei* con las tres regiones intactas se determinó con la resolución Angstrom 1.8 en el presente documento (véase el ejemplo 13). Al utilizar las coordenadas la estructura se alineó con las coordenadas del dominio catalítico de la cepa *Aspergillus awamori* X100 que se determinó previamente (Aleshin, A.E., Hoffman, C., Firsov, L.M., and Honzatko, R.B. 1994 Refined crystal structures of glucoamylase from *Aspergillus awamori* var. X100. J Mol Biol 238: 575-591.). La estructura de cristal *Aspergillus awamori* incluía únicamente el dominio catalítico. Como puede observarse en las figuras 12 y 13, la estructura de los dominios catalíticos se superpone muy estrechamente y permite la identificación de los residuos equivalentes según esta superposición estructural. Se cree que todas las glucoamilasas comparten la estructura básica representada en las figuras 12 y 13.

[0107] La figura 12 es una comparación de las estructuras tridimensionales de la glucoamilasa *Trichoderma* (negra) de SEQ ID NO: 1 (véase la figura 1 para la secuencia de aminoácidos) y de *Aspergillus awamori* (gris) vistas de lado. En esta vista, puede verse la relación entre el dominio catalítico y la región de enlace y el dominio de unión a almidón.

[0108] La figura 13 es una comparación de las estructuras tridimensionales de la glucoamilasa *Trichoderma* (negra) y de *Aspergillus awamori* (gris) vistas desde la parte superior. Las glucoamilasas mostradas en el presente documento y de hecho glucoamilasas conocidas a día de hoy comparten esta homología estructural. La conservación de la estructura se corresponde con la conservación de la actividad y un mecanismo conservado de acción para todas las glucoamilasas. Teniendo en cuenta esta alta homología, los cambios que resultan de las variantes específicas del lugar de la glucoamilasa *Trichoderma* que tienen como resultado una función modificada tendrían consecuencias estructurales y, por tanto, funcionales similares en otras glucoamilasas. Por lo tanto, la información dada a conocer sobre qué variantes presentan beneficios aconsejables puede aplicarse a otras glucoamilasas.

[0109] Se produjo una estructura de cristal adicional utilizando las coordenadas para el dominio de unión a almidón (SBD). El SBD para TrGA se alineó con el SBD para *A. niger*. Como se muestra en la figura 13, la estructura de los SBD de *A. niger* y TrGA se superpone muy estrechamente. Se cree que, aunque todos los dominios de unión a almidón comparten al menos parte de la estructura básica representada en la figura 13, algunos SBD se parecen más estructuralmente que otros. Por ejemplo, el SBD de TrGA puede clasificarse como dentro de la familia del módulo de unión de carbohidrato 20 dentro de la base de datos CAZY (cazy.org). La base de datos CAZY describe las familias de módulos de unión de carbohidrato catalíticos y estructuralmente relacionados (o dominios funcionales) de enzimas que degradan, modifican o crean enlaces glicosídicos. Dada una homología estructural alta, las variantes específicas del sitio del SBD de TrGA que tienen como resultado una función modificada también tendrían consecuencias estructurales y, por tanto, funcionales similares en otras glucoamilasas que tengan SBD con estructura similar a la del SBD de TrGA, especialmente aquellos clasificados dentro de la familia de módulo 20 de unión de carbohidratos. Por lo tanto, la información dada a conocer sobre qué variantes presentan beneficios aconsejables puede aplicarse a otros SBD que presentan similitud estructural.

[0110] Por lo tanto, los números de posición de aminoácidos detallados en el presente documento hacen referencia a aquellos asignados a la secuencia de glucoamilasa *Trichoderma reesei* madura presentada en la figura 1. Sin embargo, la presente exposición no está limitada a las variantes de la glucoamilasa *Trichoderma*, sino que se extiende a glucoamilasas que contienen residuos de aminoácidos en posiciones que son "equivalentes" a los residuos identificados específicos en la glucoamilasa *Trichoderma reesei* (SEQ ID NO: 2). En una forma de realización de la presente exposición, la glucoamilasa precursora es la glucoamilasa *Taleromyces* y las sustituciones se realizan en las posiciones de residuo de aminoácidos equivalentes en la glucoamilasa *Taleromyces* como las descritas en el presente documento. En otras partes de la exposición, la glucoamilasa precursora es una de las enumeradas en la tabla 1. En partes adicionales de la exposición, la glucoamilasa precursora es una glucoamilasa *Penicillium*, tal como *Penicillium chrysogenum*.

[0111] La identidad estructural determina si los residuos de aminoácidos son equivalentes. La identidad estructural es un equivalente topológico directo cuando las dos estructuras (estructuras de aminoácidos y tridimensionales) se alinean. Una posición de residuos (aminoácidos) de una glucoamilasa es equivalente a un residuo de glucoamilasa *T. reesei* si es tanto homóloga (es decir, se corresponde en posición ya sea en estructura primaria o terciaria) como análoga con un residuo específico o parte de ese residuo en la glucoamilasa *T. reesei* (que tenga la misma capacidad funcional o similar para combinar, reaccionar o interactuar de forma química).

[0112] Con el fin de establecer la identidad con la estructura primaria, la secuencia de aminoácidos de una glucoamilasa puede compararse directamente con la secuencia primaria de la glucoamilasa *Trichoderma reesei* y

específicamente con un conjunto de residuos que se sabe son invariables en glucoamilasas por cuya secuencia son conocidos. Por ejemplo, las figuras 4A y 4B en el presente documento muestran los residuos conservados entre glucoamilasas. Las figuras 4C y 4D muestran una alineación de dominios de unión a almidón a partir de diferentes glucoamilasas. Una vez se han alineado los residuos conservados, y permitiendo las inserciones y 5 deleciones necesarias con el fin de mantener la alineación (es decir, evitando la eliminación de residuos conservados a través de la deleción e inserción arbitrarias), se definen los residuos equivalentes a aminoácidos específicos en la secuencia primaria de glucoamilasa *Trichoderma reesei*. La alineación de residuos conservados normalmente debería conservar el 100 % de tales residuos. Sin embargo, la alineación de más del 75 % o tan solo un 50 % de residuos conservados también es adecuada para definir residuos equivalentes. Además, la identidad 10 estructural puede utilizarse junto con la identidad de secuencia para identificar residuos equivalentes.

[0113] Por ejemplo, en las figuras 4A y 4B, se alinean los dominios catalíticos de glucoamilasas de seis organismos para proporcionar la máxima cantidad de homología entre secuencias de aminoácidos. Una comparación de estas 15 secuencias muestra que existe un número de residuos conservados contenidos en cada secuencia como se indica mediante un asterisco. Por consiguiente, estos residuos conservados pueden usarse para definir los correspondientes residuos de aminoácidos equivalentes de la glucoamilasa *Trichoderma reesei* en otras glucoamilasas como la glucoamilasa de *Aspergillus niger*. De forma similar, las figuras 4C y 4D muestran los dominios de unión a almidón de glucoamilasas de siete organismos alineados para identificar residuos equivalentes.

[0114] La identidad estructural supone la identificación de residuos equivalentes entre las dos estructuras. Los "residuos equivalentes" pueden definirse mediante la determinación de homología en el nivel de la estructura terciaria (identidad estructural) para una enzima cuya estructura terciaria ha sido determinada mediante técnicas 25 de cristalografía de rayos X. Los residuos equivalentes se definen como aquellos para cuyas coordenadas atómicas de dos o más de los átomos de la cadena principal de un residuo de aminoácido específico de la glucoamilasa *Trichoderma reesei* (N en N, CA en CA, C en C y O en O) están dentro de 0,13 nm y opcionalmente 0,1 nm tras la alineación. La alineación se consigue después de que el mejor modelo se haya orientado y colocado para proporcionar la máxima superposición de coordenadas atómicas de átomos de proteína no hidrógenos de la glucoamilasa en cuestión con la glucoamilasa *Trichoderma reesei*. El mejor modelo es el modelo cristalográfico 30 que proporciona el factor R más bajo para los datos de difracción experimental con la resolución más alta disponible.

$$\text{factor } R = \frac{\sum_h |F_o(h)| - |F_c(h)|}{\sum_h |F_o(h)|}$$

[0115] Los residuos equivalentes que son análogos funcionalmente a un residuo específico de la glucoamilasa 35 *Trichoderma reesei* se definen como aquellos aminoácidos de la enzima que pueden adoptar una conformación de tal forma que bien cambian, modifican o contribuyen a la estructura proteica, unión de sustrato o catálisis de forma definida y atribuida a un residuo específico de la glucoamilasa *Trichoderma reesei*. Además, se trata de aquellos residuos de la enzima (para los que se ha obtenido una estructura terciaria mediante cristalografía de rayos X) que ocupan una posición análoga hasta el punto de que, aunque los átomos de la cadena principal del 40 residuo determinado puedan no cumplir los criterios de equivalencia sobre la base de ocupar una posición homóloga, las coordenadas atómicas de al menos dos de los átomos de la cadena lateral del residuo quedan con 0,13 nm de los correspondientes átomos de la cadena lateral de la glucoamilasa *Trichoderma reesei*. Las coordenadas de la estructura tridimensional de la glucoamilasa *Trichoderma reesei* están establecidas en la tabla 15 y pueden utilizarse como se detalla anteriormente con el fin de determinar residuos equivalentes en el nivel de 45 estructura terciaria.

[0116] Algunos de los residuos identificados para la sustitución son residuos conservados mientras que otros no lo son. En el caso de residuos que no se conservan, la sustitución de uno o más aminoácidos está limitada a las 50 sustituciones que producen una variante que presenta una secuencia de aminoácidos que no se corresponde con una hallada en la naturaleza. En el caso de residuos conservados, tales sustituciones no deberían tener como resultado una secuencia que se da de forma natural.

4. Variantes

[0117] Las variantes según la presente exposición incluyen al menos una sustitución, deleción o inserción en la 55 secuencia de aminoácidos de una glucoamilasa precursora que hace que la variante sea diferente de la glucoamilasa precursora por la secuencia. En algunas partes de la exposición, las variantes de la exposición presentarán al menos un 20 %, al menos un 40 %, al menos un 60 %, al menos un 80 %, al menos un 85 %, al

menos un 90 %, al menos un 95 % o al menos un 100 % de la actividad de glucoamilasa de la actividad de TrGA de SEQ ID NO: 2.

5 **[0118]** En algunas partes de la exposición, las variantes de acuerdo con la exposición comprenderán una sustitución, deleción o inserción en al menos una posición de aminoácidos de la TrGA precursora (SEQ ID NO: 2), o en una posición equivalente en la secuencia de otra glucoamilasa precursora con al menos un 80 % de identidad de secuencia con la secuencia de TrGA, incluyendo sin carácter limitativo al menos un 90 %, al menos un 93 %, al menos un 95 %, al menos un 97 % y al menos un 99 % de identidad de secuencia.

10 **[0119]** En otras partes de la exposición, la variante de acuerdo con la exposición comprenderá una sustitución, deleción o inserción en al menos una posición de aminoácidos de un fragmento de la TrGA precursora, donde el fragmento comprende el dominio catalítico de la secuencia de TrGA (SEQ ID NO: 3), o en una posición equivalente en un fragmento que comprende el dominio catalítico de una glucoamilasa precursora con al menos un 80 % de identidad de secuencia con el fragmento de la secuencia de TrGA, incluyendo sin carácter limitativo al menos un 90 %, al menos un 95 %, al menos un 97 % y al menos un 99 %. En algunas partes de la exposición, el fragmento comprenderá al menos 400, 425, 450 y/o 500 residuos de aminoácidos.

15 **[0120]** En otras partes de la exposición, la variante de acuerdo con la exposición comprenderá una sustitución, deleción o inserción en al menos una posición de aminoácidos de un fragmento de la TrGA precursora, donde el fragmento comprende el dominio de unión a almidón de la secuencia de TrGA (SEQ ID NO: 161), o en una posición equivalente en un fragmento que comprende el dominio de unión a almidón de una glucoamilasa precursora con al menos un 80 % de identidad de secuencia con el fragmento de la secuencia de TrGA, incluyendo sin carácter limitativo al menos un 90 %, al menos un 95 %, al menos un 97 % y al menos un 99 %. En algunas formas de realización, el fragmento comprenderá al menos 40, 50, 60, 70, 80, 90, 100 y/o 109 residuos de aminoácidos del dominio de unión a almidón de TrGA (SEQ ID NO: 161).

20 **[0121]** En algunas partes de la exposición, cuando la glucoamilasa precursora incluye un dominio catalítico, una región de enlace y un dominio de unión a almidón, la variante comprenderá una sustitución, deleción o inserción en al menos una posición de aminoácido de un fragmento que comprende parte de la región de enlace. En algunas partes de la exposición, la variante comprenderá una sustitución, deleción o inserción en la secuencia de aminoácidos de un fragmento de la secuencia de TrGA (SEQ ID NO: 2).

25 **[0122]** La identidad estructural con respecto a una sustitución de aminoácidos indica que la sustitución tiene lugar en la posición de aminoácidos equivalente en la glucoamilasa homóloga o glucoamilasa precursora. El término posición equivalente se refiere a una posición que es común a dos secuencias precursoras que se basa en una alineación de la secuencia de aminoácidos de la glucoamilasa precursora en cuestión así como una alineación de la estructura tridimensional de la glucoamilasa precursora en cuestión con la secuencia de aminoácidos de la glucoamilasa de referencia TrGA y la secuencia tridimensional. Por ejemplo, en referencia a la figura 5, la posición 24 en TrGA (SEQ ID NO: 2 o 3) es D24 y la posición equivalente para *Aspergillus niger* (SEQ ID NO: 6) es la posición D25 y la posición equivalente para *Aspergillus oryzae* (SEQ ID NO: 7) es la posición D26. Véanse las figuras 12 y 13 para una alineación de ejemplo de la secuencia tridimensional.

30 **[0123]** En algunas partes de la exposición, la glucoamilasa incluirá al menos una sustitución en la secuencia de aminoácidos de una precursora. En partes adicionales de la exposición, la variante puede presentar más de una sustitución (p. ej., dos, tres o cuatro sustituciones).

35 **[0124]** En algunas partes de la exposición, una variante de glucoamilasa comprende una sustitución, deleción o inserción y normalmente un sustitución en al menos una posición de aminoácidos en una posición correspondiente a las regiones de aminoácidos no conservados como se ilustra en la figura 5 (p. ej., posiciones de aminoácidos correspondientes a aquellas posiciones que no están indicadas con "*" en la figura 5).

40 **[0125]** Aunque las variantes pueden estar en cualquier posición en la secuencia de proteína madura (SEQ ID NO: 2), en una parte de la exposición, una variante de glucoamilasa comprende una o más sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en la SEQ ID NO: 2: 4, 5, 12, 24, 29, 43, 44, 45, 46, 47, 49, 51, 61, 70, 73, 75, 76, 94, 100, 108, 114, 116, 119, 122, 124, 125, 137, 143, 146, 148, 169, 171, 172, 175, 178, 180, 181, 208, 211, 228, 242, 243, 245, 292, 294, 297, 309, 310, 313, 314, 315, 316, 317, 321, 340, 341, 350, 353, 356, 363, 368, 369, 375, 376, 395, 398, 401, 408, 409, 412, 415, 417, 418, 421, 430, 431, 433, 436, 451, 503, 511, 535, 539 o 563 o en una posición equivalente en una glucoamilasa precursora. En algunas partes de la exposición, la glucoamilasa precursora presentará al menos un 90 %, al menos un 95 %, al menos un 96 %, al menos un 97 %, al menos un 98 % y al menos un 99 % de identidad con SEQ ID NO: 2. En otras formas de realización, la glucoamilasa precursora será un homólogo de glucoamilasa *Trichoderma*.

[0126] En algunas partes de la exposición, la variante de glucoamilasa comprende una o más sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en la SEQ ID NO: 2:

5 D4, F5, I12, D24, F29, I43, D44, P45, D46, Y47, Y49, W51, N61, Y70, G73, Q75, R76, P94, D100, K108, K114, F116, N119, R122, Q124, R125, G137, N146, Q148, Y169, N171, Q172, F175, W178, E180, V181, Q208, S211, W228, N242, E243, R245, I292, G294, K297, R309, Y310, D313, V314, Y315, Y316, N317, W321, K340, K341, T350, Q356, T363, S368, S369, N376, Y395, A398, S401, R408, N409, T412, L417, H418, W421, T430, A431, R433, I436, S451, E503, Q511, A535, A539, o N563; o una posición equivalente en una glucoamilasa precursora (p. ej., un homólogo de glucoamilasa *Trichoderma*).

10 **[0127]** En otras partes de la exposición, la variante de una precursora de glucoamilasa comprende una o más sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO:2: 4, 5, 24, 29, 43, 44, 49, 61, 70, 73, 75, 76, 100, 108, 119, 124, 137, 146, 148, 169, 171, 172, 175, 178, 181, 208, 211, 243, 292, 294, 297, 314, 316, 317, 340, 341, 350, 356, 363, 368, 369, 376, 395, 401, 409, 412, 417, 430, 431, 433, 436, 451, 503, 511, 535, 539 o 563 o una posición equivalente en una glucoamilasa precursora (p. ej., un homólogo de glucoamilasa *Trichoderma*).

20 **[0128]** En partes adicionales de la exposición, la variante de una precursora de glucoamilasa comprende al menos una de las siguientes sustituciones en las siguientes posiciones en una secuencia de aminoácidos establecida en SEQ ID NO: 2: D4L/E/R/S/C/A/Q/W, F5C/M/N/R/S/TN/W, I12L/R, D24E/L/Y/T, F29L/I/D/C/S/V/W, I43F/R/D/Y/S/Q, D44E/H/K/S/N/Y/F/R/C, Y47W, Y49N, N61D/I/L/Q/V/W, Q70R/K/M/P/G/L/F, G73F/C/L/W, Q75R/K/A, R76L/M/K/T/P, P94L, D100W/I/Q/M/P/A/N, N119P/T/Y/D/E, N146S/G/C/H/E/D/T/W/L/F/M, Q148V/Y/H/A/C/D/G/M/R/S/T, Y169D/F, Q172C/A/D/R/E/F/H/V/L/M/N/S/T/V, F175H/A/G/R/S/T/C/W/Y, W178A/C/D/E/F/G/H/K/N/R/S/T/V/Y, E180A/C/G/H/I/L/N/P/Q/R/S/T/V/Y, V181E/C/D/G/H/I/P/T/Y/S/L/K/F/A, Q208L/A/C/E/N/F/H/T, S211C/R/E/A/Y/W/M/H/L/I/R/Q/T, E243S/R/N/M/Y/A/L, R245A/E/M/I/P/V, I292D/H/P/R/T/N/V/F/L, G294C/D/E/T/Q/I/A, K297F/L/P/T/M/D/N/Q/A/Y/H/S/R/W, R309A/C/G/H/I/N/P/Q/S/T/W/Y/L, Y310E/G/L/P/S/W/R/Q, D313Q, V314A/R/N/D/C/E/Q/G/H/I/L/K/M/F/P/S/T/W/Y, Y315F, Y316Q/R, N317T/H, K340D/T, K341F/D/P/N/G/S, T350S/E/A/N, Q356H/D/E, T363L/R/C/H/W, S368W/D/F/L, S369F, N376Q/T/H/S/V, Y395Q/R/S, A398S/I/T, S401C/V, R408S, N409W/T/K, T412A/H/K/G, L417A/D/E/F/G/I/K/Q/R/S/T/V/W/Y, T430A/E/F/G/H/I/K/M/N/Q/R/V, A431C/E/H/I/L/M/Q/R/S/W/Y, R433H/Q, 25 I436A/T, S451M/T/H, E503A/C/D/H/S/V/W, Q511C/G/H/I/K/T/V, A535E/F/G/K/LN/P/R/S/TN/W/Y, A539E/H/M/R/S/W, o N563/A/C/E/I/K/L/Q/T/V; o una sustitución en una posición equivalente en un homólogo de glucoamilasa precursora.

35 **[0129]** En algunas partes de la exposición, la variante de glucoamilasa comprende al menos una sustitución en una posición correspondiente a la posición de residuo de aminoácido establecida en SEQ ID NO: 2: 5, 24, 43, 44, 49, 61, 70, 73, 75, 76, 94, 119, 146, 148, 172, 175, 178, 180, 181, 208, 211, 245, 294, 353, 315, 375, 409, 309, 314, 369, 412, 417, 430, 431, 503, 511, 535, 539 o 563; o una posición equivalente en una glucoamilasa precursora homóloga.

40 **[0130]** En algunas partes representativas de la exposición, la variante de glucoamilasa comprende al menos una sustitución elegida del grupo consistente en F5W, D24E, I43R, I43Y, I43Q, I43S, I43F, D44C, D44R, Y47W, Y49N, N61I, Q70K, G73F, Q75R, R76L, P94L, N119P/T/Y/D, N146S/D/T/E/M/L, Q148V N171D, Q172C/D/R/E/F/V/L/T, F175R/W/Y, W178K/N/Y, E180H/N/V/R, V181E/F/G/I/H, Q208A/T/N, S211H/M/L/R, R245E, R245M, G294C, R309W, V314F/G/H/K/P/R/Y, Y315F, S369F, T412K, L417R, L417V, T430A, T430M, A431L, A431Q, E503A, 45 E503V, Q511H, A535R, A539R, N563I y N563K correspondientes a la posición establecida en SEQ ID NO: 2 o una posición equivalente en una glucoamilasa precursora homóloga.

50 **[0131]** En partes concretas adicionales de la exposición, la variante de glucoamilasa comprende al menos una sustitución de un residuo de aminoácido elegido de entre las posiciones correspondientes a la posición 5, 43, 44, 61, 73, 75, 76, 94, 108, 119, 124, 146, 148, 171, 172, 175, 178, 180, 181, 208, 211, 294, 297, 314, 316, 412, 417, 430, 431, 503, 511, 535, 539 o 563 de SEQ ID NO: 2 o una posición equivalente en un homólogo de glucoamilasa *Trichoderma*. En algunas partes de la exposición, la sustitución se encuentra en una posición correspondiente a la posición número 43, 44, 61, 73, 148, 172, 175, 178, 180, 208, 211, 294, 297, 314, 412, 417, 430, 431, 503, 511, 535, 539 o 563 de SEQ ID NO: 2 o una posición equivalente en un homólogo de glucoamilasa *Trichoderma*.

55 **[0132]** En algunas partes representativas de la exposición, la sustitución está en una posición correspondiente a la posición número 43, 44, 61, 73, 108, 124, 171, 172, 208, 211, 294, 314, 316, 417, 430, 431, 503, 511, 535, 539, o 563 de SEQ ID NO: 2, o una glucoamilasa precursora homóloga (p. ej., homólogo de glucoamilasa *Trichoderma*).

60 **[0133]** En algunas partes de la exposición, las variantes de glucoamilasa comprenden múltiples sustituciones. Algunas de las sustituciones múltiples incluirán una sustitución en una o más de las posiciones equivalentes y que incluyan las posiciones 24, 43, 44, 108, 124, 171, 175, 181, 208, 243, 292, 294, 297, 310, 314, 363, 417, 430, 431,

ES 2 795 989 T3

503, 511, 535, 539, o 563 de SEQ ID NO: 2. Entre algunas de las típicas sustituciones múltiples se incluirán una o más de las posiciones equivalentes y correspondientes a las posiciones 43, 44, 61, 73, 08, 124, 171, 208, 211, 294, 314 417, 430, 431, 503, 511, 535, 539, o 563 de SEQ ID NO: 2.

- 5 **[0134]** Algunos ejemplos de variantes con sustituciones múltiples incluyen sustituciones en las posiciones:
D24/I43/D44/F175/V181/V314/T363;
D24/Q208/I292/G294/K297/Y310;
V181/E243/I292/k297/N317/Y395;
D24/V181/Q208/G294/T363/N376/N409;
- 10 D24/V181/I292/G294/E243/N409;
I43R/E24 3/I292/G294/K297;
I43/D44/N61/L417/E503/Q511/A539;
I43/D44/L417/E503/Q511/A539;
I43/N61/L417/T430/Q511/A539;
- 15 I43/N61/L417/E503/Q511/A539;
I43/N61/T430/A431/Q511/A539;
I43/N61/T430/Q511;
I43/N61/T430/Q511/A539;
I43/N61/Q511;
- 20 I43/N61/Q511/A539;
I43/G73/T430;
I43/L417/E503/Q511/A539;
I43/L417/Q511;
I43/L417/T430/A431/Q511/A539;
- 25 I43/L417/T430/Q511;
I43/L417/T430/Q511/A539;
I43/L417/E503/A539;
I43/L417/E503/Q511/A539;
I43/T430;
- 30 I43/T430/A431/E503/Q511;
I43/T430/A431/Q511;
I43/T430/A431/Q511/A539;
I43/T430/E503/Q511;
I43/T430/Q511;
- 35 I43Q/T430/Q511/A539;
I43/A431/Q511;
I43/T430/E503/Q511/N563;
I43/T430/E503/A535/N563;
I43/E503/Q511/A539;
- 40 I43/Q511/A539;
D44/G73/L417/N563;
D44/G73/E503/Q511;
D44/G73/N563;
D44/L417/N563;
- 45 D44/T430/Q511/A535;
D44/E503/Q511/N563;
G73/T430/E503/Q511;

- G73/T430/Q511;
 G294/L417/A431;
 G294/L417/A431;
 G294/L417/A431/Q511;
 5 L417/T430/A431/Q511/A535/A539/N563;
 L417/A431/Q511;
 L417/T430/Q511/A535/N563;
 L417/T430/Q511/A539/N563 y
 E503/N563;
- 10 de SEQ ID NO: 2 o posiciones equivalentes en glucoamilasas precursoras y especialmente homólogos de glucoamilasa *Trichoderma*.

[0135] Algunas variantes con múltiples sustituciones pueden incluir las sustituciones en las posiciones:

- Y47F/W,Y315F/W;
- 15 D24E,L/I43F,R/D44H,N/F175H/V181K,L/V314D,H,K/T363R;
 D24L,W,Y/Q208F/I292F,N,V/G294A,I,Q/K297A/Y310F,Q,R;
 V181F,K,L/E243A,N,M,R,Y/I292F,L,N,V/K297A,D,H,M,N,Q/N317H/Y395 Q,R;
 D24E,L,Y/V181F,K,L/Q208C,F/G294A,I,Q/T363R/N376Q/N409K,W;
 D24E,L,YN181F,K,L/I292F,L,N,V/G294A,I,Q/E243A,M,N,R,Y/N409K, W;
- 20 I43R/E243A,M,N,R,Y/I292F,L,N,V/G294A/K297A,D,H,M,N,Q,S,R,W,Y;
 I43Q/D44C/N61I/L417V/E503A/Q511H/A539R;
 I43Q/D44C/L417V/E503A/Q511H/A539R;
 I43Q/N61I/L417V/E503A/Q511H/A539R;
 I43Q/N61I/L417V/T430M/Q511H/A539R;
- 25 I43R/N61I/L417R,V/E503A/Q511H/A539R;
 I43Q/N61I/T430A/A431L/Q511H/A539R;
 I43Q/N61I/T430A/Q511H;
 I43Q/N61I/T430A/Q511H/A539R;
 I43Q/N61I/T430M/Q511H/A539RI43Q/N61I/Q511H;
- 30 I43Q/N61I/Q511H/A539R;
 I43R/G73F/T430A;
 I43Q/L417V/T430A/A431L/Q511H/A539R;
 I43Q/L417V/T430A/Q511H;
 I43Q/L417V/T430A/Q511H/A539R;
- 35 I43R/L417R/E503A/A539R;
 I43R,Q/L417V/E503A/Q511H/A539R;
 I43Q/L417V/Q511H;
 I43R,Q/T430A;
 I43Q/T430A/A431L/E503A/Q511H;
- 40 I43Q/T430A/A431L/Q511H;
 I43Q/T430A/A431L/Q511H/A539R;
 I43Q/T430A/E503A/Q511H;
 I43Q/T430A/Q511H;
 I43Q/T430A,M/Q511H/A539R;
- 45 I43R/T430A/E503A,V/Q511H/N563K;
 I43Q/A431L/Q511H;

- I43Q/E503A/Q511H/A539R;
- I43Q/Q511H/A539R;
- D44C/G73F/E503V/Q511H;
- D44C/G73F/L417R/N 563K;
- 5 D44C/G73F/N563K;
- D44C/L417R/N563K;
- D44R/E503A/Q511H/N563I;
- D44R/T430A/Q511H/A535R;
- G73F/T430A/E503V/Q511H;
- 10 G73F/T430A/Q511H;
- G294C/L417R/A431L;
- G294C/L417R/A431L,Q/Q511H;
- G294C/L417V/A431Q;
- L417R,V/A431L,Q/Q511H;
- 15 L417V/T430A/A431L,Q/Q511H/A535R/A539R/N563I;
- L417V/T430A/Q511H/A535R/N563I;
- L417V/T430A/Q511H/A539R/N563I y
- E503A/N563I

20 de SEQ ID NO: 2 o posiciones equivalentes en glucoamilasas precursoras y especialmente homólogos de glucoamilasa *Trichoderma*.

[0136] Se ha alineado un número de glucoamilasas precursoras con la secuencia de aminoácidos de TrGa. Las figuras 4A y 4B incluyen el dominio catalítico de las siguientes glucoamilasas precursoras: *Aspergillus awamori* (AaGA) (SEQ ID NO: 5); *Aspergillus niger* (AnGA) (SEQ ID NO: 6); *Aspergillus oryzae* (AoGA) (SEQ ID NO: 7); *Humicola grisea* (HgGA) (SEQ ID NO: 8); e *Hypocrea vinosa* (HvGA) (SEQ ID NO: 9). El porcentaje de identidad de los dominios catalíticos se representa en la tabla 1 a continuación. Las figuras 4C y 4D incluyen el dominio de unión a almidón de las siguientes glucoamilasas precursoras: *Trichoderma reesei* (TrGA) (SEQ ID NO: 161), *Humicola grisea* (HgGA) (SEQ ID NO: 162), *Thielavia terrestris* (TtGA) (SEQ ID NO: 163), *Thermomyces lanuginosus* (ThGA) (SEQ ID NO: 164), *Talaromyces emersonii* (TeGA) (SEQ ID NO: 165), *Aspergillus niger* (AnGA) (SEQ ID NO: 166), y *Aspergillus awamori* (AaGA) (SEQ ID NO: 167). Por ejemplo, la glucoamilasa variante se derivará de una glucoamilasa precursora que es una glucoamilasa *Aspergillus* y la variante incluirá al menos una sustitución en una posición equivalente a una posición establecida en SEQ ID NO: 2, y en concreto en una posición correspondiente a D4, F5, I12, D24, F29, I43, D44, P45, D46, Y47, Y49, W51, N61, Y70, G73, Q75, R76, P94, D100, K108, K114, F116, N119, R122, Q124, R125, G137, N146, Q148, Y169, N171, Q172, F175, W178, E180, V181, Q208, S211, W228, N242, E243, R245, I292, G294, K297, R309, Y310, D313, V314, Y315, Y316, N317, W321, K340, K341, T350, Q356, T363, S368, S369, N376, Y395, A398, S401, R408, N409, T412, L417, H418, W421, T430, A431, R433, I436, S451, E503, Q511, A535, A539, o N563.

[0137] La eliminación con Endo-H de azúcares enlazados a N en la glucoamilasa *Trichoderma reesei* presentaban un efecto estabilizador (cuando se observaba la T_m). Por lo tanto, las variantes con una sustitución N171D pueden presentar termoestabilidad aumentada en comparación con la precursora de tipo silvestre. En algunas partes de la exposición, se presentan variantes con una o más sustituciones en sitios con azúcares enlazados a N, entre las que se incluye N171D en *Trichoderma reesei* (SEQ ID NO: 2).

Tabla 1

	AaGA	AnGA	AoGA	HgGA	HvGA	TrGA
AaGA	100	95	58	53	57	56
AnGA		100	59	53	57	56
AoGA			100	55	56	56
HgGA				100	61	63
HvGA					100	91
TrGA						100

5 **[0138]** La presente exposición también proporciona variantes de glucoamilasa con al menos una propiedad modificada (p. ej., propiedad mejorada) en comparación con una glucoamilasa precursora y en concreto con la TrGA. En algunas partes de la exposición, al menos una propiedad modificada (p. ej., propiedad mejorada) se elige del grupo consistente en estabilidad del ácido, estabilidad térmica y actividad específica. Normalmente, la propiedad modificada es estabilidad del ácido aumentada, estabilidad térmica aumentada y/o actividad específica aumentada. La estabilidad térmica aumentada se da normalmente a temperaturas superiores. En una forma de realización, la estabilidad del pH aumentada se da con un pH alto. En una forma de realización adicional, la estabilidad del pH aumentada se da con un pH bajo.

10 **[0139]** Las variantes de glucoamilasa de la exposición pueden también proporcionar índices superiores de hidrólisis de almidón con concentraciones de sustrato bajas en comparación con la glucoamilasa precursora. La variante puede tener una V_{max} superior o K_m inferior a una glucoamilasa precursora cuando se realiza el análisis con las mismas condiciones. Por ejemplo, la glucoamilasa variante puede tener una V_{max} superior con un intervalo de temperaturas de 25 °C a 70 °C (p. ej., de 25 °C a 35 °C; de 30 °C a 35 °C; de 40 °C a 50 °C; de 50 °C a 55 °C y
15 de 55 °C a 62 °C). Los valores K_m , V_{max} y la constante Michaelis-Menten pueden determinarse fácilmente utilizando procedimientos conocidos convencionales.

5. Estabilidad térmica (variantes termoestables)

20 **[0140]** En un aspecto, la exposición hace referencia a una glucoamilasa variante con estabilidad térmica modificada con temperaturas modificadas en comparación con una precursora o de tipo silvestre. Las temperaturas modificadas incluyen temperaturas aumentadas o disminuidas. En algunas partes de la exposición, la variante de glucoamilasa presentará termoestabilidad mejorada tal como la retención de al menos un 50 %, 60 %, 70 %, 75 %
25 %, 80 %, 85 %, 90 %, 92 %, 95 %, 96 %, 97 %, 98 % o 99 % de actividad enzimática tras la exposición a temperaturas modificadas durante un periodo de tiempo determinado, por ejemplo, al menos 60 minutos, 120 minutos, 180 minutos, 240 minutos, 300 minutos, etc. En algunas partes de la exposición, la variante presenta estabilidad térmica aumentada en comparación con la glucoamilasa precursora con temperaturas elegidas dentro del intervalo de 40 a 80 °C, también dentro del intervalo de 50 a 75 °C y dentro del intervalo de 60 a 70 °C, y normalmente con un intervalo de pH de 4,0 a 6,0. En algunas partes de la exposición, la termoestabilidad se
30 determina tal y como se describe en los ejemplos.

[0141] En algunas partes de la exposición, variantes especialmente interesantes en relación con una mejora en la termoestabilidad incluyen una o más deleciones, sustituciones o inserciones y concretamente sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO: 2: D4, F5, I12, D24, F29, 143,
35 D44, P45, D46, Y47, Y49, W51, N61, Y70, G73, Q75, R76, P94, D100, K108, K114, F116, N119, R122, Q124, R125, G137, N146, Q148, Y169, N171, Q172, F175, W178, E180, V181, Q208, S211, W228, N242, E243, R245, 1292, G294, K297, R309, Y310, D313, V314, Y315, Y316, N317, W321, K340, K341, T350, Q356, T363, S368, S369, N376, Y395, A398, S401, R408, N409, T412, L417, H418, W421, T430, A431, R433, 1436, S451, E503, Q511, A535, A539, o N563; o una posición equivalente en una glucoamilasa precursora. En algunas formas de
40 realización, la glucoamilasa precursora será un homólogo de glucoamilasa *Trichoderma* y en formas de realización típicas, la glucoamilasa precursora presentará al menos un 90 %, al menos un 95 % o al menos un 98 % de identidad de secuencia con SEQ ID NO: 2.

6. Glucoamilasas quiméricas

45 **[0142]** Las variantes de glucoamilasa de la presente exposición pueden también incluir glucoamilasas híbridas o quiméricas con, por ejemplo, un dominio de unión a almidón (SBD) de una glucoamilasa y un dominio catalítico y enlazador de otra. Por ejemplo, una glucoamilasa híbrida puede formarse mediante el intercambio del SBD de AnGA con el SBD de TrGA, realizando un híbrido con el SBD de AnGA y el dominio catalítico y enlazador de TrGA.
50 De forma alternativa, el SBD y el enlazador de AnGA pueden intercambiarse por el SBD y el enlazador de TrGA.

7. Actividad específica

55 **[0143]** En otro aspecto, la presente exposición hace referencia a una glucoamilasa variante con actividad específica modificada en comparación con una glucoamilasa de tipo silvestre o precursora.

[0144] En algunas partes de la exposición, variantes especialmente interesantes en relación con una mejora en la actividad específica incluyen una o más deleciones, sustituciones o inserciones y concretamente sustituciones en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO: 2: D4, F5, I12, D24, F29,
60 143, D44, P45, D46, Y47, Y49, W51, N61, Y70, G73, Q75, R76, P94, D100, K108, K114, F116, N119, R122, Q124, R125, G137, N146, Q148, Y169, N171, Q172, F175, W178, E180, V181, Q208, S211, W228, N242, E243, R245, 1292, G294, K297, R309, Y310, D313, V314, Y315, Y316, N317, W321, K340, K341, T350, Q356, T363, S368,

S369, N376, Y395, A398, S401, R408, N409, T412, L417, T430, A431, H418, W421, R433, 1436, S451, E503, Q511, A535, A539, o N563; o una posición equivalente en una glucoamilasa precursora. En algunas partes de la exposición, las variantes de la presente exposición con actividad específica mejorada incluyen una sustitución en las siguientes posiciones en la secuencia de aminoácidos establecida en SEQ ID NO: 2: D4, D24, I43, D44, N61, Y70, G73, Q75, R76, D100, K108, N119, Q124, N146, Q148, N171, Q172, F175, V181, Q208, S211, E243, R245, 1292, G294, K297, V314, Y316, N317, K340, K341, T350, Q356, T363, S368, N376, Y395, A398, S401, N409, T412, L417, T430, A431, I436, S451, E503, Q511, A535, A539, o N563; o una posición equivalente en una glucoamilasa precursora. En algunas formas de realización, la glucoamilasa precursora comprenderá una secuencia con al menos un 90 % o 95 % de identidad de secuencia con la secuencia de SEQ ID NO: 2.

8. Polinucleótidos

[0145] La presente exposición también hace referencia a polinucleótidos aislados que codifican una glucoamilasa variante de la presente exposición. Los polinucleótidos que codifican una glucoamilasa variante pueden prepararse mediante técnicas consolidadas conocidas en la técnica. Los polinucleótidos pueden prepararse de forma sintética, tal como mediante un sintetizador de ADN automático. La secuencia de ADN puede ser de origen sintético y genómico combinados (o ADNc) preparada mediante la ligación conjunta de fragmentos. Los polinucleótidos pueden también prepararse mediante la reacción en cadena de la polimerasa (PCR por sus siglas en inglés) utilizando cebadores específicos. En general, se hace referencia a Minshull J., *et al.*, (2004), Engineered protein function by selective amino acid diversification, *Methods* 32(4):416-427. Asimismo, una serie de empresas ahora sintetizan ADN, como por ejemplo Genearth AG, Regensburg, Alemania.

[0146] La presente exposición también proporciona polinucleótidos aislados que comprenden una secuencia de nucleótidos (i) con al menos un 70 % de identidad con SEQ ID NO: 4, o (ii) capaz de hibridar con una sonda derivada de la secuencia de nucleótidos establecida en SEQ ID NO: 4, con condiciones de astringencia intermedia a alta, o (iii) complementaria a una secuencia de nucleótidos con al menos un 90 % de identidad de secuencia con la secuencia establecida en SEQ ID NO: 4. Sondas útiles según la presente exposición pueden incluir al menos 50, 100, 150, 200, 250, 300 o más nucleótidos contiguos de SEQ ID NO: 4.

[0147] La presente exposición proporciona además polinucleótidos aislados que codifican glucoamilasas variantes que comprenden una secuencia de aminoácidos que comprende al menos un 80 % de identidad de secuencia de aminoácidos con SEQ ID NO: 2. En algunas formas de realización, las glucoamilasas variantes presentan al menos un 93 % de identidad de secuencia de aminoácidos con SEQ ID NO: 2. En algunas formas de realización, las glucoamilasas variantes presentan al menos un 95 % de identidad de secuencia de aminoácidos con SEQ ID NO: 2. En algunas formas de realización, las glucoamilasas variantes presentan al menos un 97 % de identidad de secuencia de aminoácidos con SEQ ID NO: 2. En algunas formas de realización, las glucoamilasas variantes presentan al menos un 98 % de identidad de secuencia de aminoácidos con SEQ ID NO: 2. En algunas formas de realización, las glucoamilasas variantes presentan al menos un 99 % de identidad de secuencia de aminoácidos con SEQ ID NO: 2. La presente exposición también proporciona vectores de expresión que comprenden cualquiera de los polinucleótidos proporcionados anteriormente.

[0148] La presente exposición también proporciona fragmentos (es decir, partes) del ADN que codifica las glucoamilasas variantes proporcionadas en el presente documento. Estos fragmentos son útiles a la hora de obtener fragmentos de ADN de longitud parcial capaces de ser usados para aislar o identificar polinucleótidos que codifican enzimas de glucoamilasa maduras descritas en el presente documento a partir de células fúngicas filamentosas (p. ej., *Trichoderma*, *Aspergillus*, *Fusarium*, *Penicillium*, *Schizosaccharomyces* y *Humicola*), o un segmento de estas con actividad de glucoamilasa. En algunas partes de la exposición, los fragmentos del ADN pueden comprender al menos 50, 100, 150, 200, 250, 300 o más nucleótidos contiguos. En algunas partes de la exposición, partes del ADN proporcionadas en SEQ ID NO: 4 pueden utilizarse a la hora de obtener homólogos de glucoamilasa precursora y concretamente homólogos de glucoamilasa *Trichoderma* de otras especies, tal como hongos filamentosos que codifican una glucoamilasa.

9. Construcciones de ADN y vectores

[0149] Según una forma de realización de la presente exposición, una construcción de ADN que comprende un polinucleótido según se describe anteriormente y que codifica una glucoamilasa variante incluida en la presente exposición y unida de forma operativa a una secuencia promotora se junta para transferirse a una célula huésped.

[0150] La construcción de ADN puede introducirse en una célula huésped utilizando un vector. El vector puede ser cualquier vector que, cuando se introduce en una célula huésped, se integra normalmente en el genoma de la célula huésped y se duplica. Entre los vectores se incluyen vectores de clonación, vectores de expresión, vectores transportadores, plásmidos, partículas fágicas, casetes y similares. En algunas partes de la exposición, el vector

es un vector de expresión que comprende secuencias reguladoras unidas de forma operativa a la secuencia que codifica la glucoamilasa.

5 **[0151]** Ejemplos de vectores de expresión y/o integración adecuados se presentan en Sambrook *et al.*, (1989) *supra*, y Ausubel (1987) *supra*, y van den Hondel *et al.* (1991) en Bennett y Lasure (Eds.) MORE GENE MANIPULATIONS IN FUNGI, Academic Press pp. 396-428 y la patente estadounidense n.º 5874276. El catálogo de cepas del Centro de Almacenamiento genético fúngico (FGSC, < www.fgsc.net>) también expone vectores útiles. Entre los vectores especialmente útiles se incluyen vectores obtenidos de, por ejemplo, Invitrogen y Promega.

10

[0152] Vectores específicos adecuados para utilizarse en células huésped fúngicas incluyen vectores tales como pFB6, pBR322, pUC18, pUC100, pDONR™201, pDONR™221, pENTR™, pGEM@3Z y pGEM@4Z. Un vector de expresión con fin general útil en *Aspergillus* incluye pRAX con un promotor *glaA*, y en *Hypocrea/Trichoderma* incluye pTrex3g con un promotor *cbh1*.

15

[0153] Plásmidos adecuados para utilizarse en células bacterianas incluyen pBR322 y pUC19 que permiten la replicación en *E. coli* y pE194 por ejemplo que permite la replicación en *Bacillus*.

20

[0154] En algunas partes de la exposición, el promotor muestra actividad transcripcional en una célula huésped fúngica o bacteriana y puede derivarse de genes que codifican proteínas ya sean homólogas o heterólogas a la célula huésped. El promotor puede ser un promotor híbrido, truncado y/o mutante. Los promotores mencionados anteriormente son conocidos en la técnica.

25

[0155] Ejemplos de promotores adecuados útiles en células fúngicas y especialmente en células fúngicas filamentosas tal como células *Trichoderma* o *Aspergillus* incluyen tales promotores de ejemplo como los promotores *cbh1*, *cbh2*, *egl1*, *egl2*, *eg5*, *xln1* y *xln2* de *T. reesei*. Otros ejemplos de promotores útiles incluyen promotores de genes de glucoamilasa *A. awamori* y *A. niger* (*glaA*) (véase, Nunberg *et al.*, (1984) Mol. Cell Biol. 4:2306-2315 y Boel *et al.*, (1984) EMBO J. 3:1581-1585), promotor de TAKA amilasa *A. oryzae*, el promotor TPI (triosa fosfato isomerasa) de *S. cerevisiae*, el promotor de genes acetamidasa *Aspergillus nidulans* y genes lipasa *Rhizomucor miehei*.

30

[0156] Ejemplos de promotores adecuados útiles en células bacterianas incluyen aquellos obtenidos del operón *lac E. coli*; gen alfa-amilasa *Bacillus licheniformis* (*amyL*), gen amilasa *B. stearothermophilus* (*amyM*); genes *xyIA* y *xyIB* *Bacillus subtilis*, el gen beta-lactamasa y el promotor *tac*.

35

[0157] En una parte de la exposición, el promotor es uno que es natural a la célula huésped. Por ejemplo, cuando el huésped es *T. reesei*, el promotor es un promotor natural de *T. reesei*. En otras formas de realización, el promotor es uno que es heterólogo a la célula huésped fúngica. En algunas formas de realización, el promotor será el promotor de la glucoamilasa precursora tal como el promotor de TrGA.

40

[0158] En algunas partes de la exposición, la construcción de ADN incluye ácidos nucleicos que codifican una secuencia señal que es una secuencia de aminoácidos enlazada al amino terminal del polipéptido que dirige al polipéptido codificado a la trayectoria secretora de la célula. El extremo 5' de la secuencia codificante de la secuencia de ácido nucleico puede incluir de forma natural una región codificante de péptido señal que está enlazada de forma natural en el marco de lectura de traducción con el segmento de la secuencia codificante de glucoamilasa que codifica la glucoamilasa secretada o el extremo 5' de la secuencia codificante de la secuencia de ácido nucleico puede incluir un péptido señal que es ajeno a la secuencia codificante. En algunas partes de la exposición, la construcción de ADN incluye una secuencia señal que está asociada de forma natural a un gen de glucoamilasa precursora a partir del cual se obtiene la glucoamilasa variante. En algunas partes de la exposición, la secuencia señal será la secuencia representada en SEQ ID NO: 1 o una secuencia con al menos un 90 %, al menos un 94 % y al menos un 98 % de identidad de secuencia con la misma. Secuencias señales eficaces pueden incluir las secuencias señales obtenidas de glucoamilasas de otras enzimas fúngicas filamentosas, tal como de *Trichoderma* (glucoamilasa *T. reesei*), *Humicola* (celulasa *H. insolens* o glucoamilasa *H. grisea*), *Aspergillus* (glucoamilasa *A. niger* y TAKA amilasa *A. oryzae*) y *Rhizopus*.

45

50

55

[0159] En partes adicionales de la exposición, una construcción de ADN o vector que comprende una secuencia señal y una secuencia promotora que han de introducirse en una célula huésped se derivan de la misma fuente. En algunas formas de realización, se puede utilizar la secuencia señal de glucoamilasa natural de un homólogo de glucoamilasa *Trichoderma*, tal como una secuencia señal de una cepa de *Hypocrea*.

[0160] En algunas partes de la exposición, el vector de expresión también incluye una secuencia de terminación. En la presente exposición, puede utilizarse cualquier secuencia de terminación funcional en la célula huésped. En una forma de realización, la secuencia de terminación y la secuencia promotora se derivan de la misma fuente. En otra forma de realización, la secuencia de terminación es homóloga a la célula huésped. Entre las secuencias de terminación útiles se incluyen secuencias de terminación obtenidas a partir de los genes de *cbh1 Trichoderma reesei*; glucoamilasa *A. niger* o *A. awamori* (Nunberg *et al.* (1984) *supra*, y Boel *et al.*, (1984) *supra*), antranilato sintasa *Aspergillus nidulans*, Taka amilasa *Aspergillus oryzae* o *A. nidulans trpC* (Punt *et al.*, (1987) Gene 56:117-124).

[0161] En algunas partes de la exposición, un vector de expresión incluye un marcador seleccionable. Entre los ejemplos de típicos marcadores seleccionables se incluyen unos que confieren resistencia antimicrobiana (p. ej., higromicina y fleomicina). Los marcadores selectivos nutricionales también son útiles en la presente exposición incluyendo aquellos marcadores conocidos en la técnica como *amdS* (acetamidasa), *argB* (ornitina carbamoiltransferasa) y *pyrG* (orotidina-5' fosfato descarboxilasa). Los marcadores útiles en sistemas de vectores para la transformación de *Trichoderma* son conocidos en la técnica (véase, p. ej., Finkelstein, capítulo 6 en BIOTECHNOLOGY OF FILAMENTOUS FUNGI, Finkelstein *et al.* Eds. Butterworth-Heinemann, Boston, MA (1992); Kinghom *et al.* (1992) APPLIED MOLECULAR GENETICS OF FILAMENTOUS FUNGI, Blackie Academic and Professional, Chapman and Hall, Londres; Berges and Barreau (1991) Curr. Genet. 19:359-365; y van Hartingsveldt *et al.*, (1987) Mol. Gen. Genet. 206:71-75). En una parte típica de la exposición, el marcador selectivo es el gen *amdS*, que codifica la enzima acetamidasa, lo que permite que las células transformadas crezcan en acetamida como una fuente de nitrógeno. El uso del gen *amdS A. nidulans* como marcador selectivo se describe en Kelley *et al.*, (1985) EMBO J. 4:475-479 y Penttilä *et al.*, (1987) Gene 61:155-164.

[0162] Los métodos utilizados para ligar la construcción de ADN que comprende una secuencia de ácido nucleico que codifica una glucoamilasa variante, un promotor, un terminador y otras secuencias y para insertarlos en un vector adecuado son conocidos en la técnica. El enlace se consigue generalmente mediante la ligación en los sitios de restricción convenientes. Si no existen estos sitios, los enlazadores de oligonucleótidos sintéticos se utilizan según la práctica convencional. (Véase, Sambrook (1989) *supra*, y Bennett and Lasure, MORE GENE MANIPULATIONS IN FUNGI, Academic Press, San Diego (1991) pp 70-76.) De forma adicional, los vectores pueden construirse utilizando técnicas de recombinación conocidas (p. ej., Invitrogen Life Technologies, Gateway Technology).

10. Células huésped

[0163] La presente invención también hace referencia a células huésped que comprenden un polinucleótido que codifica una glucoamilasa variante de la exposición, que se utilizan para producir las glucoamilasas de la exposición. En algunas partes de la exposición, las células huésped se eligen de células bacterianas, fúngicas, vegetales y de levadura. El término célula huésped incluye tanto las células como la progenie de las células y protoplastos creados a partir de las células que se utilizan para producir una glucoamilasa variante según la presente exposición.

[0164] En algunas formas de realización, las células huésped son células huésped fúngicas y normalmente células huésped fúngicas filamentosas. El término "hongos filamentosos" hace referencia a las formas filamentosas de la subdivisión Eumycotina (Véase Alexopoulos, C. J. (1962), INTRODUCTORY MYCOLOGY, Wiley, Nueva York). Estos hongos se caracterizan por un micelio vegetativo con una pared celular compuesta de quitina, celulosa y otros polisacáridos complejos de la exposición. Los hongos filamentosos de la presente exposición son distintos morfológica, fisiológica y genéticamente de las levaduras. El crecimiento vegetativo por hongos filamentosos se realiza mediante elongación de las hifas y el catabolismo de carbono es aerobio obligado. En la presente exposición, la célula precursora fúngica filamentosa puede ser una célula de una especie de entre, sin carácter limitativo, *Trichoderma*, (p. ej., *Trichoderma reesei*, la mutación asexual de *Hypocrea jecorina*, clasificada anteriormente como *T. longibrachiatum*, *Trichoderma viride*, *Trichoderma koningii*, *Trichoderma harzianum*) (Sheir-Neirs *et al.*, (1984) Appl. Microbiol. Biotechnol 20:46-53; ATCC n.º 56765 y ATCC n.º 26921); *Penicillium sp.*, *Humicola sp.* (p. ej., *H. insolens*, *H. lanuginosa* y *H. grisea*); *Chrysosporium sp.* (p. ej., *C. lucknowense*), *Gliocladium sp.*, *Aspergillus sp.* (p. ej., *A. oryzae*, *A. niger*, *A. sojae*, *A. japonicus*, *A. nidulans*, y *A. awamori*) (Ward *et al.*, (1993) Appl. Microbiol. Biotechnol. 39:738-743 y Goedegebuur *et al.*, (2002) Genet 41:89-98), *Fusarium sp.*, (p.ej., *F. roseum*, *F. gramineum*, *F. cerealis*, *F. oxysporum* y *F. venenatum*), *Neurospora sp.*, (*N. crassa*), *Hypocrea sp.*, *Mucor sp.*, (*M. miehei*), *Rhizopus sp.* y *Emericella sp.* (véase también, Innis *et al.*, (1985) Sci. 228:21-26). Los términos "*Trichoderma*" o "*Trichoderma sp.*" o "*Trichoderma spp.*" hacen referencia a cualquier género fúngico clasificado previamente o actualmente como *Trichoderma*.

[0165] En algunas formas de realización, las células huésped serán células bacterianas gram positivas. Entre los ejemplos sin carácter limitativo se incluyen cepas de *Streptomyces*, (p. ej., *S. lividans*, *S. coelicolor* y *S. griseus*)

y *Bacillus*. Tal y como se utiliza en el presente documento, "el género *Bacillus*" incluye todas las especies dentro del género "*Bacillus*", según lo conocen los expertos en la técnica, entre las que se incluyen sin carácter limitativo *B. subtilis*, *B. licheniformis*, *B. lentus*, *B. brevis*, *B. stearothermophilus*, *B. alkalophilus*, *B. amyloliquefaciens*, *B. clausii*, *B. halodurans*, *B. megaterium*, *B. coagulans*, *B. circulans*, *B. lautus*, y *B. thuringiensis*. Se ha reconocido que el género *Bacillus* sigue experimentando reorganización taxonómica. Por lo tanto, se pretende que el género incluya especies que han sido reclasificadas, incluyendo sin carácter limitativo tales organismos como *B. stearothermophilus*, que ahora se denomina "*Geobacillus stearothermophilus*".

[0166] En algunas partes de la exposición, la célula huésped es una cepa bacteriana gram negativa, tal como *E. coli* o *Pseudomonas sp.* En otras partes de la exposición, las células huésped pueden ser células de levadura tal como *Saccharomyces sp.*, *Schizosaccharomyces sp.*, *Pichia sp.* o *Candida sp.*

[0167] En otras partes de la exposición, la célula huésped será una célula huésped diseñada genéticamente en la que los genes naturales se han desactivado, por ejemplo, mediante delección en células bacterianas o fúngicas. Cuando se desee obtener una célula huésped fúngica con uno o más genes desactivados, se pueden utilizar métodos conocidos (p. ej., los métodos expuestos en la patente estadounidense n.º 5246853, la patente estadounidense n.º 5475101, y WO 92/06209). La desactivación génica puede conseguirse mediante una delección completa o parcial, mediante desactivación por inserción o mediante cualquier otro medio que proporcione un gen no funcional para el fin destinado (de forma que se impida la expresión del gen de una proteína funcional). En algunas formas de realización, cuando la célula huésped es una célula *Trichoderma* y especialmente una célula huésped *T. reesei*, los genes *cbh1*, *cbh2*, *egl1* y *egl2* se desactivarán y/o deleccionarán normalmente. Normalmente, células huésped *Trichoderma reesei* con proteínas con delección cuádruple se establecen y describen en la patente estadounidense n.º 5847276 y WO 05/001036. En otras partes de la exposición, la célula huésped es una cepa carente de proteasa o proteasa negativa.

11. Transformación de células huésped

[0168] La introducción de una construcción de ADN o vector en una célula huésped incluye técnicas tales como transformación, electroporación, microinyección nuclear, transducción, transfección (p. ej., transfección mediada por lipofección y mediada por DEAE-Dextrina); incubación con precipitado de ADN fosfato calcio; bombardeo de alta velocidad con microproyectiles revestidos de ADN y fusión de protoplastos. Las técnicas de transformación generales se conocen en la técnica (véase, p. ej., Ausubel *et al.*, (1987), *supra*, capítulo 9; y Sambrook (1989) *supra*, y Campbell *et al.*, (1989) *Curr. Genet.* 16:53-56).

[0169] Los métodos de transformación para *Bacillus* se exponen en numerosas referencias entre las que se incluyen Anagnostopoulos C y J. Spizizen (1961) *J. Bacteriol.* 81:741-746 y WO 02/14490.

[0170] Los métodos de transformación para *Aspergillus* se describen en Yelton *et al.* (1984) *Proc. Natl. Acad. Sci. EE UU* 81:1470-1474; Berka *et al.*, (1991) en *Applications of Enzyme Biotechnology*, Eds. Kelly and Baldwin, Plenum Press (NY); Cao *et al.*, (2000) *Science* 9:991-1001; Campbell *et al.*, (1989) *Curr. Genet.* 16:53-56 y EP 238 023. La expresión de proteína heteróloga en *Trichoderma* se describe en la patente estadounidense n.º 6022725; la patente estadounidense n.º 6268328; Harkki *et al.* (1991); *Enzyme Microb. Technol.* 13:227-233; Harkki *et al.*, (1989) *Bio Technol.* 7:596-603; EP 244,234; EP 215,594; y Nevalainen *et al.*, "The Molecular Biology of *Trichoderma* and its Application to the Expression of Both Homologous and Heterologous Genes", en *MOLECULAR INDUSTRIAL MYCOLOGY*, Eds. Leong y Berka, Marcel Dekker Inc., NY (1992) pp. 129-148). También se hace referencia a WO96/00787 y Bajar *et al.*, (1991) *Proc. Natl. Acad. Sci. EE UU* 88:8202-8212 para la transformación de cepas *Fusarium*.

[0171] En una parte específica de la exposición, la preparación de *Trichoderma sp.* para la transformación implica la preparación de protoplastos a partir de micelios fúngicos (Véase, Campbell *et al.*, (1989) *Curr. Genet.* 16:53-56; Pentilla *et al.*, (1987) *Gene* 61:155-164). La transformación mediada por *Agrobacterium tumefaciens* de hongos filamentosos es conocida (véase de Groot *et al.*, (1998) *Nat. Biotechnol.* 16:839-842). También se hace referencia a la patente estadounidense n.º 6022725 y la patente estadounidense n.º 6268328 para los procedimientos de transformación utilizados con huéspedes fúngicos filamentosos.

[0172] Normalmente, los transformantes estables genéticamente se construyen con sistemas de vectores por los que el ácido nucleico que codifica la glucoamilasa variante se integra de forma estable en un cromosoma de la cepa huésped. A continuación, se purifican los transformantes con técnicas conocidas.

[0173] En algunas formas de realización adicionales, las células huésped son células vegetales, tal como células de una planta monocotiledónea (p. ej., maíz, trigo y sorgo) o células de una planta dicotiledónea (p. ej., soja). Los métodos para realizar construcciones de ADN útiles para la transformación de plantas y métodos para la

transformación vegetal son conocidos. Algunos de estos métodos incluyen transferencia génica mediada por *Agrobacterium tumefaciens*; bombardeo de microproyectiles, transformación de protoplastos mediada por PEG, electroporación y similares. Se hace referencia a (la patente estadounidense n.º 6803499, la patente estadounidense n.º 6777589; Fromm *et al* (1990) *Biotechnol.* 8:833-839; Potrykus *et al* (1985) *Mol. Gen. Genet.* 199:169-177.

12. Producción de proteínas

[0174] La presente invención hace referencia además a métodos para producir las glucoamilasas variantes que comprenden la transformación de una célula huésped con un vector de expresión que comprende un polinucleótido que codifica una glucoamilasa variante de acuerdo con la presente exposición, el cultivo opcional de la célula huésped en condiciones adecuadas para la producción de la glucoamilasa variante y la recuperación opcional de la glucoamilasa.

[0175] En los métodos de producción y expresión de la presente exposición, las células huésped se cultivan bajo condiciones adecuadas en cultivo con agitación del matraz, fermentaciones a pequeña o gran escala (incluyendo fermentaciones continuas, por lotes y discontinuas) en fermentadores industriales o de laboratorio, con el medio adecuado que contenga sales fisiológicas y nutrientes (véase, p. ej., Pourquie, J. *et al.*, *BIOCHEMISTRY AND GENETICS OF CELLULOSE DEGRADATION*, eds. Aubert, J. P. *et al.*, Academic Press, pp. 71-86, 1988 y Ilmen, M. *et al.*, (1997) *Appl. Environ. Microbiol.* 63:1298-1306). Los medios comunes preparados comercialmente (p. ej., caldo de extracto de malta levadura, caldo Luria Bertani y el caldo Sabouraud Dextrosa) son útiles en la presente exposición. Las condiciones de cultivo para células fúngicas filamentosas y bacterianas son conocidas en la técnica y pueden hallarse en la literatura científica y/o desde la fuente de los hongos tal como American Type Culture Collection y Fungal Genetics Stock Center. En los casos en los que una secuencia codificante de glucoamilasa esté bajo el control de un promotor inducible, el agente inductor (p. ej., un azúcar, una sal metálica o antimicrobiano) se añade al medio con una concentración eficaz para inducir la expresión de glucoamilasa.

[0176] En algunas partes de la exposición, la presente exposición hace referencia a métodos para producir la glucoamilasa variante que comprenden el cultivo de una planta o célula vegetal que comprende un polinucleótido que codifica una glucoamilasa variante de acuerdo con la exposición bajo condiciones adecuadas para la producción de la variante y la recuperación opcional de la glucoamilasa.

[0177] En algunas partes de la exposición, con el fin de evaluar la expresión de una glucoamilasa variante mediante una línea celular que ha sido transformada con un polinucleótido que codifica una glucoamilasa variante incluida en la exposición se llevan a cabo ensayos en el nivel de proteína, el nivel de ARN y/o mediante el uso de bioanálisis funcionales específicos de la actividad y/o producción de glucoamilasa. Algunos de estos ensayos incluyen la prueba de Northern blot, dot blot (análisis de ADN o ARN), RT-PCR (reacción en cadena de la polimerasa con transcriptasa inversa) o hibridación *in situ* utilizando una sonda marcada de forma adecuada (según la secuencia codificante de ácido nucleico) y la prueba de Southern blot convencional y autorradiografía.

[0178] Además, la producción y/o expresión de una glucoamilasa variante puede medirse en una muestra directamente, por ejemplo, mediante ensayos que miden directamente los azúcares reductores tal como glucosa en el medio de cultivo y mediante ensayos que miden la actividad, expresión y/o producción de glucoamilasa. En concreto, la actividad de glucoamilasa puede analizarse mediante el método del ácido 3,5-dinitrosalicílico (DNS) (Véase, Goto *et al.*, (1994) *Biosci. Biotechnol. Biochem.* 58:49-54). En formas de realización adicionales, la expresión de proteína se evalúa mediante métodos inmunológicos, tal como la coloración inmunohistoquímica de células, secciones de tejido o inmunoanálisis de un medio de cultivo de tejido (p. ej., mediante Western blot o ELISA). Tales inmunoanálisis pueden utilizarse para evaluar de forma cualitativa y cuantitativa la expresión de una glucoamilasa. Los expertos en la técnica conocen los detalles de tales métodos y muchos reactivos para practicar tales métodos están disponibles comercialmente.

[0179] Las glucoamilasas de la presente exposición pueden recuperarse o purificarse del medio de cultivo mediante una variedad de procedimientos conocidos en la técnica entre los que se incluyen la centrifugación, la filtración, la extracción, la precipitación y similares.

13. Composiciones

[0180] Las glucoamilasas variantes pueden utilizarse en composiciones de enzima entre las que se incluyen sin carácter limitativo composiciones para la hidrólisis y sacarificación de almidón, composiciones de limpieza y detergente (p. ej., detergentes para la colada, detergentes para lavar los platos y composiciones de limpieza de superficies duras), composiciones para la fermentación del alcohol y en composiciones para pienso animal. Además, las glucoamilasas variantes pueden utilizarse en aplicaciones de cocción, tal como la producción de pan

y pasteles, destilación, atención sanitaria, textil, procesos de conversión de residuos medioambientales, procesamiento de biopulpeo y aplicaciones de conversión de biomasa.

- 5 **[0181]** En algunas partes de la exposición, se utilizará de forma opcional una composición de enzima que incluye una glucoamilasa variante incluida en la exposición obtenida en un medio de cultivo o recuperada y purificada del medio de cultivo junto con cualquiera de las siguientes enzimas o una combinación de estas - alfa-amilasas, proteasas, pululanasa, isoamilasas, celulasas, hemicelulasas, xilanasas, ciclodextrina glicosiltransferasas, lipasas, fitasas, laccasas, oxidasas, esterases, cutinasas, xilanasas, enzimas hidrolizantes de almidón granular y otras glucoamilasas. En una parte de la exposición, las proteasas son proteasas fúngicas ácidas (AFP por sus siglas en inglés). En partes adicionales de la exposición, las proteasas fúngicas ácidas son de *Trichoderma* (p. ej., NSP-24, véase también US 2006/0154353, publicada el 13 de julio de 2006). En una parte adicional de la exposición, la fitasa es de *Buttiauxiella spp.* (p. ej., BP-17, véanse también las variantes expuestas en la publicación de patente PCT WO 2006/043178).
- 10
- 15 **[0182]** En algunas composiciones representativas, las glucoamilasas variantes de la exposición se combinarán con un alfa-amilasa, tal como alfa-amilasas fúngicas (p. ej., *Aspergillus sp.*) o alfa-amilasas bacterianas (p. ej., *Bacillus sp.* tal como *B. stearothermophilus*, *B. amyloliquefaciens* y *B. licheniformis*) y variantes e híbridos de estas. En algunas partes de la exposición, la alfa-amilasa es una alfa-amilasa estable ácida. En una parte de la exposición, la alfa-amilasa es una enzima hidrolizante de almidón granular (GSHE). En una parte de la exposición, la alfa-amilasa es alfa-amilasa *Aspergillus kawachi* (AKAA), véase US 7037704. Las alfa-amilasas disponibles comercialmente contempladas para utilizarse en las composiciones de la presente exposición son conocidas e incluyen GZYME G997, SPEZYME FRED, SPEZYME XTRA, STARGEN (Danisco US, Inc, Genencor Division), TERMAMYL 120-L y SUPRA (Novozymes, Biotech.) y VIRIDIUM (Diversa).
- 20
- 25 **[0183]** En otras partes de la exposición, las glucoamilasas variantes de la presente exposición pueden combinarse con otras glucoamilasas. En algunas partes de la exposición, las glucoamilasas de la presente exposición se combinarán con una o más glucoamilasas derivadas de cepas de *Aspergillus* o variantes de esta, tal como *A. oryzae*, *A. niger*, *A. kawachi*, y *A. awamori*; glucoamilasas derivadas de cepas de *Humicola* o variantes de esta, especialmente *H. grisea*, tal como la glucoamilasa con al menos un 90 %, 93 %, 95 %, 96 %, 97 %, 98 % y 99 % de identidad de secuencia con SEQ ID NO: 3 expuesta en WO 05/052148; glucoamilasas derivadas de cepas de *Talaromyces* o variantes de esta, especialmente *T. emersonii*; glucoamilasas derivadas de cepas de *Athelia* y especialmente *A. rolfsii*; glucoamilasas derivadas de cepas de *Penicillium*, especialmente *P. chrysogenum*.
- 30

14. Usos

- 35 **[0184]** En concreto, las glucoamilasas variantes pueden utilizarse para los procesos de conversión de almidón, y especialmente en la producción de dextrosa para jarabes de fructosa, azúcares especializados y en la producción de alcohol y otros productos finales (p. ej., ácido orgánico, ácido ascórbico y aminoácidos) a partir de la fermentación de sustratos que contienen almidón (G.M.A van Beynum *et al.*, Eds. (1985) STARCH CONVERSION TECHNOLOGY, Marcel Dekker Inc. NY). Las dextrinas producidas utilizando composiciones de glucoamilasa variante de la presente exposición pueden tener como resultado una producción de glucosa de al menos un 80 %, al menos un 85 %, al menos un 90 % y al menos un 95 %. La producción de alcohol a partir de la fermentación de sustratos de almidón con glucoamilasas de la presente exposición puede incluir la producción de carburante alcohólico y alcohol de boca. En algunas partes de la exposición, la producción de alcohol será mayor cuando la glucoamilasa variante se utilice con las mismas condiciones que la glucoamilasa precursora. En algunas partes de la exposición, la producción de alcohol será entre aproximadamente un 0,5 % y 2,5 % mayor, incluyendo sin carácter limitativo 0,6 %; 0,7 %; 0,8 %; 0,9 %; 1,0 %; 1,1 %; 1,2 %; 1,3 %; 1,4 %; 1,5 %; 1,6 %; 1,7 %; 1,8 %; 1,9 %; 2,0 %; 2,1 %; 2,2 %; 2,3 % y 2,4 % más de alcohol que la glucoamilasa precursora.
- 40
- 45
- 50 **[0185]** En una parte representativa de la exposición, las glucoamilasas variantes de la exposición serán útiles en la hidrólisis de almidón de diferentes sustratos con base vegetal, que se utilizan para la producción de alcohol. En algunas partes de la exposición, los sustratos con base vegetal incluirán maíz, trigo, cebada, centeno, sorgo, arroz, caña de azúcar, patatas y combinaciones de estos. En algunas partes de la exposición, el sustrato con base vegetal se fraccionará en material vegetal, por ejemplo, un grano de cereal, tal como maíz, que se fracciona en componentes tal como fibra, germen, proteína y almidón (endospermo) (patente estadounidense n.º 6254914 y patente estadounidense n.º 6899910). Los métodos para las fermentaciones de alcohol se describen en THE ALCOHOL TEXTBOOK, A REFERENCE FOR THE BEVERAGE, FUEL AND INDUSTRIAL ALCOHOL INDUSTRIES, 3ª Ed., Eds K.A. Jacques *et al.*, 1999, Nottingham University Press, Reino Unido. En determinadas partes de la exposición, el alcohol será etanol. En concreto, los procesos de producción de fermentación de alcohol se caracterizan como procesos de molienda húmeda o seca. En algunas partes de la exposición, la glucoamilasa variante se utilizará en un proceso de fermentación de molienda húmeda y en otras partes de la exposición la glucoamilasa variante será útil en un proceso de molienda seca.
- 55
- 60

[0186] La molienda de grano seco implica un número de etapas básicas, que incluyen generalmente: triturado, cocción, licuefacción, sacarificación, fermentación y separación de líquidos y sólidos con el fin de producir alcohol y otros subproductos. Se tritura material vegetal y especialmente granos de cereales integrales, tal como maíz, trigo o centeno. En algunos casos, el grano se puede fraccionar primero en partes componentes. El material vegetal triturado puede molerse para obtener una partícula fina o gruesa. El material vegetal triturado se mezcla con líquido (p. ej., agua y/o finos residuos de la destilería) en un tanque de suspensión. La suspensión está sujeta a altas temperaturas (p. ej., de 90 °C a 105 °C o superiores) en una caldera de inyección a vapor junto con enzimas de licuefacción (p. ej., alfa-amilasas) para solubilizar e hidrolizar el almidón en el grano a dextrinas. La pulpa se enfría y se trata además con enzimas de sacarificación, tal como glucoamilasas incluidas en la presente exposición, para producir glucosa. A continuación, la pulpa que contiene glucosa puede fermentarse durante aproximadamente un intervalo de 24 a 120 horas en presencia de microorganismos de fermentación, tal como microorganismos que producen etanol y especialmente levadura (*Saccharomyces spp*). Los sólidos en la mezcla se separan de la fase líquida y se obtiene alcohol tal como etanol y subproductos útiles tal como granos de destilación.

[0187] En algunas partes de la exposición, la etapa de sacarificación y la etapa de fermentación se combinan y se hace referencia al proceso como sacarificación y fermentación simultáneas o sacarificación, propagación de levadura y fermentación simultáneas.

[0188] En otras partes de la exposición, la glucoamilasa variante se utiliza en un proceso para la hidrólisis de almidón en el que la temperatura del proceso se lleva a cabo con una temperatura de entre 30 °C y 75 °C y también con una temperatura de entre 40 °C y 65 °C con un intervalo de pH de entre pH 3,0 y pH 6,5. Los procesos de fermentación en algunas partes de la exposición incluyen la molienda de un grano de cereal o grano fraccionado y la combinación del grano de cereal molido con líquido con el fin de formar una suspensión que a continuación se mezcla en un único recipiente con una glucoamilasa variante según la presente exposición y de forma opcional otras enzimas tal como, sin carácter limitativo, alfa-amilasas, otras glucoamilasas, fitasas, proteasas, pululanasa, isoamilasas u otras enzimas con actividad hidrolizante de almidón granular y levadura para producir etanol y otros subproductos (patente estadounidense n.º 4514496, WO 04/081193 y WO 04/080923).

[0189] En algunas partes de la exposición, la presente exposición pertenece a un método para sacarificar una solución de almidón líquida, que comprende una etapa de sacarificación enzimática con una glucoamilasa variante de la presente exposición.

[0190] La presente exposición también presenta pienso animal que comprende al menos una glucoamilasa variante incluida en la presente exposición. Los métodos para usar una enzima de glucoamilasa en la producción de piensos que comprenden almidón se presentan en WO 03/049550, presentada el 13 de diciembre de 2002. En resumen, la variante de glucoamilasa se mezcla con un pienso que comprende almidón. La glucoamilasa es capaz de degradar almidón resistente para ser usado por el animal.

[0191] Otros objetos y ventajas de la presente exposición resultan evidentes a partir de la presente memoria.

EJEMPLOS

[0192] Los siguientes ejemplos se presentan con el fin de demostrar e ilustrar además determinadas partes representativas de la exposición y aspectos de la presente exposición y no han de considerarse como limitadores del alcance de la misma.

[0193] En la exposición y la sección experimental que se presenta a continuación, se aplican las siguientes abreviaturas: GA (glucoamilasa); GAU (unidad glucoamilasa); % p (porcentaje en peso); °C (grados centígrados); rpm (revoluciones por minuto); H₂O (agua); dH₂O (agua desmineralizada), dIH₂O (agua desmineralizada, filtración Milli-Q) aa o AA (aminoácido); pb (pares de bases); kb (pares de kilobases); kD (kilodaltons); g o gm (gramos); µg (microgramos); mg (miligramos); µL (microlitros); ml y mL (mililitros); mm (milímetros), µm (micrómetro); M (molar); mM (milimolar); µM (micromolar); U (unidades); V (voltios); PM (peso molecular); s (segundo/segundos); min(s) o m(s) (minuto/minutos); hr(s) o h(s) (hora/horas); OD (oxígeno disuelto); ABS (absorbancia); EtOH (etanol); SF (suero fisiológico); m/v (masa/volumen) y PMT (placa de microtitulación); N (Normal); DP1 (monosacáridos); DP2 (disacáridos); DP>3 (oligosacáridos, azúcares con un grado de polimerización superior a 3); ppm (partes por millón).

[0194] A continuación, se describen los métodos utilizados para proporcionar variantes. Sin embargo, cabe destacar que pueden utilizarse diferentes métodos para proporcionar variantes de una molécula precursora y la invención no está limitada a los métodos utilizados en los ejemplos. Se pretende que se pueda utilizar cualquier medio adecuado para realizar variantes y para la selección de variantes.

Análisis de la actividad glucoamilasa pNPG para placas de microtitulación de 96 pocillos:

[0195] Las soluciones de reactivos fueron: tampón NaAc: tampón acetato de sodio 200 mM pH 4,5; Sustrato: p-nitrofenil- α -D-glucopiranosido 50 mM (Sigma N-1377) en tampón NaAc (0,3 g/20ml) y solución de cese de reacción: tampón glicina-NaOH 800 mM pH 10. Se colocaron 30 μ l de sobrenadante filtrado en una nueva PMT de 96 pocillos de base plana. A cada pocillo se añadió 50 μ l de tampón NaAc y 120 μ l de sustrato y se incubó durante 30 minutos a 50 °C (incubadora/agitador HT iEMS de Thermolab systems). Se finalizó la reacción añadiendo 100 μ l de solución de cese de reacción. Se midió la absorbancia a 405 nm en un lector de PMT (Spectramax 384 plus de Molecular Devices) y la actividad se calculó utilizando un coeficiente de extinción molar de 0,011 μ M/cm.

Análisis de estabilidad térmica:

[0196] Se añadió sobrenadante crudo (100 μ l) a 100 μ l tampón NaAc 50 mM pH 4,5. La muestra se dividió equitativamente sobre 2 PMT. Una PMT (placa inicial) se incubó durante 1 hr a 4 °C y la otra PMT (placa residual) se incubó a 60 °C (incubadora/agitador HT iEMS de Thermolab systems) durante 1 hr. La placa residual se enfrió durante 15 min en hielo. Se midió la actividad de ambas placas utilizando el análisis de la aplicación de etanol descrito a continuación, con la siguiente modificación: la cantidad de muestra tomada para el análisis de termoestabilidad es 25 μ l y la cantidad de tampón NaAc 30 mM pH 4,0 es 35 μ l.

[0197] Se calculó la termoestabilidad como % de actividad residual como se muestra a continuación:

$$\frac{ABS(340)_{residual} - vacío}{ABS(340)_{inicial} - vacío} \times 100\%$$

[0198] Se analizó el material sobrenadante crudo para detectar la glucosa restante en el medio de cultivo tras el periodo de crecimiento. Si se hallaba glucosa restante, el valor de absorbancia se restaba de los valores de absorbancia medidos tanto de la actividad inicial como de la actividad residual.

Análisis Bradford para la cuantificación de proteína en placas de microtitulación de 96 pocillos:

[0199] La solución de reactivo era la solución de trabajo Bradford Quickstart (BioRad cat n.º 500-0205). Se colocaron 100 μ l de sobrenadante filtrado 10 kD en una placa nueva con base plana de 96 pocillos. Se añadieron a cada pocillo 200 μ l de reactivo y se incubaron durante 5 minutos a temperatura ambiente. Se midió la absorbancia a 595 nm en un lector de PMT (Spectramax 384 plus de Molecular Devices). Las concentraciones de proteína se calcularon según una curva estándar de albúmina de suero bovino (BSA) (0-50 μ l/ml).

Análisis de actividad de hexoquinasa:

[0200] Cóctel de hexoquinasa: 10-15 minutos antes de su uso, se añadieron 90 ml de agua a un recipiente BoatiL de glucosa HK R1 (kit de análisis de glucosa IL (HK), Instrumental Laboratory n.º 182507-40) y se mezcló gradualmente. Se añadieron 85 μ l de cóctel de hexoquinasa a 100 μ l de dH₂O. Se añadieron 15 μ l de muestra a las mezclas y se incubaron durante 5 minutos en la oscuridad a temperatura ambiente. Se leyó la absorbancia a 340 nm en un lector de PMT. Las concentraciones de glucosa se calcularon según una curva estándar de glucosa (0-1 mg/ml).

Condiciones de análisis de la aplicación de etanol:

[0201] Para preparar el 8 % de solución madre: se suspendieron 8 g de almidón de maíz soluble (Sigma n.º S4180) en 40 ml dH₂O a temperatura ambiente. Se añadieron cincuenta mililitros de dH₂O hirviendo a la suspensión en un matraz de 250 ml y se coció durante 5 minutos. La solución de almidón se enfrió a 25 °C y se reguló el volumen a 100 ml con dH₂O. La solución de trabajo de almidón soluble 4 % (m/v) se preparó mediante la dilución (1:1) de la solución madre con tampón acetato de sodio 100 mM, pH 3,7.

[0202] Para el análisis de detección, se diluyeron 5 μ l de sobrenadante crudo con 175 μ l de tampón NaAc 50 mM pH 4,5 en una PMT de 96 pocillos de base plana. Se añadieron sesenta microlitros de esta dilución a 120 μ l almidón de maíz soluble 4 % y se incubaron a 900 rpm durante 2 horas a 32 °C (incubadora/agitador HT iEMS de Thermolab systems). Se finalizó la reacción añadiendo 90 μ l de solución de cese de reacción enfriada a 4 °C (tampón glicina-NaOH 800 mM, pH 10). La muestra se colocó en hielo. El almidón se centrifugó a 1118 \times g a 15 °C durante 5 minutos (SIGMA 6K15) y se utilizaron 15 μ l de sobrenadante en el análisis de actividad de Hexoquinasa descrito anteriormente con el fin de determinar el contenido de glucosa.

[0203] Se analizó el material sobrenadante crudo para detectar la glucosa restante en el medio de cultivo tras el periodo de crecimiento. Si se hallaba glucosa restante, la cantidad de glucosa producida por la glucoamilasa no se calculaba.

5 Condiciones de análisis de la aplicación del edulcorante:

[0204] Para preparar la solución madre 8 %: se suspendieron 8 g de almidón soluble (Sigma nº S4180) en 40 ml de agua a temperatura ambiente. Se añadieron 50 ml de dH₂O hirviendo a la suspensión en un matraz de 250ml y se coció durante 5 minutos. La solución de almidón se enfrió a 25 °C y se reguló el volumen a 100 ml con dH₂O. La solución de trabajo de almidón soluble 4 % (m/v) se preparó mediante la dilución de la solución madre 1:1 con tampón acetato de sodio 100 mM, pH 4,5.

[0205] Se colocaron cincuenta microlitros de tampón NaAc 80 mM pH 4,5 en una placa de 96 pocillos de base plana. Se añadieron a cada pocillo 120 µl almidón de maíz soluble 4 % y 5 µl de sobrenadante filtrado 10 kD y se incubaron durante 1 hora a 60 °C. Se finalizó la reacción añadiendo 90 µl de solución de cese de reacción enfriada a 4°C (tampón glicina-NaOH 800 mM, pH 10,0). La muestra se colocó en hielo durante 30 minutos. El almidón se centrifugó a 716 rpm a 15 °C durante 5 minutos (Sigma 6K15, centrifugadora) y se utilizaron 15 µl del sobrenadante del análisis de actividad de Hexoquinasa descrito anteriormente con el fin de determinar el contenido de glucosa.

20 Ejemplo 1

Construcción del vector pREP3Y-TrGA

[0206] El casete de expresión TrGA compuesto de la secuencia de ADN (SEQ ID NO: 4) que codifica el péptido señal TrGA, la prosequencia y la proteína madura, que incluye el dominio catalítico, la región de enlace y el dominio de unión a almidón, se clonó en pDONR™201, un vector de entrada Gateway® (Invitrogen, Carlsbad, CA, EE UU). El casete de expresión TrGA se clonó en el vector de destino compatible Gateway pREP3Y-DEST (fig. 3) mediante la reacción de recombinación Gateway®.

[0207] El vector de expresión pRep3Y-TrGA (figura 3B) permitió la expresión de la proteína TrGA (SEQ ID NO: 2) en *Schizosaccharomyces pombe*.

[0208] Se crearon sesenta y cinco bibliotecas de mutagénesis por saturación del sitio (SSM por sus siglas en inglés) de TrGA utilizando el vector de entrada pDONR-TrGA (fig. 2) como plantilla y los cebadores enumerados en la tabla 2. Todos los cebadores de mutagénesis utilizados en los experimentos contienen el código de secuencia de ADN triple NNS (N= A,C,T,G y S=C o G) en la posición que se corresponde con el codón de la secuencia de TrGA que ha de mutarse (SEQ ID NO: 2 e iniciaron una incorporación aleatoria de nucleótidos en esa posición. La construcción de cada biblioteca SSM se inició con dos amplificaciones por PCR utilizando el Gateway forward (pDONR201-FW) y un cebador de mutagénesis inverso específico (tabla 3) y el cebador Gateway reversed (pDONR201-RV) y un cebador de mutagénesis directo específico (tabla 2) (mismas posiciones para los cebadores de mutagénesis). La polimerasa de ADN Phusion de alta fidelidad (Finnzymes OY, Espoo, Finlandia) se utilizó en una amplificación por PCR (0,2 µM cebadores, 25 ciclos) según el protocolo proporcionado por Finnzymes. En resumen, se añadió 1 µl (SEQ ID NO: 1), fragmento de ADN de ambas mezclas de PCR, ambas dirigidas al mismo codón a 48 µl de solución de reacción por PCR nueva junto con los cebadores Gateway FW y Gateway RV (Invitrogen) y se mezcló. Esta amplificación por PCR de fusión (22 ciclos) tuvo como resultado un fragmento de ADN de casete de expresión lineal con un codón de TrGA específico mutado de forma aleatoria y sitios de recombinación Gateway únicos en ambos extremos. La purificación de este fragmento de ADN (Limpieza PCR ChargeSwitch®, Invitrogen, Carlsbad EEUU) y una reacción de recombinación BP (Invitrogen, Carlsbad, USA) con pDONR201 (Invitrogen) generaron un ADN multimérico circular (vector de entrada) que se transformó posteriormente en *E. coli* Max efficiency DH5α (Invitrogen) y se recubrió en un medio 2xTY [Bacto Tryptone (Difco) 16 g/L, Bacto Yeast Extract (Difco) 10 g/L, NaCl 5 g/L] complementado con 50 µg/mL kanamicina.

Tabla 2. Cebadores directos utilizados para generar bibliotecas SSM de TrGA

Cebadores	SECUENCIA ADN 5' A 3'	SEQ ID NO:
pDONR201-FW	TCGCGTTAACGCTAGCATGGATCTC	10
D4F	TCTGTTGACNNS TTCATCAGCACCGAGACGC	11
F5F	TCTGTTGACGACNNSATCAGCACCGAGACGCCTA	12
I12F	ATCAGCACCGAGACGCCTNNSGCACTGAACAATCTTCTTT	13
D24F	CTTTGCAATGTTGGTCTNNSGGATGCCGTGCATTTCGGCA	14

ES 2 795 989 T3

Cebadores	SECUENCIA ADN 5' A 3'	SEQ ID NO:
F29F	CCTGATGGATGCCGTGCANNSSGGCACATCAGCTGGTGCGG	15
I43F	ATTGCATCTCCCAGCACANNSSGACCCGGACTACTATTACA	16
D44F	GCATCTCCCAGCACAAATTNNSCCGGACTACTATTACATGT	17
P45F	TCTCCCAGCACAAATTGACNNSGACTACTATTACATGTGGA	18
D46F	CCCAGCACAAATTGACCCGNNSACTATTACATGTGGACGC	19
Y47F	AGCACAAATTGACCCGGACNNSATTACATGTGGACGCGAG	20
Y49F	ATTGACCCGGACTACTATNNSATGTGGACGCGAGATAGCG	21
W51F	CCGGACTACTATTACATGNNSACGCGAGATAGCGCTTTG	22
Y70F	GACCGCTTACCGAAACGNNSGATGCGGGCCTGCAGCGCC	23
Q75F	ACGTACGATGCGGGCCTGNNSCGCCGCATCGAGCAGTACA	24
R76F	TACGATGCGGGCCTGCAGNNSCGCATCGAGCAGTACATTA	25
P94F	CTCCAGGGCCTCTCTAACNNSTCGGGCTCCCTCGCGGACG	26
D100F	CCCTCGGGCTCCCTCGCGNNSGGCTCTGGTCTCGGCGAGC	27
K114F	AAGTTTGAGTTGACCCTGNNSCCTTTCACCGGCAACTGGG	28
F116F	GAGTTGACCCTGAAGCCTNNSACCGGCAACTGGGGTCGAC	29
N119F	CTGAAGCCTTTCACCGGCNNSGTTGGGTCGACCGCAGCGGG	30
R122F	TTCACCGGCAACTGGGGTNNSCCGCAGCGGGATGGCCCAG	31
R125F	AACTGGGGTCGACCGCAGNNSGATGGCCCAGCTCTGCGAG	32
N146F	AAGTGGCTCATCAACAACNNSSTATCAGTCGACTGTGTCCA	33
Q148F	CTCATCAACAACAACATNNSTCGACTGTGTCCAACGTCA	34
Y169F	CTCAACTATGTTGCCAGNNSGGAACCAAACCGGCTTTG	35
Q172F	GTTGCCAGTACTGGAACNNSACCGGCTTTGACCTCTGGG	36
F175F	TACTGGAACCAAACCGGCNNSGACCTCTGGGAAGAAGTCA	37
W178F	CAAACCGGCTTTGACCTCNNSGAAGAAGTCAATGGGAGCT	38
E180F	GGCTTTGACCTCTGGGAANNSTGCAATGGGAGCTCATTCT	39
V181F	TTTGACCTCTGGGAAGAANNSTGCAATGGGAGCTCATTCTTA	40
Q208F	CTTGCTGCCACTCTTGGCNNSTCGGGAAGCGCTTATTCAT	41
S211F	ACTCTTGCCAGTCGGGANNSGCTTATTCATCTGTTGCTC	42
W228F	TGCTTTCTCCAACGATTCNNSGTGTGCTCTGGTGGATACG	43
N242F	GACTCCAACATCAACACCNNSGAGGGCAGGACTGGCAAGG	44
E243F	TCCAACATCAACACCAACNNSGGCAGGACTGGCAAGGATG	45
R245F	ATCAACACCAACGAGGGCNNSACTGGCAAGGATGTCAACT	46
I292F	GTCGACTCCTTCCGCTCCNNSACGGCGTGAACAAGGGCA	47
G294F	TCCTTCCGCTCCATCTACNNSGTGAACAAGGGCATTCTTG	48
K297F	TCCATCTACGGCGTGAACNNSGGCATTCTGCGGGTGCTG	49
R309F	GCTGCCGTGCCATTGGCNNSSTATGCAGAGGATGTGTACT	50
Y310F	GCCGTCGCCATTGGCCGNNSGCAGAGGATGTGTACTACA	51
D313F	ATTGGCCGGTATGCAGAGNNSGTGTACTACAACGGCAACC	52
V314F	GGCCGGTATGCAGAGGATNNSACTACAACGGCAACCCTT	53
Y315F	CGGTATGCAGAGGATGTGNNSATACAACGGCAACCCTTGGT	54
Y316F	TATGCAGAGGATGTGTACNNSAACGGCAACCCTTGGTATC	55
N317F	GCAGAGGATGTGTACTACNNSGGCAACCCTTGGTATCTTG	56
W321F	TACTACAACGGCAACCCTNNSSTATCTTGCTACATTTGCTG	57

ES 2 795 989 T3

Cebadores	SECUENCIA ADN 5' A 3'	SEQ ID NO:
K340F	GATGCCATCTACGTCTGGNNSAAGACGGGCTCCATCACGG	58
K341F	GCCATCTACGTCTGGAAGNNSACGGGCTCCATCACGGTGA	59
T350F	TCCATCACGGTGACCGCCNNSGCCCTTCTTCCAGG	60
Q356F	ACCTCCCTGGCCTTCTTCNNSGAGCTTGTTCTGGCGTGA	61
T363F	GAGCTTGTTCTGGCGTGNNSGCCGGGACCTACTCCAGCA	62
S368F	GTGACGGCCGGGACCTACNNSAGCAGCTCTTCGACCTTTA	63
S369F	ACGGCCGGGACCTACTCCNNSAGCTCTTCGACCTTTACCA	64
N376F	AGCTCTTCGACCTTTACCNNSATCATCAACGCCGTCTCGA	65
Y395F	CTCAGCGAGGCTGCCAAGNNSGTCCCCGCCGACGGTTCCG	66
A398F	GCTGCCAAGTACGTCCCCNNSGACGGTTCGCTGGCCGAGC	67
S401F	TACGTCCCCGCCGACGGTNNSCTGGCCGAGCAGTTTGACC	68
R408F	CTGGCCGAGCAGTTTGACNNSAACAGCGGCACTCCGCTGT	69
N409F	GCCGAGCAGTTTGACCGCNNSAGCGGCACTCCGCTGTCTG	70
T412F	TTGACCGCAACAGCGGCNNSCCGCTGTCTGCGCTTACC	71
H418F	ACTCCGCTGTCTGCGCTTNNSCTGACGTGGTCGTACGCCT	72
W421F	TCTGCGCTTACCTGACGNNSCTCGTACGCCTCGTTCTTGA	73
R433F	TTGACAGCCACGGCCCGTNNSGCTGGCATCGTGCCCCCT	74
I436F	ACGGCCCGTCCGGCTGGCNNSGTGCCCCCTCGTGGGCCA	75
S451F	AGCGCTAGCACGATCCCCNNSACGTGCTCCGGCGCGTCCG	76

Tabla 3. Cebadores utilizados para generar bibliotecas SSM de TrGA

Cebador	Secuencia ADN 5' a 3'	SEQ ID NO:
pDON201-RV	GTAACATCAGAGATTTTGAGACAC	77
D4R	GCGTCTCGGTGCTGATGAASNNGTCAACAGA	78
F5R	TAGGCGTCTCGGTGCTGATSNNGTCTCAACAGA	79
I12R	AAAGAAGATTGTTCAAGTGCSSNAGGCGTCTCGGTGCTGAT	80
D24R	TGCCGAATGCACGGCATCCSNNAGGACCAACATTGCAAAG	81
F29R	CCGCACCAGCTGATGTGCCSNTGCACGGCATCCATCAGG	82
I43R	TGTAATAGTAGTCCGGTCSNNTGTGCTGGGAGATGCAAT	83
D44R	ACATGTAATAGTAGTCCGGSNNAATTGTGCTGGGAGATGC	84
P45R	TCCACATGTAATAGTAGTCSNNGTCAATTGTGCTGGGAGA	85
D46R	GCGTCCACATGTAATAGTASNNGGGTCAATTGTGCTGGG	86
Y47R	CTCGCGTCCACATGTAATASNNGTCCGGGTCAATTGTGCT	87
Y49R	CGCTATCTCGCGTCCACATSNNATAGTAGTCCGGGTCAAT	88
W51R	CAAGAGCGCTATCTCGCGTSNNCATGTAATAGTAGTCCGG	89
Y70R	GGCGCTGCAGGCCCGCATCSNNGTTCGGTGAAGCGGTC	90
Q75R	TGTAAGTCTCGATGCGGCGSNNCAGGCCCGCATCGTACGT	91
R76R	TAATGTAAGTCTCGATGCGSNNGTGCAGGCCCGCATCGTA	92
P94R	CGTCCGCGAGGGAGCCCGASNNGTTAGAGAGGCCCTGGAG	93
D100R	GCTCGCCGAGACCAGAGCCSNNCGGAGGGAGCCCGAGGG	94
K114R	CCCAGTTGCCGGTGAAGGSNNCAGGGTCAACTCAAATT	95
F116R	GTCGACCCAGTTGCCGGTSNNAGGCTCAGGGTCAACTC	96
N119R	CCCGCTGCGGTGACCCASNNGCCGGTGAAGGCTTCAG	97

ES 2 795 989 T3

Cebador	Secuencia ADN 5' a 3'	SEQ ID NO:
R122R	CTGGGCCATCCCGCTGCGGSNNACCCAGTTGCCGGTGAA	98
R125R	CTCGCAGAGCTGGGCCATCSNNCTGCGGTGACCCAGTT	99
N146R	TGGACACAGTCGACTGA T ASNNGTTGTTGA TGAGCCACTT	100
Q148R	TGACGTTGGACACAGTCGASNNATAGTTGTTGTTGATGAG	101
Y169R	CAAAGCCGGTTTGGTTCCASNNCTGGGCAACATAGTTGAG	102
Q172R	CCCAGAGGTCAAAGCCGGTSNNGTTCCAGTACTGGGCAAC	103
F175R	TGACTTCTTCCCAGAGGTCSNNGCCGGTTTGGTTCCAGTA	104
W178R	AGTCCCATTGACTTCTTCSNNGAGGTCAAAGCCGGTTTG	105
E180R	AGAATGAGCTCCCATTGACSNNTCCCAGAGGTCAAAGCC	106
V181R	TAAAGAATGAGCTCCCATTSNNTTCTTCCCAGAGGTCAA	107
Q208R	ATGAATAAGCGCTTCCCGASNNGCCAAGAGTGGCAGCAAG	108
S211R	GAGCAACAGATGAATAAGCSNNTCCCGACTGGCCAAGAGT	109
W228R	CGTATCCACCAGACGACACSNNGAATCGTTGGAGAAAGCA	110
N242R	CCTTGCCAGTCTGCCCTCSNNGGTGTTGATGTTGGAGTC	111
E243R	CATCCTTGCCAGTCTGCCSNNGTGGTGTGATGTTGGA	112
R245R	AGTTGACATCCTTGCCAGTSNNGCCCTCGTTGGTGTGAT	113
I292R	TGCCCTTGTTACGCCGTASNNGGAGCGGAAGGAGTCGAC	114
G294R	CAGGAATGCCCTTGTTACSNNGTAGATGGAGCGGAAGGA	115
K297R	CAGCACCGGCAGGAATGCCSNNGTTCACGCCGTAGATGGA	116
R309R	AGTACACATCCTCTGCATASNNGCCAATGGCGACGGCAGC	117
Y310R	TGTAGTACACATCCTCTGCSNCCGGCCAATGGCGACGGC	118
D313R	GGTTGCCGTTGT AGT ACACSNNTCTGCA T ACCGGCCAA T	119
V314R	AAGGGTTGCCGTTGTAGTASNATCCTCTGCATACCGGCC	120
Y315R	ACCAAGGGTTGCCGTTGTASNACATCCTCTGCATACCG	121
Y316R	GATACCAAGGGTTGCCGTTSNNGTACACATCCTCTGCATA	122
N317R	CAAGATACCAAGGGTTGCCSNNGTAGTACACATCCTCTGC	123
W321R	CAGCAAATGTAGCAAGATASNAGGGTTGCCGTTGTAGTA	124
K340R	CCGTGATGGAGCCCGTCTTSSNCCAGACGTAGATGGCATC	125
K341R	TCACCGTGATGGAGCCCGTSNNTTCCAGACGTAGATGGC	126
T350R	CCTGGAAGAAGGCCAGGGASNNGCGGTACCGTGATGGA	127
Q356R	TCACGCCAGGAACAAGCTCSNNGAAGAAGGCCAGGGAGGT	128
T363R	TGCTGGAGTAGGTCCCGGCSNNCACGCCAGGAACAAGCTC	129
S368R	TAAAGGTCGAAGAGCTGCTSNNGTAGGTCCCGGCCGTAC	130
S369R	TGGTAAAGGTCTGAAGAGCTSNNGGAGTAGGTCCCGGCCGT	131
N376R	TCGAGACGGCGTTGATGATSNNGGTAAAGGTCTGAAGAGCT	132
Y395R	GCGAACCGTCGGCGGGGACSNNTTGGCAGCCTCGCTGAG	133
A398R	GCTCGGCCAGCGAACCGTCSNNGGGGACGTA CTGGCAGC	134
S401R	GGTCAAAGTCTCGGCCAGSNACCGTCCGGCGGGACGTA	135
R408R	ACAGCGGAGTGCCGCTGTTSNNGTCAAAGTCTCGGCCAG	136
N409R	CAGACAGCGGAGTGCCGCTSNNGCGGTCAAAGTCTCGGC	137
T412R	GGTGAAGCGCAGACAGCGGSNNGCCGCTGTTGCGGTCAA	138
H418R	AGGCGTACGACCACGTACGSNNAAGCGCAGACAGCGGAGT	139
W421R	TCAAGAACGAGGCGTACGASNNGTCAAGTGAAGCGCAGA	140

Cebador	Secuencia ADN 5' a 3'	SEQ ID NO:
R433R	AGGGGGGCACGATGCCAGCSNNACGGGCCGTGGCTGTCAA	141
I436R	TGGCCCACGAGGGGGGCACSNNGCCAGCCCGACGGGCCGT	142
S451R	CGGACGCGCCGGAGCACGTSNNGGGGATCGTGCTAGCGCT	143

5 **[0209]** Para cada biblioteca, tras la incubación durante la noche a 37 °C, se agruparon colonias mediante resuspensión de los clones en SF. A partir de los transformantes *E. coli* agrupados, se aisló el plásmido total (Qiaqen) utilizando técnicas convencionales. En resumen, se añadió 1 µl de la solución de plásmido a 1 µl de la solución de vector de destino pRep3Y (figura 1 A) y se añadió a la mezcla de enzima LR CLON- ASET™ II según el protocolo proporcionado por Invitrogen. Se generó un ADN multimérico circular y se transformó en *E. coli* Max EfficiencyDH5a como describe el proveedor.

10 **[0210]** Tras la incubación durante la noche a 37 °C, se recogieron 96 colonias únicas de cada biblioteca de las placas de agar 2xTY con 100µg/ml ampicilina y se cultivaron durante 24 hrs a 37 °C en una PMT con 200 µl de medio 2xYT con 100 µg/ml ampicilina. Los cultivos se utilizaron para los análisis de secuencia (BaseClear B.V., Leiden, Holanda).

15 **[0211]** Los números de biblioteca variaban desde 1 a 65 con una adición que hacía referencia al codón de la secuencia TrGA que se mutó de forma aleatoria. Tras la selección, cada biblioteca incluía un máximo de 19 variantes TrGA. Estas variantes se transfirieron de forma individual a *Schizosaccharomyces pombe* según las instrucciones del fabricante. (Zymo Research, Orange CA. EE UU).

20 **[0212]** Las transformaciones de *S. pombe* se cubrieron en un medio selectivo (Agar EMM, Qbiogene, Irvine, EE UU n.º Cat. 4110-232) y se incubaron a 28 °C durante 4 días. Los transformantes se purificaron a partir de la placa de transformación mediante el rayado de las colonias en agar EMM.

Ejemplo 2

25 Descripción de las condiciones de crecimiento y el pretratamiento de la muestra

30 **[0213]** Los transformantes *S.pombe* se inocularon en placas de microtitulación de 96 pocillos (PMT) con medio selectivo (2xEMM- C) [64,4 g/L Caldo EMM (Qbiogene n.º Cat. 4110-032), 0,62 g/L Complete Supplement Mixture (CSM-HIS-LEU-TRP, Qbiogene, n.º Cat. 4530-122)] y se incubaron durante la noche a 28 °C. A partir de la placa de microtitulación incubada durante la noche, se inocularon 200 µl del cultivo *S. pombe* en 20 ml de medio líquido 2xEMM-C en un matraz de agitación de 100 ml y se incubaron durante 4 días a 26 °C con 280 rpm en un incubadora de agitación Multitron (Infors AG, Bottmingen, Suiza). A partir del cultivo, se tomaron 2 ml de muestras del cultivo *S. pombe* y se centrifugaron durante 5 min a 14.000 rpm (Sigma). El sobrenadante se transfirió a una configuración de filtro 500 HT 10 kD Vivaspin (VivaScience AG, Hannover, Alemania) y se centrifugó durante 25 min a 1000 g.

35 El material retenido se diluyó de nuevo a volumen inicial original con NaAc 50 mM pH 4,5 complementado con 0,015% Tween-80. Esta solución se utilizó en diferentes ensayos.

Ejemplo 3

40 Construcción de una biblioteca combinatoria de 4 variantes de TrGA

[0214] (A) Se llevaron a cabo experimentos para la construcción de variantes TrGA portando combinaciones de las siguientes mutaciones de sitios únicos: Q172F; Q208N; S211R y V314H. A continuación se muestra un análisis de las variantes:

- 45 a) Q172F; Q208N
 b) Q172F; S211R
 c) Q172F; V314H
 d) Q208N; S211R
 e) Q208N V314H
 50 f) S211R; V314H
 g) Q172F; Q208N; S211R
 h) Q172F; Q208N; V314H
 i) Q172F; S211R; V314H

j) Q208N; S211R; V314H

k) Q172F; Q208N; S211R; V314H

5 [0215] Se utilizó el kit de mutagénesis dirigida en múltiples sitios Quikchange® (QCMS) (Stratagene) para construir la biblioteca. Los cebadores fosforilados 5' utilizados para crear la biblioteca se muestran en la tabla 4. Los resultados óptimos en cuanto a incorporación de los cebadores de longitud completa así como reducción significativa en errores derivados de cebadores se obtuvieron mediante el uso de HPLC PAGE o cualquier otro tipo de cebadores purificados (Invitrogen).

10 **Tabla 4. Cebadores utilizados para construir variantes combinatorias seleccionadas.**

Cebador	Secuencia	SEQ ID NO:
I43R	GCATCTCCCAGCACAC CG AGACCCGGACT ACTA T	144
I43Y	GCATCTCCCAGCACAT AC GACCCGGACTACTAT	145
R76L	GATGCGGGCCTGCAG CTG CGCATCGAGCAGTAC	146
N119T	CTGAAGCCTTTTACC GGC ACT GGGGTTCGACCGCAGCGGG	147
N119Y	TGAAGCCTTTTACC GGCT ACT GGGGTTCGACCGCAGCGGG	148
N119D	CTGAAGCCTTTTACC GGCG ACT GGGGTTCGACCGCAGCGGG	149
N146	AGTGGCTCATCAACAAC CG ASTATCAGTCGACTGTGT	150
N146T	AGTGGCTCATCAACAAC ACCT ATCAGTCGACTGTGT	151
N146W	GTGGCTCATCAACAAT GG TATCAGTCGACTGTGT	152
N146L	AGTGGCTCATCAACAAC CTG TATCAGTCGACTGTGT	153
N146S	AGTGGCTCATCAACAAC TC TATCAGTCGACTGTGT	154
Q172D/E	TTGCCAGTACTGGAAC CG ASACCGCTTTGACCTCTGG	155
Q172V/L	TTGCCAGTACTGGAAC STG ACCGCTTTGACCTCTGG	156
Q172T	TTGCCAGTACTGGAAC ACC ACCGCTTTGACCTCTGG	157
Q172R	TTGCCAGTACTGGAAC CGA ACCGCTTTGACCTCTGG	158
Q172C	TTGCCAGTACTGGAAC TGC ACCGCTTTGACCTCTGG	159

15 [0216] Se utilizó el plásmido plantilla pDONR-TrGA (Figura 2) para crear la biblioteca combinatoria utilizando el kit QCMS de Stratagene. La biblioteca se creó según lo descrito por el proveedor con concentraciones de cebadores modificados utilizadas en las reacciones. De forma específica, se mezclaron 4 µl pDONR-TrGA (25-50 ng) con 11 µl de agua destilada estéril; 1,5 µl de dNTP; 2,5 µl de 10xtampón QCMS; 1 µl de mezcla de enzima y 1 µl de cada mezcla de cebador mutante proporcionando un total de 100 ng de cebadores en cada reacción. Las condiciones de PCR fueron 95 °C durante 1 min, seguido de 30 ciclos de 95 °C durante 1 min, 55 °C durante 1 min y 65 °C durante 6 min en un termociclador MJ Research utilizando tubos de PCR de 0,2 ml de pared fina. El producto de reacción se digirió con 1 µl de DpnI del kit QCMS mediante incubación a 37 °C durante la noche. Se utilizó un kit de purificación PCR (Qiagen) para la purificación de la muestra y se llevó a cabo una segunda ronda de digestión con DpnI (Stratagene) durante 1 hora a 37 °C.

25 [0217] La mezcla de reacción se transformó en *E. coli* Max efficiency DH5α (Invitrogen) y se cubrió en agar selectivo (2xTY complementado con 50 µl kanamicina/ml). Tras la incubación durante la noche a 37 °C, se recogieron 96 colonias únicas para un análisis de secuencias ((BaseClear B.V., Leiden, Holanda). Las variantes combinatorias se clonaron y expresaron en una cepa huésped *T. reesei* según se describe a continuación y en WO 06/060062.

30 [0218] (B) Se realizaron de forma sintética seis bibliotecas combinatorias adicionales (tabla 5) mediante Genart (Regensburg, Alemania) y se analizaron para determinar la estabilidad térmica y en ensayos de aplicación de edulcorante y etanol según se describe en el presente documento.

Tabla 5. Bibliotecas combinatorias

1	D24E,L/ I43F,R/ D44H,N/ F175H/ V181K,L/ V314D,H,K/ T363R
2	D24L,W,Y/ Q208F/ I292F,L,N,V/ G294A,I,Q/ K297A/ Y310F,Q,R
3	V181 F,K,L/ E243A,N,M,R,Y/ I292F,L,N,V/ K297A,D,H,M,N,Q/ N317H/ Y395Q,R

4	D24E,L,Y/ V181F,K,L/ Q208C/F/ G294A,I,Q/ T363R/ N376Q/ N409K,W
5	D24E,L,Y/ V181F,K,L/ I292F,L,N,V/ G294A,I,Q/ E243A,M,N,R,Y/ N409K,W
6	I43R/ E243A,M,N,R,Y/ I292F,L,N,V/ G294A/ K297A,D,H,M,N,Q,S,R,W,Y

Ejemplo 4**Variantes con estabilidad térmica mejorada**

5

[0219] La molécula TrGA precursora bajo las condiciones descritas presentaba una actividad residual entre un 15 y un 44 % (variación diaria). El índice de rendimiento se calculó según la termoestabilidad de TrGA del mismo lote. Los índices de rendimiento son los cocientes $PI = (\text{actividad residual variante})/(\text{actividad residual TrGA})$. Un índice de rendimiento >1 indica una estabilidad mejorada. Las variantes que presentaban un índice de rendimiento de estabilidad térmica de más del 1,0 se muestran en la siguiente tabla 6.

10

Tabla 6. Detección termoestabilidad

Variante	PI Estabilidad Térmica, 60 °C, pH 4,5	Variante	PI Estabilidad Térmica, 60 °C, pH 4,5
D4P	1,05	S368Q	1,10
I12E	1,09	S368R	1,14
I12Y	1,40	S368T	1,15
D24L	1,09	S368W	1,16
D24W	1,13	S369A	1,22
D24Y	1,03	S369D	1,05
I43R	1,28	S369F	1,20
D44N	1,06	S369G	1,05
D44Q	1,10	S369K	1,12
Q75N	1,09	S369L	1,49
R76K	1,03	S369M	1,36
N146D	1,20	S369N	1,25
N146E	1,24	S369P	1,16
N146L	1,10	S369R	1,12
N146V	1,28	S369T	1,25
N146W	1,17	N376F	1,12
Q148D	1,02	N376G	1,26
F175I	1,02	N376H	1,21
F175Y	1,06	N376K	1,40
E180A	1,41	N376L	1,34
E180D	1,02	N376P	1,05
E180G	1,13	N376Q	1,11
E180I	1,41	N376S	1,09
E180L	1,38	N376V	1,19
E180M	1,10	N376W	1,12
E180N	1,27	N376Y	1,05
E180Q	1,72	Y395A	1,05
E180R	1,59	Y395C	1,02
E180V	1,08	Y395F	1,03
E180W	1,30	Y395G	1,13
E180Y	1,31	Y395H	1,10

ES 2 795 989 T3

Variante	PI Estabilidad Térmica, 60 °C, pH 4,5	Variante	PI Estabilidad Térmica, 60 °C, pH 4,5
V181I	1,20	Y395L	1,50
V181K	1,12	Y395N	1,20
V181L	1,06	Y395Q	1,18
V181Q	1,09	Y395R	1,14
V181R	1,07	Y395S	1,13
Q208F	1,06	Y395T	1,04
Q208T	1,17	A398C	1,10
Q208V	1,15	A398D	1,39
S211D	1,10	A398F	1,05
S211E	1,02	A398G	1,17
S211I	1,31	A398H	1,33
S211M	1,90	A398I	1,41
E243A	1,19	A398K	1,47
E243H	1,04	A398L	1,44
E243M	1,53	A398N	1,23
E243N	1,35	A398P	1,38
E243P	1,06	A398Q	1,43
E243R	1,21	A398R	1,59
E243S	1,09	A398S	1,14
E243T	1,48	A398T	1,25
E243Y	1,43	A398V	1,29
I292F	1,17	A398W	1,45
I292L	1,10	A398Y	1,38
I292N	1,31	S401A	1,12
I292V	1,02	S401E	1,08
G294A	1,30	S401I	1,05
G294C	1,41	S401N	1,12
G294D	1,31	S401P	1,15
G294E	1,34	S401R	1,25
G294H	1,17	S401T	1,26
G294I	2,15	S401V	1,18
G294L	2,01	R408A	1,14
G294P	1,13	R408E	1,41
G294Q	1,91	R408G	1,15
G294R	1,34	R408H	1,12
G294V	1,10	R408I	1,19
K297A	1,47	R408K	1,80
K297C	1,10	R408L	1,55
K297D	1,50	R408N	1,09
K297F	1,24	R408Q	1,23
K297G	1,25	R408S	1,17
K297H	1,63	N409A	1,25
K297L	1,62	N409C	1,18

ES 2 795 989 T3

Variante	PI Estabilidad Térmica, 60 °C, pH 4,5	Variante	PI Estabilidad Térmica, 60 °C, pH 4,5
K297M	1,62	N409D	1,21
K297N	1,87	N409E	1,27
K297Q	1,82	N409F	1,32
K297R	1,29	N409G	1,14
K297S	1,22	N409H	1,29
K297T	1,33	N409I	1,56
K297V	1,10	N409K	1,44
K297W	1,85	N409L	1,57
K297Y	1,71	N409M	1,17
R309S	1,08	N409Q	1,03
Y310C	1,06	N409R	1,29
Y310F	1,35	N409V	1,11
Y310L	1,11	N409W	1,58
Y310Q	1,40	T412L	1,10
Y310R	1,61	S451R	1,01
Y315E	1,24		
Y315H	1,48		
Y315L	1,35		
Y315N	1,17		
Y315P	1,19		
Y315Q	1,43		
Y315T	1,34		
Y316D	1,06		
N317H	1,26		
N317Q	1,09		
K340H	1,02		
K340R	1,09		
K341I	1,10		
K341V	1,07		
T350G	1,08		
T350P	1,08		
T350S	1,33		
Q356L	1,20		
T363N	1,30		
S368C	1,12		
S368E	1,07		
S368F	1,16		
S368H	1,26		
S368I	1,15		
S368L	1,33		
S368N	1,21		
S368P	1,05		

Ejemplo 5**Variantes de alto rendimiento a partir de un análisis de detección de etanol**

- 5 **[0220]** Se analizaron las variantes en un análisis de detección de etanol utilizando los análisis descritos anteriormente. La tabla 7 muestra los resultados del análisis de detección para las variantes con un índice de rendimiento (PI) > 1,0 en comparación con el PI de la TrGA precursora. El PI es una medida de actividad específica (actividad/mg enzima). El PI de la actividad específica es el cociente "actividad específica variante/actividad específica WT". El PI de la actividad específica es 1,0 y una variante con un PI > 1,0 presenta una actividad específica que es mayor que la TrGA precursora. La actividad específica es la actividad medida mediante el análisis de detección de etanol dividida por los resultados obtenidos en el análisis Bradford descrito anteriormente.
- 10

Tabla 7. Detección etanol

Variante	P.I. 32 °C, pH 4	Variante	P.I. 32 °C, pH 4
D4A	1,07	K297W	1,70
D4C	1,08	K297Y	1,80
D4E	1,23	R309L	1,43
D4L	1,34	Y310F	1,05
D4R	1,18	Y310Q	1,16
D4S	1,17	Y310R	1,24
F5C	1,35	V314D	1,10
I12L	1,19	V314F	1,04
I12R	1,13	V314H	1,31
D24E	1,60	V314K	1,08
D24L	1,19	V314L	1,02
D24W	1,03	V314N	1,05
D24Y	1,14	V314R	1,06
F29A	1,05	Y316R	1,42
F29C	1,12	Y316W	1,05
F29D	1,20	N317H	1,14
F29E	1,05	N317K	1,02
F29I	1,26	N317S	1,03
F29L	1,42	N317T	1,23
F29Q	1,01	K340D	1,33
F29S	1,07	K340T	1,16
F29V	1,06	K341D	1,04
I43D	1,14	K341F	1,64
I43F	1,33	K341G	1,64
I43R	1,21	K341L	1,04
I43Y	1,05	K341N	1,05
D44E	1,37	K341S	1,06
D44F	1,07	T350A	1,56
D44G	1,03	T350D	1,04
D44H	1,11	T350E	1,59
D44K	1,09	T350H	1,03
D44N	1,07	T350N	1,06
D44S	1,08	T350Q	1,05
D44Y	1,07	T350R	1,02

ES 2 795 989 T3

Variante	P.I. 32 °C, pH 4	Variante	P.I. 32 °C, pH 4
Y70E	1,02	Q356D	1,69
Y70G	1,06	Q356E	1,07
Y70K	1,01	Q356H	1,03
Y70M	1,36	Q356K	1,03
Y70P	1,15	T363A	1,04
Y70R	1,40	T363C	1,54
Y70S	1,04	T363G	1,02
Q75A	1,10	T363H	1,09
Q75K	1,77	T363N	1,02
R76K	1,06	T363R	1,61
R76L	1,11	T363V	1,05
R76M	1,13	T363W	1,08
R76N	1,02	S368D	1,11
R76T	1,04	S368F	1,08
R76V	1,05	S368H	1,04
R76W	1,02	S368L	1,07
R76Y	1,05	S368M	1,03
D100A	1,08	S368N	1,02
D100I	1,14	S368W	1,24
D100L	1,03	S369F	1,68
D100M	1,12	S369M	1,04
D100N	1,06	S369T	1,05
D100P	1,09	N376G	1,05
D100Q	1,14	N376H	1,10
D100T	1,06	N376Q	1,16
D100W	1,19	N376S	1,06
D100Y	1,05	N376T	1,12
N119E	1,02	N376V	1,64
N119F	1,03	Y395A	1,02
N119Y	1,28	Y395C	1,05
N146C	1,11	Y395G	1,02
N146E	1,02	Y395Q	1,63
N146G	1,11	Y395R	1,20
N146H	1,07	Y395S	1,09
N146K	1,06	A398D	1,05
Q148H	1,10	A398P	1,03
Q148N	1,05	S401A	1,04
Q148V	1,18	S401D	1,01
Q148W	1,05	S401G	1,04
Q148Y	1,16	S401N	1,02
Y169D	1,18	S401V	1,06
Y169F	1,10	N409K	1,30
Y169H	1,05	N409L	1,04

ES 2 795 989 T3

Variante	P.I. 32 °C, pH 4	Variante	P.I. 32 °C, pH 4
Y169R	1,02	N409W	1,31
Q172E	1,08	T412A	1,04
Q172G	1,05	T412G	1,06
Q172R	1,22	T412K	1,05
Q172S	1,03	R433Q	1,16
F175C	1,18	I436A	1,32
F175H	1,26	I436H	1,02
F175T	1,28	I436T	1,03
F175W	1,16	S451A	1,03
F175Y	1,05	S451M	1,28
V181F	1,28	S451T	1,09
V181K	1,35	S451Y	1,03
V181L	1,37		
V181R	1,01		
Q208A	1,22		
Q208C	1,17		
Q208F	1,12		
Q208H	1,02		
Q208I	1,02		
Q208L	1,32		
S211A	1,30		
S211E	1,30		
S211G	1,05		
S211L	1,04		
S211M	1,05		
S211R	1,34		
S211W	1,07		
S211Y	1,08		
E243A	1,23		
E243L	1,20		
E243M	1,26		
E243N	1,28		
E243R	1,31		
E243Y	1,25		
I292F	1,23		
I292H	1,04		
I292L	1,21		
I292N	1,27		
I292R	1,02		
I292V	1,24		
G294A	1,91		
G294I	1,92		
G294Q	1,99		

Variante	P.I. 32 °C, pH 4	Variante	P.I. 32 °C, pH 4
K297A	1,82		
K297D	1,87		
K297H	1,79		
K297M	1,91		
K297N	1,87		
K297Q	1,85		
K297R	1,71		
K297S	1,72		

Ejemplo 6

Variantes de alto rendimiento a partir de un análisis de detección de edulcorante

5

[0221] Se analizaron las variantes en un análisis de detección de edulcorante según se describe anteriormente en el presente documento. La tabla 8 muestra los resultados del análisis de detección en el que se muestran las variantes con un índice de rendimiento (PI) > 1,00 en comparación con el PI de la TrGA precursora. El PI es una medida de actividad específica (actividad/mg enzima). El PI de la actividad específica es el cociente "actividad específica variante/actividad específica WT". El PI de la actividad específica es 1,0 y una variante con un PI > 1,0 presenta una actividad específica que es mayor que la TrGA precursora.

10

Tabla 8. Detección edulcorante

Variante	P.I. 60 °C, pH 4,5	Variante	P.I. 60 °C, pH 4,5
D4E	1,09	G294Q	1,02
D4L	1,03	K297A	1,04
D4S	1,07	K297D	1,10
D24E	1,45	K297Q	1,07
D24L	1,31	V314D	1,22
D24Y	1,01	V314H	1,85
I43D	1,05	V314K	1,34
I43F	1,31	V314L	1,13
I43R	1,28	V314N	1,08
D44E	1,09	V314R	1,20
D44H	1,12	V314Y	1,05
D44N	1,31	Y316R	1,20
Y70F	1,26	N317H	1,25
Y70L	1,22	N317K	1,03
Q75K	1,12	K340D	1,21
R76K	1,11	K340E	1,05
R76M	1,03	K341D	1,08
R76P	1,13	K341G	1,22
R76T	1,11	K341L	1,08
R76W	1,07	K341N	1,08
D100Y	1,04	K341S	1,12
N119E	1,12	T350H	1,03
N119Y	1,01	T350L	1,04
N146D	1,05	Q356D	1,31

N146E	1,11	Q356E	1,04
Q148D	1,02	Q356K	1,05
Q148W	1,05	T363C	1,08
Q172H	1,05	T363G	1,04
Q172Y	1,03	T363N	1,02
F175H	1,42	T363R	1,50
F175Y	1,11	S368G	1,04
V181A	1,10	S368M	1,03
V181F	1,01	N376G	1,02
V181K	1,43	N376Q	1,07
V181L	1,42	Y395Q	1,01
Q208C	1,08	A398H	1,03
Q208F	1,20	A398S	1,03
Q208H	1,11	S401A	1,01
Q208L	1,03	N409K	1,19
Q208N	1,03	N409T	1,02
Q208S	1,06	N409W	1,01
Q208T	1,12	T412G	1,06
S211H	1,16	T412S	1,03
S211M	1,16	I436D	1,02
S211R	1,34	I436Q	1,06
S211W	1,09	I436T	1,16
E243A	1,06	S451D	1,01
E243F	1,01	S451E	1,09
E243N	1,05	S451F	1,02
E243R	1,14	S451H	1,11
E243S	1,09	S451T	1,11
E243Y	1,07		
R245A	1,01		
I292N	1,04		
I292V	1,12		
G294A	1,06		

Ejemplo 7

Construcción de vectores y transformación en células huésped *Trichoderma reesei*.

5

A. Construcción de vectores de expresión que comprenden un polinucleótido que codifica una GA variante.

10

[0222] El casete de expresión TrGA que comprende la secuencia de ADN (SEQ ID NO: 4 se clonó en pDONR™201, un vector de entrada Gateway® (Invitrogen, Carlsbad, CA). El casete de expresión TrGA se clonó en el vector de destino compatible Gateway pTrex3g-DEST (figura 5), que también se describe en WO 06/060062 mediante la reacción de recombinación Gateway® LR. El vector de expresión pTrex3g-TrGA (figura 5) permitió la expresión de la proteína TrGA (SEQ ID NO: 2) en un huésped *Trichoderma reesei*. Se construyeron los vectores, que incluían ADNc de TrGA modificado que codificaba al menos las siguientes variantes (1) V314H; (2) S211R; (3) Q208N y (4) Q172F.

15

B. Transformación

[0223] Un vector de expresión que contiene una GA variante se transformó en una cepa huésped *T. reesei* derivada de RL- P37 (IA52) y con diferentes deleciones de genes ($\Delta cbh1$, $\Delta cbh2$, $\Delta eg1$, $\Delta eg2$) utilizando un bombardeo de partículas mediante PDS-1000/Helium System (BioRad n.º Cat. 165-02257). El protocolo se resume a continuación y también se hace referencia a los ejemplos 6 y 11 de WO 05/001036.

5

[0224] Se preparó una suspensión de esporas (aproximadamente 5×10^8 esporas/ml) a partir de la cepa de *T. reesei*. Se extendieron de cien a doscientos microlitros de suspensión de esporas en el centro de las placas de medio de acetamida de medio mínimo (MM). El medio de acetamida MM presentaba la siguiente composición: 0,6 g/L acetamida; 1,68 g/L CsCl; 20 g/L glucosa; 20 g/L KH_2PO_4 ; 0,6 g/L $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$; 1 ml/L 1000X solución de oligoelementos: 20 g/L agar; y pH 5,5; 1 ml/L 400X solución salina de oligoelementos; ácido cítrico 175 g/L, $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 200g/L, $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 16 g/L, $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 3,2 g/L, $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ 1,4 g/L, H_3BO_3 0,8 g/L. Se dejó que la suspensión de esporas se secase en la superficie del medio de acetamida MM.

10

[0225] La transformación siguió las instrucciones de los fabricantes. En resumen, se colocaron 60 mg de partículas de tungsteno M10 en un tubo de microcentrifugación. Se añadió 1 mL de etanol y se dejó reposar durante 15 minutos. Las partículas se centrifugaron a 15 000 rpm durante 15 segundos. Se extrajo el etanol y las partículas se lavaron tres veces con dH_2O estéril antes de añadir 1 mL de glicerol estéril 50 % (v/v). Se colocaron 25 μl de suspensión de partículas de tungsteno en un tubo de microcentrifugación. Mientras se mezclaba de forma continua con agitador vortex, se añadió lo siguiente: 0,5-5 μl (100-200 ng/ μl) de ADN plasmídico, 25 μl de 2,5 M CaCl_2 y 10 μl de 0,1 M espermidina. Las partículas se centrifugaron durante 3 segundos. Se extrajo el sobrenadante y se lavaron las partículas con 200 μl de etanol 70 % (v/v) y se centrifugaron durante 3 segundos. Se extrajo el sobrenadante y se añadieron 24 μl de etanol 100 %, se mezcló mediante pipeteo y el tubo se colocó en un baño ultrasónico, se extrajeron 8 μl partes alícuotas de partículas y se colocaron en el centro de los discos microportadores que se fijaron en un desecador. Una vez se secó la suspensión de tungsteno/ADN, el disco microportador se colocó en la cámara de bombardeo junto con la placa de acetamida MM con esporas y el proceso de bombardeo se llevó a cabo según las instrucciones de los fabricantes. Tras el bombardeo de las esporas cubiertas con las partículas de tungsteno/ADN, las placas se incubaron a 28 °C. Las colonias transformadas se recogieron en placas limpias de acetamida MM tras 4 días (Penttillä *et al.* (1987) Gene 61: 155-164). C. Demostración de la actividad GA a partir de la TrGA variante expresada en células transformadas.

15

20

25

30

[0226] Tras 5 días de cultivo en placas de acetamida MM, los transformantes que mostraban morfología estable se inocularon en matraces de agitación de 250 ml con 30 ml de medio Proflo. El medio Proflo contenía: 30 g/L α -lactosa; 6,5 g/L $(\text{NH}_4)_2\text{SO}_4$; 2 g/L KH_2PO_4 ; 0,3 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$; 0,2 g/L $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$; 1 ml/L 400X solución salina de oligoelementos: ácido cítrico 175 g/L, $\text{FeSO}_4 \cdot 7\text{H}_2\text{O}$ 200g/L, $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$ 16 g/L, $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$ 3,2 g/L, $\text{MnSO}_4 \cdot \text{H}_2\text{O}$ 1,4 g/L, H_3BO_3 0,8 g/L; 2 ml/L 10 % Tween 80; 22,5 g/L harina de semilla de algodón ProFlo (Traders protein, Memphis, TN); 0,72 g/L CaCO_3 . Tras dos días de cultivo a 28 °C y 140 rpm, 10 % del cultivo Proflo se transfirió a un matraz de agitación de 250 ml con 30 ml de Medio definido de lactosa. 5 g/L $(\text{NH}_4)_2\text{SO}_4$; 33 g/L tampón 1,4-Piperazinebis (ácido propanosulfónico); 9 g/L casamino ácidos, 4,5 g/L KH_2PO_4 , 1,0 g/L $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 5 ml/L antiespumante Mazu DF60-P (Mazur Chemicals, IL), 1 ml/L de 1000X solución de oligoelementos; pH 5,5; 40 ml/L de solución lactosa 40 % (p/v) se añadió al medio tras la esterilización. Los matraces de agitación con el medio definido de lactosa se incubaron a 28 °C, 140 rpm durante 4-5 días.

35

40

[0227] Las muestras del sobrenadante de cultivo se mezclaron con un volumen adecuado de 2X tampón de carga de muestra con agente reductor. Se extrajo el micelio mediante centrifugación y se analizó el sobrenadante para detectar el contenido de proteína total (BCA Protein Assay Kit, Pierce Cat. n.º 23225).

45

[0228] La actividad GA se midió utilizando el análisis p-nitrofenil- α -D-glucopiranosido (pNPG) con pNPG como un sustrato (Sigma N-1377). En este análisis se mide la capacidad de la glucoamilasa de catalizar la hidrólisis de p-nitrofenil- α -D-glucopiranosido (pNPG) a glucosa y p-nitrofenol. Con un pH alcalino, el nitrofenol forma un color amarillo que es proporcional a la actividad de glucoamilasa y se controla a 405 nm y se compara frente a la norma de enzima medida como una GAU (Elder, M. T. y Montgomery R. S., Glucoamylase activity in industrial enzyme preparations using colorimetric enzymatic method, Journal of AOAC International, vol. 78(2), 1995). Una GAU se define como la cantidad de enzima que producirá 1 gm de azúcar reductora calculada como glucosa por hora a partir de un sustrato de almidón soluble (4 % ds) con pH 4,2 y 60 °C.

50

55

[0229] El perfil proteínico se determinó mediante electroforesis PAGE sobre NuPAGE® Novex 10% Bis-Tris Gel con tampón de migración MES SDS (Invitrogen, Carlsbad, CA, EE UU).

Ejemplo 8

60

Ensayo de aplicaciones a pequeña escala de variantes elegidas sobre almidón soluble

[0230] Las cepas huésped *Trichoderma reesei* que expresan las variantes únicas a) V314H, b) S211R, c) Q172F y d) Q208N se cultivaron en fermentadores de 14 L de alimentación discontinua a 34 °C, pH 3,5 en medios de nutrientes que incluían glucosa (Cerelose DE99), KH₂PO₄, MgSO₄•7H₂O, (NH₄)₂SO₄, CaCl₂•2H₂O, oligoelementos y antiespumante Mazu (DF6000K). Tras la disminución de glucosa, la temperatura del cultivo y el pH se cambiaron a 28 °C y 4,0 respectivamente. Se extrajo el material celular mediante filtración y se recogieron los sobrenadantes de cultivo y se concentraron para contener más del 90 % de glucoamilasa como proteína total.

[0231] Se determinaron diferentes propiedades cinéticas para la producción de glucosa sobre almidón de patata soluble con pH 4,3 a 32 °C y a 60 °C y se compararon con la TrGA de tipo silvestre. Cada una de las cuatro variantes mostraron valores de V_{max} aumentada (μM glucosa/s) en comparación con el tipo silvestre (TrGA) lo que indica índices catalíticos elevados (k_{cat} (sec⁻¹)). La figura 6 ilustra la V_{max} de dos réplicas para cada temperatura analizada.

Ejemplo 9

Método para determinar el rendimiento sobre la producción de EtOH de variantes

[0232] La validación de la detección se llevó a cabo sobre las variantes que se identificaron con un índice de rendimiento superior en comparación con la TrGA precursora (véase tabla 7/8) utilizando un ensayo de aplicación de etanol de pequeña escala novedoso. Se derivaron veinticuatro variantes de la evaluación del sitio y se eligieron bibliotecas combinatorias (tabla 9) y se transformaron directamente en *T. reesei* para su expresión y análisis a mayor escala. Se analizaron las variantes para detectar el despliegue térmico utilizando la calorimetría diferencial de barrido (análisis DSC descrito a continuación en el presente documento) y el rendimiento utilizando un análisis de aplicación de etanol a pequeña escala secundario y novedoso. El método consistía en dos etapas: 1) inyección de variantes sobre una columna de intercambio aniónico para determinar de forma precisa la concentración proteica; y 2) titulación de variantes con tres concentraciones de TrGA diferentes (0,3-0,15-0,075g/ 28 g ds) con el fin de calcular su rendimiento sobre la producción de etanol en relación con la molécula de tipo silvestre.

Tabla 9. Lista de variantes combinatorias

variante	mutación
LR8	Q172F/Q208N
LR6	Q172F/Q208N/V314H
LR12	Q172F/S211R
SW3-1	D24E/I43R/D44N/F175H/V181L/V314H/T353R
SW3-2	D24L/I43F/D44N/F175H/V181L/V314H/T353R
ET4-1	D24L/Q208Q/I292V/G294A/K297A/Y310R
ET4-2	D24W/Q208F/I292V/G294Q/K297A/Y310R
ET5-1	V181L/E243A/I292N/K297N/N317N/Y395Q
ET5-2	V181L/E243R/I292F/K297A/N317N/Y395Q
ET7-1	D24Y/V181L/Q208C/G294A/T353R/N375N/N409W
ET7-2	D24L/V181L/Q208C/G294A/T353R/N375Q/N409W
ET8-1	D24E/V181K/E243Y/I292V/G294Q/N409K
ET8-2	D24E/V181F/E243R/I292N/G294I/N409W
ET9-1	I043R/E243R/I292F/G294A/K297A
ET9-2	I043R/E243R/I292L/G294A/K297M

Determinación y purificación proteica

[0233] Se purificó una preparación de enzima cruda utilizando un sistema AKTA explorer 100 FPLC (Amersham Biosciences, Piscataway, NJ). β -Ciclodextrina (Sigma-Aldrich, Zwijndrecht, Holanda; 85,608-8) se unió a esferas de Sepharose activadas con epoxy (GE Healthcare, Diegem, Bélgica; 17-0480-01). La columna se utilizó para capturar glucoamilasas a partir de la preparación de enzimas. La enzima se eluyó de las esferas utilizando tampón Tris 25 mM pH 7,5 o tampón acetato de sodio 50 mM pH 4,3 con α -ciclodextrina 10 mM (Sigma, 28705). Las muestras purificadas se analizaron mediante SDS-PAGE. Con el fin de determinar de forma precisa la concentración de proteína de las variantes se desarrolló un método de determinación de proteína basado en FPLC.

La concentración proteica de la molécula TrGA del marcador purificado se determinó en primer lugar utilizando un protocolo Bradford convencional (Bio-Rad cat n.º 500-0205). Posteriormente, las muestras purificadas se inyectaron en una columna de 1 ml ResourceQ (GE Healthcare) y la enzima se eluyó con tampón Tris 25 mM pH 5 con NaCl 500 mM. Se determinó el área pico y se calculó la concentración proteica en relación al área pico de la TrGA convencional con concentración conocida.

Aplicación de EtOH a pequeña escala

[0234] La tabla 10 resume la producción de etanol y azúcares (DP1, DP2, DP>3) mediante diferentes variantes combinatorias. Se obtuvo una muestra de sustrato licuado de pulpa de maíz y se diluyó a 26 % DS utilizando finos residuos de la destilería. El pH de la suspensión se reguló a pH 4,3 utilizando 4N ácido sulfúrico. Una parte alícuota de 100 g de pulpa se colocó en un baño de agua a 32 °C y se dejó que se equilibrara. Tras una adición de 100 µl 400 ppm urea, se añadió 1 ml de muestra de enzima TrGA variante purificada (150 µg/ml) o TrGA purificada (300, 150, 75 µg/ml) a cada muestra de pulpa de maíz. Finalmente, se añadieron a cada muestra 333 µl de levadura Red Star Red hidratada 30 minutos 15 g en 45 ml de solución de agua DI (Lesaffre yeast Corp. Milwaukee, WI). Las muestras se tomaron tras 5, 21, 28, 48 y 52 horas y se analizaron mediante HPLC utilizando una columna 9 Aminex HPX-87H (Bio-Rad).

Determinaciones de etanol y carbohidratos

[0235] Se llenó un tubo de centrifugación eppendorf de 2 ml con cerveza del fermentador y se enfrió en hielo durante 10 minutos. La muestra se centrifugó durante 3 minutos a 14,000 x g y 500 µl del sobrenadante se transfirieron a un tubo de ensayo con 50 µl de solución de desactivación (1,1 N ácido sulfúrico) y se dejó reposar durante 5 minutos. Se añadieron 5,0 ml de agua al tubo de ensayo y, a continuación, se filtró en una placa de filtro de 0,22 µm (multipantalla, Millipore, Amsterdam, Holanda) y se analizó en HPLC. Temperatura de la columna: 60 °C; fase móvil: 0,01 N ácido sulfúrico; índice de flujo 0,6 ml/min; detector: RI; volumen de inyección: 20 µl. La columna separa las moléculas según la carga y el peso molecular; DP1 (monosacáridos); DP2 (disacáridos); DP3 (trisacáridos); DP>3 (azúcares oligosacáridas con un grado de polimerización superior a 3); ácido succínico; ácido láctico; glicerol; metanol; etanol.

Análisis DSC

[0236] La temperatura de fusión de las muestras de enzima purificadas (0,2-0,4 mg/ml) se determinó utilizando la calorimetría diferencial de barrido (DSC).

Tabla 10. Producción de etanol y sacáridos

	horas	DP>3 (m/v)%	DP2 (m/v)%	DP1 (m/v)%	Etanol (v/v)%
TrGA (0,3 mg)	5,5	3,46	2,70	0,91	1,02
	21,5	3,40	0,50	0,06	6,80
	28,5	1,68	1,46	0,07	8,13
	46	0,04	0,71	0,06	10,21
	52,5	0,04	0,45	0,03	10,96
TrGA (0,150 mg)	5,5	3,40	2,43	0,15	1,00
	21,5	3,78	0,21	0,03	4,23
	28,5	3,86	0,20	0,03	5,07
	46	2,73	0,52	0,06	7,86
	52,5	1,70	0,87	0,04	7,92
TrGA (0,075 mg)	5,5	3,43	2,16	-0,01	0,94
	21,5	3,54	0,20	0,03	3,10
	28,5	3,43	0,18	0,03	3,14
	46	3,93	0,18	0,05	4,65
	52,5	4,01	0,18	0,03	4,79
ET7-1	5,5	3,45	2,53	0,21	1,00
	21,5	3,94	0,22	0,04	4,77
	28,5	3,89	0,23	0,04	5,58

ES 2 795 989 T3

	horas	DP>3 (m/v)%	DP2 (m/v)%	DP1 (m/v)%	Etanol (v/v)%
	46	1,58	1,22	0,06	8,64
	52,5	0,62	1,50	0,04	9,14
LR8	5,5	3,43	2,50	0,17	1,00
	21,5	3,96	0,22	0,04	4,79
	28,5	3,86	0,21	0,04	6,21
	46	1,27	1,11	0,07	9,17
	52,5	0,45	1,24	0,04	8,73
LR12	5,5	3,47	2,51	0,16	1,05
	21,5	3,86	0,22	0,04	4,44
	28,5	3,94	0,22	0,04	5,30
	46	2,09	1,08	0,07	8,56
	52,5	0,99	1,52	0,04	9,16
LR6	5,5	3,37	2,44	0,18	0,96
	21,5	3,88	0,21	0,04	4,44
	28,5	3,90	0,20	0,04	5,10
	46	2,44	0,64	0,08	8,59
	52,5	1,27	1,01	0,04	8,97
ET8-1	5,5	3,46	2,53	0,22	0,99
	21,5	3,99	0,21	0,04	4,86
	28,5	3,90	0,21	0,04	5,76
	46	1,29	1,11	0,08	8,94
	52,5	0,47	1,25	0,04	9,56
ET7-2	5,5	3,57	2,46	0,17	1,02
	21,5	4,26	0,21	0,03	4,21
	28,5	4,37	0,20	0,04	5,14
	46	3,87	0,27	0,05	7,21
	52,5	3,27	0,33	0,03	8,07

[0237] La tabla 11 representa la producción de etanol final y el rendimiento de las variantes con 0,15 mg de dosis. El rendimiento se calculó mediante interpolación de los valores 0,3 mg y 0,15 mg de la TrGA por los valores de las variantes.

5

Tabla 11: Producción etanol

variante	EtOH % (v/v)	Rendimiento en relación con TrGA
TrGA 0,3 mg	10,21	
TrGA 0,15 mg	7,86	1,00
TrGA 0,075 mg	4,65	
ET7-1	8,64	1,33
LR8	9,17	1,56
LR12	8,56	1,30
LR6	8,59	1,31
ET8-1	8,94	1,46
ET7-2	7,21	0,72

[0238] Todas las variantes combinatorias excepto ET7-2 funcionaron mejor que el tipo silvestre de TrGA. LR8 tuvo el mejor resultado con un rendimiento mejorado de 1,56.

[0239] La tabla 12 muestra un resumen de todas las variantes combinatorias y de sitio único sometidas a prueba por medio del análisis de aplicación de etanol a pequeña escala. Las variantes que están sombreadas en la tabla 12 funcionaron mejor que la TrGA y también presentaron una temperatura de despliegue térmico superior (dT_m).

Tabla 12. Rendimiento y despliegue térmico de variantes en relación con TrGA

Variante	X TrGA	dT_m
LR8	1.56	0.30
ET8-1	1.46	1.60
ET7-1	1.33	0.90
LR6	1.31	0.73
LR12	1.30	-0.13
ET5-2	1.29	-0.31
ET4-2	1.27	0.71
Y213C	1.27	0.03
ET4-1	1.21	-2.64
SW3-1	1.21	0.29
Q172F	1.18	-0.01
V314H	1.17	-0.44
G294I	1.16	-0.22
S211R	1.12	-0.44
Q208N	1.08	-0.22
ET9-1	1.04	-1.33
K297A	1.03	-1.11
SW3-2	1.00	0.88
TrGA	1.00	0.00
G294Q	0.99	-0.86
ET8/2	0.93	-0.01
P94N	0.76	-5.17
ET5-1	0.76	-3.28
ET7-2	0.72	-3.59
S214L/C222F	0.70	-3.98

5

[0240] Los resultados mostraron que la cromatografía (FPLC) era una herramienta útil para determinar de forma precisa la concentración de proteína. Los resultados también mostraron que la titulación de las variantes con tres concentraciones TrGA era un método valioso para determinar el rendimiento de variantes a pequeña escala. Diferentes variantes funcionaron mejor que la TrGA de tipo silvestre (véase la tabla 12) y también presentaban una temperatura de despliegue térmico superior y las variantes que no funcionaron tan bien como la TrGA también presentaban una T_m inferior.

10

Ejemplo 10: Determinación de actividad específica de un conjunto seleccionado de variantes combinatorias y de sitio único y especificidad de sustrato de LR8

15

[0241] Se analizó la actividad específica de un conjunto de variantes combinatorias y diferentes variantes de sitio único que se utilizaron para construir variantes combinatorias (tabla 13). LR8 (PI 1,56 determinado con análisis de aplicación a pequeña escala) se estudió además con respecto a la especificidad de sustrato. Esto se realizó mediante la configuración de un análisis PMT con el fin de determinar los índices de producción de glucosa de variantes GA y para determinar la especificidad de sustrato de la variante LR8. Se halló que el análisis de PMT discriminaba entre variantes y todas las variantes excepto ET7-1 mostraron índices superiores a la glucoamilasa *Trichoderma reesei* de tipo salvaje (wt). Además, diferentes variantes (LR8/ET8/Q172F) funcionaron un 20-30 % mejor que TrGA. LR8 funcionó mejor sobre almidón de maíz soluble y dos muestras diferentes de sustrato licuado de pulpa de maíz en comparación con la de tipo salvaje.

20

25

[0242] Los sustratos utilizados en los siguientes experimentos fueron la solución madre de almidón de maíz soluble preparada como se muestra a continuación: 8 g de almidón de maíz soluble (Sigma n.º S4180) se disolvieron en 100 ml de agua milliQ y se calentaron en un microondas durante 1 minuto. La dispersión se hirvió durante 5 minutos y tras el enfriado se reguló el volumen a 100 ml. Se preparó almidón de maíz soluble 4 % mediante la dilución de la solución madre 1:1 con tampón NaAC 100 mM, pH 4. En un experimento, se preparó un sustrato licuado de maíz (NE) utilizando un analizador de humedad para medir el % ds, a continuación el sustrato se diluyó 7,5 x con

30

NaAc 50 mM para conseguir finalmente 4 % ds. El sustrato se centrifugó durante 5' a 2000 x g y el sobrenadante se filtró con un filtro 0,22 µm. En otro experimento, se preparó un sustrato licuado de maíz (BSE) del mismo modo, excepto que el sustrato se diluyó 10x antes de la centrifugación.

5 [0243] La enzima se diluyó utilizando la solución madre de 150 µg enzima/ml (3 µg/180 µl mezcla de reacción). Las soluciones se diluyeron además con NaAc 50 mM pH 4,0 como se muestra a continuación: 300 ng (10 x), 200 ng, 150 ng, 100 ng, 75 ng, 50 ng, 25 ng, 10 ng/180 µl mezcla de reacción.

10 [0244] El análisis se llevó a cabo como se muestra a continuación: 40 µl NaAc 50 mM pH 4,0, 120 µl almidón de maíz soluble 4 % y 20 µl de enzima se añadieron a cada pocillo. Las muestras se incubaron durante 2 h a 32 °C 900 rpm y se terminaron en hielo tras la adición de 90 µl tampón glicina-NaOH 800 mM pH 10 durante 5 min. La placa se centrifugó durante 5 min a 2000 rpm a 15°C. Se añadieron 85 µl agua milliQ y 100 µl cóctel hexoquinasa (kit análisis de glucosa II HK), Instrumental Laboratory nº 182507-40) y 20 µl de sobrenadante a una placa nueva. Para una línea de calibración de glucosa (0-1 mg/ml) se añadió en su lugar un almacenamiento de glucosa 20 µl.

15 Las placas se incubaron durante 10 min a temperatura ambiente en la oscuridad seguido de la medición de absorción a 340 nm utilizando el Spectramax.

Tabla 13
Rendimiento relativo a peso

gráfico 1	300	200	150	100	75	50	25	10 ng GA
TrGA	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
ET7-1	0.94	0.98	0.91	0.92	0.63	0.38	3.50	0.53
LR8	0.93	1.00	1.22	0.91	0.64	0.06	-0.89	-3.99
LR12	0.94	0.99	1.17	0.83	0.66	-0.15	-0.99	0.62
LR6	0.94	1.01	1.16	0.99	0.55	0.17	-1.67	-4.36
ET8-1	0.87	0.95	1.29	1.02	0.62	-0.05	-1.64	-3.42
ET7-2	0.94	1.00	1.24	0.93	0.50	0.22	-0.62	-1.40
gráfico 2								
V314H	0.95	0.97	1.24	0.91	0.78	0.54	0.74	1.89
G294Q	0.98	0.99	1.03	1.21	1.10	0.19	-0.34	-1.12
S211R	0.97	0.99	1.21	1.05	0.94	0.27	-1.32	0.53
Q208N	0.97	1.01	1.13	0.95	0.90	-0.05	-1.01	-2.81
Q172F	0.99	1.04	1.31	1.23	1.32	0.38	-1.59	-2.20
G294I	0.91	0.96	1.07	1.25	0.80	0.09	-1.20	-3.66
P94N	0.98	1.00	1.24	1.09	1.05	0.29	6.29	-3.80
TrGA	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

20 [0245] Los resultados del análisis para determinar los índices de producción de glucosa de las variantes GA se muestran en la figura 7 y 8. En estas figuras, el rendimiento relativo a TrGA se calculó por cantidad de enzima añadida. Se extrajeron conclusiones de la región lineal del gráfico a 150 ng de enzima. Los resultados en las figuras 7, 8 y la tabla 13 mostraron que LR8, ET8, ET7-2, S211R, Q172F y P94N funcionaron mejor que el tipo silvestre sobre una gama lineal.

25 [0246] Con el fin de determinar la especificidad de sustrato de LR8, se analizó el rendimiento de LR8 y TrGA de tipo silvestre sobre sustratos (almidón de maíz soluble y los dos sustratos de pulpa de maíz producidos en el Ejemplo 10) utilizados en la detección y aplicación. Cuando se analizaron mediante HPLC, los sustratos mostraron una diferencia en el grado del patrón de polimerización (DP) (véanse las figuras 9-11). En NE y BSE DP1->=DP4 está presente mientras que el almidón de maíz soluble consiste en al menos cuatro o más moléculas de glucosa.

30 En todos los sustratos LR8 funcionó mejor que el tipo silvestre (véanse las figuras 9, 10 y 11).

Ejemplo 11: Detección y caracterización de variantes expresadas en *Trichoderma reesei* utilizando bibliotecas de evaluación del sitio (SEL) en el vector pTTT

35 [0247] Se crearon variantes de TrGA adicionales, especialmente variantes con sustituciones en el SBD y se detectaron directamente en *Trichoderma reesei*. De forma similar al Ejemplo 1, se crearon otras diez bibliotecas de mutagénesis por saturación del sitio (SSM) de TrGA utilizando el vector de entrada pDONR-TrGA como plantilla y los cebadores enumerados en la tabla 14. Los sitios incluyen: N61, G73, L417, T430, A431, E503, Q511, A535, A539 y N563. Entre estos sitios, E503, Q511, A535, A539, y N563 se encuentran dentro del dominio de unión a almidón de la TrGA. Por consiguiente, la recombinación se llevó a cabo con el vector pTTT-Dest (figura 14) utilizando la mezcla de enzima LR CLONASE™ II según el protocolo proporcionado por Invitrogen. Los productos de recombinación se transformaron en *E. coli* Max efficiency DH5α (Invitrogen) y se cubrieron sobre medio 2xTY [Bacto Tryptone (Difco) 16 g/L, Bacto Yeast Extract (Difco) 10 g/L, NaCl 5 g/L] complementado con 100 µg/ml ampicilina. Tras la incubación durante la noche a 37 °C, se recogieron 96 colonias únicas de cada biblioteca de las

40 placas de agar 2xTY con 100 µg/ml ampicilina y se cultivaron durante 24 hrs a 37 °C en una PMT con 200 µl de

45

medio 2xYT con 100 µg/ml ampicilina. Los cultivos se utilizaron para los análisis de secuencia (analyzer de secuencia ABI3100, Applied Biosystems). Cada biblioteca contenía de 15 a 19 variantes de TrGA diferentes en el vector de expresión final. Estas variantes se transformaron de forma individual en *T. reesei*.

5 **Tabla 14 Cebadores utilizados para generar bibliotecas SSM TrGA adicionales**

Posición AA	D/I	Secuencia ADN 5' a 3'
61	D	AGATAGCGCTCTTGTCTTCAAGNNSCTCATCGACCGC
61	I	CTTGAAGACAAGAGCGCTATC
73	D	CTTCACCGAAACGTACGATGCGNNSCTGCAGCGCCGC
73	I	CGCATCGTACGTTTCGGTGAA
417	D	CAGCGGCACTCCGCTGTCTGCGNNSCACCTGACGTGGT
417	I	CGCAGACAGCGGAGTGCCGCT
430	D	GTACGCCTCGTTCTTGACAGCCNNSGCCCGTCGGGCT
430	I	GGCTGTCAAGAACGAGGCGTA
431	D	CGCCTCGTTCTTGACAGCCACGNNSCGTCGGGCTGGC
431	I	CGTGGCTGTCAAGAACGAGGC
503	D	CTCCGTGGCCGTACCTTCCACNNSCTCGTGTCGACACA
503	I	GTGGAAGGTGACGGCCACGGA
511	D	GCTCGTGTCGACACAGTTTGGCNSACGGTCAAGGTG
511	I	GCCAAACTGTGTCGACACGAG
535	D	GAGCGCCGCCGTGGCTCTGGACNNSGTCAACTATGCCGATA
535	I	GTCCAGAGCCACGGCGGCGCTCGTGCT
539	D	GGCTCTGGACGCCGTCAACTATNNSGATAACCACCCCTGT
539	I	ATAGTTGACGGCGTCCAGAGCCACGGC
563	D	CGTCGTGGAGTACAAGTACATCNSGTGGGCCAAGATGGCTCC
563	I	GATGTACTTGTACTCCACGACGTCTCC

- 10 **[0248]** Los SEL se transformaron en *T. reesei* utilizando el método PEG-protoplasto (véase p. ej., Pentillä *et al.* (1987) Gene 61: 155-164). El huésped *T. reesei* es una cepa derivada de RL-P37 (IA52) con cuatro deleciones génicas ($\Delta cbh1$, $\Delta cbh2$, $\Delta eg1$, $\Delta eg2$; es decir, "con deleción cuádruple"; véase la patente estadounidense n.º 5847276, WO 92/06184 y WO 05/001036). Las mezclas de transformación con un contenido de hasta 600 ng de ADN y $1-5 \times 10^5$ protoplastos en un volumen total de 25 µl se trataron con 200 ml de 25 % solución PEG, diluida con 2 volúmenes de 1,2 M solución sorbitol, mezclado con MM agarosa superior selectiva 3 % con acetamida y se vertió en agarosa selectiva 2 % con acetamida bien en placas de microtitulación de 24 pocillos. Las placas se
- 15 regeneradas en cada pocillo individual se recogieron de las placas utilizando una solución de 0,85 % NaCl, 0,015% Tween 80. Las suspensiones de esporas se utilizaron para inocular fermentaciones en PMT de 96 pocillos. En el caso de PMT de 24 pocillos, se introdujo una etapa adicional de recubrimiento sobre una PMT de 24 pocillos nueva con MM acetamida selectiva con el fin de enriquecer el número de esporas.
- 20 **[0249]** Los transformantes se fermentaron en PMT y se utilizaron para los análisis los sobrenadantes de cultivo con variantes de proteína expresadas. En resumen, las PMT con 200 µl de medio LD-GSM se inocularon por cuadruplicado con suspensiones de esporas de transformantes de *T. reesei* que expresan variantes de TrGA (más de 10^4 esporas por pocillo). Las placas se incubaron a 28 °C con agitación a 230 rpm y 80 % de humedad durante 6 días. Los sobrenadantes de cultivo se recolectaron mediante filtración al vacío. Los sobrenadantes se utilizaron
- 25 en diferentes análisis para la detección de variantes con propiedades mejoradas.

[0250] Las variantes que muestran un índice de rendimiento superior a 1,0 para la estabilidad térmica, actividad específica y tanto estabilidad térmica como actividad específica se muestran en las tablas 15-17.

Tabla 15. Detección de estabilidad térmica de variantes de TrGA adicionales

Variantes	PI de estabilidad térmica	Variantes	PI de estabilidad térmica
N061V	1,06	E503A	1,43
G073F	1,44	E503C	1,39
G073M	1,01	E503S	1,02
G073N	1,10	E503T	1,04
G073W	1,36	E503V	1,68
L417I	1,04	Q511A	1,11
L417K	1,20	Q511D	1,12
L417M	1,05	Q511H	1,33
L417Q	1,04	Q511N	1,14
L417R	1,20	Q511S	1,15
L417V	1,07	A535D	1,12
L417Y	1,01	A535K	1,24
T430A	1,05	A535N	1,37
T430E	1,02	A535P	1,59
T430F	1,06	A535R	1,26
T430H	1,10	A539E	1,32
T430I	1,04	A539H	1,17
T430K	1,08	A539M	1,05
T430M	1,17	A539R	1,36
T430N	1,13	A539S	1,30
T430Q	1,05	N563A	1,14
T430R	1,13	N563C	1,49
T430S	1,17	N563E	1,44
T430V	1,05	N563I	1,65
A431I	1,03	N563K	1,77
A431N	1,03	N563L	1,60
A431P	1,08	N563Q	1,29
A431R	1,08	N563T	1,31
A431V	1,03	N563V	1,53

Tabla 16. Detección de actividad específica de variantes de TrGA adicionales

Variantes	PI de actividad específica	Variantes	PI de actividad específica
N061D	1,05	A4311	1,20
N0611	1,21	A431L	1,21
N061L	1,18	A431M	1,12
N061Q	1,08	A431Q	1,22
N061V	1,11	A431R	1,11
N061W	1,02	A431S	1,09
G073C	1,02	A431W	1,04
G073L	1,07	A431Y	1,13
G073W	1,03	E503C	1,05
L417A	1,12	E503D	1,06

Variantes	PI de actividad específica		Variantes	PI de actividad específica
L417D	1,19		E503H	1,01
L417E	1,10		E503S	1,10
L417F	1,08		E503W	1,04
L417G	1,19		Q511C	1,07
L417I	1,10		Q511G	1,06
L417K	1,02		Q511H	1,05
L417Q	1,04		Q511I	1,10
L417R	1,30		Q511K	1,09
L417S	1,05		Q511T	1,04
L417T	1,10		Q511V	1,04
L417V	1,21		A535E	1,19
L417W	1,05		A535F	1,06
L417Y	1,10		A535G	1,02
T430A	1,19		A535K	1,07
T430E	1,15		A535L	1,02
T430F	1,09		A535N	1,04
T430G	1,16		A535P	1,14
T430H	1,15		A535R	1,22
T430I	1,06		A535S	1,06
T430K	1,24		A535T	1,04
T430M	1,16		A535V	1,04
T430N	1,07		A535W	1,09
T430Q	1,15		A535Y	1,13
T430R	1,04		A539E	1,08
T430V	1,09		A539M	1,03
A431C	1,04		A539S	1,02
A431E	1,08		A539W	1,06
A431H	1,11		A539R	1,22

Tabla 17. Variantes de TrGA adicionales que muestran tanto estabilidad térmica como actividad específica aumentadas

Variantes	PI de actividad específica	PI de estabilidad térmica	Variantes	PI de actividad específica	PI de estabilidad térmica
N061V	1,11	1,06	T430Q	1,15	1,05
G073W	1,03	1,36	T430R	1,04	1,13
L417I	1,10	1,04	T430V	1,09	1,05
L417K	1,02	1,20	A431I	1,20	1,03
L417Q	1,04	1,04	A431R	1,11	1,08
L417R	1,30	1,20	E503C	1,05	1,39
L417V	1,21	1,07	E503S	1,10	1,02
L417Y	1,10	1,01	Q511H	1,05	1,33
T430A	1,19	1,05	A535K	1,07	1,24
T430E	1,15	1,02	A535N	1,04	1,37

T430F	1,09	1,06		A535P	1,14	1,59
T430H	1,15	1,10		A535R	1,22	1,26
T430I	1,06	1,04		A539E	1,08	1,32
T430K	1,24	1,08		A539M	1,03	1,05
T430M	1,16	1,17		A539R	1,22	1,36
T430N	1,07	1,13		A539S	1,02	1,30

Ejemplo 12: Caracterización de un conjunto seleccionado de variantes combinatorias y de sitio único

- 5 **[0251]** Según los resultados de los Ejemplos 4-6 y 11, se caracterizó además un conjunto de variantes combinatorias y variantes de sitio único para determinar propiedades modificadas. El conjunto seleccionado incluye variantes combinatorias y de sitio único con sustituciones en: 143, D44, N61, G73, G294, L417, T430, A431, E503, Q511, A535, A539 y/o N563. Las variantes se purificaron a partir de fermentación a gran escala y se determinaron los PI de estabilidad térmica y actividad específica. De forma específica, las actividades específicas se determinaron utilizando diferentes sustratos, entre los que se incluyen DP7, almidón de maíz y sustrato licuado.
- 10 Los resultados se muestran en la tabla 18 y 19.

Tabla 18. PI de un conjunto seleccionado de variantes de sitio único, cada una de las cuales viene de una fermentación de 500 ml

Variantes	P.I. DP7-FPLC	P.I. PI Maicena-FPLC	P.I. Estabilidad térmica	P.I. Sustrato licuado-FPLC
N61I	1,16	1,35	1,00	1,66
A431L	1,15	1,38	1,18	1,51
L417V	1,18	1,32	1,02	1,40
A431Q	1,06	1,20	0,92	1,24
G294C	1,01	0,84	0,94	1,23
N563K	1,07	1,12	1,97	1,15
Q511H	1,05	1,09	1,52	1,13
T430M	1,05	1,15	0,89	1,09
E503A	1,08	1,16	1,40	1,09
I43Q	1,11	1,24	0,94	1,08
A539R	1,15	1,37	1,43	1,08
I43R	1,03	1,07	1,41	1,07
L417R	1,23	1,27	1,51	1,04
T430A	1,13	1,35	1,23	1,04
G73F	1,06	1,06	1,45	1,03
D44R	0,97	1,06	1,46	0,98
N563I	1,09	1,22	2,06	0,92
D44C	0,80	0,82	0,96	0,91
E503V	1,17	1,07	1,66	0,88
A535R	1,09	1,44	1,47	0,85

Tabla 19. PI de un conjunto seleccionado de variantes combinatorias

Ferm.	Variante	Sustituciones	P.I. DP7-FPLC	P.I. Maicena-FPLC	P.I. Estabilidad térmica	P.I. Sustrato licuado-FPLC
100ml	C4	G73F/T430A/E503V/Q511H;	1,05	1,06	2,57	1,01
100ml	C9	D44C/G73F/L417R/N563K	0,95	1,00	2,57	1,22
100ml	C11	D44C/G73F/N563K	0,83	1,04	2,42	
500ml	ALL5	I43R/L417V/E503A/Q511H/A539R	1,22	1,48	2,37	1,87
100ml	C15	D44C/L417R/N563K	0,95	1,09	2,30	0,99
100ml	C8	D44C/G73F/N563K	0,87	1,05	2,28	
500ml	C5	D44C/G73F/N563K	0,90	0,71	2,16	0,97
500ml	TS1	I43R/T430A/E503V/A535R/N563K			2,13	
500ml	C7	D44C/G73F/E503V/Q511H;	0,90	0,79	2,10	0,89
500ml	ALL1	I43Q/D44C/L417V/E503A/Q511H/A539R	1,00	1,27	1,95	1,59
500ml	CS4	L417V/T430A/Q511H/A539R/N563I	1,21	1,53	1,95	1,73
100ml	C2	I43R/G73F/T430A	1,09	1,25	1,91	1,18
500ml	RB7	I43Q/T430A/Q511H/L417V/A539R	1,11	1,36	1,85	2,17
500ml	C1	G73F/T430A/Q511H	1,20	1,07	1,84	0,96
500ml	ALL2	I43Q/L417V/E503A/Q511H/A539R	1,16	1,52	1,84	1,96
500ml	ALL6	I43R/N61I/L417V/E503A/Q511H/A 539R	1,06	1,53	1,84	2,24
500ml	ALL8	I43R/N61I/L417R/E503A/Q511H/A 539R	1,10	1,53	1,84	2,11
500ml	RB20	I43Q/Q511H/A539R/T430M/N61I/L 417V	1,01	1,55	1,79	1,95
500ml	ALL4	I43Q/N61I/L417V/E503A/Q511H/A 539R	1,08	1,52	1,79	1,73
500ml	CS3	L417VAT430A/Q511H/A535R/N563I	0,98	1,40	1,69	1,78
500ml	RB9	I43Q/T430A/Q511H/A431L/E503A	0,93	1,39	1,69	1,55
500ml	RB17	I43Q/Q511H/A539R/E503A	1,12	1,38	1,68	1,40
500ml	ALL7	I43R/L417R/E503A/A539R	1,04	1,54	1,67	1,62
500ml	TS4	I43R/T430A/E503A/Q511H/N563K	0,98	1,33	1,66	1,33
500ml	RB18	I43Q/Q511H/A539R/T430M	1,06	1,44	1,64	1,40
100ml	C12	I43R/T430A/	1,15	1,35	1,60	1,26
500ml	RB16	I43Q/Q511H/A539R/N61I	1,11	1,43	1,60	1,34
500ml	TS5	D44R/T430A/Q511H/A535R	0,89	1,22	1,59	1,21
500ml	RB8	I43Q/T430A/Q511H/A431L/A539R	1,02	1,38	1,59	2,18
500ml	RB11	I43Q/T430A/Q511H/L417V/A539R/ A431L	1,05	1,36	1,58	2,03
500ml	CS1	L417V/T430A/A431L/Q511H/A535R/A539R/N563I	1,08	1,71	1,56	2,35
500ml	LQ5	L417R/A431L/Q511H	0,95	1,21	1,56	1,88
500ml	CS2	L417V/T430A/A431Q/Q511H/A535R/A539R/N563I	1,07	1,71	1,56	2,32
500ml	TS2	D44R/E503A/Q511H/N563I	0,90	1,14	1,53	1,23
500ml	RB15	I43Q/Q511H/A539R	1,19	1,42	1,52	1,69
500ml	RB13	I43Q/Q511H/L417V	1,06	1,38	1,50	1,50
500ml	RB5	I43Q/T430A/Q511H/A539R	0,99	1,45	1,47	1,59
500ml	RB4	I43Q/T430A/Q511H/E503A	1,07	1,29	1,47	1,50
500ml	RB19	I43Q/Q511H/A539R/T430M/N61I	1,00	1,50	1,47	1,88

Ferm.	Variante	Sustituciones	P.I. DP7-FPLC	P.I. Maicena-FPLC	P.I. Estabilidad térmica	P.I. Sustrato licuado-FPLC
500ml	ALL3	I43Q/D44C/N61I/L417V/E503A/Q511H/A539R	0,82	1,30	1,43	1,64
500ml	LQ4	G294C/L417R/A431Q/Q511H	1,02	1,19	1,43	1,65
500ml	RB2	I43Q/T430A/Q511H/L417V	1,09	1,38	1,42	1,87
500ml	TS3	E503A/N563I/	0,94	1,17	1,39	1,04
500ml	RB12	I43Q/Q511H/N611	0,96	1,25	1,37	1,56
500ml	RB14	I43Q/Q511H/A431L	1,01	1,31	1,37	1,55
500ml	LQ6	L417V/A431Q/Q511H	0,94	1,33	1,35	2,05
100ml	C14	I43Q/T430A/Q511H	1,09	1,34	1,33	1,46
500ml	LQ3	G294C/L417V/A431L/Q511H	0,80	1,21	1,29	2,07
500ml	RB3	I43Q/T430A/Q511H/A431L	1,20	1,25	1,29	1,70
500ml	RB10	I43Q/T430A/Q511H/N61I/A539R/A431L	0,87	1,47	1,29	1,73
500ml	RB1	I43Q/T430A/Q511H/N61I	1,19	1,15	1,27	1,47
100ml	C13	I43Q/T43A/	1,07	1,29	1,10	1,45
500ml	LQ1	G294C/L417R/A431L	0,90	1,08	1,04	1,77
500ml	LQ2	G294C/L417V/A431Q	0,80	1,08	1,02	2,08
500ml	RB6	I43Q/T430A/Q511H/N61I/A539R			0,94	

Ejemplo 11: Estructura de cristal de TrGA

5 **[0252]** La estructura tridimensional completa de la glucoamilasa *Trichoderma reesei* (*Hypocrea jecorina*) (TrGA) se determinó con la resolución 1,9 Å. TrGA se cristalizó de forma intacta con 599 residuos y con todas las modificaciones traduccionales posteriores que normalmente tendrían lugar en el huésped natural. La estructura de cristal se produjo y se analizó de la siguiente manera:

10 **[0253]** Para la expresión proteica y purificación, el gen codificante de GA *H. jecorina* se clonó y se expresó según los protocolos descritos en la patente estadounidense n.º 7413887.

15 **[0254]** El material de proteína TrGA usado para todos los experimentos de cristalización se purificó inicialmente en una etapa mediante cromatografía de intercambio de aniones como se muestra a continuación: se prepararon sobrenadantes de cultivo concentrados de TrGA expresada, consistentes en 180 mg/ml de proteína total, mediante la dilución de la muestra 1:10 en un tampón de Tris-HCl 25 mM, pH 8,0. Se empleó una columna FF Sepharose Q HiPrep 16/10 (GE Healthcare) para la purificación de intercambio aniónico. La columna HiPrep se equilibró con 4 volúmenes de columna (VC) de tampón de inicio (Tris-HCl 25 mM, pH 8,0) seguido de una aplicación de 10 ml de la muestra proteica diluida. Se aplicó un gradiente lineal 8 VC de 0 a 140 mM NaCl en el tampón de migración (Tris-HCl 25 mM, pH 8,0) con el fin de eluir la proteína unida. Se eluyó la TrGA unida de la columna sepharose Q 20 HiPrep con una concentración de sal de aproximadamente NaCl 80 mM. Las fracciones con proteína TrGA pura se acumularon y concentraron a 50 mg/ml utilizando un tubo de concentración centrífuga Vivaspin de 25 ml (Viva Science) con un límite de peso molecular (LPM) de 10 kD. Se cambió el tampón del material de TrGA purificado y concentrado utilizando una columna desaladora DG-10 (Bio-Rad) equilibrada con tampón acetato de sodio 50 mM, pH 4,3. Las concentraciones de proteína se determinaron mediante la medición de la absorbancia a 280 nm. El contenido proteico de TrGA concentrado y purificado inicialmente se almacenó a -20 °C. 25

30 **[0255]** Se introdujeron dos etapas de purificación adicionales, una purificación mediante intercambio de aniones adicional y una purificación por exclusión de tamaño, con el fin de mejorar la capacidad de formar cristales del material proteico TrGA. Estas dos etapas de purificación adicionales se llevaron a cabo de la siguiente forma. En la primera etapa de purificación por intercambio de aniones se utilizó una columna MonoQ 10 ml (GE Healthcare). Una muestra de 1 ml del material TrGA congelado y purificado inicialmente (50 mg proteína) se descongeló y el tampón se cambió a Tris-HCl 20 mM, pH 8,0, mediante dilución repetida de la muestra a 6 ml en el nuevo tampón, seguido de una concentración de la muestra de nuevo a 0,5 ml utilizando un tubo de concentración de 6 ml LPM 5 kD. La muestra TrGA se diluyó tras la última etapa de concentración en agua destilada hasta que se alcanzó una

conductividad de la muestra proteica que correspondía a la conductividad del tampón inicial de la purificación de aniones, es decir, Tris-HCl 25 mM, pH 8,0. La columna MonoQ se equilibró primero con tampón de inicio 4 volúmenes de columna (VC), seguido de la aplicación de la muestra proteica diluida a la columna. Se eluyó la proteína unida de la columna MonoQ mediante dos gradientes diferentes. En el primero, se aplicó un gradiente pH lineal 4 VC donde el pH del tampón de inicio se disminuyó de 8,0 a 6,0. En el segundo gradiente, se aplicó un gradiente de sal largo 8 VC en el que la concentración de sal aumentó de 0 a 350 mM NaCl en el tampón de migración (Tris-HCl 25 mM, pH 6,0). Se halló que la TrGA unida se eluía de la columna durante el segundo gradiente de sal con una concentración de NaCl aproximada de 150 mM. Las fracciones con TrGA se acumularon y concentraron a 2 ml utilizando un tubo de concentración Vivaspin 6 ml LPM 5 kD. A continuación, la muestra TrGA concentrada se aplicó a una columna de exclusión de tamaño Superdex 200 16/60 (GE Healthcare) equilibrada con 4 VC de Tris-Cl 20 mM, pH 8,0 y NaCl 50 mM, que también se utilizó como tampón de migración. Las fracciones del principal pico de elución tras la purificación por exclusión de tamaño se acumularon y concentraron a una concentración de proteína aproximada de 7,5 mg/ml utilizando un tubo de concentración Vivaspin de 6 ml LPM 5kD.

[0256] Para la cristalización de proteína, la muestra proteica que se utilizó para hallar las condiciones de cristalización de TrGA iniciales fue una muestra del material TrGA que se purificó una vez mediante purificación por intercambio de aniones y a partir de ahí se almacenó a -20 °C. La muestra proteica TrGA se descongeló y se diluyó con tampón acetato de sodio 50 mM, pH 4,3 a aproximadamente 12 mg/ml antes de los experimentos de cristalización iniciales. El conjunto de datos de rayos X ortorrómbicos se utilizó para solucionar la estructura TrGA mediante reemplazo molecular (RM) y el conjunto de datos ortorrómbicos de alta resolución se utilizó para el modelo de estructura TrGA del grupo espacial ortorrómbico final. Se halló que los cristales TrGA ortorrómbicos se cultivaban en solución consistente en 25 % PEG 3350, 0,20M acetato de amonio, 0,10M Bis-Tris pH 5,5 (solución de reserva), utilizando el método de difusión por vapor con gotas colgantes (McPherson 1982), a 20 °C. Las gotas de cristalización se prepararon mediante la mezcla de iguales cantidades de solución proteica (12 mg/ml) y solución de reserva para un volumen final de 10 μ l. Se halló que los cristales TrGA pertenecen al grupo espacial ortorrómbico P212121 con las dimensiones celulares aproximadas: $a = 52.2 \text{ \AA}$, $b = 99.2 \text{ \AA}$, $c = 121.2 \text{ \AA}$, y presentan una V_m calculada de 2,3 (Matthews 1968) con una molécula en la unidad asimétrica.

[0257] Para la recogida de datos de rayos X, los dos conjuntos de datos TrGA ortorrómbicos se recogieron de cristales únicos fijados en tubos capilares sellados, a temperatura ambiente. El conjunto de datos de rayos X TrGA ortorrómbicos de baja resolución iniciales, utilizado para solucionar la estructura mediante métodos de reemplazo molecular (RM), se recogió en una fuente de rayos X doméstica, un detector de placa de imagen Raxis IV++ MSC/Rigaku (Molecular Structures Corp., The Woodlands, Texas) con espejos de enfoque con una radiación Cu $K\alpha$ de un generador de ánodos giratorio Rigaku RU200. Este conjunto de datos se procesó, se organizó y se realizó la media utilizando el software d*trek suministrado por MSC/Rigaku. El conjunto de datos monoclinicos centrados C se recogió de un único cristal TrGA congelado a 100 K equilibrado en un agente crioprotector compuesto de 25 % PEG 3350, 15% Glicerol 50 mM CaCl_2 y 0,1 M Bis-Tris pH 5,5 como crioprotector, fijado en bucles de fibra rayón y sumergido congelado en nitrógeno líquido antes de su transporte al sincrotrón. Tanto el conjunto de datos ortorrómbicos de alta resolución (1.9 Å) como el conjunto de datos monoclinicos centrados C (1.8 Å) se recogieron en una fuente de sincrotrón, línea de haz 911:5 en MAX LAB en Lund, Suecia. Ambos conjuntos de datos recogidos en una fuente de sincrotrón se procesaron con MOSFLM y se organizaron con el programa SCALA incluido en el paquete de programas CCP4 (Collaborative Computational Project Number 4 1994). El procesamiento de datos posterior se llevó a cabo utilizando el paquete de programas CCP4 (Collaborative Computational Project Number 4 1994), a menos que se indique lo contrario. Un conjunto del 5 % de las reflexiones de cada conjunto de datos se deja apartado y se utiliza para controlar el libre R (Brünger, A (1992) Nature, 355:472-475).

[0258] La estructura TrGA se solucionó inicialmente mediante RM con el programa de reemplazo automático MOLREP (Collaborative Computational Project Number 4 1994), incluido en el paquete de programas CCP4, utilizando el conjunto de datos ortorrómbicos de resolución baja inicial y utilizando las coordenadas de la variante *Aspergillus awamori* GA (AaGA) X100 (pdb entrada 1GLM (Aleshin *et al.* (1994) J. Mol. Biol. 238: 575-591) como modelo de investigación. Se editó el modelo de investigación A. *awamori* GA para eliminar todas las fracciones de glicosilación fijadas a la molécula de proteína como glicosilaciones N- y O- y todas las moléculas de solventes antes de llevar a cabo los experimentos de RM. Todas las reflexiones entre 36,8 y 2,8 Å de resolución, del conjunto de datos de TrGA de baja resolución inicial, se utilizaron para la solución RM. El programa de RM halló una única solución de función de rotación, con un máximo de 11,1 σ por encima de lo ya registrado, el siguiente valor más alto fue 3,8 σ por encima de lo ya registrado. La solución de función de traducción proporcionó un factor R de 48,7 % y presentó un factor de contraste de 17,4. La solución MR se refinó durante 10 ciclos de refinamiento de cuadrados mínimos limitado utilizando el programa Refmac 5.0 (Murshudov *et al.* (1997) Acta Crystallogr. D53: 240-255). Esto redujo el factor R cristalográfico a 31,1% mientras que el valor libre R cayó desde 42,2 % a 41,1 %.

[0259] El modelo de solución RM refinado se utilizó para calcular un mapa de densidad inicial a partir del conjunto de datos TrGA ortorrómbicos de baja resolución. La densidad de electrones para un puente disulfuro entre los

residuos 19 y 26 de TrGA, un puente disulfuro no presente en el modelo de estructura de la variante *A. awamori* X100, podría identificarse fácilmente en este mapa de densidad de electrones. Esto se tomó como una indicación de que el mapa de densidad de electrones presentaba la suficiente calidad para utilizarse para la construcción de un modelo de estructura de TrGA a partir de su secuencia de aminoácidos. El modelo de estructura TrGA inicial, según el conjunto de datos de baja resolución, se refinó con ciclos alternos de construcción de modelo mediante Coot (Emsley y Cowtan, (2004) Acta Crystallogr. D boil Crystallogr. 60: 2126-2132), y la máxima probabilidad de refinamiento utilizando Refmac 5.0.

[0260] La resolución del modelo de estructura TrGA inicial se extendió a la resolución del conjunto de datos ortorrómbicos de alta resolución (1,9Å) mediante el refinamiento del modelo de estructura TrGA inicial frente al conjunto de datos de alta resolución durante 10 ciclos de refinamiento limitado utilizando el programa Refmac 5.0. La mayoría de las moléculas de agua en los modelos de estructura se ubicaron de forma automática mediante el uso de protocolos de recogida de agua en los programas de refinamiento y, a continuación, de forma manual se eligieron o descartaron mediante inspección visual. Todas las comparaciones estructurales se realizaron con Coot (Emsley y Cowtan (2004) *supra*) u O (Jones *et al.* (1991) Acta Crystallogr. A47: 110-119), y las figuras se prepararon con PyMOL (Delano W.L. (2002) The PyMOL Molecular Graphics System. Palo Alto, CA, EE UU; Delano Scientific).

[0261] A partir de estos resultados, puede verse que el segmento principal catalítico TrGA seguía la misma topología de 6 barriles (α/α) descrita por Aleshin *et al.* 1992 para AaGA, consistente en un doble barril de hélices alfa con el C-terminal de la hélice externa hacia el N-terminal de una hélice interna. Fue posible identificar diferencias clave en la densidad de electrones tal como el puente disulfuro entre los residuos 19 y 26 y una inserción (residuos 257-260) en relación con AaGA. El segmento que comprende 80-100 también experimentó una reconstrucción de modelo extensiva. Se identificó un sitio de glicosilación principal en Asn 171 que presentaba hasta cuatro fracciones de glucósidos fijadas. Se identificó un sitio de glicosilación similar en AaGa. De forma adicional, el núcleo catalítico con tres cis péptidos entre los residuos 22-23, 44-45 y 122-123 se conservó entre TrGA y AaGA. En general, existía una variación rms de 0,535 Å entre 409 de 453 átomos C α cuando se compararon las coordenadas de los núcleos catalíticos de TrGA y AaGA.

30 Ejemplo 14: Homología entre TrGA y AaGA

[0262] La estructura de cristal de la TrGA identificada en el Ejemplo 13 se superpuso sobre la estructura de cristal identificada previamente de la GA *Aspergillus awamori* (AaGA). La estructura de cristal AaGA se obtuvo a partir de la base de datos proteica (BDP) y la forma de AaGA que se cristalizó fue la forma con solo un dominio catalítico. La estructura de la glucoamilasa *Trichoderma reesei* con las tres regiones intactas se determinó con la resolución 1,8 Angstrom en el presente documento (véase la tabla 15 y el Ejemplo 12). Al utilizar las coordenadas la estructura se alineó con las coordenadas del dominio catalítico de la cepa *Aspergillus awamori* X100 que se determinó previamente (Aleshin, A.E., Hoffman, C., Firsov, L.M., y Honzatko, R.B. 1994 Refined crystal structures of glucoamylase from *Aspergillus awamori* var. X100. J Mol Biol 238: 575-591 y la PDB). Como puede observarse en las figuras 12 y 13, la estructura del dominio catalítico se superpuso muy estrechamente y permitió la identificación de los residuos equivalentes según esta superposición estructural.

[0263] Según el presente análisis, se identificaron los sitios que podían mutarse en TrGA y tener como resultado una estabilidad y/o actividad específica aumentada. Estos sitios incluyen 108, 124, 175 y 316 en el sitio activo. También se identificaron variantes por pares específicas Y47W/Y315F y Y47F/Y315W. Otros sitios identificados fueron 143, D44, P45, D46, R122, R125, V181, E242, Y310, D313, V314, N317, R408 y N409. Debido a la alta homología estructural se espera que las variantes beneficiosas halladas en los sitios en TrGA tengan consecuencias similares en *Aspergillus awamori* y otras glucoamilasas homólogas.

50 LISTADO DE SECUENCIAS

[0264]

<110> DANISCO US, INC.

<120> VARIANTES DE GLUCOAMILASA CON PROPIEDADES MODIFICADAS

<130> CSC/FP7297674

<140> TBC

< 141> 2007-11-20

<150> 14177958.7

< 151> 2007-11-20

<150> 08851977.2

ES 2 795 989 T3

< 151> 2007-11-20

<150> 60/989,426

< 151> 2007-11-20

<160> 168

5 <170> PatentIn version 3.5

<210> 1

< 211> 632

< 212> PRT

< 213> Trichoderma reesei

10 <400> 1

Met His Val Leu Ser Thr Ala Val Leu Leu Gly Ser Val Ala Val Gln
1 5 10 15

Lys Val Leu Gly Arg Pro Gly Ser Ser Gly Leu Ser Asp Val Thr Lys
20 25 30

Arg Ser Val Asp Asp Phe Ile Ser Thr Glu Thr Pro Ile Ala Leu Asn
35 40 45

Asn Leu Leu Cys Asn Val Gly Pro Asp Gly Cys Arg Ala Phe Gly Thr
50 55 60

Ser Ala Gly Ala Val Ile Ala Ser Pro Ser Thr Ile Asp Pro Asp Tyr
65 70 75 80

Tyr Tyr Met Trp Thr Arg Asp Ser Ala Leu Val Phe Lys Asn Leu Ile
85 90 95

Asp Arg Phe Thr Glu Thr Tyr Asp Ala Gly Leu Gln Arg Arg Ile Glu
100 105 110

Gln Tyr Ile Thr Ala Gln Val Thr Leu Gln Gly Leu Ser Asn Pro Ser
115 120 125

Gly Ser Leu Ala Asp Gly Ser Gly Leu Gly Glu Pro Lys Phe Glu Leu

ES 2 795 989 T3

130 135 140

Thr Leu Lys Pro Phe Thr Gly Asn Trp Gly Arg Pro Gln Arg Asp Gly
 145 150 155 160

Pro Ala Leu Arg Ala Ile Ala Leu Ile Gly Tyr Ser Lys Trp Leu Ile
 165 170 175

Asn Asn Asn Tyr Gln Ser Thr Val Ser Asn Val Ile Trp Pro Ile Val
 180 185 190

Arg Asn Asp Leu Asn Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Phe
 195 200 205

Asp Leu Trp Glu Glu Val Asn Gly Ser Ser Phe Phe Thr Val Ala Asn
 210 215 220

Gln His Arg Ala Leu Val Glu Gly Ala Thr Leu Ala Ala Thr Leu Gly
 225 230 235 240

Gln Ser Gly Ser Ala Tyr Ser Ser Val Ala Pro Gln Val Leu Cys Phe
 245 250 255

Leu Gln Arg Phe Trp Val Ser Ser Gly Gly Tyr Val Asp Ser Asn Ile
 260 265 270

Asn Thr Asn Glu Gly Arg Thr Gly Lys Asp Val Asn Ser Val Leu Thr
 275 280 285

Ser Ile His Thr Phe Asp Pro Asn Leu Gly Cys Asp Ala Gly Thr Phe
 290 295 300

Gln Pro Cys Ser Asp Lys Ala Leu Ser Asn Leu Lys Val Val Val Asp
 305 310 315 320

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

Ala Val Ala Ile Gly Arg Tyr Ala Glu Asp Val Tyr Tyr Asn Gly Asn
 340 345 350

Pro Trp Tyr Leu Ala Thr Phe Ala Ala Ala Glu Gln Leu Tyr Asp Ala
 355 360 365

Ile Tyr Val Trp Lys Lys Thr Gly Ser Ile Thr Val Thr Ala Thr Ser
 370 375 380

ES 2 795 989 T3

Leu Ala Phe Phe Gln Glu Leu Val Pro Gly Val Thr Ala Gly Thr Tyr
 385 390 395 400

Ser Ser Ser Ser Ser Thr Phe Thr Asn Ile Ile Asn Ala Val Ser Thr
 405 410 415

Tyr Ala Asp Gly Phe Leu Ser Glu Ala Ala Lys Tyr Val Pro Ala Asp
 420 425 430

Gly Ser Leu Ala Glu Gln Phe Asp Arg Asn Ser Gly Thr Pro Leu Ser
 435 440 445

Ala Leu His Leu Thr Trp Ser Tyr Ala Ser Phe Leu Thr Ala Thr Ala
 450 455 460

Arg Arg Ala Gly Ile Val Pro Pro Ser Trp Ala Asn Ser Ser Ala Ser
 465 470 475 480

Thr Ile Pro Ser Thr Cys Ser Gly Ala Ser Val Val Gly Ser Tyr Ser
 485 490 495

Arg Pro Thr Ala Thr Ser Phe Pro Pro Ser Gln Thr Pro Lys Pro Gly
 500 505 510

Val Pro Ser Gly Thr Pro Tyr Thr Pro Leu Pro Cys Ala Thr Pro Thr
 515 520 525

Ser Val Ala Val Thr Phe His Glu Leu Val Ser Thr Gln Phe Gly Gln
 530 535 540

Thr Val Lys Val Ala Gly Asn Ala Ala Ala Leu Gly Asn Trp Ser Thr
 545 550 555 560

Ser Ala Ala Val Ala Leu Asp Ala Val Asn Tyr Ala Asp Asn His Pro
 565 570 575

Leu Trp Ile Gly Thr Val Asn Leu Glu Ala Gly Asp Val Val Glu Tyr
 580 585 590

Lys Tyr Ile Asn Val Gly Gln Asp Gly Ser Val Thr Trp Glu Ser Asp
 595 600 605

Pro Asn His Thr Tyr Thr Val Pro Ala Val Ala Cys Val Thr Gln Val
 610 615 620

Val Lys Glu Asp Thr Trp Gln Ser
 625 630

<210> 2
 < 211> 599
 < 212> PRT
 < 213> Trichoderma reesei
 <400> 2

ES 2 795 989 T3

Ser Val Asp Asp Phe Ile Ser Thr Glu Thr Pro Ile Ala Leu Asn Asn
1 5 10 15

Leu Leu Cys Asn Val Gly Pro Asp Gly Cys Arg Ala Phe Gly Thr Ser
20 25 30

Ala Gly Ala Val Ile Ala Ser Pro Ser Thr Ile Asp Pro Asp Tyr Tyr
35 40 45

Tyr Met Trp Thr Arg Asp Ser Ala Leu Val Phe Lys Asn Leu Ile Asp
50 55 60

Arg Phe Thr Glu Thr Tyr Asp Ala Gly Leu Gln Arg Arg Ile Glu Gln
65 70 75 80

Tyr Ile Thr Ala Gln Val Thr Leu Gln Gly Leu Ser Asn Pro Ser Gly
85 90 95

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

Leu Lys Pro Phe Thr Gly Asn Trp Gly Arg Pro Gln Arg Asp Gly Pro
115 120 125

Ala Leu Arg Ala Ile Ala Leu Ile Gly Tyr Ser Lys Trp Leu Ile Asn
130 135 140

Asn Asn Tyr Gln Ser Thr Val Ser Asn Val Ile Trp Pro Ile Val Arg
145 150 155 160

Asn Asp Leu Asn Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Phe Asp
165 170 175

Leu Trp Glu Glu Val Asn Gly Ser Ser Phe Phe Thr Val Ala Asn Gln
180 185 190

His Arg Ala Leu Val Glu Gly Ala Thr Leu Ala Ala Thr Leu Gly Gln
195 200 205

Ser Gly Ser Ala Tyr Ser Ser Val Ala Pro Gln Val Leu Cys Phe Leu
210 215 220

ES 2 795 989 T3

Gln Arg Phe Trp Val Ser Ser Gly Gly Tyr Val Asp Ser Asn Ile Asn
 225 230 235 240

Thr Asn Glu Gly Arg Thr Gly Lys Asp Val Asn Ser Val Leu Thr Ser
 245 250 255

Ile His Thr Phe Asp Pro Asn Leu Gly Cys Asp Ala Gly Thr Phe Gln
 260 265 270

Pro Cys Ser Asp Lys Ala Leu Ser Asn Leu Lys Val Val Val Asp Ser
 275 280 285

Phe Arg Ser Ile Tyr Gly Val Asn Lys Gly Ile Pro Ala Gly Ala Ala
 290 295 300

Val Ala Ile Gly Arg Tyr Ala Glu Asp Val Tyr Tyr Asn Gly Asn Pro
 305 310 315 320

Trp Tyr Leu Ala Thr Phe Ala Ala Ala Glu Gln Leu Tyr Asp Ala Ile
 325 330 335

Tyr Val Trp Lys Lys Thr Gly Ser Ile Thr Val Thr Ala Thr Ser Leu
 340 345 350

Ala Phe Phe Gln Glu Leu Val Pro Gly Val Thr Ala Gly Thr Tyr Ser
 355 360 365

Ser Ser Ser Ser Thr Phe Thr Asn Ile Ile Asn Ala Val Ser Thr Tyr
 370 375 380

Ala Asp Gly Phe Leu Ser Glu Ala Ala Lys Tyr Val Pro Ala Asp Gly
 385 390 395 400

Ser Leu Ala Glu Gln Phe Asp Arg Asn Ser Gly Thr Pro Leu Ser Ala
 405 410 415

Leu His Leu Thr Trp Ser Tyr Ala Ser Phe Leu Thr Ala Thr Ala Arg
 420 425 430

Arg Ala Gly Ile Val Pro Pro Ser Trp Ala Asn Ser Ser Ala Ser Thr
 435 440 445

Ile Pro Ser Thr Cys Ser Gly Ala Ser Val Val Gly Ser Tyr Ser Arg
 450 455 460

Pro Thr Ala Thr Ser Phe Pro Pro Ser Gln Thr Pro Lys Pro Gly Val
 465 470 475 480

ES 2 795 989 T3

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

Val Ala Val Thr Phe His Glu Leu Val Ser Thr Gln Phe Gly Gln Thr
500 505 510

Val Lys Val Ala Gly Asn Ala Ala Leu Gly Asn Trp Ser Thr Ser
515 520 525

Ala Ala Val Ala Leu Asp Ala Val Asn Tyr Ala Asp Asn His Pro Leu
530 535 540

Trp Ile Gly Thr Val Asn Leu Glu Ala Gly Asp Val Val Glu Tyr Lys
545 550 555 560

Tyr Ile Asn Val Gly Gln Asp Gly Ser Val Thr Trp Glu Ser Asp Pro
565 570 575

Asn His Thr Tyr Thr Val Pro Ala Val Ala Cys Val Thr Gln Val Val
580 585 590

Lys Glu Asp Thr Trp Gln Ser
595

<210> 3

< 211> 453

< 212> PRT

5

< 213> Trichoderma reesei

<400> 3

Ser Val Asp Asp Phe Ile Ser Thr Glu Thr Pro Ile Ala Leu Asn Asn
1 5 10 15

Leu Leu Cys Asn Val Gly Pro Asp Gly Cys Arg Ala Phe Gly Thr Ser
20 25 30

Ala Gly Ala Val Ile Ala Ser Pro Ser Thr Ile Asp Pro Asp Tyr Tyr
35 40 45

Tyr Met Trp Thr Arg Asp Ser Ala Leu Val Phe Lys Asn Leu Ile Asp
50 55 60

Arg Phe Thr Glu Thr Tyr Asp Ala Gly Leu Gln Arg Arg Ile Glu Gln
65 70 75 80

Tyr Ile Thr Ala Gln Val Thr Leu Gln Gly Leu Ser Asn Pro Ser Gly
85 90 95

ES 2 795 989 T3

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

Leu Lys Pro Phe Thr Gly Asn Trp Gly Arg Pro Gln Arg Asp Gly Pro
115 120 125

Ala Leu Arg Ala Ile Ala Leu Ile Gly Tyr Ser Lys Trp Leu Ile Asn
130 135 140

Asn Asn Tyr Gln Ser Thr Val Ser Asn Val Ile Trp Pro Ile Val Arg
145 150 155 160

Asn Asp Leu Asn Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Phe Asp
165 170 175

Leu Trp Glu Glu Val Asn Gly Ser Ser Phe Phe Thr Val Ala Asn Gln
180 185 190

His Arg Ala Leu Val Glu Gly Ala Thr Leu Ala Ala Thr Leu Gly Gln
195 200 205

Ser Gly Ser Ala Tyr Ser Ser Val Ala Pro Gln Val Leu Cys Phe Leu
210 215 220

Gln Arg Phe Trp Val Ser Ser Gly Gly Tyr Val Asp Ser Asn Ile Asn
225 230 235 240

Thr Asn Glu Gly Arg Thr Gly Lys Asp Val Asn Ser Val Leu Thr Ser
245 250 255

Ile His Thr Phe Asp Pro Asn Leu Gly Cys Asp Ala Gly Thr Phe Gln
260 265 270

Pro Cys Ser Asp Lys Ala Leu Ser Asn Leu Lys Val Val Val Asp Ser
275 280 285

Phe Arg Ser Ile Tyr Gly Val Asn Lys Gly Ile Pro Ala Gly Ala Ala
290 295 300

Val Ala Ile Gly Arg Tyr Ala Glu Asp Val Tyr Tyr Asn Gly Asn Pro
305 310 315 320

Trp Tyr Leu Ala Thr Phe Ala Ala Ala Glu Gln Leu Tyr Asp Ala Ile
325 330 335

Tyr Val Trp Lys Lys Thr Gly Ser Ile Thr Val Thr Ala Thr Ser Leu
340 345 350

ES 2 795 989 T3

Ala Phe Phe Gln Glu Leu Val Pro Gly Val Thr Ala Gly Thr Tyr Ser
 355 360 365

Ser Ser Ser Ser Thr Phe Thr Asn Ile Ile Asn Ala Val Ser Thr Tyr
 370 375 380

Ala Asp Gly Phe Leu Ser Glu Ala Ala Lys Tyr Val Pro Ala Asp Gly
 385 390 395 400

Ser Leu Ala Glu Gln Phe Asp Arg Asn Ser Gly Thr Pro Leu Ser Ala
 405 410 415

Leu His Leu Thr Trp Ser Tyr Ala Ser Phe Leu Thr Ala Thr Ala Arg
 420 425 430

Arg Ala Gly Ile Val Pro Pro Ser Trp Ala Asn Ser Ser Ala Ser Thr
 435 440 445

Ile Pro Ser Thr Cys
 450

<210> 4

< 211> 1899

< 212> ADN

5 < 213> Trichoderma reesei

<400> 4

atgcacgtcc tgtcgactgc ggtgctgctc ggctccggtg ccgttcaaaa ggtcctggga 60
 agaccaggat caagcggctc gtccgacgtc accaagaggt ctggtgacga cttcatcagc 120
 accgagacgc ctattgcact gaacaatctt ctttgcaatg ttggtcctga tggatgccgt 180
 gcattcggca catcagctgg tgcggtgatt gcatctccca gcacaattga cccggactac 240
 tattacatgt ggacgcgaga tagcgctctt gtcttcaaga acctcatcga ccgcttcacc 300
 gaaacgtacg atgcgggcct gcagcggcgc atcgagcagt acattactgc ccaggctact 360
 ctccagggcc tctctaacc ctcgggctcc ctgcgggacg gctctggtct cggcgagccc 420
 aagtttgagt tgacctgaa gcctttcacc ggcaactggg gtcgaccgca gcgggatggc 480
 ccagctctgc gagccattgc cttgattgga tactcaaagt ggctcatcaa caacaactat 540
 cagtcgactg tgtccaacgt catctggcct attgtgcgca acgacctcaa ctatgttgcc 600
 cagtactgga accaaaccgg ctttgacctc tgggaagaag tcaatgggag ctcattcttt 660
 actggtgcca accagcaccg agcacttgtc gagggcgcca ctcttgctgc cactcttggc 720
 cagtcgggaa gcgcttattc atctggtgct cccaggttt tgtgctttct ccaacgattc 780
 tgggtgctgt ctggtggata cgtcgactcc aacatcaaca ccaacgaggg caggactggc 840

ES 2 795 989 T3

aaggatgtca actccgtcct gacttccatc cacaccttcg atcccaacct tggctgtgac 900
 gcaggcacct tccagccatg cagtgacaaa gcgctctcca acctcaaggt tgttgtcgac 960
 tccttccgct ccatctacgg cgtgaacaag ggcattoctg ccggtgctgc cgtcgccatt 1020
 ggccggtatg cagaggatgt gtactacaac ggcaaccctt ggtatcttgc tacatttgc 1080
 gctgccgagc agctgtacga tgccatctac gtctggaaga agacgggctc catcacggtg 1140
 accgccacct ccctggcctt cttccaggag cttgttctctg gcgtgacggc cgggacctac 1200
 tccagcagct cttcgacctt taccaacatc atcaacgccg tctcgacata cgccgatggc 1260
 ttctcagcg aggtgccea gtacgtcccc gccgacgggt cgctggccga gcagtttgac 1320
 cgcaacagcg gactccgct gtctgcgctt cacctgacgt ggtcgtacgc ctcggtcttg 1380
 acagccacgg cccgtcgggc tggcatcgtg cccccctcgt gggccaacag cagcgttagc 1440
 acgatcccct cgacgtgctc cggcgcgtcc gtggtcggat cctactcgcg tcccaccgcc 1500
 acgtcattcc ctccgtcgca gacgcccaag cctggcgtgc cttccggtac tccctacacg 1560
 ccctgcctt gcgcgacccc aacctccgtg gccgtcacct tccacgagct cgtgtcgaca 1620
 cagtttgccc agacgggtcaa ggtggcgggc aacgccgcgg ccctgggcaa ctggagcacg 1680
 agcgcgcccg tggctctgga cgccgtcaac tatgcccata accaccccct gtggattggg 1740
 acggccaacc tcgaggctgg agacgtcgtg gactacaagt acatcaatgt gggccaagat 1800
 ggctccgtga cctgggagag tgatcccaac cacacttaca cggttcctgc ggtggcttgt 1860
 gtgacgcagg ttgtcaagga ggacacctgg cagtcgtaa 1899

<210> 5

< 211> 448

< 212> PRT

5

< 213> Aspergillus awamori

<400> 5

Ala Thr Leu Asp Ser Trp Leu Ser Asn Glu Ala Thr Val Ala Arg Thr
 1 5 10 15

Ala Ile Leu Asn Asn Ile Gly Ala Asp Gly Ala Trp Val Ser Gly Ala
 20 25 30

Asp Ser Gly Ile Val Val Ala Ser Pro Ser Thr Asp Asn Pro Asp Tyr
 35 40 45

Phe Tyr Thr Trp Thr Arg Asp Ser Gly Leu Val Ile Lys Thr Leu Val
 50 55 60

Asp Leu Phe Arg Asn Gly Asp Thr Asp Leu Leu Ser Thr Ile Glu Asn
 65 70 75 80

ES 2 795 989 T3

Tyr Ile Ser Ser Gln Ala Ile Val Gln Gly Ile Ser Asn Pro Ser Gly
 85 90 95
 Asp Leu Ser Ser Gly Gly Leu Gly Glu Pro Lys Phe Asn Val Asp Glu
 100 105 110
 Thr Ala Tyr Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro Ala
 115 120 125
 Leu Arg Ala Thr Ala Met Ile Gly Phe Arg Gln Trp Leu Leu Asp Asn
 130 135 140
 Gly Tyr Thr Ser Ala Ala Thr Glu Ile Val Trp Pro Leu Val Arg Asn
 145 150 155 160
 Asp Leu Ser Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Tyr Asp Leu
 165 170 175
 Trp Glu Glu Val Asn Gly Ser Ser Phe Phe Thr Ile Ala Val Gln His
 180 185 190
 Arg Ala Leu Val Glu Gly Ser Ala Phe Ala Thr Ala Val Gly Ser Ser
 195 200 205
 Cys Ser Trp Cys Asp Ser Gln Ala Pro Gln Ile Leu Cys Tyr Leu Gln
 210 215 220
 Ser Phe Trp Thr Gly Glu Tyr Ile Leu Ala Asn Phe Asp Ser Ser Arg
 225 230 235 240
 Ser Gly Lys Asp Thr Asn Thr Leu Leu Gly Ser Ile His Thr Phe Asp
 245 250 255
 Pro Glu Ala Gly Cys Asp Asp Ser Thr Phe Gln Pro Cys Ser Pro Arg
 260 265 270
 Ala Leu Ala Asn His Lys Glu Val Val Asp Ser Phe Arg Ser Ile Tyr
 275 280 285
 Thr Leu Asn Asp Gly Leu Ser Asp Ser Glu Ala Val Ala Val Gly Arg
 290 295 300
 Tyr Pro Lys Asp Ser Tyr Tyr Asn Gly Asn Pro Trp Phe Leu Cys Thr
 305 310 315 320
 Leu Ala Ala Ala Glu Gln Leu Tyr Asp Ala Leu Tyr Gln Trp Asp Lys
 325 330 335

ES 2 795 989 T3

Gln Gly Ser Leu Glu Ile Thr Asp Val Ser Leu Asp Phe Phe Gln Ala
 340 345 350

Leu Tyr Ser Asp Ala Ala Thr Gly Thr Tyr Ser Ser Ser Ser Ser Thr
 355 360 365

Tyr Ser Ser Ile Val Asp Ala Val Lys Thr Phe Ala Asp Gly Phe Val
 370 375 380

Ser Ile Val Glu Thr His Ala Ala Ser Asn Gly Ser Leu Ser Glu Gln
 385 390 395 400

Tyr Asp Lys Ser Asp Gly Asp Glu Leu Ser Ala Arg Asp Leu Thr Trp
 405 410 415

Ser Tyr Ala Ala Leu Leu Thr Ala Asn Asn Arg Arg Asn Ser Val Met
 420 425 430

Pro Pro Ser Trp Gly Glu Thr Ser Ala Ser Ser Val Pro Gly Thr Cys
 435 440 445

<210> 6

< 211> 449

< 212> PRT

5

< 213> Aspergillus niger

<400> 6

Ala Thr Leu Asp Ser Trp Leu Ser Asn Glu Ala Thr Val Ala Arg Thr
 1 5 10 15

Ala Ile Leu Asn Asn Ile Gly Ala Asp Gly Ala Trp Val Ser Gly Ala
 20 25 30

Asp Ser Gly Ile Val Val Ala Ser Pro Ser Thr Asp Asn Pro Asp Tyr
 35 40 45

Phe Tyr Thr Trp Thr Arg Asp Ser Gly Leu Val Leu Lys Thr Leu Val
 50 55 60

Asp Leu Phe Arg Asn Gly Asp Thr Ser Leu Leu Ser Thr Ile Glu Asn
 65 70 75 80

Tyr Ile Ser Ala Gln Ala Ile Val Gln Gly Ile Ser Asn Pro Ser Gly
 85 90 95

Asp Leu Ser Ser Gly Ala Gly Leu Gly Glu Pro Lys Phe Asn Val Asp
 100 105 110

ES 2 795 989 T3

Glu Thr Ala Tyr Thr Gly Ser Trp Gly Arg Pro Gln Arg Asp Gly Pro
 115 120 125

Ala Leu Arg Ala Thr Ala Met Ile Gly Phe Gly Gln Trp Leu Leu Asp
 130 135 140

Asn Gly Tyr Thr Ser Thr Ala Thr Asp Ile Val Trp Pro Leu Val Arg
 145 150 155 160

Asn Asp Leu Ser Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Tyr Asp
 165 170 175

Leu Trp Glu Glu Val Asn Gly Ser Ser Phe Phe Thr Ile Ala Val Gln
 180 185 190

His Arg Ala Leu Val Glu Gly Ser Ala Phe Ala Thr Ala Val Gly Ser
 195 200 205

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

Gln Ser Phe Trp Thr Gly Ser Phe Ile Leu Ala Asn Phe Asp Ser Ser
 225 230 235 240

Arg Ser Gly Lys Asp Ala Asn Thr Leu Leu Gly Ser Ile His Thr Phe
 245 250 255

Asp Pro Glu Ala Ala Cys Asp Asp Ser Thr Phe Gln Pro Cys Ser Pro
 260 265 270

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

Tyr Thr Leu Asn Asp Gly Leu Ser Asp Ser Glu Ala Val Ala Val Gly
 290 295 300

Arg Tyr Pro Glu Asp Thr Tyr Tyr Asn Gly Asn Pro Trp Phe Leu Cys
 305 310 315 320

Thr Leu Ala Ala Ala Glu Gln Leu Tyr Asp Ala Leu Tyr Gln Trp Asp
 325 330 335

Lys Gln Gly Ser Leu Glu Val Thr Asp Val Ser Leu Asp Phe Phe Lys
 340 345 350

Ala Leu Tyr Ser Asp Ala Ala Thr Gly Thr Tyr Ser Ser Ser Ser Ser
 355 360 365

ES 2 795 989 T3

Thr Tyr Ser Ser Ile Val Asp Ala Val Lys Thr Phe Ala Asp Gly Phe
 370 375 380

Val Ser Ile Val Glu Thr His Ala Ala Ser Asn Gly Ser Met Ser Glu
 385 390 395 400

Gln Tyr Asp Lys Ser Asp Gly Glu Gln Leu Ser Ala Arg Asp Leu Thr
 405 410 415

Trp Ser Tyr Ala Ala Leu Leu Thr Ala Asn Asn Arg Arg Asn Ser Val
 420 425 430

Val Pro Ala Ser Trp Gly Glu Thr Ser Ala Ser Ser Val Pro Gly Thr
 435 440 445

Cys

<210> 7

< 211> 450

< 212> PRT

5 < 213> Aspergillus oryzae

<400> 7

Gln Ser Asp Leu Asn Ala Phe Ile Glu Ala Gln Thr Pro Ile Ala Lys
 1 5 10 15

Gln Gly Tyr Leu Asn Asn Ile Gly Ala Asp Gly Lys Leu Val Glu Gly
 20 25 30

Ala Ala Ala Gly Ile Val Tyr Ala Ser Pro Ser Lys Ser Asn Pro Asp
 35 40 45

Tyr Phe Tyr Thr Trp Thr Arg Asp Ala Gly Leu Thr Met Glu Glu Tyr
 50 55 60

Ile Glu Gln Phe Ile Gly Gly Asp Ala Thr Leu Glu Ser Thr Ile Gln
 65 70 75 80

Asn Tyr Val Asp Ser Gln Ala Asn Glu Gln Ala Val Ser Asn Pro Ser
 85 90 95

Gly Gly Leu Ser Asp Gly Ser Gly Leu Ala Glu Pro Lys Phe Tyr Tyr
 100 105 110

Asn Ile Ser Gln Phe Thr Asp Ser Trp Gly Arg Pro Gln Arg Asp Gly
 115 120 125

ES 2 795 989 T3

Pro Ala Leu Arg Ala Ser Ala Leu Ile Ala Tyr Gly Asn Ser Leu Ile
130 135 140

Ser Ser Asp Lys Gln Ser Val Val Lys Ala Asn Ile Trp Pro Ile Tyr
145 150 155 160

Gln Asn Asp Leu Ser Tyr Val Gly Gln Tyr Trp Asn Gln Thr Gly Phe
165 170 175

Asp Leu Trp Glu Glu Val Gln Gly Ser Ser Phe Phe Thr Val Ala Val
180 185 190

Gln His Lys Ala Leu Val Glu Gly Asp Ala Phe Ala Lys Ala Leu Gly
195 200 205

Glu Glu Cys Gln Ala Cys Ser Val Ala Pro Gln Ile Leu Cys His Leu
210 215 220

Gln Asp Phe Trp Asn Gly Ser Ala Val Leu Ser Asn Leu Pro Thr Asn
225 230 235 240

Gly Arg Ser Gly Leu Asp Thr Asn Ser Leu Leu Gly Ser Ile His Thr
245 250 255

Phe Asp Pro Ala Ala Ala Cys Asp Asp Thr Thr Phe Gln Pro Cys Ser
260 265 270

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

Val Tyr Gly Ile Asn Asn Gly Arg Gly Ala Gly Lys Ala Ala Ala Val
290 295 300

Gly Pro Tyr Ala Glu Asp Thr Tyr Gln Gly Gly Asn Pro Trp Tyr Leu
305 310 315 320

Thr Thr Leu Val Ala Ala Glu Leu Leu Tyr Asp Ala Leu Tyr Gln Trp
325 330 335

Asp Lys Gln Gly Gln Val Asn Val Thr Glu Thr Ser Leu Pro Phe Phe
340 345 350

Lys Asp Leu Ser Ser Asn Val Thr Thr Gly Ser Tyr Ala Lys Ser Ser
355 360 365

Ser Ala Tyr Glu Ser Leu Thr Ser Ala Val Lys Thr Tyr Ala Asp Gly
370 375 380

ES 2 795 989 T3

Phe Ile Ser Val Val Gln Glu Tyr Thr Pro Asp Gly Gly Ala Leu Ala
385 390 395 400

Glu Gln Tyr Ser Arg Asp Gln Gly Thr Pro Val Ser Ala Ser Asp Leu
405 410 415

Thr Trp Ser Tyr Ala Ala Phe Leu Ser Ala Val Gly Arg Arg Asn Gly
420 425 430

Thr Val Pro Ala Ser Trp Gly Ser Ser Thr Ala Asn Ala Val Pro Ser
435 440 445

Gln Cys
450

<210> 8

< 211> 441

< 212> PRT

5 < 213> Humicola grisea

<400> 8

Ala Ala Val Asp Thr Phe Ile Asn Thr Glu Lys Pro Ile Ala Trp Asn
1 5 10 15

Lys Leu Leu Ala Asn Ile Gly Pro Asn Gly Lys Ala Ala Pro Gly Ala
20 25 30

Ala Ala Gly Val Val Ile Ala Ser Pro Ser Arg Thr Asp Pro Pro Tyr
35 40 45

Phe Phe Thr Trp Thr Pro Asp Ala Ala Leu Val Leu Thr Gly Ile Ile
50 55 60

Glu Ser Leu Gly His Asn Tyr Asn Thr Thr Leu Gln Gln Val Ser Asn
65 70 75 80

Pro Ser Gly Thr Phe Ala Asp Gly Ser Gly Leu Gly Glu Ala Lys Phe
85 90 95

Asn Val Asp Leu Thr Ala Phe Thr Gly Glu Trp Gly Arg Pro Gln Arg
100 105 110

Asp Gly Pro Pro Leu Arg Ala Ile Ala Leu Ile Gln Tyr Ala Lys Trp
115 120 125

Leu Ile Ala Asn Gly Tyr Lys Ser Thr Ala Lys Ser Val Val Trp Pro
130 135 140

ES 2 795 989 T3

Val Val Lys Asn Asp Leu Ala Tyr Thr Ala Gln Tyr Trp Asn Glu Thr
145 150 155 160

Gly Phe Asp Leu Trp Glu Glu Val Pro Gly Ser Ser Phe Phe Thr Ile
165 170 175

Ala Ser Ser His Arg Ala Leu Thr Glu Gly Ala Tyr Leu Ala Ala Gln
180 185 190

Leu Asp Thr Glu Cys Pro Pro Cys Thr Thr Val Ala Pro Gln Val Leu
195 200 205

Cys Phe Gln Gln Ala Phe Trp Asn Ser Lys Gly Asn Tyr Val Val Ser
210 215 220

Thr Ser Thr Ala Gly Glu Tyr Arg Ser Gly Lys Asp Ala Asn Ser Ile
225 230 235 240

Leu Ala Ser Ile His Asn Phe Asp Pro Glu Ala Gly Cys Asp Asn Leu
245 250 255

Thr Phe Gln Pro Cys Ser Glu Arg Ala Leu Ala Asn His Lys Ala Tyr
260 265 270

Val Asp Ser Phe Arg Asn Leu Tyr Ala Ile Asn Lys Gly Ile Ala Gln
275 280 285

Gly Lys Ala Val Ala Val Gly Arg Tyr Ser Glu Asp Val Tyr Tyr Asn
290 295 300

Gly Asn Pro Trp Tyr Leu Ala Asn Phe Ala Ala Ala Glu Gln Leu Tyr
305 310 315 320

Asp Ala Ile Tyr Val Trp Asn Lys Gln Gly Ser Ile Thr Val Thr Ser
325 330 335

Val Ser Leu Pro Phe Phe Arg Asp Leu Val Ser Ser Val Ser Thr Gly
340 345 350

Thr Tyr Ser Lys Ser Ser Ser Thr Phe Thr Asn Ile Val Asn Ala Val
355 360 365

Lys Ala Tyr Ala Asp Gly Phe Ile Glu Val Ala Ala Lys Tyr Thr Pro
370 375 380

Ser Asn Gly Ala Leu Ala Glu Gln Tyr Asp Arg Asn Thr Gly Lys Pro
385 390 395 400

ES 2 795 989 T3

Asp Ser Ala Ala Asp Leu Thr Trp Ser Tyr Ser Ala Phe Leu Ser Ala
 405 410 415

Ile Asp Arg Arg Ala Gly Leu Val Pro Pro Ser Trp Arg Ala Ser Val
 420 425 430

Ala Lys Ser Gln Leu Pro Ser Thr Cys
 435 440

<210> 9

< 211> 452

< 212> PRT

5 < 213> Hypocrea vinosa

<400> 9

Ser Val Asp Asp Phe Ile Asn Thr Gln Thr Pro Ile Ala Leu Asn Asn
 1 5 10 15

Leu Leu Cys Asn Val Gly Pro Asp Gly Cys Arg Ala Phe Gly Thr Ser
 20 25 30

Ala Gly Ala Val Ile Ala Ser Pro Ser Thr Thr Asp Pro Asp Tyr Tyr
 35 40 45

Tyr Met Trp Thr Arg Asp Ser Ala Leu Val Phe Lys Asn Ile Val Asp
 50 55 60

Arg Phe Thr Gln Gln Tyr Asp Ala Gly Leu Gln Arg Arg Ile Glu Gln
 65 70 75 80

Tyr Ile Ser Ala Gln Val Thr Leu Gln Gly Ile Ser Asn Pro Ser Gly
 85 90 95

Ser Leu Ser Asp Gly Ser Gly Leu Gly Glu Pro Lys Phe Glu Leu Thr
 100 105 110

Leu Ser Gln Phe Thr Gly Asn Trp Gly Arg Pro Gln Arg Asp Gly Pro
 115 120 125

Ala Leu Arg Ala Ile Ala Leu Ile Gly Tyr Ser Lys Trp Leu Ile Asn
 130 135 140

Asn Asn Tyr Gln Ser Thr Val Ser Asn Ile Ile Trp Pro Ile Val Arg
 145 150 155 160

Asn Asp Leu Asn Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Phe Asp
 165 170 175

ES 2 795 989 T3

Leu Trp Glu Glu Val Asn Gly Ser Ser Phe Phe Thr Val Ala Asn Gln
 180 185 190
 His Arg Ala Leu Val Glu Gly Ala Thr Leu Ala Ala Thr Leu Gly Gln
 195 200 205
 Ser Gly Ser Thr Tyr Ser Ser Val Ala Pro Gln Ile Leu Cys Phe Leu
 210 215 220
 Gln Arg Phe Trp Val Ser Gly Gly Tyr Ile Asp Ser Asn Ile Asn Thr
 225 230 235 240
 Asn Glu Gly Arg Thr Gly Lys Asp Ala Asn Ser Leu Leu Ala Ser Ile
 245 250 255
 His Thr Phe Asp Pro Ser Leu Gly Cys Asp Ala Ser Thr Phe Gln Pro
 260 265 270
 Cys Ser Asp Lys Ala Leu Ser Asn Leu Lys Val Val Val Asp Ser Phe
 275 280 285
 Arg Ser Ile Tyr Gly Val Asn Lys Gly Ile Pro Ala Gly Ser Ala Val
 290 295 300
 Ala Ile Gly Arg Tyr Pro Glu Asp Val Tyr Phe Asn Gly Asn Pro Trp
 305 310 315 320
 Tyr Leu Ala Thr Phe Ala Ala Ala Glu Gln Leu Tyr Asp Ser Val Tyr
 325 330 335
 Val Trp Lys Lys Thr Gly Ser Ile Thr Val Thr Ser Thr Ser Ser Ala
 340 345 350
 Phe Phe Gln Glu Leu Val Pro Gly Val Ala Ala Gly Thr Tyr Ser Ser
 355 360 365
 Ser Gln Ser Thr Phe Thr Ser Ile Ile Asn Ala Ile Ser Thr Tyr Ala
 370 375 380
 Asp Gly Phe Leu Ser Glu Ala Ala Lys Tyr Val Pro Ala Asp Gly Ser
 385 390 395 400
 Leu Ala Glu Gln Phe Asp Arg Asn Thr Gly Thr Pro Leu Ser Ala Val
 405 410 415
 His Leu Thr Trp Ser Tyr Ala Ser Phe Leu Thr Ala Ala Ala Arg Arg
 420 425 430
 Ala Gly Val Val Pro Pro Ser Trp Ala Ser Ser Gly Ala Asn Thr Val
 435 440 445
 Pro Ser Ser Cys
 450
 <210> 10
 <211> 25

< 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 5 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <400> 10
 tcgcgtaac gctagcatgg atctc 25
 <210> 11
 < 211> 31
 10 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 15 <220>
 < 221> base_modificada
 < 222> (10)..(11)
 < 223> a, c, g, t, desconocido u otro
 <400> 11
 20 tctgttgacn nsttcacag caccgagacg c 31
 <210> 12
 < 211> 34
 < 212> ADN
 < 213> Secuencia Artificial
 25 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 30 < 222> (13)..(14)
 < 223> a, c, g, t, desconocido u otro
 <400> 12
 tctgttgacg acnnsatcag caccgagacg ccta 34
 <210> 13
 35 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 40 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 45 <400> 13
 atcagcaccg agacgcctnn sgactgaac aatctcttt 40

<210> 14
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 5 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 10 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 14
 ctttgcaatg ttgtcctnn sggatgccgt gcattcggca 40
 <210> 15
 15 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 20 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 25 <400> 15
 cctgatggat gccgtgcann sggcacatca gctggtgcgg 40
 <210> 16
 < 211> 40
 < 212> ADN
 30 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 35 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 16
 attgcatctc ccagcacann sgaccggac tactattaca 40
 40 <210> 17
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 45 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"

<220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 5 <400> 17
 gcatctccca gcacaattnn scggactac tattacatgt 40
 <210> 18
 < 211> 40
 < 212> ADN
 10 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 15 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 18
 tctccagca caattgacnn sgactactat tacatgtgga 40
 20 <210> 19
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 25 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 30 < 223> a, c, g, t, desconocido u otro
 <400> 19
 ccagcaciaa ttgaccggnn stactattac atgtggacgc 40
 <210> 20
 < 211> 40
 35 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 40 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 20
 45 agcacaattg acccgacnn stattacatg tggacgag 40
 <210> 21

< 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 5 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 10 < 223> a, c, g, t, desconocido u otro
 <400> 21
 attgaccg actactatnn satgtggacg cgagatagcg 40
 <210> 22
 < 211> 40
 15 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 20 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 22
 25 ccgactact attacatgnn sacgcgagat agcgctcttg 40
 <210> 23
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 30 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 35 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 23
 gaccgcttca ccgaaacgnn s gatgcgggc ctgcagcgcc 40
 <210> 24
 40 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 45 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>

< 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 24
 5 acgtacgatg cgggcctggn scgccgcatc gagcagtaca 40
 <210> 25
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 10 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 15 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 25
 tagcatgctg gcctgcaggn scgcatcgag cagtacatta 40
 <210> 26
 20 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 25 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 30 <400> 26
 ctccagggcc tctctaacnn stcgggctcc ctcgcgagc 40
 <210> 27
 < 211> 40
 < 212> ADN
 35 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 40 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 27
 ccctcgggct ccctcgcggn sggctctggt ctcgcgagc 40
 45 <210> 28
 < 211> 40

< 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 5 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 10 <400> 28
 aagtttgagt tgaccctggn sccttcacc ggcaactggg 40
 <210> 29
 < 211> 40
 < 212> ADN
 15 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 20 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 29
 gagttgacct tgaagcctnn saccggaac tggggtcgac 40
 25 <210> 30
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 30 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 35 < 223> a, c, g, t, desconocido u otro
 <400> 30
 ctgaagcctt tcaccgcnn stggggtcga cgcagcggg 40
 <210> 31
 < 211> 40
 40 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 45 <220>
 < 221> base_modificada

< 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 31
 5 ttcaccggca actggggtnn sccgcagcgg gatggcccag 40
 <210> 32
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 10 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 15 < 223> a, c, g, t, desconocido u otro
 <400> 32
 aactggggtc gaccgcagnn sgcagccca gctctgcgag 40
 <210> 33
 < 211> 40
 20 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 25 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 33
 30 aagtggctca tcaacaacnn statcagtcg actgtgtcca 40
 <210> 34
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 35 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 40 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 34
 ctcatcaaca acaactatnn stcgactgtg tccaacgtca 40
 <210> 35
 45 < 211> 40
 < 212> ADN

< 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 5 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 35
 10 ctcaactatg ttgccagann stggaaccaa accggctttg 40
 <210> 36
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 15 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 20 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 36
 gttgccagtg actggaacnn saccggcttt gacctctggg 40
 <210> 37
 25 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 30 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 35 <400> 37
 tactggaacc aaaccggcnn sgacctctgg gaagaagtca 40
 <210> 38
 < 211> 40
 < 212> ADN
 40 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 45 < 221> base_modificada
 < 222> (19)..(20)

< 223> a, c, g, t, desconocido u otro
 <400> 38
 caaacggct ttgacctcnn sgaagaagtc aatgggagct 40
 <210> 39
 5 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 10 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 15 <400> 39
 ggcttgacc tctgggaann sgtcaatggg agctcattct 40
 <210> 40
 < 211> 40
 < 212> ADN
 20 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 25 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 40
 30 ttgacctct gggaagaann saatgggagc tcattctta 40
 <210> 41
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 35 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 40 < 223> a, c, g, t, desconocido u otro
 <400> 41
 cttgctgcca ctctggcnn stcggaagc gcttattcat 40
 <210> 42
 < 211> 40
 45 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 5 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 42
 actcttgcc agtcgggann sgcttattca tctgttgctc 40
 10 <210> 43
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 15 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 20 < 223> a, c, g, t, desconocido u otro
 <400> 43
 tgctttctcc aacgattcnn sgtgtcgtct ggtggatagc 40
 <210> 44
 < 211> 40
 25 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 30 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 44
 35 gactccaaca tcaacaccnn sgagggcagg actggcaagg 40
 <210> 45
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 40 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 45 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro

<400> 45
 tccaacatca acaccaacnn sggcaggact ggcaaggatg 40
 <210> 46
 < 211> 40
 5 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 10 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 46
 15 atcaacacca acgagggcnn sactggcaag gatgtcaact 40
 <210> 47
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 20 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 25 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 47
 gtcgactcct tccgctcnn stacggcgtg aacaaggga 40
 <210> 48
 30 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 35 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 40 <400> 48
 tcctccgct ccatctacnn sgtgaacaag ggcattcctg 40
 <210> 49
 < 211> 40
 < 212> ADN
 45 < 213> Secuencia Artificial
 <220>

< 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 5 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 49
 tccatctacg gcgtagaacnn sggcattcct gccggtgctg 40
 <210> 50
 10 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 15 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 20 <400> 50
 gctgccgtcg ccattggcnn statgcagag gatgtgtact 40
 <210> 51
 < 211> 40
 < 212> ADN
 25 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 30 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 51
 gccgtcgcca ttggccggnn sgcagaggat gtgtactaca 40
 35 <210> 52
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 40 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 45 < 223> a, c, g, t, desconocido u otro
 <400> 52

attggccggt atgcagagnn sgtgtactac aacggcaacc 40
 <210> 53
 < 211> 40
 < 212> ADN
 5 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 10 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 53
 ggccggtatg cagaggatnn stactacaac ggcaaccctt 40
 15 <210> 54
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 20 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 25 < 223> a, c, g, t, desconocido u otro
 <400> 54
 cggtatgcag aggatgtggn stacaacggc aacccttggt 40
 <210> 55
 < 211> 40
 30 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 35 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 55
 40 tatgcagagg atgtgtacnn saacggcaac ccttggtatc 40
 <210> 56
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 45 <220>
 < 221> fuente

< 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 5 < 223> a, c, g, t, desconocido u otro
 <400> 56
 gcagaggatg tgtactacnn sggcaaccct tggatcttg 40
 <210> 57
 < 211> 40
 10 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 15 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 57
 20 tactacaacg gcaaccctnn statctgct acattgctg 40
 <210> 58
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 25 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 30 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 58
 gatgccatct acgtctggnn saagacgggc tccatcacgg 40
 <210> 59
 35 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 40 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 45 <400> 59
 gccatctacg tctggaagnn sacgggctcc atcacggtga 40

<210> 60
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 5 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 10 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 60
 tccatcacgg tgaccgccnn stccctggcc ttctccagg 40
 <210> 61
 15 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 20 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 25 <400> 61
 acctcctgg ccttctcnn sgagctgtt cctggcgtga 40
 <210> 62
 < 211> 40
 < 212> ADN
 30 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 35 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 62
 gagctgttc ctggcgtgnn sgccgggacc tactccagca 40
 40 <210> 63
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 45 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"

<220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 5 <400> 63
 gtgacggccg ggacctacnn sagcagctct tcgacctta 40
 <210> 64
 < 211> 40
 < 212> ADN
 10 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 15 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 64
 acggccggga cctactccnn sagctctcg accttacca 40
 20 <210> 65
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 25 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 30 < 223> a, c, g, t, desconocido u otro
 <400> 65
 agctctcga cctttaccnn satcatcaac gccgtctcga 40
 <210> 66
 < 211> 40
 35 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 40 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 66
 45 ctgagcgagg ctgccaagnn sgtccccgcc gacggttcgc 40
 <210> 67

< 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 5 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 10 < 223> a, c, g, t, desconocido u otro
 <400> 67
 gctgccaagt acgtcccn sgacggttcg ctggccgagc 40
 <210> 68
 < 211> 40
 15 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 20 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 68
 25 tagtccccg cgcacggttn sctggccgag cagttgacc 40
 <210> 69
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 30 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 35 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 69
 ctggccgagc agttgacnn saacagcggc actccgctgt 40
 <210> 70
 40 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 45 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>

< 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 70
 5 gccgagcagt ttgaccgcn sagcggcact ccgctgtctg 40
 <210> 71
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 10 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 15 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 71
 ttgaccgca acagcggcnn scgctgtct gcgcttcacc 40
 <210> 72
 20 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 25 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 30 <400> 72
 actccgctgt ctgcgctnn scgacgtgg tcgtacgct 40
 <210> 73
 < 211> 40
 < 212> ADN
 35 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 40 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 73
 45 tctgcgctc acctgacgcn stcgtacgcc tcgttctga 40
 <210> 74
 < 211> 40

< 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 5 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 10 <400> 74
 ttgacagcca cggcccgttn sgctggcatc gtgccccct 40
 <210> 75
 < 211> 40
 < 212> ADN
 15 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 20 < 221> base_modificada
 < 222> (19)..(20)
 < 223> a, c, g, t, desconocido u otro
 <400> 75
 acgccccgtc gggctggcnn sgtgcccccc tcgtgggcca 40
 25 <210> 76
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 30 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (19)..(20)
 35 < 223> a, c, g, t, desconocido u otro
 <400> 76
 agcgctagca cgatccccnn sacgtgctcc ggcgctccg 40
 <210> 77
 < 211> 24
 40 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 45 <400> 77
 gtaacatcag agatttgag acac 24

<210> 78
 < 211> 31
 < 212> ADN
 < 213> Secuencia Artificial
 5 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 10 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 78
 gcgtctcggg gctgatgaas nngtcaacag a 31
 <210> 79
 15 < 211> 34
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 20 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 25 <400> 79
 taggcgtctc ggtgctgats nngtcgtaa caga 34
 <210> 80
 < 211> 40
 < 212> ADN
 30 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 35 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 80
 aaagaagatt gttcagtgc nnggcgtct cgggtctgat 40
 40 <210> 81
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 45 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"

<220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 5 <400> 81
 tgccgaatgc acggcatccs nnaggaccaa cattgcaaag 40
 <210> 82
 < 211> 40
 < 212> ADN
 10 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 15 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 82
 ccgcaccagc tgatgtgccs nntgcacggc atccatcagg 40
 20 <210> 83
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 25 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 30 < 223> a, c, g, t, desconocido u otro
 <400> 83
 tgtaatagta gtccgggtcs nntgtgctgg gagatgcaat 40
 <210> 84
 < 211> 40
 35 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 40 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 84
 45 acatgtaata gtagtccggs nnaattgtgc tgggagatgc 40
 <210> 85

< 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 5 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 10 < 223> a, c, g, t, desconocido u otro
 <400> 85
 tccacatgta atagtagtcs nngtcaattg tgctgggaga 40
 <210> 86
 < 211> 40
 15 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 20 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 86
 25 ggcgccat gtaatagtas nncgggtcaa ttgtgctggg 40
 <210> 87
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 30 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 35 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 87
 ctcgcgtcca catgtaatas nngtccgggt caattgtgct 40
 <210> 88
 40 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 45 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>

< 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 88
 5 cgctatctcg cgtccacats nnatagtagt ccgggtcaat 40
 <210> 89
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 10 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 15 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 < 400> 89
 caagagcgct atctcgcgts nncatgtaat agtagtccgg 40
 <210> 90
 20 < 211> 40
 <212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 25 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 30 <400> 90
 ggcgctgcag gcccgcatcs nncggttcgg tgaagcggtc 40
 <210> 91
 < 211> 40
 < 212> ADN
 35 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 40 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 91
 tgtactgctc gatcggcgs nncaggcccg catcgtacgt 40
 45 <210> 92
 < 211> 40

< 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 5 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 10 <400> 92
 taatgtactg ctcgatgcgs nnctgcaggc ccgcatcgta 40
 <210> 93
 < 211> 40
 < 212> ADN
 15 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 20 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 93
 cgtccgcgag ggagcccgas nngttagaga ggccctggag 40
 25 <210> 94
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 30 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 35 < 223> a, c, g, t, desconocido u otro
 <400> 94
 gctcgccgag accagagccs nncgcgaggg agcccgaggg 40
 <210> 95
 < 211> 40
 40 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 45 <220>
 < 221> base_modificada

< 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 95
 cccagttgcc ggtgaaaggs nncagggtca actcaaact 40
 5 <210> 96
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 10 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 15 < 223> a, c, g, t, desconocido u otro
 <400> 96
 gtcgacccca gttgccgts nnaggctca gggtaactc 40
 <210> 97
 < 211> 40
 20 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 25 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 97
 30 cccgctgctg tcgaccccas nngccggtga aaggctcag 40
 <210> 98
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 35 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 40 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 98
 ctgggccatc cgcgtgcggs nnacccagtg tgcgggtgaa 40
 <210> 99
 45 < 211> 40
 < 212> ADN

< 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 5 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 99
 10 ctgcagagc tggccatcs nnctgctggtc gaccccagtt 40
 <210> 100
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 15 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 20 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 100
 tggacacagt cgactgatas nngttgttga tgagccact 40
 <210> 101
 25 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 30 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 35 <400> 101
 tgacgttga cacagtgas nnatagttgt tgttgatgag 40
 <210> 102
 < 211> 40
 < 212> ADN
 40 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 45 < 221> base_modificada
 < 222> (21)..(22)

< 223> a, c, g, t, desconocido u otro
 <400> 102
 caaagccggt ttggtccas nnctgggcaa catagttgag 40
 <210> 103
 5 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 10 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 15 <400> 103
 ccagaggtc aaagccggt nngttccagt actgggcaac 40
 <210> 104
 < 211> 40
 < 212> ADN
 20 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 25 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 104
 tgactcttc ccagaggtcs nngccggtt ggtccagta 40
 30 <210> 105
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 35 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 40 < 223> a, c, g, t, desconocido u otro
 <400> 105
 agctccatt gactcttcs nngaggtcaa agccggttg 40
 <210> 106
 < 211> 40
 45 < 212> ADN
 < 213> Secuencia Artificial

<220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 5 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 106
 agaatgagct cccattgacs nntcccaga ggtcaaagcc 40
 10 <210> 107
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 15 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 20 < 223> a, c, g, t, desconocido u otro
 <400> 107
 taaagaatga gctcccatts nnttctccc agaggtcaaa 40
 <210> 108
 < 211> 40
 25 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 30 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 108
 35 atgaataagc gctcccgas nngccaagag tggcagcaag 40
 <210> 109
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 40 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 45 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro

<400> 109
 gagcaacaga tgaataagcs nntcccgact ggccaagagt 40
 <210> 110
 < 211> 40
 5 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 10 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 110
 15 cgtatccacc agacgacacs nngaatacgtt ggagaaagca 40
 <210> 111
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 20 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 25 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 111
 ccttgccagt cctgccctcs nnggtgtga tgttgagtc 40
 <210> 112
 30 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 35 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 40 <400> 112
 catcctgcc agtccctgccs nngtgggtgt tgatgttga 40
 <210> 113
 < 211> 40
 < 212> ADN
 45 < 213> Secuencia Artificial
 <220>

< 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 5 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 113
 agttgacatc cttgccagts nngccctcgt tgggttgat 40
 <210> 114
 10 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 15 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 20 <400> 114
 tggccttgtt cacgccgtas nnggagcggga aggagtcgac 40
 <210> 115
 < 211> 40
 < 212> ADN
 25 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 30 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 115
 caggaatgcc cttgtcacs nngtagatgg agcgggaagga 40
 35 <210> 116
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 40 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 45 < 223> a, c, g, t, desconocido u otro
 <400> 116

cagcaccggc aggaatgccs nngttcacgc cgtagatgga 40
 <210> 117
 < 211> 40
 < 212> ADN
 5 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 10 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 117
 agtacacatc ctctgcatas nngccaatgg cgacggcagc 40
 15 <210> 118
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 20 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 25 < 223> a, c, g, t, desconocido u otro
 <400> 118
 tgtagtacac atcctctgcs nncggccaa tggcgacggc 40
 <210> 119
 < 211> 40
 30 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 35 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 119
 40 ggttgccgtt gtagtacacs nnctctgcat accggccaat 40
 <210> 120
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 45 <220>
 < 221> fuente

< 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 5 < 223> a, c, g, t, desconocido u otro
 <400> 120
 aagggtgcc gttgtagtas nnatcctctg cataccggcc 40
 <210> 121
 < 211> 40
 10 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 15 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 121
 20 accaagggtt gccgttgtagtas nncacatcct ctgcataccg 40
 <210> 122
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 25 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 30 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 122
 gataccaagg gttgccgtagtas nngtacacat cctctgcata 40
 <210> 123
 35 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 40 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 45 <400> 123
 caagatacca agggtagccs nngtagtagaca catcctctgc 40

<210> 124
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 5 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 10 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 124
 cagcaaatgt agcaagatas nnagggtgc cgtgtagta 40
 <210> 125
 15 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 20 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 25 <400> 125
 ccgtgatgga gcccgctts nccagacgt agatggcatc 40
 <210> 126
 < 211> 40
 < 212> ADN
 30 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 35 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 126
 tcaccgtgat ggagcccgts nnctccaga cgtagatggc 40
 40 <210> 127
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 45 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"

<220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 5 <400> 127
 cctggaagaa ggccagggas nnggcggtca ccgtgatgga 40
 <210> 128
 < 211> 40
 < 212> ADN
 10 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 15 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 128
 tcaccgagg aacaagctcs nngaagaagg ccagggaggt 40
 20 <210> 129
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 25 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 30 < 223> a, c, g, t, desconocido u otro
 <400> 129
 tgctggagta ggtcccggcs nncaccgag gaacaagctc 40
 <210> 130
 < 211> 40
 35 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 40 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 130
 45 taaaggtcga agagctgcts nngtaggtcc cggccgtcac 40
 <210> 131

< 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 5 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 10 < 223> a, c, g, t, desconocido u otro
 <400> 131
 tggtaaaggt cgaagagcts nnggagtagg tcccggccgt 40
 <210> 132
 < 211> 40
 15 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 20 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 132
 25 tcgagacggc gttgatgats nnggtaaagg tcgaagagct 40
 <210> 133
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 30 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 35 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 133
 gcgaaccgtc ggcggggacs nnctggcag cctcgctgag 40
 <210> 134
 40 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 45 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>

< 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 134
 5 gctcggccag cgaaccgtcs nnggggacgt acttggcagc 40
 <210> 135
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 10 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 15 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 135
 ggtcaaactg ctcggccags nnaccgtcgg cggggacgta 40
 <210> 136
 20 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 25 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 30 <400> 136
 acagcggagt gccgctgts nngtcaaact gctcggccag 40
 <210> 137
 < 211> 40
 < 212> ADN
 35 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 40 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 137
 cagacagcgg agtgccgcts nngcgggtcaa actgctcggc 40
 45 <210> 138
 < 211> 40

< 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 5 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 10 <400> 138
 ggtgaagcgc agacagcggg nngccgctgt tgcggtcaaa 40
 <210> 139
 < 211> 40
 < 212> ADN
 15 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 20 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 139
 aggcgtacga ccacgtcags nnaagcgcag acagcggagt 40
 25 <210> 140
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 30 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 35 < 223> a, c, g, t, desconocido u otro
 <400> 140
 tcaagaacga ggcgtacgas nncgtcaggt gaagcgcaga 40
 <210> 141
 < 211> 40
 40 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 45 <220>
 < 221> base_modificada

< 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 141
 aggggggcac gatgccagcs nnacgggccg tggctgtcaa 40
 5 <210> 142
 < 211> 40
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 10 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 15 < 223> a, c, g, t, desconocido u otro
 <400> 142
 tggcccacga ggggggcacs nngccagccc gacgggccgt 40
 <210> 143
 < 211> 40
 20 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 25 <220>
 < 221> base_modificada
 < 222> (21)..(22)
 < 223> a, c, g, t, desconocido u otro
 <400> 143
 30 cggacgcgcc ggagcagcgs nnggggatcg tgctagcgct 40
 <210> 144
 < 211> 33
 < 212> ADN
 < 213> Secuencia Artificial
 35 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <400> 144
 gcatctccca gcacacgaga cccggactac tat 33
 40 <210> 145
 < 211> 33
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 45 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"

ES 2 795 989 T3

<400> 145
gcatctccca gcacatacga cccggactac tat 33
<210> 146
< 211> 33
5 < 212> ADN
< 213> Secuencia Artificial
<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
10 <400> 146
gatgcgggcc tgcagctgcg catcgagcag tac 33
<210> 147
< 211> 40
< 212> ADN
15 < 213> Secuencia Artificial
<220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
20 <400> 147
ctgaagcctt tcaccggcac ctggggtcga cgcagcggg 40
<210> 148
< 211> 39
< 212> ADN
< 213> Secuencia Artificial
25 <220>
< 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
<400> 148
tgaagccttt caccggctac tggggtcgac cgcagcggg 39
30 <210> 149
< 211> 40
< 212> ADN
< 213> Secuencia Artificial
<220>
35 < 221> fuente
< 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
<400> 149
ctgaagcctt tcaccggcga ctggggtcga cgcagcggg 40
<210> 150
40 < 211> 36
< 212> ADN
< 213> Secuencia Artificial
<220>
< 221> fuente
45 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
<400> 150

agtggctcat caacaacgas tatcagtcga ctgtgt 36
 <210> 151
 < 211> 36
 < 212> ADN
 5 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <400> 151
 10 agtggctcat caacaacacc tatcagtcga ctgtgt 36
 <210> 152
 < 211> 34
 < 212> ADN
 < 213> Secuencia Artificial
 15 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <400> 152
 20 gtggctcatc aacaatggta tcagtcgact gtgt 34
 <210> 153
 < 211> 36
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 25 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <400> 153
 agtggctcat caacaacctg tatcagtcga ctgtgt 36
 <210> 154
 30 <211> 36
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 35 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <400> 154
 agtggctcat caacaactcc tatcagtcga ctgtgt 36
 <210> 155
 < 211> 38
 40 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 45 <400> 155
 ttgccagta ctggaacgas accggcttg acctctgg 38

<210> 156
 < 211> 38
 < 212> ADN
 < 213> Secuencia Artificial
 5 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <400> 156
 ttgccagta ctggaacstg accggcttg acctctgg 38
 10 <210> 157
 < 211> 38
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 15 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <400> 157
 ttgccagta ctggaacacc accggcttg acctctgg 38
 20 <210> 158
 < 211> 38
 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 25 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 <400> 158
 ttgccagta ctggaaccga accggcttg acctctgg 38
 <210> 159
 < 211> 38
 30 < 212> ADN
 < 213> Secuencia Artificial
 <220>
 < 221> fuente
 < 223> /nota="Descripción de Secuencia Artificial: Cebador sintético"
 35 <400> 159
 ttgccagta ctggaactgc accggcttg acctctgg 38
 <210> 160
 < 211> 6
 < 212> PRT
 40 < 213> Trichoderma reesei
 <400> 160
Ser Val Asp Asp Phe Ile
 1 5
 <210> 161
 < 211> 599
 45 < 212> PRT

ES 2 795 989 T3

< 213> Secuencia Artificial

<220>

< 221> fuente

< 223> /nota="Descripción de Secuencia Artificial: Synthetic polypeptide"

5

<400> 161

Ser Val Asp Asp Phe Ile Ser Thr Glu Thr Pro Ile Ala Leu Asn Asn
 1 5 10 15

Leu Leu Cys Asn Val Gly Pro Asp Gly Cys Arg Ala Phe Gly Thr Ser
 20 25 30

Ala Gly Ala Val Ile Ala Ser Pro Ser Thr Ile Asp Pro Asp Tyr Tyr
 35 40 45

Tyr Met Trp Thr Arg Asp Ser Ala Leu Val Phe Lys Asn Leu Ile Asp
 50 55 60

Arg Phe Thr Glu Thr Tyr Asp Ala Gly Leu Gln Arg Arg Ile Glu Gln
 65 70 75 80

Tyr Ile Thr Ala Gln Val Thr Leu Gln Gly Asn Ser Asn Pro Ser Gly
 85 90 95

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

Leu Lys Pro Phe Thr Gly Asn Trp Gly Arg Pro Gln Arg Asp Gly Pro
 115 120 125

Ala Leu Arg Ala Ile Ala Leu Ile Gly Tyr Ser Lys Trp Leu Ile Asn
 130 135 140

Asn Asn Tyr Gln Phe Thr Val Ser Asn Val Ile Trp Pro Ile Val Arg
 145 150 155 160

ES 2 795 989 T3

Asn Asp Leu Asn Tyr Val Ala Gln Tyr Trp Asn Gln Thr Gly Phe Asp
 165 170 175
 Leu Trp Glu Glu Val Asn Gly Ser Ser Phe Phe Thr Val Ala Asn Gln
 180 185 190
 His Arg Ala Leu Val Glu Gly Ala Thr Leu Ala Ala Thr Leu Gly Gln
 195 200 205
 Ser Gly Ser Ala Tyr Ser Ser Val Ala Pro Gln Val Leu Cys Phe Leu
 210 215 220
 Gln Arg Phe Trp Val Ser Ser Gly Gly Tyr Val Asp Ser Asn Ile Asn
 225 230 235 240
 Thr Asn Glu Gly Arg Thr Gly Lys Asp Val Asn Ser Val Leu Thr Ser
 245 250 255
 Ile His Thr Phe Asp Pro Asn Leu Gly Cys Asp Ala Gly Thr Phe Gln
 260 265 270
 Pro Cys Ser Asp Lys Ala Leu Ser Asn Leu Lys Val Val Val Asp Ser
 275 280 285
 Phe Arg Ser Ile Tyr Gly Val Asn Lys Gly Ile Pro Ala Gly Ala Ala
 290 295 300
 Val Ala Ile Gly Arg Tyr Ala Glu Asp Val Tyr Tyr Asn Gly Asn Pro
 305 310 315 320
 Trp Tyr Leu Ala Thr Phe Ala Ala Ala Glu Gln Leu Tyr Asp Ala Ile
 325 330 335
 Tyr Val Trp Lys Lys Thr Gly Ser Ile Thr Val Thr Ala Thr Ser Leu
 340 345 350
 Ala Phe Phe Gln Glu Leu Val Pro Gly Val Thr Ala Gly Thr Tyr Ser
 355 360 365
 Ser Ser Ser Ser Thr Phe Thr Asn Ile Ile Asn Ala Val Ser Thr Tyr
 370 375 380
 Ala Asp Gly Phe Leu Ser Glu Ala Ala Lys Tyr Val Pro Ala Asp Gly
 385 390 395 400
 Ser Leu Ala Glu Gln Phe Asp Arg Asn Ser Gly Thr Pro Leu Ser Ala
 405 410 415

ES 2 795 989 T3

Leu His Leu Thr Trp Ser Tyr Ala Ser Phe Leu Thr Ala Thr Ala Arg
 420 425 430

Arg Ala Gly Ile Val Pro Pro Ser Trp Ala Asn Ser Ser Ala Ser Thr
 435 440 445

Ile Pro Ser Thr Cys Ser Gly Ala Ser Val Val Gly Ser Tyr Ser Arg
 450 455 460

Pro Thr Ala Thr Ser Phe Pro Pro Ser Gln Thr Pro Lys Pro Gly Val
 465 470 475 480

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

Val Ala Val Thr Phe His Glu Leu Val Ser Thr Gln Phe Gly Gln Thr
 500 505 510

Val Lys Val Ala Gly Asn Ala Ala Ala Leu Gly Asn Trp Ser Thr Ser
 515 520 525

Ala Ala Val Ala Leu Asp Ala Val Asn Tyr Ala Asp Asn His Pro Leu
 530 535 540

Trp Ile Ala Thr Val Asn Leu Glu Ala Gly Asp Val Val Glu Tyr Lys
 545 550 555 560

Tyr Ile Asn Val Gly Gln Asp Gly Ser Val Thr Trp Glu Ser Asp Pro
 565 570 575

Asn His Thr Tyr Thr Val Pro Ala Val Ala Cys Val Thr Gln Val Val
 580 585 590

Lys Glu Asp Thr Trp Gln Ser
 595

<210> 162

<211> 109

<212> PRT

5

<213> Trichoderma reesie

<400> 162

Cys Ala Thr Pro Thr Ser Val Ala Val Thr Phe His Glu Leu Val Ser
 1 5 10 15

Thr Gln Phe Gly Gln Thr Val Lys Val Ala Gly Asn Ala Ala Ala Leu
 20 25 30

ES 2 795 989 T3

Gly Asn Trp Ser Thr Ser Ala Ala Val Ala Leu Asp Ala Val Asn Tyr
 35 40 45

Ala Asp Asn His Pro Leu Trp Ile Gly Thr Val Asn Leu Glu Ala Gly
 50 55 60

Asp Val Val Glu Tyr Lys Tyr Ile Asn Val Gly Gln Asp Gly Ser Val
 65 70 75 80

Thr Trp Glu Ser Asp Pro Asn His Thr Tyr Thr Val Pro Ala Val Ala
 85 90 95

Cys Val Thr Gln Val Val Lys Glu Asp Thr Trp Gln Ser
 100 105

<210> 163

<211> 112

<212> PRT

5 <213> Humicola grisea

<400> 163

Cys Ala Asp Ala Ser Glu Val Tyr Val Thr Phe Asn Glu Arg Val Ser
 1 5 10 15

Thr Ala Trp Gly Glu Thr Ile Lys Val Val Gly Asn Val Pro Ala Leu
 20 25 30

Gly Asn Trp Asp Thr Ser Lys Ala Val Thr Leu Ser Ala Ser Gly Tyr
 35 40 45

Lys Ser Asn Asp Pro Leu Trp Ser Ile Thr Val Pro Ile Lys Ala Thr
 50 55 60

Gly Ser Ala Val Gln Tyr Lys Tyr Ile Lys Val Gly Thr Asn Gly Lys
 65 70 75 80

Ile Thr Trp Glu Ser Asp Pro Asn Arg Ser Ile Thr Leu Gln Thr Ala
 85 90 95

Ser Ser Ala Gly Lys Cys Ala Ala Gln Thr Val Asn Asp Ser Trp Arg
 100 105 110

<210> 164

<211> 108

<212> PRT

10 <213> Thielavia terrestris

<400> 164

Cys Ser Thr Pro Thr Ala Val Ala Val Thr Phe Asn Glu Arg Val Thr
 1 5 10 15

ES 2 795 989 T3

Thr Gln Trp Gly Gln Thr Ile Lys Val Val Gly Asp Ala Ala Ala Leu
 20 25 30

Gly Gly Trp Asp Thr Ser Lys Ala Val Pro Leu Ser Ala Ala Gly Tyr
 35 40 45

Thr Ala Ser Asp Pro Leu Trp Ser Gly Thr Val Asp Leu Pro Ala Gly
 50 55 60

Leu Ala Val Gln Tyr Lys Tyr Ile Asn Val Ala Ala Asp Gly Gly Val
 65 70 75 80

Thr Trp Glu Ala Asp Pro Asn His Ser Phe Thr Val Pro Ala Ala Cys
 85 90 95

Gly Thr Thr Ala Val Thr Arg Asp Asp Thr Trp Gln
 100 105

<210> 165

< 211> 109

< 212> PRT

5 < 213> Thermomyces lanuginosus

<400> 165

Cys Thr Pro Pro Ser Glu Val Thr Leu Thr Phe Asn Ala Leu Val Asp
 1 5 10 15

Thr Ala Phe Gly Gln Asn Ile Tyr Leu Val Gly Ser Ile Pro Glu Leu
 20 25 30

Gly Ser Trp Asp Pro Ala Asn Ala Leu Leu Met Ser Ala Lys Ser Trp
 35 40 45

Thr Ser Gly Asn Pro Val Trp Thr Leu Ser Ile Ser Leu Pro Ala Gly
 50 55 60

Thr Ser Phe Glu Tyr Lys Phe Ile Arg Lys Asp Asp Gly Ser Ser Asp
 65 70 75 80

Val Val Trp Glu Ser Asp Pro Asn Arg Ser Tyr Asn Val Pro Lys Asp
 85 90 95

Cys Gly Ala Asn Thr Ala Thr Val Asn Ser Trp Trp Arg
 100 105

<210> 166

< 211> 108

< 212> PRT

10 < 213> Talaromyces emersonii

<400> 166

ES 2 795 989 T3

Cys Thr Thr Pro Thr Ser Val Ala Val Thr Phe Asp Glu Ile Val Ser
1 5 10 15

Thr Ser Tyr Gly Glu Thr Ile Tyr Leu Ala Gly Ser Ile Pro Glu Leu
20 25 30

Gly Asn Trp Ser Thr Ala Ser Ala Ile Pro Leu Arg Ala Asp Ala Tyr
35 40 45

Thr Asn Ser Asn Pro Leu Trp Tyr Val Thr Val Asn Leu Pro Pro Gly
50 55 60

Thr Ser Phe Glu Tyr Lys Phe Phe Lys Asn Gln Thr Asp Gly Thr Ile
65 70 75 80

Val Trp Glu Asp Asp Pro Asn Arg Ser Tyr Thr Val Pro Ala Tyr Cys
85 90 95

Gly Gln Thr Thr Ala Ile Leu Asp Asp Ser Trp Gln
100 105

<210> 167

< 211> 108

< 212> PRT

5 < 213> Aspergillus niger

<400> 167

Cys Thr Thr Pro Thr Ala Val Ala Val Thr Phe Asp Leu Thr Ala Thr
1 5 10 15

Thr Thr Tyr Gly Glu Asn Ile Tyr Leu Val Gly Ser Ile Ser Gln Leu
20 25 30

Gly Asp Trp Glu Thr Ser Asp Gly Ile Ala Leu Ser Ala Asp Lys Tyr
35 40 45

Thr Ser Ser Asp Pro Leu Trp Tyr Val Thr Val Thr Leu Pro Ala Gly
50 55 60

Glu Ser Phe Glu Tyr Lys Phe Ile Arg Ile Glu Ser Asp Asp Ser Val
65 70 75 80

Glu Trp Glu Ser Asp Pro Asn Arg Glu Tyr Thr Val Pro Gln Ala Cys
85 90 95

Gly Thr Ser Thr Ala Thr Val Thr Asp Thr Trp Arg
100 105

<210> 168

< 211> 108

< 212> PRT

10 < 213> Aspergillus awamori

<400> 168

ES 2 795 989 T3

Cys Thr Thr Pro Thr Ala Val Ala Val Thr Phe Asp Leu Thr Ala Thr
 1 5 10 15
 Thr Thr Tyr Gly Glu Asn Ile Tyr Leu Val Gly Ser Ile Ser Gln Leu
 20 25 30
 Gly Asp Trp Asp Thr Ser Asp Gly Ile Ala Leu Ser Ala Asp Lys Tyr
 35 40 45
 Thr Ser Ser Asn Pro Leu Trp Tyr Val Thr Val Thr Leu Pro Ala Gly
 50 55 60
 Glu Ser Phe Glu Tyr Lys Phe Ile Arg Ile Glu Ser Asp Asp Ser Val
 65 70 75 80
 Glu Trp Glu Ser Asp Pro Asn Arg Glu Tyr Thr Val Pro Gln Ala Cys
 85 90 95
 Gly Glu Ser Thr Ala Thr Val Thr Asp Thr Trp Arg
 100 105

REIVINDICACIONES

1. Una variante de glucoamilasa que comprende al menos:
- (i) una sustitución de aminoácidos correspondiente a la posición 431 en SEQ ID NO: 2 o una posición equivalente basada en la alineación de la secuencia de aminoácidos con la secuencia de SEQ ID NO: 2; y
- (ii) una o más sustituciones de aminoácidos en posiciones correspondientes a las posiciones 61, 73, 417, 430, 503, 511, 535, 539 o 563 de SEQ ID NO: 2 o posiciones equivalentes basadas en la alineación de la secuencia de aminoácidos con la secuencia de SEQ ID NO: 2;
- en la que la variante de glucoamilasa presenta al menos un 93 % de identidad de secuencia con SEQ ID NO: 2,
- y en la que la variante de glucoamilasa muestra termoestabilidad aumentada y/o actividad específica aumentada en comparación con la glucoamilasa de SEQ ID NO: 2.
2. Variante de glucoamilasa según la reivindicación 1, en la que la sustitución en la posición 431 en SEQ ID NO: 2 se elige de: A431L/Q.
3. Variante de glucoamilasa según la reivindicación 1 o la reivindicación 2, en la que la variante es una variante de una glucoamilasa precursora que presenta un dominio catalítico con al menos un 80 % de identidad de secuencia con SEQ ID NO: 1 o 2, o que presenta un dominio de unión a almidón con al menos un 80 % de identidad de secuencia con SEQ ID NO: 1 o 2.
4. Variante de glucoamilasa según la reivindicación 1 o la reivindicación 2, en la que la variante de glucoamilasa presenta al menos un 95 % o al menos un 99,5 % de identidad de secuencia con SEQ ID NO: 2, o en la que la variante es una variante de una glucoamilasa precursora que es SEQ ID NO: 1 o 2
5. Variante de glucoamilasa según la reivindicación 1 o la reivindicación 2, en la que la variante comprende sustituciones de aminoácidos correspondientes a A431L y uno o más de N61I, G73F, L417R/V, T430A/M, E503A/V, Q511H, A535R, A539R, o N563I/K en SEQ ID NO: 2.
6. Variante de glucoamilasa según la reivindicación 5, en la que la variante comprende además una o más de las siguientes sustituciones: D4L/E/R/S/C/A/Q/W, F5C/M/N/R/S/T/V/W, I12L/R, D24E/L/Y/T, F29L/I/D/C/S/M/W, I43F/R/D/Y/S/Q, D44E/H/K/S/N/Y/F/R/C, Y47W, Y49N, Q70R/K/M/P/G/L/F, Q75R/K/A, R76L/M/K/T/P, P94L, D100W/I/Q/M/P/A/N, N119P/T/Y/D/E, N146S/G/C/H/E/D/T/W/L/F/M, Q148V/Y/H/A/C/D/G/M/R/S/T, Y169D/F, Q172C/A/D/R/E/F/H/V/L/M/N/S/T/V, F175H/A/G/R/S/T/C/W/Y, W178A/C/D/E/F/G/H/K/N/R/S/T/V/Y, E180A/C/G/H/I/L/N/P/Q/R/S/T/V/Y, VI81E/C/D/G/H/I/P/T/Y/S/L/K/F/A, Q208L/A/C/E/N/F/H/T, S211C/R/E/A/Y/W/M/H/L/I/R/Q/T, E243S/R/N/M/Y/A/L, R245A/E/M/I/P/V, I292D/H/P/R/T/N/V/F/L, G294C/D/E/T/Q/I/A, K297F/L/P/T/M/D/N/Q/A/Y/H/S/R/W, R309A/C/G/H/I/N/P/Q/S/T/W/Y/L, Y310E/G/L/P/S/W/R/Q, D313Q, V314A/R/N/D/C/E/Q/G/H/I/L/K/M/F/P/S/T/W/Y, Y315F, Y316Q/R, N317T/H, K340D/T, K341F/D/P/V/G/S, T350S/E/A/N, Q356H/D/E, T363L/R/C/H/W, S368W/D/F/L, S369F, N376Q/T/H/S/V, Y395Q/R/S, A398S/I/T, S401C/V, R408S, N409W/T/K, T412A/H/K/G, R433H/Q, 1436A/T, o S451 M/T/H en SEQ ID NO: 2.
7. Variante de glucoamilasa según cualquiera de las reivindicaciones anteriores que comprende sustituciones múltiples en las posiciones:
- I43/N61/T430/A431/Q511/A539;
- I43/L417/T430/A431/Q511/A539;
- I43/T430/A431/E503/Q511;
- I43/T430/A431/Q511;
- I43/T430/A431/Q511/A539;
- I43/A431/Q511;
- G294/L417/A431;
- G294/L417/A431/Q511;
- L417/T430/A431/Q511/A535/A539/N563; y
- L417/A431/Q511.

8. Variante de glucoamilasa según cualquiera de las reivindicaciones anteriores que comprende uno de los conjuntos de sustitución:
- 5 I43Q/N61I/T430A/**A431L**/Q511H/A539R;
 143Q/L417V/T430A/**A431L**/Q511H/A539R;
 I43Q/T430A/**A431L**/E503A/Q511H;
 I43Q/T430A/**A431L**/Q511H;
 I43Q/T430A/**A431L**/Q511H/A539R;
 I43Q/**A431L**/Q511H;
 G294C/L417R/**A431L**;
 10 G294C/L417R/**A431L,Q**/Q511H;
 G294C/L417V/**A431Q**;
 L417R,V/**A431L,Q**/Q511H; y
 L417V/T430A/**A431L,Q**/Q511H/A535R/A539R/N563I,
 de SEQ ID NO: 2.
- 15
9. Variante de glucoamilasa según cualquiera de las reivindicaciones 1 a 8, en la que la glucoamilasa precursora se elige de una glucoamilasa obtenida a partir de una *Trichoderma* spp., una *Aspergillus* spp., una *Humicola* spp., una *Penicillium* spp., una *Talaromyces* spp., o una *Schizosaccharomyces* spp.
- 20
10. Un polinucleótido que codifica la variante de glucoamilasa de cualquiera de las reivindicaciones 1 a 9.
11. Un vector que comprende el polinucleótido de la reivindicación 10.
12. Una célula huésped que comprende el vector de la reivindicación 11.
- 25
13. Una composición de enzima que comprende la variante de glucoamilasa de cualquiera de las reivindicaciones 1 a 9.
- 30
14. Un método para producir una variante de glucoamilasa en una célula huésped que comprende el cultivo de la célula huésped de la reivindicación 12 con las condiciones adecuadas para la expresión y producción de la variante de glucoamilasa y la producción de la variante de glucoamilasa.
15. Método de acuerdo con la reivindicación 14, que comprende además la recuperación de la variante de glucoamilasa del cultivo.

Proteína precursora TrGA (632 Aminoácidos) (SEQ ID NO:1)

1 MHVLSTAVLL GSVAVQKVLG RPGSSGLSDV TKRSVDDFIS TETPIALNNL
 51 LCNVGPDGCR AFGTSAGAVI ASPSTIDPDY YMWTRDSAL VFKNLIDRFT
 101 ETYDAGLQRR IEQYITAQVT LQGLSNPSGS LADGSGLGEP KFELTLKPFT
 151 GNWGRPQRDG PALRAIALIG YSKWLINNNY QSTVSNVIWP IVRNDLNYVA
 201 QYWNQTGFDL WEEVNGSSFF TVANQHRALV EGATLAATLG QSGSAYSSVA
 251 PQVLCFLQRF WVSSGGYVDS NINTNEGRTG KDVNSVLTSI HTFDPNLGCD
 301 AGTFQPCSDK ALSNLKVVVD SFRSIYGVNK GIPAGAAVAI GRYAEDVYYN
 351 GNPWYLATFA AAEQLYDAIY VWKKTGSITV TATSLAFFQE LVPGVTAGTY
 401 SSSSSTFTNI INAVSTYADG FLSEAAKYVP ADGSLAEQFD RNSGTPLSAL
 451 HLTWSYASFL TATARRAGIV PPSWANSAS TIPSTCSGAS VVGSYSRPTA
 501 TSFPPSQTPK PGVPSGTPYT PLPCATPTSV AVTFHELVST QFGQTVKVAG
 551 NAAALGNWST SAAVALDAVN YADNHPLWIG TVNLEAGDVV EYKYINVGQD
 601 GSVTWESDPN HTYTVPAVAC VTQVVKEDTW QS

FIG. 1A

Secuencia codificante ADN de TrGA (1899 pb) (SEQ ID NO:4)

```

1 ATGCACGTCC TGTCGACTGC GGTGCTGCTC GGCTCCGTTG CCGTTCAAAA
51 GGTCCCTGGGA AGACCAGGAT CAAGCGGTCT GTCCGACGTC ACCAAGAGGT
101 CTGTTGACGA CTTCATCAGC ACCGAGACGC CTATTGCACT GAACAATCTT
151 CTTTGCAATG TTGGTCCTGA TGGATGCCGT GCATTGCGCA CATCAGCTGG
201 TGCGGTGATT GCATCTCCCA GCACAATTGA CCCGGACTAC TATTACATGT
251 GGACGCGAGA TAGCGCTCTT GTCTTCAAGA ACCTCATCGA CCGCTTCACC
301 GAAACGTACG ATGCGGGCCT GCAGCGCCGC ATCGAGCAGT ACATTACTGC
351 CCAGGTCACT CTCCAGGGCC TCTCTAACCC CTCGGGCTCC CTCGCGGACG
401 GCTCTGGTCT CGGCGAGCCC AAGTTTGAGT TGACCCTGAA GCCTTTCACC
451 GGCAACTGGG GTCGACC GCGGGATGGC CCAGCTCTGC GAGCCATTGC
501 CTTGATTGGA TACTCAAAGT GGCTCATCAA CAACAACAT CAGTCGACTG
551 TGTCCAACGT CATCTGGCCT ATTGTGCGCA ACGACCTCAA CTATGTTGCC
601 CAGTACTGGA ACCAAACCGG CTTTGACCTC TGGGAAGAAG TCAATGGGAG
651 CTCATTCTTT ACTGTTGCCA ACCAGCACCG AGCACTTGTC GAGGGCGCCA
701 CTCTTGCTGC CACTCTTGGC CAGTCGGGAA GCGCTTATTC ATCTGTTGCT
751 CCCCAGGTTT TGTGCTTTCT CCAACGATTC TGGGTGTTCG CTGGTGGATA
801 CGTCGACTCC AACATCAACA CCAACGAGGG CAGGACTGGC AAGGATGTCA
851 ACTCCGTCCT GACTTCCATC CACACCTTCG ATCCCAACCT TGGCTGTGAC
901 GCAGGCACCT TCCAGCCATG CAGTGACAAA GCGCTCTCCA ACCTCAAGGT
951 TGTTGTCGAC TCCTTCCGCT CCATCTACGG CGTGAACAAG GGCATTCCCTG
1001 CCGGTGCTGC CGTCGCCATT GGCCGGTATG CAGAGGATGT GTACTACAAC
1051 GGCAACCCTT GGTATCTTGC TACATTTGCT GCTGCCGAGC AGCTGTACGA
1101 TGCCATCTAC GTCTGGAAGA AGACGGGCTC CATCACGGTG ACCGCCACCT
1151 CCCTGGCCTT CTTCCAGGAG CTTGTTCCCTG GCGTGACGGC CGGGACCTAC
1201 TCCAGCAGCT CTTGACCTT TACCAACATC ATCAACGCCG TCTCGACATA
1251 CGCCGATGGC TTCCTCAGCG AGGCTGCCAA GTACGTCCCC GCCGACGGTT
1301 CGCTGGCCGA GCAGTTTGAC CGCAACAGCG GCACTCCGCT GTCTGCGCTT
1351 CACCTGACGT GGTCGTACGC CTCGTTCTTG ACAGCCACGG CCCGTCGGGC
1401 TGGCATCGTG CCCCCCTCGT GGGCCAACAG CAGCGCTAGC ACGATCCCCT
1451 CGACGTGCTC CGGCGCGTCC GTGGTCGGAT CCTACTCGCG TCCCACCGCC
1501 ACGTCATTCC CTCCGTGCGA GACGCCAAG CCTGGCGTGC CTTCCGGTAC
1551 TCCCTACACG CCCCTGCCCT GCGCGACCCC AACCTCCGTG GCCGTACCT
1601 TCCACGAGCT CGTGTGACA CAGTTTGCC AGACGGTCAA GGTGGCGGGC
1651 AACGCCGCGG CCCTGGGCAA CTGGAGCACG AGCGCCGCCG TGGCTCTGGA
1701 CGCCGTCAAC TATGCCGATA ACCACCCCCT GTGGATTGGG ACGGTCAACC
1751 TCGAGGCTGG AGACGTCGTG GAGTACAAGT ACATCAATGT GGGCCAAGAT
1801 GGCTCCGTGA CCTGGGAGAG TGATCCCAAC CACTTACA CCGTTCCTGC
1851 GGTGGCTTGT GTGACGCAGG TTGTCAAGGA GGACACCTGG CAGTCGTAA

```

FIG. 1B

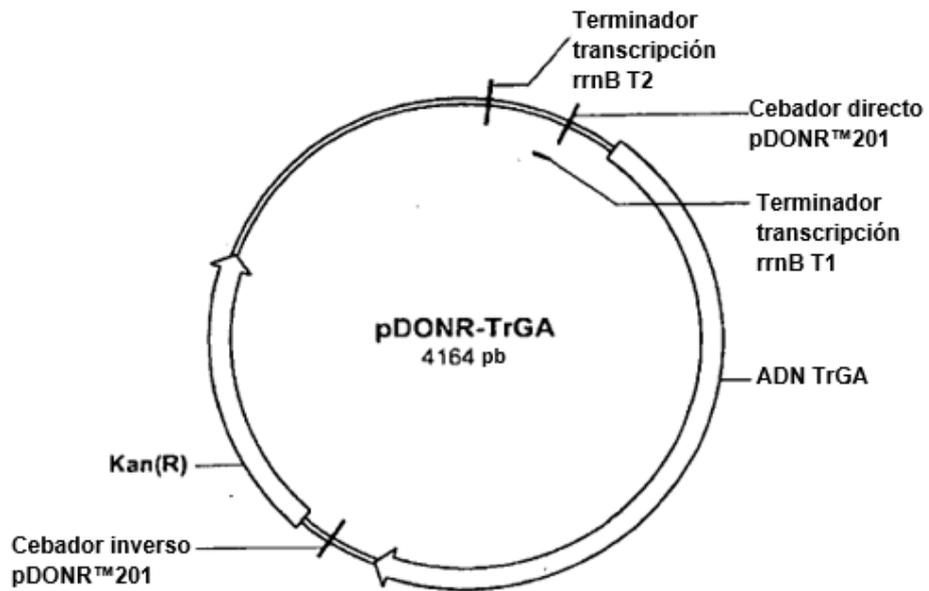


FIG. 2

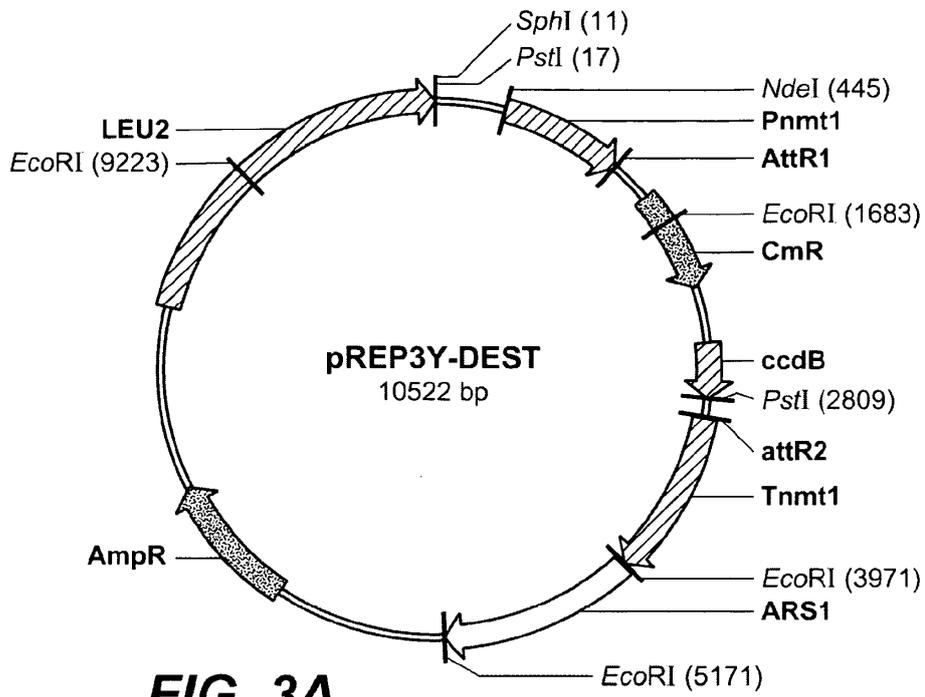


FIG. 3A

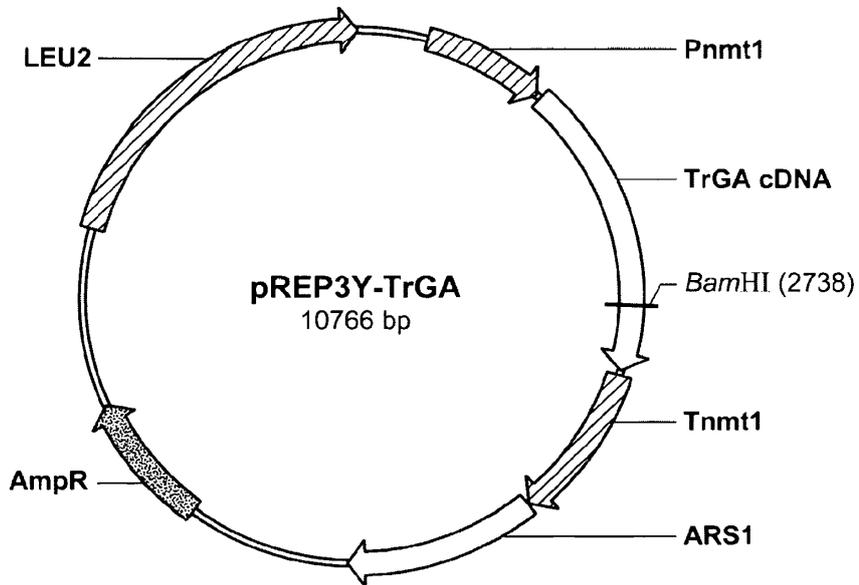


FIG. 3B

AaGA (294) LSDSEAVAVGRYPKDSYYNGNPWFLCTLAAAEQLYDALYQWDKQGSLEIT
 AnGA (295) LSDSEAVAVGRYPEDTYNGNPWFLCTLAAAEQLYDALYQWDKQGSLEVT
 AoGA (296) RGAGKAAAVGPPYAEDTYQGGNPWYLTTLVAAELLYDALYQWDKQGVNVT
 HgGA (286) IAQ GKAVAVGRYSEDVYYNGNPWYLANFAAAEQLYDAIYVWNKQGSITVT
 HvGA (298) IPAGSAVAIGRYPEDVYFNGNPWYLATFAAAEQLYDSVYVWKKTGSIIVT
 TrGA (299) IPAGA AVAIGRYAEDVYYNGNPWYLATFAAAEQLYDAIYVWKKTGSIIVT
 *

AaGA (344) DVSLDFFQALYSDAATGTYSSSSSTYSSIVDAVKTFADGFVSIIVETHAAS
 AnGA (345) DVSLDFFKALYSDAATGTYSSSSSTYSSIVDAVKTFADGFVSIIVETHAAS
 AoGA (346) ETSLPFFKDLSSNVTGTSYAKSSSAYESLTSVAVKTYADGFISVVQYETPD
 HgGA (336) SVSLPFFRDLVSSVSTGTYSKSSSTFTNIVNAVKAYADGFIEVAAKYTPS
 HvGA (348) STSSAFFQELVPGVAAGTYSSSQSTFTSIINAISTYADGFLSEAAKYVPA
 TrGA (349) ATSLAFFQELVPGVTAGTYSSSSSTFTNIINAVSTYADGFLSEAAKYVPA
 * ** * * * * * * * * * * * * * * * *

AaGA (394) NGSLSEQYDKSDGDELSARDLTWSYAALLTANNRRNSVMPSPWGETSAS-
 AnGA (395) NGSMSEQYDKSDGEQLSARDLTWSYAALLTANNRRNSVVPASWGETSAS-
 AoGA (396) GGALAEQYSRDQGTPVSASDLTWSYAFLSAVGRNRNGTVPASWGSSTAN-
 HgGA (386) NGALAEQYDRNTGKPD SAADLTWSYSAFLSAIDRRAGLVPSPWRASVAKS
 HvGA (398) DGSLAEQFDRNTGTPLSAVHLTWSYASFLTAARRAGVPPSWASSGAN-
 TrGA (399) DGSLAEQFDRNSGTPLSALHLTWSYASFLTATARRAGIVPPSWANSSAS-
 * ** * * * * * * * * * * * * * * * *

AaGA (443) SVPGTC
 AnGA (444) SVPGTC
 AoGA (445) AVPSQC
 HgGA (436) QLPSTC
 HvGA (447) TVPSSC
 TrGA (448) TIPSTC
 * *

FIG. 4B

FIG. 4C-1
 FIG. 4C-2

FIG. 4C

FIG. 4C-1

HOJA SUSTITUCIÓN (NORMA 26)

Trichoderma reesei	Posición	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520
	Aminoácido	C	A	T	P	T	S	V	A	V	T	F	H	E	L	V	S	T	Q	F	G	Q	T	V	K	V	A	G	N	A	A
fumicola grisea	Posición	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522
	Aminoácido	C	A	D	A	S	E	V	Y	V	T	F	N	E	R	V	S	T	A	W	G	E	T	I	K	V	V	G	N	V	P
hielavii terrestris	Posición	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518
	Aminoácido	C	S	T	P	T	A	V	A	V	T	F	N	E	R	V	T	T	Q	W	G	Q	T	I	K	V	V	G	D	A	A
heromyces anuginosus	Posición	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504
	Aminoácido	C	T	P	P	S	E	V	T	L	T	F	N	A	L	V	D	T	A	F	G	Q	N	I	Y	L	V	G	S	I	P
alaromyces emersonii	Posición	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511
	Aminoácido	C	T	T	P	T	S	V	A	V	T	F	D	E	I	V	S	T	S	Y	G	E	T	I	Y	L	A	G	S	I	P
Aspergillus niger	Posición	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538
	Aminoácido	C	T	T	P	T	A	V	A	V	T	F	D	L	T	A	T	T	Y	G	E	N	I	Y	L	V	G	S	I	S	
Aspergillus awamori	Posición	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537
	Aminoácido	C	T	T	P	T	A	V	A	V	T	F	D	L	T	A	T	T	Y	G	E	N	I	Y	L	V	G	S	I	S	

HOJA SUSTITUCIÓN (NORMA 26)

Trichoderma reesei	Posición	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550
	Aminoácido	A	L	G	N	W	S	T	S	A	A	V	A	L	D	A	V	N	Y	A	D	N	H	P	L	W	I	G	T	V	N
Humicola grisea	Posición	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552
	Aminoácido	A	L	G	N	W	D	T	S	K	A	V	T	L	S	A	S	G	Y	K	S	N	D	P	L	W	S	I	T	V	P
Hielavia terrestris	Posición	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541	542	543	544	545	546	547	548
	Aminoácido	A	L	G	G	W	D	T	S	K	A	V	P	L	S	A	A	G	Y	T	A	S	D	P	L	W	S	G	T	V	D
hemomyces truginosus	Posición	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534
	Aminoácido	E	L	G	S	W	D	P	A	N	A	L	L	M	S	A	K	S	W	T	S	G	N	P	V	W	T	L	S	I	S
alaromyces emersonii	Posición	512	513	514	515	516	517	518	519	520	521	522	523	524	525	526	527	528	529	530	531	532	533	534	535	536	537	538	539	540	541
	Aminoácido	E	L	G	N	W	S	T	A	S	A	I	P	L	R	A	G	A	Y	T	N	S	N	P	L	W	Y	V	T	V	N
Aspergillus niger	Posición	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568
	Aminoácido	Q	L	G	D	W	E	T	S	D	G	I	A	L	S	A	D	K	Y	T	S	S	D	P	L	W	Y	V	T	V	T
Aspergillus awamori	Posición	536	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567
	Aminoácido	Q	L	G	D	W	D	T	S	D	G	I	A	L	S	A	D	K	Y	T	S	S	N	P	L	W	Y	V	T	V	T

FIG. 4C-2

FIG. 4D-1

FIG. 4D

FIG. 4D-2

FIG. 4D-1

HOJA SUSTITUCIÓN (NORMA 26)

Trichoderma reesei	Posición	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	
	Aminoácido	L	E	A	.	G	D	V	V	E	Y	K	Y	I	N	V	G	Q	D	G	S	V	T	.	W	E	S	D	P	N
Humicola grisea	Posición	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	577	578	579	580	581
	Aminoácido	I	K	A	T	G	S	A	V	Q	Y	K	Y	I	K	V	G	T	N	G	K	I	T	.	W	E	S	D	P	N
Hielavia terrestris	Posición	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	570	571	572	573	574	575	576	
	Aminoácido	L	P	A	.	G	L	A	V	Q	Y	K	Y	I	N	V	A	A	D	G	G	V	T	.	W	E	A	D	P	N
hemomyces truginosus	Posición	535	536	537	538	539	540	541	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563
	Aminoácido	L	P	A	.	G	T	S	F	E	Y	K	F	I	R	K	D	D	G	S	S	D	V	V	W	E	S	D	P	N
alaromyces emersonii	Posición	542	543	544	545	546	547	548	549	550	551	552	553	554	555	556	557	558	559	560	561	562	563	564	565	566	567	568	569	
	Aminoácido	L	P	P	.	G	T	S	F	E	Y	K	F	F	K	N	Q	T	D	G	T	I	V	.	W	E	D	D	P	N
Aspergillus niger	Posición	566	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	596	
	Aminoácido	L	P	A	.	G	E	S	F	E	Y	K	F	I	R	I	E	S	D	D	S	V	E	.	W	E	S	D	P	N
Aspergillus awamori	Posición	568	569	570	571	572	573	574	575	576	577	578	579	580	581	582	583	584	585	586	587	588	589	590	591	592	593	594	595	
	Aminoácido	L	P	A	.	G	E	S	F	E	Y	K	F	I	R	I	E	S	D	D	S	V	E	.	W	E	S	D	P	N

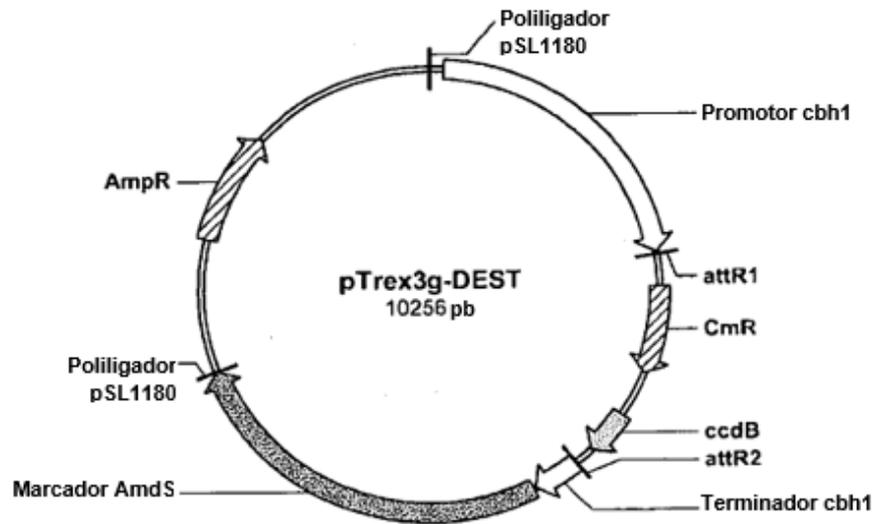


FIG. 5A

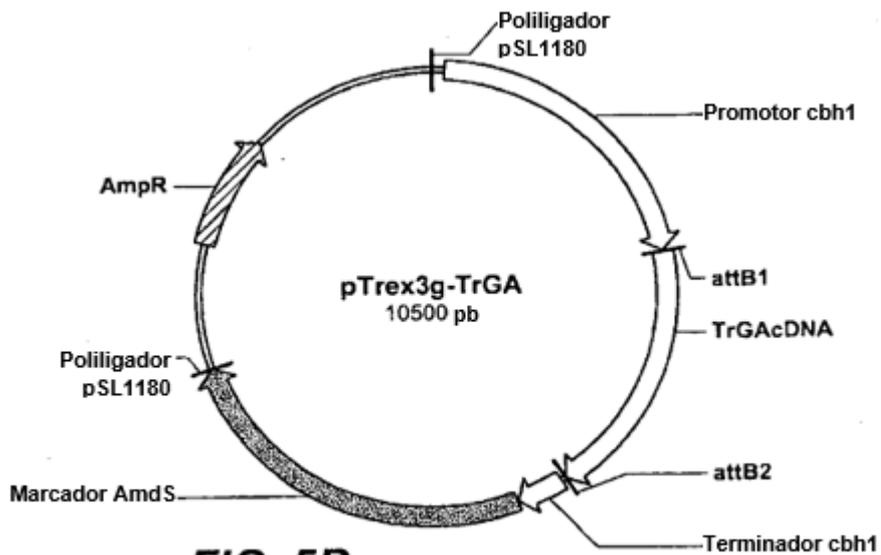


FIG. 5B

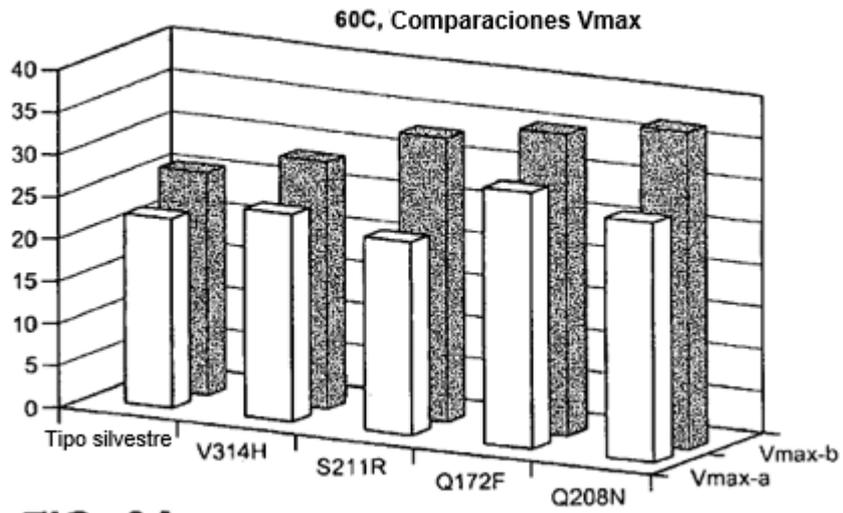


FIG. 6A

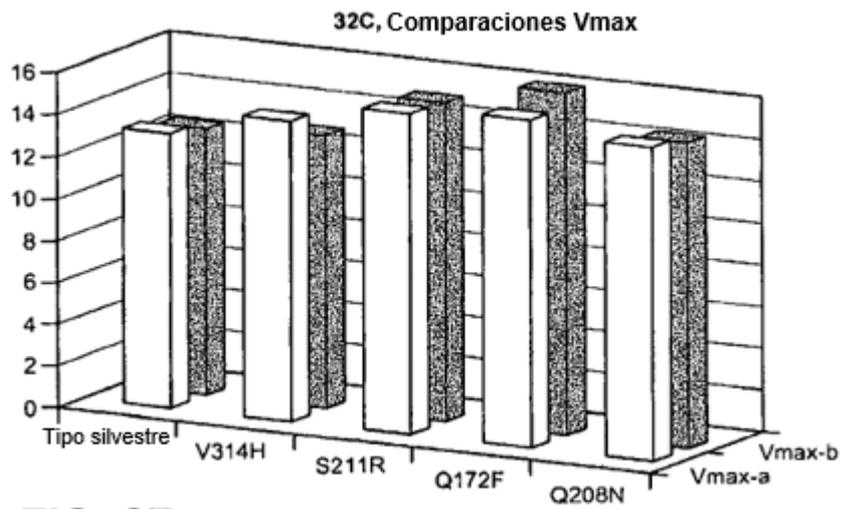
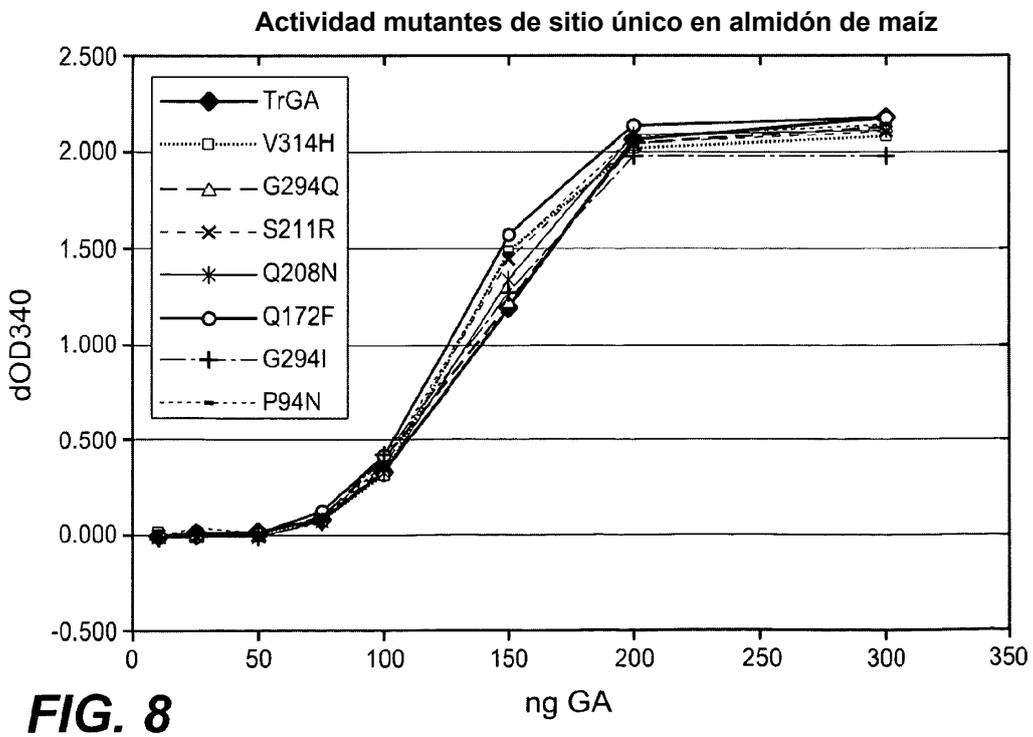
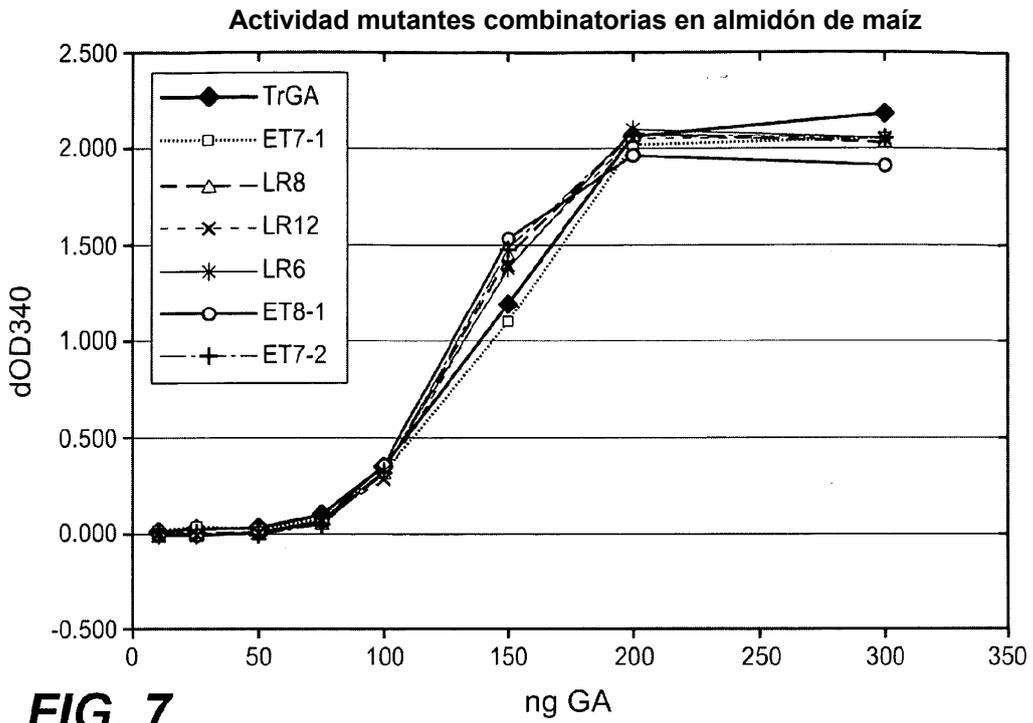
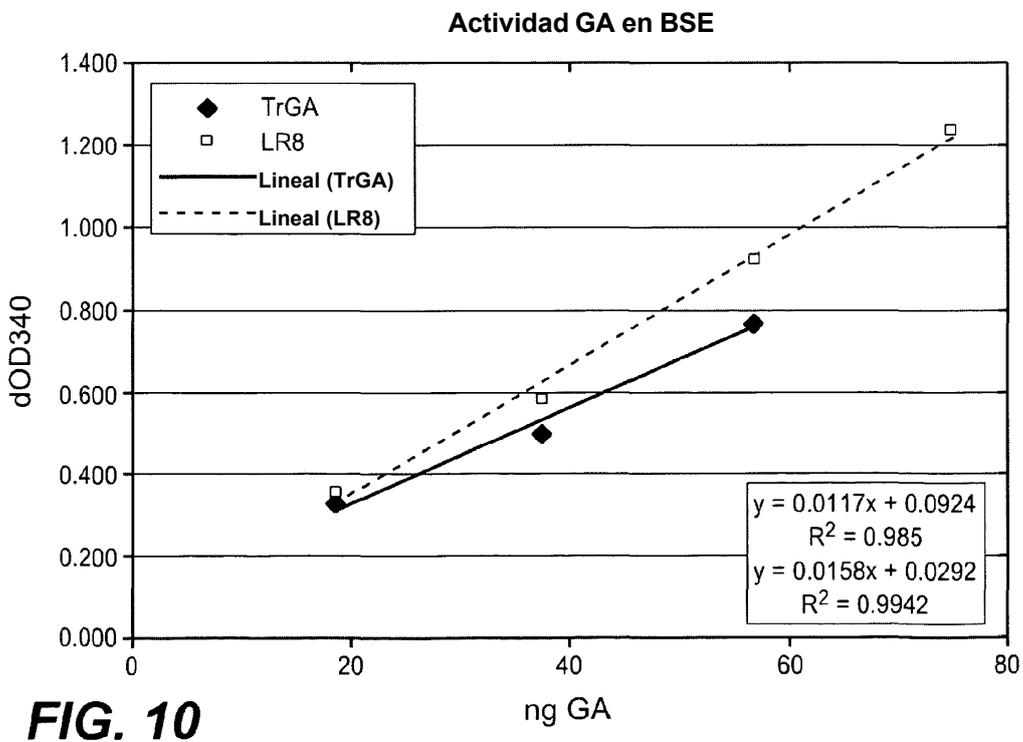
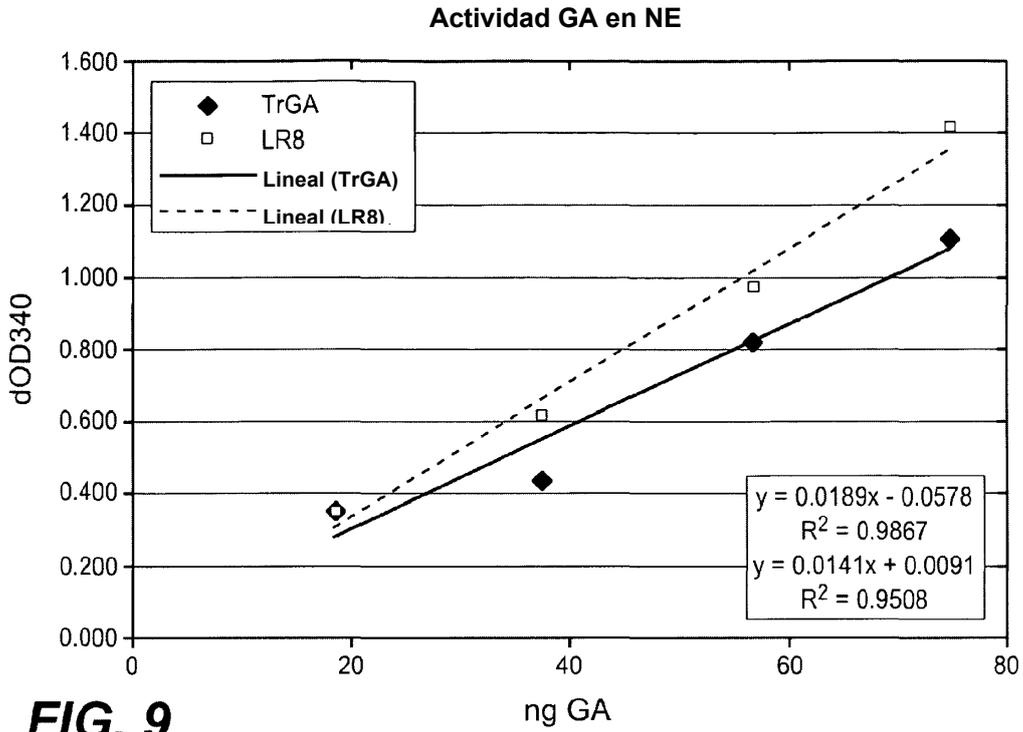


FIG. 6B





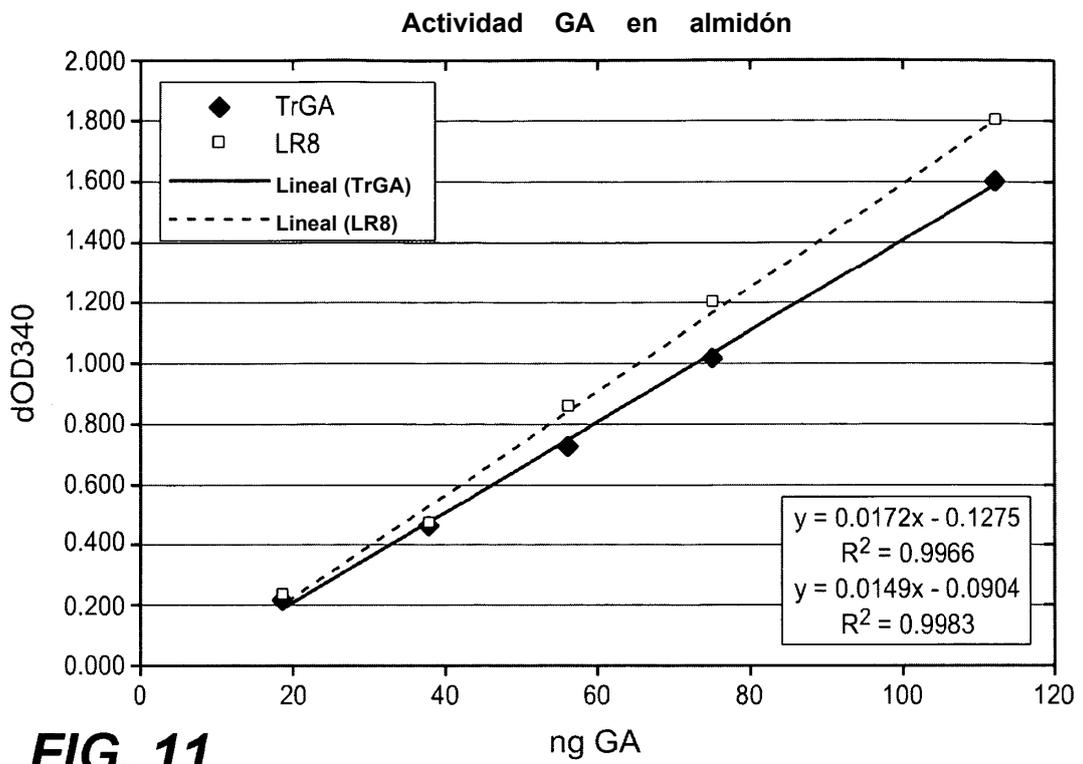


FIG. 11

