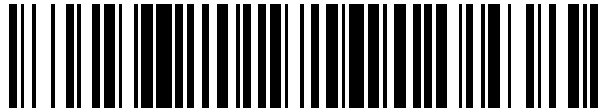


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 551 141**

51 Int. Cl.:

C07K 14/37 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.09.2010 E 10765711 (6)**

97 Fecha y número de publicación de la concesión europea: **29.07.2015 EP 2483296**

54 Título: **Polipéptidos derivados de Thermoascus crustaceus con actividad de aumento celulolítico y polinucleótidos que codifican los mismos**

30 Prioridad:

30.09.2009 US 247250 P

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

16.11.2015

73 Titular/es:

**NOVOZYMES INC. (50.0%)
1445 Drew Avenue
Davis, CA 95618, US y
NOVOZYMES A/S (50.0%)**

72 Inventor/es:

**TANG, LAN;
LIU, YE;
DUAN, JUNXIN;
WU, WENPING y
KRAMER, RANDALL**

74 Agente/Representante:

TOMAS GIL, Tesifonte Enrique

ES 2 551 141 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín europeo de patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre concesión de Patentes Europeas).

DESCRIPCIÓN

- Polipéptidos derivados de *Thermoascus crustaceus* con actividad de aumento celulolítico y polinucleótidos que codifican los mismos
- 5 Balance en cuanto a derechos sobre invenciones hechas a continuación
- Investigación y desarrollo financiado federalmente
- 10 [0001] Esta invención fue llevada a cabo parcialmente con el apoyo del gobierno bajo el acuerdo cooperativo DE-FC36-08GO18080 otorgado por el Departamento de Energía. El gobierno tiene ciertos derechos sobre esta invención.
- Referencia a un listado de secuencias
- 15 [0002] Esta aplicación contiene un listado de secuencias en forma legible informáticamente.
- Referencia a un Depósito de Material Biológico
- [0003] Esta aplicación contiene una referencia a un depósito de material biológico.
- 20 Antecedentes de la invención
- Campo de la invención
- 25 [0004] La presente invención se refiere a polipéptidos aislados con actividad de aumento celulolítico y polinucleótidos aislados que codifican los polipéptidos. La invención también se refiere a métodos de producción y uso de los polipéptidos.
- 30 Descripción de las técnicas relacionadas
- [0005] La celulosa es un polímero de la glucosa simple de azúcar enlazada por enlaces beta-1,4. Muchos microorganismos producen enzimas que hidrolizan glucanos beta-enlazados. Estas enzimas incluyen endoglucanasas, celobiohidrolasas y beta-glucosidasas. Las endoglucanasas digieren el polímero de celulosa en ubicaciones aleatorias, abriéndolo al ataque por parte de celobiohidrolasas. Las celobiohidrolasas liberan consecutivamente moléculas de celobiosa de los extremos del polímero de celulosa. La celobiosa es un dímero de glucosa beta-1,4-enlazado hidrosoluble. Las beta-glucosidasas hidrolizan celobiosa en glucosa.
- 35 [0006] La conversión de materias primas lignocelulósicas en etanol tiene las ventajas de la disponibilidad preparada de grandes cantidades de materia prima, la conveniencia de evitar la quema o desecho de los materiales, y la limpieza del etanol combustible. Madera, residuos agrícolas, plantas de cultivo herbáceas, y residuos sólidos municipales han sido considerados como materias primas para la producción de etanol. Estos materiales principalmente consisten en celulosa, hemicelulosa y lignina. Una vez la celulosa se convierte en glucosa, la glucosa es fácilmente fermentada por levadura en etanol.
- 40 [0007] Sería ventajoso en la técnica mejorar la capacidad de degradar enzimáticamente materias primas lignocelulósicas.
- [0008] El documento WO 2005/074647 divulga polipéptidos aislados con actividad de aumento celulolítico y polinucleótidos de los mismos de *Thielavia terrestris*. El documento WO 2005/074656 divulga un polipéptido aislado con actividad de aumento celulolítico y su polinucleótido de *Thermoascus aurantiacus*. El documento WO 2007/089290 divulga un polipéptido aislado con actividad de aumento celulolítico y su polinucleótido de *Trichoderma reesei*.
- 50 [0009] La presente invención proporciona polipéptidos con actividad de aumento celulolítico y polinucleótidos que codifican los polipéptidos.
- 55 Resumen de la invención
- [0010] La presente invención se refiere a polipéptidos aislados con actividad de aumento celulolítico seleccionados del grupo consistente en:
- 60 (a) un polipéptido que incluye una secuencia de aminoácidos con al menos el 90% de identidad con el polipéptido

ES 2 551 141 T3

maduro mostrado como los aminoácidos 23-251 de la SEC ID nº 2;

(b) un polipéptido codificado por un polinucleótido que comprende una secuencia de nucleótidos con al menos el 90% de identidad a la secuencia codificante del polipéptido maduro mostrado como los nucleótidos 67-868 de la SEC ID nº 1.

5 [0011] La presente invención también se refiere a polinucleótidos aislados que comprenden una secuencia de nucleótidos que codifica el polipéptido de la invención.

[0012] La presente invención también se refiere a métodos de producción del polipéptido de la invención, que comprenden: (a) cultivo de una célula de *Thermoascus crustaceus*, que en su forma de tipo natural produce el polipéptido, bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.

10 [0013] La invención también se refiere a métodos de producción del polipéptido con actividad de aumento celulolítico, que comprenden: (a) cultivo de una célula huésped recombinante que comprende el polinucleótido de la invención bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.

15 [0014] Además, la invención también se refiere a composiciones detergentes que comprenden el polipéptido de la invención y un tensioactivo.

[0015] La presente invención también se refiere a métodos para la degradación o conversión de un material celulósico, que comprenden: tratar el material celulósico con una composición enzimática en presencia de un polipéptido de la invención con actividad de aumento celulolítico. En un aspecto preferido, el método comprende además la recuperación del material celulósico degradado o convertido.

20 [0016] La presente invención también se refiere a métodos de producción de un producto de fermentación, que comprenden: (a) sacarificación de un material celulósico con una composición enzimática en presencia de un polipéptido de la invención con actividad de aumento celulolítico; (b) fermentación del material celulósico sacarificado con uno o más microorganismos de fermentación (diferentes) para producir el producto de fermentación; y (c) recuperación del producto de fermentación a partir de la fermentación.

25 Breve descripción de las figuras

[0017]

30 La Figura 1 muestra la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un gen que codifica un polipéptido GH61A de *Thermoascus crustaceus* CBS 181.67 con actividad de aumento celulolítico (SEC ID nº 1 y 2, respectivamente).

La Figura 2 muestra la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un gen que codifica un polipéptido GH61B de *Thermoascus crustaceus* CBS 181.67 con actividad de aumento celulolítico (SEC ID nº 3 y 4, respectivamente).

35 La Figura 3 muestra la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un gen que codifica un polipéptido GH61C de *Thermoascus crustaceus* CBS 181.67 con actividad de aumento celulolítico (SEC ID nº 5 y 6, respectivamente).

La Figura 4 muestra un mapa de restricción de pGH61a51486.

La Figura 5 muestra un mapa de restricción de pGH61 D14YF.

40 La Figura 6 muestra un mapa de restricción de pGH61 D14YH.

La Figura 7 muestra el efecto del polipéptido GH61A de *Thermoascus crustaceus* CBS 181.67 con actividad de aumento celulolítico en la hidrólisis PCS.

La Figura 8 muestra el efecto del polipéptido GH61B de *Thermoascus crustaceus* CBS 181.67 con actividad de aumento celulolítico en la hidrólisis PCS.

45 Definiciones

[0018] **Polipéptido con actividad de aumento celulolítico:** el término "polipéptido con actividad de aumento celulolítico" se refiere a un polipéptido GH61 que mejora la hidrólisis de un material celulósico mediante enzimas con actividad celulolítica. Para fines de la presente invención, la actividad de aumento celulolítico es determinada midiendo el aumento en azúcares reductores o el aumento del total de celobiosa y glucosa a partir de la hidrólisis de un material celulósico mediante enzima celulolítica bajo las siguientes condiciones: 1-50 mg de proteína total/g de celulosa en PCS, donde la proteína total está compuesta por el 50-99,5% p/p de proteína enzimática celulolítica y el 0,5-50% p/p de proteína de un polipéptido GH61 con actividad de aumento celulolítico durante 1-7 días a 50°C en comparación con una hidrólisis de control con igual carga de proteína total sin actividad de aumento celulolítico (1-50 mg de proteína celulolítica/g de celulosa en PCS).

ES 2 551 141 T3

En un aspecto preferido, como la fuente de la actividad celulolítica se usa una mezcla de CELLULCLAST® 1.5L (Novozymes A/S, Bagsvaerd, Dinamarca) en presencia del 2-3% del peso de proteína total de beta-glucosidasa de *Aspergillus oryzae* (producida recombinantemente en *Aspergillus oryzae* según el documento WO 02/095014) o el 2-3% del peso de proteína total de beta-glucosidasa de *Aspergillus fumigatus* (producida recombinantemente en *Aspergillus oryzae* como se describe en el documento WO 2002/095014) de carga de proteína de celulasa.

[0019] Los polipéptidos con actividad de aumento celulolítico tienen al menos el 20%, preferiblemente al menos el 40%, de forma más preferible al menos el 50%, de forma más preferible al menos el 60%, de forma más preferible al menos el 70%, de forma más preferible al menos el 80%, incluso de forma más preferible al menos el 90%, de la forma más preferible al menos el 95%, e incluso de la forma más preferible al menos el 100% de la actividad de aumento celulolítico del polipéptido maduro de la SEC ID nº 2.

[0020] Los polipéptidos GH61 con actividad de aumento celulolítico mejoran la hidrólisis de un material celulósico catalizado por enzima con actividad celulolítica reduciendo la cantidad de enzima celulolítica necesaria para alcanzar el mismo grado de hidrólisis preferiblemente al menos 1,01 veces, de forma más preferible al menos 1,05 veces, de forma más preferible al menos 1,10 veces, de forma más preferible al menos 1,25 veces, de forma más preferible al menos 1,5 veces, de forma más preferible al menos 2 veces, de forma más preferible al menos 3 veces, de forma más preferible al menos 4 veces, de forma más preferible al menos 5 veces, incluso de forma más preferible al menos 10 veces, y de la forma más preferible al menos 20 veces.

[0021] **Glucósido hidrolasa de familia 61:** el término "glucósido hidrolasa de familia 61" o "familia GH61" se define por la presente como un polipéptido incluido en la familia 61 de glucósido hidrolasas según Henrissat B., 1991, A classification of glycosyl hydrolases based on aminoacid sequence similarities, Biochem. J. 280: 309-316, y Henrissat B., y Bairoch A., 1996, Updating the sequence-based classification of glycosyl hydrolases, Biochem. J. 316: 695-696.

[0022] **Enzima celulolítica o celulasa:** el término "enzima celulolítica" o "celulasa" se refiere a una o varias enzimas (diferentes) que hidrolizan un material celulósico. Tales enzimas incluyen endoglucanasa(s), celobiohidrolasa(s), beta-glucosidasa(s), o combinaciones de las mismas. Los dos métodos básicos para medir la actividad celulolítica incluyen: (1) medición de la actividad celulolítica total, y (2) medición de las actividades celulolíticas individuales (endoglucanasas, celobiohidrolasas, y beta-glucosidasas) como se comprueba en Zhang et al., Outlook for cellulase improvement: Screening and selection strategies, 2006, Biotechnology Advances 24: 452-481. La actividad celulolítica total se mide normalmente usando sustratos insolubles, incluyendo papel de filtro Whatman Nº1, celulosa microcristalina, celulosa bacteriana, celulosa algal, algodón, lignocelulosa pretratada, etc. El ensayo de actividad celulolítica total más común es el ensayo de papel de filtro usando papel de filtro Whatman Nº1 como sustrato. El ensayo fue establecido por la Unión Internacional de Química Pura y Aplicada (IUPAC) (Ghose, 1987, Measurement of cellulase activities, Pure Appl. Chem. 59: 257-68).

[0023] Para fines de la presente invención, la actividad enzimática celulolítica se determina midiendo el aumento en la hidrólisis de un material celulósico mediante enzima(s) celulolíticas bajo las siguientes condiciones: 1-20 mg de proteína de enzima celulolítica/g de celulosa en PCS durante 3-7 días a 50°C en comparación con una hidrólisis de control sin adición de proteína enzimática celulolítica. Las condiciones típicas son reacciones de 1 ml, PCS lavado o sin lavar, 5% sólidos insolubles, 50 mM de acetato sódico a pH 5, 1 mM MnSO₄, 50°C, 72 horas, análisis de azúcar mediante columna AMINEX® HPX-87H (Bio-Rad Laboratories, Inc., Hércules, CA, EE.UU).

[0024] **Endoglucanasa:** el término "endoglucanasa" se define por la presente como una endo-1,4-(1,3;1,4)-beta-D-glucano-4-glucanohidrolasa (E.C. 3.2.1.4), que cataliza la endohidrólisis de los enlaces 1,4-beta-D-glucosídicos en la celulosa, derivados de la celulosa (tales como carboximetilcelulosa e hidroxietilcelulosa), liquenina, enlaces beta-1,4 en beta-1,3 glucanos mezclados tales como beta-D-glucanos o xiloglucanos de cereales, y otro material vegetal que contiene componentes celulósicos. La actividad de endoglucanasa se puede determinar en base a la reducción en la viscosidad del sustrato o al aumento en la reducción de los extremos determinado mediante un ensayo de azúcar reductor (Zhang et al., 2006, Biotechnology Advances 24: 452-481). Para fines de la presente invención, la actividad de endoglucanasa es determinada usando hidrólisis de carboximetilcelulosa (CMC) según el procedimiento de Ghose, 1987, Pure and Appl. Chem. 59: 257-268.

[0025] **Celobiohidrolasa:** el término "celobiohidrolasa" se define por la presente como una 1,4-beta-D-glucano celobiohidrolasa (E.C. 3.2.1.91), que cataliza la hidrólisis de enlaces 1,4-beta-D-glucosídicos en celulosa, celooligosacáridos, o cualquier polímero que contenga glucosa beta-1,4-enlazada, liberando celobiosa a partir de los extremos reducidos o no-reducidos de la cadena (Teeri, 1997, Crystalline cellulose degradation: New insight into the function of cellobiohydrolases, Trends in Biotechnology 15: 160-167; Teeri et al., 1998, Trichoderma reesei cellobiohydrolases: why so efficient on crystalline cellulose?, Biochem. Soc. Trans. 26: 173-178). Para fines de la presente invención, la actividad de celobiohidrolasa se determina sobre un derivado de disacárido fluorescente 4-

metilumbeliferil-p-β-lactosida según los procedimientos descritos por van Tilbeurgh et al., 1982, FEBS Letters 149: 152-156 y van Tilbeurgh y Claeysens, 1985, FEBS Letters 187: 283-288.

5 [0026] **Beta-glucosidasa:** el término "beta-glucosidasa" se define por la presente como una beta-D-glucósido glucohidrolasa (E.C. 3.2.1.21), que cataliza la hidrólisis de residuos de beta-D-glucosa extremos no reductores con la liberación de beta-D-glucosa. Para fines de la presente invención, la actividad de beta-glucosidasa se determina según el procedimiento básico descrito por Venturi et al., 2002, Extracellular beta-D-glucosidase from *Chaetomium thermophilum* var. *coprophilum*: production, purification and some biochemical properties, *J. Basic Microbiol.* 42: 55-66, excepto en que se emplearon condiciones diferentes como se describe en este caso. Una unidad de actividad de beta-glucosidasa se define como 1,0 μmol de *p*-nitrofenol producido por minuto a 40°C, pH 5 a partir de 1 mM de *p*-nitrofenil-beta-D-glucopiranosido como sustrato en 100 mM de citrato sódico a pH 5 que contiene 0,01% de TWEEN® 20.

15 [0027] **Actividad de degradación de xilano:** los términos "actividad de degradación de xilano" o "actividad xilanolítica" se definen por la presente como una actividad biológica que hidroliza material que contiene xilano. Los dos métodos básicos para medir la actividad xilanolítica incluyen: (1) medición de la actividad xilanolítica total, y (2) medición de las actividades xilanolíticas individuales (endoxilanasas, beta-xilosidasas, arabinofuranosidasas, alfa-glucuronidasas, acetilxilan estererasas, estererasas de ácido ferúlico, y alfa-glucuronil estererasas). El progreso reciente en ensayos de enzimas xilanolíticas está resumido en diferentes publicaciones incluyendo Biely y Puchard, Recent progress in the assays of xylanolytic enzymes, 2006, *Journal of the Science of Food and Agriculture* 86(11): 1636-1647; Spanikova y Biely, 2006, Glucuronoyl esterase-Novel carbohydrate esterase produced by *Schizophyllum commune*, *FEBS Letters* 580(19): 4597-4601; y Herrmann, Vrsanska, Jurickova, Hirsch, Biely y Kubicek, 1997, The beta-D-xylosidase of *Trichoderma reesei* is a multifunctional beta-D-xylan xylohydrolase, *Biochemical Journal* 321: 375-381.

25 [0028] La actividad total de degradación de xilano se puede medir mediante la determinación de los azúcares reductores formados a partir de tipos varios de xilano, incluyendo xilanos de avena, espelta, madera de haya y madera de alerce, o mediante la determinación fotométrica de fragmentos de xilano teñidos liberados de varios xilanos teñidos de manera covalente. El ensayo de actividad xilanolítica total más común se basa en la producción de azúcares reductores de 4-O-metil glucuronoxilano polimérico como se describe en Bailey, Biely, Poutanen, 1992, Interlaboratory testing of methods for assay of xylanase activity, *Journal of Biotechnology* 23(3): 257-270.

30 [0029] Para fines de la presente invención, la actividad de degradación de xilano se determina mediante la medición del aumento en la hidrólisis de xilano de madera de abedul (*Sigma Chemical Co., Inc., St. Louis, MO, EE.UU.*) mediante enzima(s) que degradan xilano bajo las siguientes condiciones típicas: 1 ml reacciones, 5 mg/ml sustrato (sólidos totales), 5 mg de proteína xilanolítica/g de sustrato, 50 mM acetato sódico a pH 5, a 50°C, 24 horas, análisis de azúcar usando un ensayo de hidrazida de ácido *p*-hidroxibenzoico (PHBAH por sus siglas en inglés) como describe Lever, 1972, A new reaction for colorimetric determination of carbohydrates, *Anal. Biochem* 47: 273- 279.

40 [0030] **Xilanasa:** el término "xilanasa" se define por la presente como una 1,4-beta-D-xilan-xilohidrolasa (E.C. 3.2.1.8) que cataliza la endo-hidrólisis de enlaces 1,4-beta-D-xilosídicos en xilanos. Para fines de la presente invención, la actividad de xilanasa es determinada usando el xilano de madera de abedul como sustrato. Una unidad de actividad de xilanasa se define como 1,0 μmol de azúcar reductor (medido en los equivalentes de glucosa como describe Lever, 1972, A new reaction for colorimetric determination of carbohydrates, *Anal. Biochem* 47: 273-279) producida por minuto durante el periodo inicial de hidrólisis a 50°C, pH 5 de 2 g de xilano de madera de abedul por litro como sustrato en 50 mM de acetato sódico a pH 5 que contiene 0,01% TWEEN® 20.

45 [0031] **Beta-xilosidasa:** el término "beta-xilosidasa" se define por la presente como una beta-D-xilósido xilohidrolasa (E.C. 3.2.1.37) que cataliza la exohidrólisis de beta-(1→4)-xilooligosacáridos cortos, para eliminar los sucesivos residuos de D-xilosa de los terminales no-reductores. Para fines de la presente invención, una unidad de actividad de beta-xilosidasa se define como 1,0 μmol de *p*-nitrofenol producido por minuto a 40°C, pH 5 a partir de 1 mM *p*-nitrofenil-beta-D-xilósido como sustrato en 100 mM citrato sódico a pH 5 que contiene 0,01% TWEEN® 20.

50 [0031] **Acetilxilan esterasa:** el término "acetilxilan esterasa" se define por la presente como una carboxilesterasa (EC 3.1.1.72) que cataliza la hidrólisis de grupos acetilo a partir de xilano polimérico, xilosa acetilada, glucosa acetilada, acetato de alfa-naptilo y acetato de *p*-nitrofenilo. Para fines de la presente invención, la actividad de la acetilxilan esterasa es determinada usando el 0,5 mM *p*-nitrofenilacetato como sustrato en 50 mM acetato sódico a pH 5,0 que contiene 0,01% TWEEN™ 20. Una unidad de actividad de acetilxilan esterasa fue definida como la cantidad de enzima capaz de liberar 1 μmol de anión de *p*-nitrofenolato por minuto a pH 5, 25°C.

60 [0032] **Feruloil esterasa:** el término "feruloil esterasa" se define por la presente como una 4-hidroxi-3-metoxicinnamoil-azúcar hidrolasa (EC 3.1.1.73) que cataliza la hidrólisis del grupo 4-hidroxi-3-metoxicinnamoilo (feruloilo) a partir de un azúcar esterificado, que es normalmente arabinosa en sustratos "naturales", para producir ferulato (4-hidroxi-3-metoxicinnamato). La feruloil esterasa es también conocida como esterasa de ácido ferúlico, hidroxicinnamoil esterasa,

ES 2 551 141 T3

F AE-III, cinnamoil éster hidrolasa, FAEA, cinnAE, FAE-I, o FAE-II. Para fines de la presente invención, la actividad de la feruloil esterasa se determina usando 0,5 mM *p*-nitrofenilferulato como sustrato en 50 mM acetato sódico a pH 5,0. Una unidad de actividad de feruloil esterasa iguala la cantidad de enzima capaz de liberar 1 μmol de anión de *p*-nitrofenolato por minuto a pH 5, 25°C.

5

[0033] **Alfa-glucuronidasa:** el término "alfa-glucuronidasa" se define por la presente como una alfa-D-glucosiduronato glucuronohidrolasa (EC 3.2.1.139) que cataliza la hidrólisis de una alfa-D-glucuronosida en D-glucuronato y un alcohol. Para fines de la presente invención, la actividad de alfa-glucuronidasa se determina según de Vries, 1998, J. Bacteriol. 180: 243-249. Una unidad de actividad de alfa-glucuronidasa iguala la cantidad de enzima capaz de liberar 1 μmol de ácido glucurónico o ácido 4-O-metilglucurónico por minuto a pH 5, 40°C.

10

[0034] **Alfa-L-arabinofuranosidasa:** el término "alfa-L-arabinofuranosidasa" se define por la presente como una alfa-L-arabinofuranosida arabinofuranohidrolasa (EC 3.2.1.55) que cataliza la hidrólisis de residuos de alfa-L-arabinofuranosida terminales no reductores en alfa-L-arabinósidos. La actividad enzimática actúa sobre alfa-L-arabinofuranósidos, alfa-L-arabinanos que contienen enlaces 1,3 y/o 1,5, arabinoxilanos y arabinogalactanos. La alfa-L-arabinofuranosidasa también es conocida como arabinosidasa, alfa-arabinosidasa, alfa-L-arabinosidasa, alfa-arabinofuranosidasa, alfa-L-arabinofuranosidasa polisacárida, alfa-L-arabinofuranosida hidrolasa, L-arabinosidasa o alfa-L-arabinanasa. Para fines de la presente invención, la actividad de alfa-L-arabinofuranosidasa se determina usando 5 mg arabinoxilano de trigo de viscosidad media (Megazyme Internacional Ireland, Ltd., Bray, Co. Wicklow, Irlanda) por ml de 100 mM acetato sódico a pH 5 en un volumen total de 200 μl durante 30 minutos a 40°C seguido del análisis de arabinosa por cromatografía en columna AMINEX® HPX-87H.

15

20

[0035] **Material celulósico:** el material celulósico puede ser cualquier material que contiene celulosa. El polisacárido predominante en la pared celular primaria de biomasa es celulosa, el segundo más abundante es hemicelulosa, y el tercero es pectina. La pared celular secundaria, producida después de que la célula haya dejado de crecer, también contiene polisacáridos y está reforzada por lignina polimérica reticulada covalentemente en hemicelulosa. La celulosa es un homopolímero de anhidrocelobiosa y por consiguiente un beta-(1-4)-D-glucano lineal, mientras que las hemicelulosas incluyen una variedad de compuestos, tales como xilanos, xiloglucanos, arabinoxilanos y mananos en estructuras ramificadas complejas con un espectro de sustituyentes. Aunque generalmente es polimorfa, la celulosa se encuentra principalmente como una matriz cristalina insoluble de cadenas de glucano paralelas en el tejido vegetal. Las hemicelulosas normalmente enlazan con hidrógeno a celulosa, al igual que a otras hemicelulosas, que ayudan a estabilizar la matriz de pared celular.

25

30

[0036] La celulosa generalmente se encuentra, por ejemplo, en los tallos, hojas, vainas, cáscaras y mazorcas de plantas u hojas, ramas y madera de árboles. El material celulósico puede ser, pero no se limita a, material herbáceo, residuo agrícola, residuo de silvicultura, residuos sólidos municipales, papel de desecho, y pulpa y residuo de fábrica de papel (véase, por ejemplo, Wiseloge et al., 1995, en Handbook on Bioethanol (Charles E. Wyman, editor), págs. 105-118, Taylor & Francis, Washington D.C.; Wyman, 1994, Bioresource Technology 50: 3-16; Lynd, 1990, Applied Biochemistry and Biotechnology 24/25: 695-719; Mosier et al., 1999, Recent Progress in Bioconversion of Lignocellulosics, en Advances in Biochemical Engineering/Biotechnology, T. Scheper, jefe de edición, volumen 65; pp. 23-40, Springer-Verlag, Nueva York). Por la presente se entiende que la celulosa puede estar en forma de lignocelulosa, un material de pared celular vegetal que contiene lignina, celulosa, y hemicelulosa en una matriz mixta. En un aspecto preferido, el material celulósico es lignocelulosa.

35

40

[0037] En un aspecto, el material celulósico es material herbáceo. En otro aspecto, el material celulósico es residuo agrícola. En otro aspecto, el material celulósico es residuo de silvicultura. En otro aspecto, el material celulósico son residuos sólidos municipales. En otro aspecto, el material celulósico es papel de desecho. En otro aspecto, el material celulósico es pulpa y residuo de fábrica de papel.

45

[0038] En otro aspecto, el material celulósico son rastrojos de maíz. En otro aspecto, el material celulósico es fibra de maíz. En otro aspecto, el material celulósico es mazorca de maíz. En otro aspecto, el material celulósico es cáscara de naranja. En otro aspecto, el material celulósico es paja de arroz. En otro aspecto, el material celulósico es paja de trigo. En otro aspecto, el material celulósico es pasto varilla. En otro aspecto, el material celulósico es miscanthus. En otro aspecto, el material celulósico es bagazo.

50

55

[0039] En otro aspecto, el material celulósico es celulosa microcristalina. En otro aspecto, el material celulósico es celulosa bacteriana. En otro aspecto, el material celulósico es celulosa algal. En otro aspecto, el material celulósico es linter de algodón. En otro aspecto, el material celulósico es celulosa amorfa tratada con ácido fosfórico. En otro aspecto, el material celulósico es papel de filtro.

60

[0040] El material celulósico se puede utilizar como si es sometido o se puede someter a pretratamiento, usando

ES 2 551 141 T3

métodos convencionales conocidos en la técnica, como se describe en este caso. En un aspecto preferido, el material celulósico es pretratado.

5 [0041] **Rastrojos de maíz pretratados:** el término "rastrojos de maíz pretratados" o "PCS" (por sus siglas en inglés) se define por la presente como un material celulósico derivado del forraje de maíz mediante tratamiento con calor y ácido sulfúrico diluido.

[0042] **Polipéptido aislado:** el término "polipéptido aislado" como se usa por la presente se refiere a un polipéptido que está aislado de una fuente.

10 En un aspecto preferido, el polipéptido es al menos 1% puro, preferiblemente al menos 5% puro, de forma más preferible al menos 10% puro, de forma más preferible al menos 20% puro, de forma más preferible al menos 40% puro, de forma más preferible al menos 60% puro, incluso de forma más preferible al menos 80% puro, y de la forma más preferible al menos 90% puro, como se determina mediante SDS-PAGE.

15 [0043] **Polipéptido sustancialmente puro:** el término "polipéptido sustancialmente puro" denota por la presente una preparación de polipéptido que contiene como mucho el 10%, preferiblemente como mucho el 8%, de forma más preferible como mucho el 6%, de forma más preferible como mucho el 5%, de forma más preferible como mucho el 4%, de forma más preferible como mucho el 3%, incluso de forma más preferible como mucho el 2%, de la forma más preferible como mucho el 1%, e incluso de la forma más preferible como mucho el 0,5% en peso de otro material de polipéptido con el cual está asociado originalmente o recombinantemente. Por lo tanto, se prefiere que el polipéptido sustancialmente puro sea al menos 92% puro, preferiblemente al menos 94% puro, de forma más preferible al menos 95% puro, de forma más preferible al menos 96% puro, de forma más preferible al menos 97% puro, de forma más preferible al menos 98%, puro incluso de forma más preferible al menos 99% puro, de la forma más preferible al menos 99,5% puro, e incluso de la forma más preferible 100% puro en peso del material de polipéptido total presente en la preparación. Los polipéptidos de la presente invención están preferiblemente en una forma sustancialmente pura, es decir, que la preparación del polipéptido está esencialmente libre de otro material de polipéptido con el cual está asociado originalmente o recombinantemente. Esto puede realizarse, por ejemplo, preparando el polipéptido mediante métodos recombinantes bien conocidos o mediante métodos de purificación tradicional.

30 [0044] **Polipéptido maduro:** el término "polipéptido maduro" se define por la presente como un polipéptido en su forma final siguiendo la traducción y cualquier modificación postraduccional, tal como tratamiento del N-terminal, truncamiento del C-terminal, glicosilación, fosforilación, etc. En un aspecto, el polipéptido maduro es los aminoácidos 23 a 251 de la SEC ID n° 2 en base al programa SignalP (Nielsen et al., 1997, Protein Engineering 10:1-6) que predice que los aminoácidos 1 a 22 de la SEC ID n° 2 son un péptido señal. En otro aspecto, el polipéptido maduro es los aminoácidos 19 a 349 de la SEC ID n° 4 en base al programa SignalP que predice que los aminoácidos 1 a 18 de la SEC ID n° 4 son un péptido señal. En otro aspecto, el polipéptido maduro es los aminoácidos 24 a 436 de la SEC ID n° 6 en base al programa SignalP que predice que los aminoácidos 1 a 23 de la SEC ID n° 6 son un péptido señal.

40 [0045] **Secuencia codificante de polipéptido maduro:** el término "secuencia codificante de polipéptido maduro" se define por la presente como una secuencia de nucleótido que codifica un polipéptido maduro con actividad de aumento celulolítico. En un aspecto, la secuencia codificante de polipéptido maduro es los nucleótidos 67 a 868 de la SEC ID n° 1 en base al programa SignalP (Nielsen et al., 1997, *supra*) que predice que los nucleótidos 1 a 66 de la SEC ID n° 1 codifican un péptido señal. En otro aspecto, la secuencia codificante de polipéptido maduro es los nucleótidos 55 a 1099 de la SEC ID n° 3 en base al programa SignalP que predice que los nucleótidos 1 a 54 de la SEC ID n° 3 codifican un péptido señal. En un aspecto, la secuencia codificante de polipéptido maduro es los nucleótidos 70 a 1483 de la SEC ID n° 5 en base al programa SignalP que predice que los nucleótidos 1 a 69 de la SEC ID n° 5 codifican un péptido señal.

[0046] **.Identidad:** la relación entre dos secuencias de aminoácido o entre dos secuencias de nucleótidos se describe con el parámetro "identidad".

50 [0047] Para fines de la presente invención, el grado de identidad entre dos secuencias de aminoácidos se determina utilizando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, J. Mol. Biol. 48: 443-453) como viene implementado en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, Trends in Genetics 16: 276- 277), preferiblemente versión 3.0.0 o posterior. Los parámetros opcionales usados son penalización por apertura de gap de 10, penalización por extensión de gap de 0,5, y la matriz de sustitución EBLOSUM62 (versión EMBOSS de BLOSUM62). La salida de Needle marcada como "identidad más larga" (obtenida utilizando la opción -nobrief) se usa como la identidad en porcentaje y es calculada de la siguiente manera:

60
$$\text{(Residuos idénticos x 100)} / (\text{Longitud de alineación} - \text{número total de espacios en la alineación})$$

[0048] Para fines de la presente invención, el grado de identidad entre dos secuencias desoxirribonucleótidas se

ES 2 551 141 T3

determina utilizando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, *supra*) como viene implementado en el programa Needle del paquete EMBOSS (EMBOSS: The European Molecular Biology Open Software Suite, Rice et al., 2000, *supra*), preferiblemente versión 3.0.0 o posteriores. Los parámetros opcionales usados son penalización por apertura de gap de 10, penalización por extensión de gap de 0,5, y la matriz de sustitución EDNAFULL (versión EMBOSS de NCBI NUC4.4). La salida de Needle marcada como "identidad más larga" (obtenida utilizando la opción -nobrief) se usa como la identidad en porcentaje y es calculada de la siguiente manera:

$(\text{Desoxiribonucleótidos idénticos} \times 100) / (\text{Longitud de alineación} - \text{número total de espacios en la alineación})$

10 [0049] **Secuencia homóloga:** el término "secuencia homóloga" se define por la presente como una proteína predicha con un valor E (o puntuación de expectativa) inferior a 0,001 en una búsqueda tfasty (Pearson, W.R., 1999, en *Bioinformatics Methods and Protocols*, S. Misener y S. A. Krawetz, ed., págs. 185-219) con el polipéptido con actividad de aumento celulolítico de *Thermoascus crustaceus* de la SEC ID nº 2 o el polipéptido maduro del mismo.

15 [0050] **Fragmento de polipéptido:** el término "fragmento de polipéptido" se define por la presente como un polipéptido con uno o más aminoácidos (diferentes) eliminados del terminal amino y/o carboxilo del polipéptido maduro de la SEC ID nº 2, SEC ID nº 4, o SEC ID nº 6; o una secuencia homóloga del mismo; donde el fragmento tiene actividad de aumento celulolítico. En un aspecto preferido, un fragmento contiene al menos 190 residuos de aminoácidos, de forma más preferible al menos 200 residuos de aminoácidos, y de la forma más preferible al menos 210 residuos de aminoácidos,
20 del polipéptido maduro de la SEC ID nº 2 o una secuencia homóloga del mismo. En otro aspecto preferido, un fragmento contiene al menos 270 residuos de aminoácidos, de forma más preferible al menos 290 residuos de aminoácidos, y de la forma más preferible al menos 310 residuos de aminoácidos, del polipéptido maduro de la SEC ID nº 4 o una secuencia homóloga del mismo. En otro aspecto preferido, un fragmento contiene al menos 340 residuos de aminoácidos, de forma más preferible al menos 360 residuos de aminoácidos, y de la forma más preferible al menos 380 residuos de aminoácidos, del polipéptido maduro de la SEC ID nº 6 o una secuencia homóloga del mismo.

[0051] **Subsecuencia:** el término "subsecuencia" se define por la presente como una secuencia de nucleótido con uno o más nucleótidos (diferentes) eliminados del extremo 5' y/o 3' de la secuencia codificante de polipéptido maduro de la SEC ID nº 1, SEC ID nº 3, o SEC ID nº 5; o una secuencia homóloga del mismo; donde la subsecuencia codifica un fragmento de polipéptido con actividad de aumento celulolítico. En un aspecto preferido, una subsecuencia contiene al menos 570 nucleótidos, de forma más preferible al menos 600 nucleótidos, y de la forma más preferible al menos 630 nucleótidos de la secuencia codificante de polipéptido maduro de la SEC ID nº 1 o una secuencia homóloga del mismo. En otro aspecto preferido, una subsecuencia contiene al menos 810 nucleótidos, de forma más preferible al menos 870 nucleótidos, y de la forma más preferible al menos 930 nucleótidos de la secuencia codificante de polipéptido maduro de
35 la SEC ID nº 3 o una secuencia homóloga del mismo. En otro aspecto preferido, una subsecuencia contiene al menos 1020 nucleótidos, de forma más preferible al menos 1080 nucleótidos, y de la forma más preferible al menos 1140 nucleótidos de la secuencia codificante de polipéptido maduro de la SEC ID nº 5 o una secuencia homóloga del mismo.

40 [0052] **Variante alélica:** el término "variante alélica" denota por la presente cualquiera de dos o más formas alternativas de un gen que ocupa la misma localización cromosómica. La variación alélica surge naturalmente a través de mutación, y puede resultar en el polimorfismo dentro de poblaciones. Las mutaciones del gen pueden ser silenciosas (sin cambios en el polipéptido codificado) o pueden codificar polipéptidos con secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

45 [0053] **Polinucleótido aislado:** el término "polinucleótido aislado" como se usa por la presente se refiere a un polinucleótido que está aislado de una fuente. En un aspecto preferido, el polinucleótido es al menos 1% puro, preferiblemente al menos 5% puro, de forma más preferible al menos 10% puro, de forma más preferible al menos 20% puro, de forma más preferible al menos 40% puro, de forma más preferible al menos 60% puro, incluso de forma más preferible al menos 80% puro, y de la forma más preferible al menos 90% puro, como se determina mediante electroforesis de agarosa.

50 [0054] **Polinucleótido sustancialmente puro:** el término "polinucleótido sustancialmente puro" como se usa por la presente se refiere a una preparación de polinucleótido libre de otros nucleótidos extraños o indeseados y en una forma adecuada para el uso dentro de sistemas de producción de proteína genéticamente modificada. Así, un polinucleótido sustancialmente puro contiene como mucho el 10%, preferiblemente como mucho el 8%, de forma más preferible como mucho el 6%, de forma más preferible como mucho el 5%, de forma más preferible como mucho el 4%, de forma más preferible como mucho el 3%, incluso de forma más preferible como mucho el 2%, de la forma más preferible como mucho el 1%, e incluso de la forma más preferible como mucho el 0,5% en peso de otro material polinucleótido con el cual está asociado originalmente o recombinantemente. Un polinucleótido sustancialmente puro puede, no obstante,
55 incluir regiones 5' y 3' no traducidas de origen natural, tales como promotores y terminadores. Se prefiere que el polinucleótido sustancialmente puro sea al menos 90% puro, preferiblemente al menos 92% puro, de forma más

ES 2 551 141 T3

- preferible al menos 94% puro, de forma más preferible al menos 95% puro, de forma más preferible al menos 96% puro, de forma más preferible al menos 97% puro, incluso de forma más preferible al menos 98% puro, de la forma más preferible al menos 99% puro, e incluso de la forma más preferible al menos 99,5% en peso puro. Los polinucleótidos de la presente invención se encuentran preferiblemente en una forma sustancialmente pura, es decir, que la preparación polinucleótida está esencialmente libre de otro material polinucleótido con el cual está originalmente o recombinantemente asociado. Los polinucleótidos pueden ser de origen genómico, ADNc, ARN, semisintético, sintético, o cualquier combinación de los mismos.
- 5
- [0055] **Secuencia codificante:** cuando se usa por la presente el término "secuencia codificante" se refiere a una secuencia de nucleótidos que especifica directamente la secuencia de aminoácidos de su producto proteínico. Los límites de la secuencia codificante son generalmente determinados por un marco de lectura abierto, que empieza normalmente con el codón de inicio ATG o los codones de inicio alternativos tales como GTG y TTG y finaliza con un codón de terminación tal como, TAA, TAG y TGA. La secuencia codificante puede ser una secuencia de nucleótidos de ADN, ADNc, sintéticos o recombinantes.
- 10
- [0056] **ADNc:** el término "ADNc" se define por la presente como una molécula de ADN que se puede preparar mediante transcripción inversa a partir de una molécula de ARNm madura empalmada obtenida de una célula eucariota. El ADNc carece de secuencias de intrón que pueden estar presentes en el ADN genómico correspondiente. La transcripción de ARN inicial primaria es un precursor del ARNm que está procesado a través de una serie de pasos antes de que aparezca como ARNm empalmado maduro. Estos pasos incluyen la eliminación de las secuencias de intrón mediante un proceso llamado empalme. El ADNc derivado del ARNm carece, por lo tanto, de cualquier secuencia intrón.
- 15
- [0057] **Constructo de ácidos nucleicos:** el término "constructo de ácidos nucleicos" como se utiliza por la presente se refiere a una molécula de ácido nucleico, bien mono- o bien bicatenaria, que está aislada de un gen de origen natural o que está modificada para contener segmentos de ácidos nucleicos de manera que de lo contrario no existirían en la naturaleza o que son sintéticos. El término constructo de ácidos nucleicos es sinónimo del término "cassette de expresión" cuando el constructo de ácidos nucleicos contiene las secuencias de control requeridas para la expresión de una secuencia codificante de la presente invención.
- 20
- [0058] **Secuencias de control:** el término "secuencias de control" se define por la presente para incluir todos los componentes necesarios para la expresión de un polinucleótido que codifica un polipéptido de la presente invención. Cada secuencia de control puede ser nativa o foránea a la secuencia de nucleótidos que codifica el polipéptido o nativa o foránea entre sí. Tales secuencias de control incluyen, pero de forma no limitativa, un líder, secuencia de poliadenilación, secuencia de propéptido, promotor, secuencia de péptido señal y terminador de transcripción. A un mínimo, las secuencias de control incluyen un promotor, y señales de parada transcripcionales y traduccionales. Las secuencias de control se pueden proporcionar con enlaces con motivo de introducir sitios de restricción específicos facilitando el ligamiento de las secuencias de control con la región codificante de la secuencia de nucleótidos que codifica un polipéptido.
- 25
- [0059] **Operativamente enlazado:** el término "operativamente enlazado" denota por la presente una configuración en la que una secuencia de control se coloca en una posición apropiada relativa a la secuencia de codificación de una secuencia polinucleótida tal que la secuencia de control dirige la expresión de la secuencia codificante de un polipéptido.
- 30
- [0060] **Expresión:** el término "expresión" incluye cualquier paso implicado en la producción de un polipéptido incluyendo, pero no limitándose a, transcripción, modificación postranscripcional, traducción, modificación postraduccional, y secreción.
- 35
- [0061] **Vector de expresión:** el término "vector de expresión" se define por la presente como una molécula de ADN lineal o circular que comprende un polinucleótido que codifica un polipéptido de la presente invención y está operativamente enlazada a más nucleótidos que ayudan a su expresión.
- 40
- [0062] **Célula huésped:** el término "célula huésped", como se utiliza por la presente, incluye cualquier tipo de célula que sea susceptible a la transformación, transfección, transducción, y similares con un constructo de ácidos nucleicos o vector de expresión que comprenda un polinucleótido de la presente invención.
- 45
- [0063] **Modificación:** el término "modificación" se refiere por la presente a cualquier modificación química del polipéptido que comprende o consiste en el polipéptido maduro de la SEC ID nº 2, SEC ID nº 4, o SEC ID nº 6; o una secuencia homóloga de los mismos; así como manipulación genética del ADN que codifica tal polipéptido. La modificación puede ser una sustitución, una delección y/o una inserción de uno o varios aminoácidos (diferentes) al igual que reemplazos de una o varias cadenas laterales de aminoácido (diferentes).
- 50
- 55
- 60

[0064] **Variante artificial:** cuando se usa por la presente, el término "variante artificial" se refiere a un polipéptido con actividad de aumento celulolítico producida por un organismo que expresa una secuencia de polinucleótidos modificada de la secuencia codificante de polipéptido maduro de la SEC ID nº 1, SEC ID nº 3, o SEC ID nº 5; o una secuencia homóloga de las mismas. La secuencia de nucleótidos modificada se obtiene a través de intervención humana mediante modificación de la secuencia de polinucleótidos descrita en la SEC ID nº 1, SEC ID nº 3, o SEC ID nº 5; o una secuencia homóloga de las mismas.

Descripción detallada de la invención

10 Polipéptidos con actividad de aumento celulolítico

[0065] En un primer aspecto, la presente invención se refiere a polipéptidos aislados con actividad de aumento celulolítico, seleccionados del grupo consistente en:

15 (a) un polipéptido que incluye una secuencia de aminoácidos con al menos el 90% de identidad al polipéptido maduro mostrado como los aminoácidos 23-251 de la SEC ID nº 2;
 (b) un polipéptido codificado por un polinucleótido que comprende una secuencia de nucleótidos con al menos el 90% de identidad a la secuencia codificante de polipéptido maduro mostrado como nucleótidos 67-868 de la SEC ID nº 1.

20 [0066] En una forma de realización las secuencias de aminoácidos tienen un grado de identidad al polipéptido maduro de la SEC ID nº 2 de preferiblemente al menos el 95%, y de la forma más preferible al menos el 96%, al menos el 97%, al menos el 98%, o al menos el 99%, con tiene actividad de aumento celulolítico (de ahora en adelante "polipéptidos homólogos"). En un aspecto preferido, los polipéptidos homólogos comprenden secuencias de aminoácidos que difieren en diez aminoácidos, preferiblemente en cinco aminoácidos, de forma más preferible en cuatro aminoácidos, incluso de forma más preferible en tres aminoácidos, de la forma más preferible en dos aminoácidos, e incluso de la forma más preferible en un aminoácido del polipéptido maduro de la SEC ID nº 2.

25 [0067] En otro aspecto preferido, el polipéptido comprende los aminoácidos 23 a 251 de la SEC ID nº 2. En otro aspecto preferido, el polipéptido consiste en los aminoácidos 23 a 251 de la SEC ID nº 2.

30 [0068] La secuencia de nucleótidos de la SEC ID nº 1; o una subsecuencia de la misma; así como la secuencia de aminoácidos de la SEC ID nº 2; o un fragmento de la misma; se puede usar para designar sondas de ácido nucleico para identificar y clonar polipéptidos codificantes de ADN con actividad de aumento celulolítico a partir de cepas de diferentes géneros o especies según métodos bien conocidos en la técnica. En particular, tales sondas se pueden usar para la hibridación con el ADN genómico o ADNc del género o especie de interés, siguiendo procedimientos de hibridación Southern estándar, para identificar y aislar el gen correspondiente en ellos. Tales sondas pueden ser considerablemente más cortas que toda la secuencia, pero deberían ser de al menos 14, preferiblemente al menos 25, de forma más preferible al menos 35, y de la forma más preferible al menos 70 nucleótidos de largo. No obstante, se prefiere que la sonda de ácido nucleico sea al menos de 100 nucleótidos de longitud. Por ejemplo, la sonda de ácido nucleico puede ser de al menos 200 nucleótidos, preferiblemente de al menos 300 nucleótidos, de forma más preferible de al menos 400 nucleótidos, o de la forma más preferible de al menos 500 nucleótidos de longitud. Se pueden utilizar sondas incluso más largas, por ejemplo, sondas de ácido nucleico que son preferiblemente de al menos 600 nucleótidos, de forma más preferible de al menos 700 nucleótidos, incluso de forma más preferible de al menos 800 nucleótidos, o de la forma más preferible de al menos 900 nucleótidos de longitud. Se pueden usar sondas tanto de ADN como de ARN. Las sondas están típicamente marcadas para la detección del gen correspondiente (por ejemplo, con ³²P, ³H, ³⁵S, biotina, o avidina). Tales sondas están comprendidas en la presente invención.

35 [0069] Una genoteca de ADN genómico o de ADNc obtenido a partir de tales otras cepas puede seleccionarse, por lo tanto, para el ADN que hibridiza con las sondas anteriormente descritas y que codifica un polipéptido con actividad de aumento celulolítico. Se puede separar ADN genómico u otro ADN a partir de tales otras cepas mediante electroforesis en gel de poliacrilamida o agarosa, u otras técnicas de separación. El ADN de las genotecas o el ADN separado se pueden transferir a e inmovilizar en nitrocelulosa u otro material portador adecuado. Para identificar un clon o ADN que es homólogo a la SEC ID nº 1, o una subsecuencia de la misma, el material portador se usa preferiblemente en una hibridación Southern.

50 [0070] La hibridación indica que la secuencia de nucleótidos hibridiza en una sonda de ácido nucleico marcada que se corresponde con la secuencia codificante de polipéptido maduro de la SEC ID nº 1; la secuencia de ADNc contenida en la secuencia codificante de polipéptido maduro de la SEC ID nº 1; su cadena complementaria en toda su longitud; o una subsecuencia de la misma; bajo condiciones de astringencia desde muy bajas a muy altas. Las moléculas a las que la sonda de ácido nucleico hibridiza bajo estas condiciones se pueden detectar usando, por ejemplo, película radiográfica.

ES 2 551 141 T3

[0071] La sonda de ácido nucleico puede ser los nucleótidos 67 a 868 de la SEC ID nº 1 o la región codificante de polipéptido maduro contenida en plásmido pGEM-T-GH61a51486 que es contenido en *E. coli* DSM 22656.

5 [0072] Para sondas largas de al menos 100 nucleótidos de longitud, se definen condiciones de astringencia de muy bajas a muy altas como prehibridación e hibridación a 42°C en 5X SSPE, 0,3% SDS, 200 µg/ml ADN de esperma de salmón cortado y desnaturalizado, y bien 25% formamida para astringencias muy bajas y bajas, bien 35% formamida para astringencias medias y medias-altas, o bien 50% formamida para astringencias altas y muy altas, siguiendo los procedimientos de hibridación Southern estándar óptimamente durante 12 a 24 horas.

10 [0073] Para sondas largas de al menos 100 nucleótidos de longitud, cada uno de los materiales portadores es lavado finalmente tres veces durante 15 minutos usando 2X SSC, 0,2% SDS preferiblemente a 45°C (astringencia muy baja), de forma más preferible a 50°C (astringencia baja), de forma más preferible a 55°C (astringencia media), de forma más preferible a 60°C (astringencia media-alta), incluso de forma más preferible a 65°C (astringencia alta), y de la forma más preferible a 70°C (astringencia muy alta).

15 [0074] Para sondas cortas de aproximadamente 15 nucleótidos hasta aproximadamente 70 nucleótidos de longitud, las condiciones de astringencia se definen como prehibridación, hibridación y post-hibridación de lavado a alrededor 5°C hasta aproximadamente 10°C por debajo de la T_m calculada utilizando el cálculo según Bolton y McCarthy (1962, Proceedings of the National Academy of Sciences USA 48:1390) en 0,9 M NaCl, 0,09 M Tris-HCl a pH 7,6, 6 mM EDTA, 0,5% NP-40, 1X solución de Denhardt, 1 mM pirofosfato de sodio, 1 mM fosfato monobásico de sodio, 0,1 mM ATP, y 0,2 mg ARN de levadura por ml siguiendo procedimientos de hibridación Southern estándar óptimamente durante 12 a 24 horas.

20 [0075] Para sondas cortas de aproximadamente 15 nucleótidos hasta aproximadamente 70 nucleótidos de longitud, el material portador se lava una vez en 6X SSC más 0,1% SDS durante 15 minutos y dos veces cada uno durante 15 minutos usando 6X SSC a 5°C hasta 10°C por debajo de la T_m calculada.

25 [0076] La presente invención se refiere a polipéptidos aislados con actividad de aumento celulolítico codificados por polinucleótidos que comprenden o consisten en secuencias de nucleótidos que tienen un grado de identidad a la secuencia codificante de polipéptido maduro mostrada como los nucleótidos 67-868 de la SEC ID nº 1 de preferiblemente al menos el 90%, de forma más preferible al menos el 95%, y de la forma más preferible al menos el 96%, al menos el 97%, al menos el 98%, o al menos el 99%, que codifican un polipéptido con actividad de aumento celulolítico. Véase sección de polinucleótido por la presente.

30 Fuentes de polipéptidos con actividad de aumento celulolítico

35 [0077] Un polipéptido con actividad de aumento celulolítico de la presente invención puede obtenerse a partir de microorganismos de cualquier género. Para fines de la presente invención, el término "obtenerse a partir de" como se utiliza en este caso en relación con una fuente dada debe significar que el polipéptido codificado por una secuencia de nucleótidos es producido por la fuente o por una cepa en la que se ha insertado la secuencia de nucleótidos de la fuente. En un aspecto preferido, el polipéptido obtenido de una fuente dada es secretado extracelularmente.

40 [0078] Un polipéptido con actividad de aumento celulolítico de la presente invención puede ser un polipéptido bacteriano. Por ejemplo, el polipéptido puede ser un polipéptido bacteriano Gram positivo tal como un polipéptido de *Bacillus*, *Streptococcus*, *Streptomyces*, *Staphylococcus*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Clostridium*, *Geobacillus*, u *Oceanobacillus* con actividad de aumento celulolítico, o un polipéptido bacteriano Gram negativo tal como un polipéptido de *E. coli*, *Pseudomonas*, *Salmonella*, *Campylobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Ilyobacter*, *Neisseria* o *Ureaplasma* con actividad de aumento celulolítico.

45 [0079] En un aspecto preferido, el polipéptido es un polipéptido de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis* o *Bacillus thuringiensis* con actividad de aumento celulolítico.

50 [0080] En otro aspecto preferido, el polipéptido es un polipéptido de *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis* o *Streptococcus equi* subesp. *zooepidemicus* con actividad de aumento celulolítico.

55 [0081] En otro aspecto preferido, el polipéptido es un polipéptido de *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus* o *Streptomyces lividans* con actividad de aumento celulolítico.

60 [0082] Un polipéptido con actividad de aumento celulolítico de la presente invención también puede ser un polipéptido

ES 2 551 141 T3

- fúngico, y de forma más preferible un polipéptido de levadura tal como un polipéptido de *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia* con actividad de aumento celulolítico; o de forma más preferible un polipéptido fúngico filamentoso tal como un polipéptido de *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Aureobasidium*, *Botryosphaeria*, *Ceriporiopsis*, *Chaetomidium*, *Chrysosporium*, *Claviceps*, *Cochliobolus*, *Coprinopsis*, *Coptotermes*, *Corynascus*, *Cryphonectria*, *Cryptococcus*, *Diplodia*, *Exidia*, *Filibasidium*, *Fusarium*, *Gibberella*, *Holomastigotoides*, *Humicola*, *Irpex*, *Lentinula*, *Leptosphaeria*, *Magnaporthe*, *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*, *Pseudoplectania*, *Pseudotriconympha*, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trichoderma*, *Trichophaea*, *Verticillium*, *Volvariella*, o *Xylaria* con actividad de aumento celulolítico.
- [0083] En un aspecto preferido, el polipéptido es un polipéptido de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, o *Saccharomyces oviformis* con actividad de aumento celulolítico.
- [0084] En otro aspecto preferido, el polipéptido es un polipéptido de *Acremonium cellulolyticus*, *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium tropicum*, *Chrysosporium merdarium*, *Chrysosporium inops*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium zonatum*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochromum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola grisea*, *Humicola insolens*, *Humicola lanuginosa*, *Irpex lacteus*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium funiculosum*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Thermoascus aurantiacus*, *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia spededonium*, *Thielavia setosa*, *Thielavia subthermophila*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* o *Trichoderma viride* con actividad de aumento celulolítico.
- [0085] En un aspecto más preferido, el polipéptido es un polipéptido de *Thermoascus crustaceus* con actividad de aumento celulolítico. En un aspecto más preferido, el polipéptido es un polipéptido de *Thermoascus crustaceus* CBS 181.67 con actividad de aumento celulolítico, por ejemplo, el polipéptido que comprende el polipéptido maduro de la SEC ID n° 2.
- [0086] Se entiende que para las especies anteriormente mencionadas la invención abarca tanto los estados perfectos como imperfectos, y otros equivalentes taxonómicos, por ejemplo, anamorfos, independientemente del nombre de especies por el que son conocidos. Aquellos expertos en la técnica reconocerán fácilmente la identidad de equivalentes apropiados.
- [0087] Las cepas de estas especies son fácilmente accesibles al público en numerosas colecciones de cultivo, tales como la American Type Culture Collection (ATCC por sus siglas en inglés), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ por sus siglas en alemán), Centraalbureau Voor Schimmelcultures (CBS por sus siglas en neerlandés), y Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL por sus siglas en inglés).
- [0088] Además, tales polipéptidos pueden identificarse y obtenerse de otras fuentes incluyendo microorganismos aislados de la naturaleza (por ejemplo, tierra, abonos, agua, etc.) utilizando las sondas mencionadas anteriormente. Las técnicas para aislar microorganismos de hábitats naturales son conocidas en la técnica. El polinucleótido puede luego obtenerse seleccionando de forma similar una genoteca genómica o de ADNc de tal microorganismo. Una vez se ha detectado con la(s) sonda(s) un polinucleótido que codifica un polipéptido, el polinucleótido se puede aislar o clonar utilizando técnicas bien conocidas por aquellos expertos en la materia en la técnica (véase, por ejemplo, Sambrook et al., 1989, *supra*).
- [0089] Los polipéptidos de la presente invención incluyen también polipéptidos fusionados o polipéptidos de fusión divisibles donde otro polipéptido se fusiona en el N-terminal o el C-terminal del polipéptido o fragmento del mismo. Un polipéptido fusionado se produce fusionando una secuencia de nucleótidos (o una parte de la misma) que codifica otro polipéptido a una secuencia de nucleótidos (o una parte de la misma) de la presente invención. En la técnica se conocen técnicas para producir polipéptidos de fusión, e incluyen ligar las secuencias codificantes que codifican los polipéptidos de modo que éstas estén en el marco y que la expresión del polipéptido fusionado esté bajo control del (de los) mismo(s) promotor(es) y terminador.

[0090] Un polipéptido de fusión puede además comprender un sitio de escisión. Tras la secreción de la proteína de fusión, el sitio es dividido liberando el polipéptido con actividad de aumento celulolítico a partir de la proteína de fusión.

5 Ejemplos de sitios de escisión incluyen, pero de forma no limitativa, un sitio Kex2 que codifica el dipéptido Lys-Arg (Martin et al., 2003, J. Ind. Microbiol. Biotechnol. 3: 568-576; Svetina et al., 2000, J. Biotechnol. 76: 245-251; Rasmussen-Wilson et al., 1997, Appl. Environ. Microbiol. 63: 3488-3493; Ward et al., 1995, Biotechnology 13: 498-503; y Contreras et al., 1991, Biotechnology 9: 378-381), un sitio Ile-(Glu o Asp)-Gly-Arg, que es escindido por una proteasa de factor Xa después del residuo de arginina (Eaton et al., 1986, Biochem. 25: 505-512); un sitio Asp-Asp-Asp-Asp-Lys, que es escindido por una enteroquinasa después de la lisina (Collins-Racie et al., 1995, Biotechnology 13: 982-987); un sitio His-Tyr-Glu o sitio His-Tyr-Asp, que es escindido por Genenase I (Carter et al., 1989, Proteins: Structure, Function, and Genetics 6: 240-248); un sitio Leu-Val-Pro-Arg-Gly-Ser, que es escindido por trombina después de la Arg (Stevens, 2003, Drug Discovery World 4: 35-48); un sitio Glu-Asn-Leu-Tyr-Phe-Gln-Gly, que es escindido por proteasa TEV después de la Gln (Stevens, 2003, supra); y un sitio Leu-Glu-Val-Leu-Phe-Gln-Gly-Pro, que es escindido por una forma genéticamente modificada de proteasa de rinovirus 3C humano después de la Gln (Stevens, 2003, supra).

15 Polinucleótidos

[0091] La presente invención también se refiere a polinucleótidos aislados que comprenden o consisten en secuencias de nucleótidos que codifican polipéptidos con actividad de aumento celulolítico de la presente invención.

20 [0092] En un aspecto preferido, la secuencia de nucleótidos comprende o consiste en la SEC ID nº 1. En otro aspecto más preferido, la secuencia de nucleótidos comprende o consiste en la secuencia contenida en el plásmido pGEM-T-GH61a51486 que está contenido en *E. coli* DSM 22656. En otro aspecto preferido, la secuencia de nucleótidos comprende o consiste en la secuencia codificante de polipéptido maduro de la SEC ID nº 1. En otro aspecto preferido, la secuencia de nucleótidos comprende o consiste en los nucleótidos 67 a 868 de la SEC ID nº 1. En otro aspecto más preferido, la secuencia de nucleótidos comprende o consiste en la secuencia codificante de polipéptido maduro contenida en el plásmido pGEM-T-GH61a51486 que está contenida en *E. coli* DSM 22656. La presente invención también abarca secuencias de nucleótidos que codifican polipéptidos que comprenden o consisten en la secuencia de aminoácidos de la SEC ID nº 2 o el polipéptido maduro de la misma, que difiere de la SEC ID nº 1 o la secuencia de codificación de polipéptido maduro de la misma en virtud de la degeneración del código genético. La presente invención también se refiere a subsecuencias de la SEC ID nº 1 que codifican fragmentos de la SEC ID nº 2 que tienen actividad de aumento celulolítico.

35 [0093] Las técnicas usadas para aislar o clonar un polinucleótido que codifica un polipéptido son conocidas en la técnica e incluyen el aislamiento de ADN genómico, preparación de ADNc, o una combinación de los mismos. La clonación de los polinucleótidos de la presente invención de tal ADN genómico puede llevarse a cabo, por ejemplo, usando la conocida reacción en cadena de polimerasa (PCR por sus siglas en inglés) o la selección de anticuerpos de bibliotecas de expresión para detectar fragmentos de ADN clonado con características estructurales compartidas. Véase, por ejemplo, Innis et al., 1990, PCR: A Guide to Methods and Application, Academic Press, Nueva York. Se pueden utilizar otros procedimientos de amplificación de ácido nucleico tales como la reacción en cadena de la ligasa (LCR por sus siglas en inglés), transcripción activada ligada (LAT por sus siglas en inglés) y amplificación basada en secuencias de nucleótidos (NASBA por sus siglas en inglés). Los polinucleótidos se pueden clonar a partir de una cepa de *Thermascus*, u otro organismo u organismo relacionado y así, por ejemplo, puede ser una variante alélica o de especie del polipéptido codificante de la región de la secuencia de nucleótidos.

45 [0094] La presente invención también se refiere a polinucleótidos aislados que comprenden o consisten en secuencias de nucleótidos con un grado de identidad a la secuencia codificante de polipéptido maduro de la SEC ID nº 1 de preferiblemente al menos el 90%, de forma más preferible al menos el 95%, y de la forma más preferible al menos el 96%, al menos el 97%, al menos el 98%, o al menos el 99%, que codifican un polipéptido con actividad de aumento celulolítico.

50 [0095] La modificación de una secuencia de nucleótidos que codifica un polipéptido de la presente invención puede ser necesaria para la síntesis de polipéptidos sustancialmente similar al polipéptido. El término "sustancialmente similar" al polipéptido se refiere a formas del polipéptido que no ocurren naturalmente. Estos polipéptidos pueden diferir de alguna manera manipulada del polipéptido aislado de su fuente nativa, por ejemplo, variantes artificiales que difieren en la actividad específica, termoestabilidad, pH óptimo, o similar. La secuencia variante se puede construir en base a la secuencia de nucleótidos presentada como la secuencia codificante de polipéptido maduro de la SEC ID nº 1, por ejemplo, una subsecuencia de la misma, y/o por introducción de sustituciones de nucleótido que no da lugar a otra secuencia de aminoácidos del polipéptido codificado por la secuencia de nucleótidos, pero que corresponden al uso del codón del organismo huésped destinado a la producción de la enzima, o por introducción de sustituciones de nucleótido que puede dar lugar a una secuencia de aminoácidos diferente. Para una descripción general de sustitución de nucleótido, véase, por ejemplo, Ford et al., 1991, Protein Expression and Purification 2: 95-107.

60

[0096] Se hará aparente a los expertos en la técnica que tales sustituciones pueden hacerse fuera de las regiones críticas para la función de la molécula y aún dar como resultado un polipéptido activo. Los residuos de aminoácidos esenciales para la actividad del polipéptido codificado por un polinucleótido aislado de la invención, y por lo tanto preferiblemente no sometido a sustitución, se pueden identificar según procedimientos conocidos en la técnica, tales como mutagénesis dirigida al sitio o mutagénesis de escáner de alanina (véase, por ejemplo, Cunningham y Wells, 1989, *supra*). En la última técnica, las mutaciones se introducen en cada residuo cargado positivamente en la molécula, y las moléculas mutantes resultantes son evaluadas para comprobar la actividad de aumento celulolítico para identificar los residuos de aminoácido que son críticos para la actividad de la molécula. Los sitios de interacción enzima-sustrato también pueden ser determinados mediante análisis de la estructura tridimensional como viene determinado por tales técnicas como análisis de resonancia magnética nuclear, cristalografía o marcaje por fotoafinidad (véase, por ejemplo, de Vos et al., 1992, *supra*; Smith et al., 1992, *supra*; Wlodaver et al., 1992, *supra*).

Constructos de ácidos nucleicos

[0097] También están descritos constructos de ácidos nucleicos que comprenden un polinucleótido aislado de la presente invención enlazado operativamente a una o varias secuencias de control (diferentes) que dirigen la expresión de la secuencia codificante en una célula huésped adecuada bajo condiciones compatibles con las secuencias de control.

[0098] Un polinucleótido aislado que codifica un polipéptido de la presente invención se puede manipular de varias maneras para ayudar a la expresión del polipéptido. La manipulación de la secuencia del polinucleótido antes de la inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar secuencias polinucleótidas utilizando métodos de ADN recombinante son bien conocidas en la técnica.

[0099] La secuencia de control puede ser una secuencia promotora apropiada, una secuencia de nucleótidos que es reconocida por una célula huésped para la expresión de un polinucleótido que codifica un polipéptido de la presente invención. La secuencia promotora contiene secuencias de control transcripcionales que median la expresión del polipéptido. El promotor puede ser cualquier secuencia de nucleótidos que muestra actividad transcripcional en la célula huésped de elección incluyendo promotores mutantes, truncados e híbridos, y se puede obtener de genes que codifican polipéptidos extracelulares o intracelulares bien homólogos o heterólogos a la célula huésped.

[0100] Ejemplos de promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente invención, especialmente en una célula huésped bacteriana, son los promotores obtenidos del operón *lac* de *E. coli*, del gen de agarasa (*dagA*) de *Streptomyces coelicolor*, gen de levansucrasa (*sacB*) de *Bacillus subtilis*, gen de alfa-amilasa (*amyL*) de *Bacillus licheniformis*, gen de amilasa maltogénica (*amyM*) de *Bacillus stearothermophilus*, gen de alfa-amilasa (*amyQ*) de *Bacillus amyloliquefaciens*, gen de penicilinas (*penP*) de *Bacillus licheniformis*, genes *xylA* y *xylB* de *Bacillus subtilis*, y gen de beta-lactamasa procariótica (Villa-Kamaroff et al., 1978, Proceedings of the National Academy of Sciences USA 75: 3727-3731), al igual que el promotor *tac* (DeBoer et al., 1983, Proceedings of the National Academy of Sciences USA 80: 21-25). Otros promotores se describen en "Useful proteins from recombinant bacteria" en Scientific American, 1980, 242: 74-94; y en Sambrook et al., 1989, *supra*.

[0101] Ejemplos de promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente invención en una célula huésped fúngica filamentosa son promotores obtenidos a partir de los genes para TAKA amilasa de *Aspergillus oryzae*, proteinasa aspártica de *Rhizomucor miehei*, alfa-amilasa neutra de *Aspergillus niger*, alfa-amilasa estable en ácido de *Aspergillus niger*, glucoamilasa (*gluA*) de *Aspergillus niger* o *Aspergillus awamori*, lipasa de *Rhizomucor miehei*, proteasa alcalina de *Aspergillus oryzae*, triosa fosfato isomerasa de *Aspergillus oryzae*, acetamidasa de *Aspergillus nidulans*, amilogucosidasa de *Fusarium venenatum* (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn (WO 00/56900), proteasa de tipo tripsina de *Fusarium oxysporum* (WO 96/00787), beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa IV de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasa I de *Trichoderma reesei*, xilanasa II de *Trichoderma reesei*, beta-xilosidasa de *Trichoderma reesei*, al igual que el promotor NA2-tpi (un promotor modificado apartir del gen que codifica alfa-amilasa neutra en *Aspergillus niger* donde el líder no traducido ha sido sustituido por un líder no traducido a partir del gen que codifica triosa fosfato isomerasa en *Aspergillus nidulans*); y promotores mutantes, truncados e híbridos de los mismos.

[0102] En un huésped de levadura, los promotores útiles se obtienen a partir de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), galactoquinasa de *Saccharomyces cerevisiae* (GAL1), deshidrogenasa de deshidrogenasa/gliceraldehído-3-fosfato de alcohol de *Saccharomyces cerevisiae* (ADH1, ADH2/GAP), triosa fosfato isomerasa de *Saccharomyces cerevisiae* (TPI), metalotioneína de *Saccharomyces cerevisiae* (CUP1), y quinasa de 3-

ES 2 551 141 T3

fosfoglicerato de *Saccharomyces cerevisiae*. Otros promotores útiles para células huésped de levadura son descritos por Romanos et al., 1992, *Yeast* 8: 423-488.

5 [0103] La secuencia de control también puede ser una secuencia terminadora de transcripción adecuada, una secuencia reconocida por una célula huésped para terminar la transcripción. La secuencia terminadora está enlazada operativamente al terminal 3' de la secuencia de nucleótidos que codifica el polipéptido. Cualquier terminador que sea funcional en la célula huésped de elección se puede utilizar en la presente invención.

10 [0104] Los terminadores preferidos para células huésped fúngicas filamentosas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae*, glucoamilasa de *Aspergillus niger*, antranilato sintasa de *Aspergillus nidulans*, alfa-glucosidasa de *Aspergillus niger* y proteasa de tipo tripsina de *Fusarium oxysporum*.

15 [0105] Los terminadores preferidos para células huésped de levadura se obtienen a partir de los genes para enolasa de *Saccharomyces cerevisiae*, citocromo C (CYC1) de *Saccharomyces cerevisiae*, y gliceraldehído-3-fosfato-deshidrogenasa de *Saccharomyces cerevisiae*. Otros terminadores útiles para células huésped de levadura son descritos por Romanos et al., 1992, *supra*.

20 [0106] La secuencia de control también puede ser una secuencia líder adecuada, una región no traducida de un ARNm que sea importante para la traducción por la célula huésped. La secuencia líder está enlazada operativamente al terminal 5' de la secuencia de nucleótidos que codifica el polipéptido. Cualquier secuencia líder que sea funcional en la célula huésped de elección se puede utilizar en la presente invención.

25 [0107] Los líderes preferidos para células huésped fúngicas filamentosas son obtenidos de los genes para TAKA amilasa de *Aspergillus oryzae* y triosa fosfato isomerasa de *Aspergillus nidulans*.

30 [0108] Los líderes adecuados para células huésped de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), quinasa de 3-fosfoglicerato de *Saccharomyces cerevisiae*, alfa-factor de *Saccharomyces cerevisiae* y deshidrogenasa de deshidrogenasa/gliceraldehído-3-fosfato de alcohol de *Saccharomyces cerevisiae* (ADH2/GAP).

35 [0109] La secuencia de control también puede ser una secuencia de poliadenilación, una secuencia enlazada operativamente al terminal 3' de la secuencia de nucleótidos y, cuando es transcrita, es reconocida por la célula huésped como una señal para añadir residuos de poliadenosina al ARNm transcrito. En la presente invención se puede utilizar cualquier secuencia de poliadenilación que sea funcional en la célula huésped de elección.

40 [0110] Las secuencias de poliadenilación preferidas para células huésped fúngicas filamentosas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae*, glucoamilasa de *Aspergillus niger*, antranilato sintasa de *Aspergillus nidulans*, proteasa de tipo tripsina de *Fusarium oxysporum* y alfa-glucosidasa de *Aspergillus niger*.

[0111] Las secuencias de poliadenilación útiles para células huésped de levadura son descritas por Guo y Sherman, 1995, *Molecular Cellular Biology* 15: 5983-5990.

45 [0112] La secuencia de control también puede ser una secuencia codificante del péptido señal que codifica un péptido señal enlazado al terminal amino de un polipéptido y dirige el polipéptido codificado en la vía secretora de la célula. El terminal 5' de la secuencia codificante de la secuencia de nucleótidos puede contener intrínsecamente una secuencia codificante del péptido señal enlazada naturalmente en el marco de lectura de la traducción con el segmento de la secuencia codificante que codifica el polipéptido segregado. Alternativamente, el terminal 5' de la secuencia codificante puede contener una secuencia codificante del péptido señal que sea foránea a la secuencia codificante. La secuencia codificante foránea del péptido señal puede ser necesaria donde la secuencia codificante no contiene naturalmente una secuencia codificante del péptido señal. Alternativamente, la secuencia codificante foránea del péptido señal puede simplemente reemplazar la secuencia codificante del péptido señal natural para mejorar la secreción del polipéptido. No obstante, en la presente invención se puede utilizar cualquier secuencia codificante del péptido señal que dirija el polipéptido expresado en la vía secretora de una célula huésped de elección, es decir, que es segregado en un medio de cultivo.

55 [0113] Las secuencias de codificación de péptido señal eficaces para células huésped bacterianas son las secuencias de codificación del péptido señal obtenidas a partir de los genes para amilasa maltogénica de *Bacillus* NCIB 11837, alfa-amilasa de *Bacillus stearothermophilus*, subtilisina de *Bacillus licheniformis*, beta-lactamasa de *Bacillus licheniformis*, proteasas neutras de (*nprT*, *nprS*, *nprM*) *Bacillus stearothermophilus*, y *prsA* de *Bacillus subtilis*. Otros péptidos señal son descritos por Simonen y Palva, 1993, *Microbiological Reviews* 57: 109-137.

60

ES 2 551 141 T3

- 5 [0114] Las secuencias de codificación de péptido señal eficaces para células huésped fúngicas filamentosas son las secuencias de codificación de péptido señal obtenidas a partir de los genes para TAKA amilasa de *Aspergillus oryzae*, amilasa neutra de *Aspergillus niger*, glucoamilasa de *Aspergillus niger*, proteinasa aspártica de *Rhizomucor miehei*, celulasa de *Humicola insolens*, endoglucanasa V de *Humicola insolens*, y lipasa de *Humicola lanuginosa*.
- 10 [0115] Los péptidos señal útiles para células huésped de levadura se obtienen a partir de los genes para factor alfa de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*. Otras secuencias de codificación de péptido señal útiles son descritas por Romanos et al., 1992, *supra*.
- [0116] El péptido señal comprende o consiste en los aminoácidos 1 a 22 de la SEC ID nº 2. La secuencia codificante del péptido señal comprende o consiste en los nucleótidos 1 a 66 de la SEC ID nº 1.
- 15 [0117] El péptido señal comprende o consiste en los aminoácidos 1 a 18 de la SEC ID nº 4. La secuencia codificante del péptido señal comprende o consiste en los nucleótidos 1 a 54 de la SEC ID nº 3.
- [0118] El péptido señal comprende o consiste en los aminoácidos 1 a 23 de la SEC ID nº 6. La secuencia codificante del péptido señal comprende o consiste en los nucleótidos 1 a 69 de la SEC ID nº 5.
- 20 [0119] La secuencia de control también puede ser una secuencia codificante de propéptido que codifica un propéptido situado en el terminal amino de un polipéptido. El polipéptido resultante es conocido como proenzima o propolipéptido (o zimógeno en algunos casos). Un propéptido está generalmente inactivo y se puede convertir en un polipéptido activo maduro por escisión catalítica o autocatalítica del propéptido a partir del propolipéptido. La secuencia codificante del propéptido se puede obtener a partir de los genes para proteasa alcalina (*aprE*) de *Bacillus subtilis*, proteasa neutra (*nprT*) de *Bacillus subtilis*, factor alfa de *Saccharomyces cerevisiae*, proteinasa aspártica de *Rhizomucor miehei*, y lacasa de *Myceliophthora thermophila* (WO 95/33836).
- 25 [0120] Donde ambas secuencias de péptido y propéptido señal están presentes en el terminal amino de un polipéptido, la secuencia de propéptido está situada junto al terminal amino de un polipéptido y la secuencia de péptido señal está situada junto al terminal amino de la secuencia de propéptido.
- 30 [0121] También puede ser deseable añadir secuencias reguladoras que permitan la regulación de la expresión del polipéptido en relación al crecimiento de la célula huésped. Ejemplos de sistemas reguladores son aquellos que causan que la expresión del gen se encienda o apague en respuesta a un estímulo químico o físico, incluyendo la presencia de un compuesto regulador. Los sistemas reguladores en sistemas procarióticos incluyen los sistemas operadores *lac*, *tac* y *trp*. En la levadura, se puede utilizar el sistema ADH2 o sistema GAL1. En hongos filamentosos, el promotor de TAKA alfa-amilasa, promotor de glucoamilasa de *Aspergillus niger*, y promotor de glucoamilasa de *Aspergillus oryzae* se puede utilizar como secuencias reguladoras. Otros ejemplos de secuencias reguladoras son aquellas que permiten la amplificación génica. En sistemas eucarióticos, estas secuencias reguladoras incluyen el gen de dihidrofolato reductasa que se amplifica en presencia de metotrexato, y los genes de metalotioneína que se amplifican con metales pesados. En estos casos, la secuencia de nucleótidos que codifica el polipéptido sería enlazada operativamente con la secuencia reguladora.
- 35 [0122] También hay descritos vectores de expresión recombinantes que comprenden un polinucleótido de la presente invención, un promotor, y señales de parada transcripcionales y traduccionales. Los diferentes ácidos nucleicos y secuencias de control descritos por la presente se pueden unir entre ellos para producir un vector de expresión recombinante que puede incluir uno o varios sitios de restricción (diferentes) convenientes para permitir la inserción o sustitución de la secuencia de nucleótidos que codifica el polipéptido en tales sitios. Alternativamente, una secuencia de polinucleótidos de la presente invención se puede expresar mediante la inserción de la secuencia de nucleótidos o un constructo de ácidos nucleicos que comprende la secuencia en un vector apropiado para la expresión. En la creación del vector de expresión, la secuencia codificante se localiza en el vector de modo que la secuencia codificante está enlazada operativamente con las secuencias de control apropiadas para la expresión.
- 40 [0123] El vector de expresión recombinante puede ser cualquier vector (por ejemplo, un plásmido o virus) que puede ser sometido convenientemente a procedimientos de ADN recombinante y puede provocar la expresión de la secuencia de nucleótidos. La elección del vector dependerá típicamente de la compatibilidad del vector con la célula huésped en la que el vector debe ser introducido. Los vectores puede ser plásmidos lineales o circulares cerrados.
- 45 [0124] El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, la replicación del cual es independiente de la replicación cromosómica, por ejemplo, un plásmido, un
- 50
- 55
- 60

ES 2 551 141 T3

elemento extracromosómico, un minicromosoma, o un cromosoma artificial. El vector puede contener cualquier medio para asegurar la autorreplicación. Alternativamente, el vector puede ser uno que, cuando es introducido en la célula huésped, se integra en el genoma y se replica con el (los) cromosoma(s) en que ha sido integrado. Además, puede(n) utilizarse un único vector o plásmido o dos o más vectores o plásmidos que juntos contengan el ADN total por introducir en el genoma de la célula huésped, o un transposón.

[0125] Los vectores preferiblemente contienen uno o varios marcadores seleccionables (diferentes) que permiten la fácil selección de células transformadas, modificadas, transducidas o similares. Una etiqueta seleccionable es un gen el producto del cual proporciona resistencia biocida o vírica, resistencia frente a metales pesados, prototrofia frente a auxótrofos, y similares.

[0126] Ejemplos de marcadores bacterianos seleccionables son los genes *dal* de *Bacillus subtilis*, *Bacillus licheniformis*, o marcadores que confieren resistencia antibiótica tales como ampicilina, canamicina, cloranfenicol, o resistencia a tetraciclina. Marcadores adecuados para células huésped de levadura son ADE2, HIS3, LEU2, LYS2, MET3, TRP1, y URA3. Marcadores seleccionables para su uso en una célula huésped fúngica filamentosa incluyen, pero de forma no limitativa, *amdS* (acetamidasa), *argB* (ornitina-carbamoyltransferasa), *bar* (fosfonitricina acetiltransferasa), *hph* (higromicina fosfotransferasa), *niaD* (nitrito-reductasa), *pyrG* (orotidina-5'-fosfato-decarboxilasa), *sC* (sulfato adeniltransferasa), y *trpC* (antranilato sintasa), al igual que equivalentes de los mismos. Para el uso en una célula de *Aspergillus* se prefieren los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces hygroscopicus*.

[0127] Los vectores del preferiblemente contienen un(os) elemento(s) que permite(n) la integración del vector en el genoma de la célula huésped o la replicación autónoma del vector en la célula independiente del genoma.

[0128] Para la integración en el genoma de la célula huésped, el vector puede depender de la secuencia del polinucleótido que codifica el polipéptido o cualquier otro elemento del vector para la integración en el genoma mediante recombinación homóloga o no homóloga. Alternativamente, el vector puede contener secuencias adicionales de nucleótidos para dirigir la integración por recombinación homóloga dentro del genoma de la célula huésped en una ubicación(es) precisa(s) en el (los) cromosoma(s). Para aumentar la probabilidad de integración en una ubicación precisa, los elementos integracionales deberían contener preferiblemente un número suficiente de ácidos nucleicos, tal como de 100 a 10.000 pares de bases, preferiblemente 400 a 10.000 pares de bases, y de la forma más preferible 800 a 10.000 pares de bases, que tienen un alto grado de identidad a la secuencia diana correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos integracionales pueden ser cualquier secuencia que sea homóloga a la secuencia diana en el genoma de la célula huésped. Además, los elementos integracionales pueden ser secuencias de nucleótidos no codificantes o codificantes. Por otro lado, el vector se puede integrar en el genoma de la célula huésped mediante recombinación no homóloga.

[0129] Para la replicación autónoma, el vector puede comprender además un origen de replicación que permite al vector replicarse de manera autónoma en la célula huésped en cuestión. El origen de la replicación puede ser cualquier replicador plásmido que medie la replicación autónoma que funciona en una célula. El término "origen de replicación" o "replicador plásmido" se define por la presente como una secuencia de nucleótidos que permite que un plásmido o vector se replique *in vivo*.

[0130] Ejemplos de orígenes bacterianos de replicación son los orígenes de replicación de plásmidos pBR322, pUC19, pACYC177, y pACYC184 que permiten la replicación en *E. coli*, y pUB110, pE194, pTA1060, y pAM β 1 que permiten la replicación en *Bacillus*.

[0131] Ejemplos de orígenes de replicación para su uso en una célula huésped de levadura son el origen de replicación de 2 micras, ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6.

[0132] Ejemplos de orígenes de replicación útiles en una célula fúngica filamentosa son AMA1 y ANS1 (Gems et al., 1991, Gene 98: 61-67; Cullen et al., 1987, Nucleic Acids Research 15: 9163-9175; WO 00/24883). El aislamiento del gen AMA1 y la construcción de plásmidos o vectores que comprenden el gen se puede conseguir según los métodos descritos en el documento WO 00/24883.

[0133] Se puede insertar más de una copia de un polinucleótido de la presente invención en una célula huésped para aumentar la producción del producto génico. Puede conseguirse un aumento en el número de copias del polinucleótido integrando al menos una copia adicional de la secuencia en el genoma de la célula huésped o incluyendo un gen marcador seleccionable amplificable con el polinucleótido donde las células que contienen copias amplificadas del gen marcador seleccionable, y por consiguiente copias adicionales del polinucleótido, se pueden seleccionar mediante el cultivo de células en presencia del agente seleccionable apropiado.

[0134] Los procedimientos usados para enlazar los elementos anteriormente descritos para la construcción de los vectores de expresión recombinantes de la presente invención son conocidos por un experto en la materia (véase, por ejemplo, Sambrook et al., 1989, *supra*).

5

Células huésped

[0135] Descritas hay células huésped recombinantes, que incluyen un polinucleótido aislado de la presente invención, que son usadas ventajosamente en la producción recombinante de los polipéptidos con actividad de aumento celulolítico. Un vector que comprende un polinucleótido de la presente invención se introduce en una célula huésped de modo que el vector es mantenido como un integrante cromosómico o como un vector extracromosómico que se autorreplica como se ha descrito anteriormente. El término "célula huésped" abarca cualquier descendiente de una célula madre que no es idéntica a la célula madre debido a mutaciones que ocurren durante la replicación. La elección de una célula huésped dependerá en gran parte del gen que codifica el polipéptido y su fuente.

10

15

[0136] La célula huésped puede ser cualquier célula útil en la producción recombinante de un polipéptido de la presente invención, por ejemplo, un procarionta o un eucariota.

20

[0137] La célula huésped procariótica puede ser cualquier bacteria Gram positiva o una bacteria Gram negativa. Las bacterias Gram positivas incluyen, pero no se limitan a, *Bacillus*, *Streptococcus*, *Streptomyces*, *Staphylococcus*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Clostridium*, *Geobacillus*, y *Oceanobacillus*. Las bacterias Gram negativas incluyen, pero no se limitan a, *E. coli*, *Pseudomonas*, *Salmonella*, *Campylobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Ilyobacter*, *Neisseria*, y *Ureaplasma*.

25

[0138] La célula huésped bacteriana puede ser cualquier célula de *Bacillus*. Células de *Bacillus* útiles en la práctica de la presente invención incluyen, pero de forma no limitativa, células de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, y *Bacillus thuringiensis*.

30

[0139] En un aspecto preferido, la célula huésped bacteriana es una célula de *Bacillus amyloliquefaciens*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus* o *Bacillus subtilis*. En un aspecto más preferido, la célula huésped bacteriana es una célula de *Bacillus amyloliquefaciens*. En otro aspecto más preferido, la célula huésped bacteriana es una célula de *Bacillus clausii*. En otro aspecto más preferido, la célula huésped bacteriana es una célula de *Bacillus licheniformis*. En otro aspecto más preferido, la célula huésped bacteriana es una célula de *Bacillus subtilis*.

35

[0140] La célula huésped bacteriana también puede ser cualquier célula de *Streptococcus*. Células de *Streptococcus* útiles en la práctica de la presente invención incluyen, pero de forma no limitativa, células de *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, y *Streptococcus equi* subesp. *zooepidemicus*.

40

[0141] En un aspecto preferido, la célula huésped bacteriana es una célula de *Streptococcus equisimilis*. En otro aspecto preferido, la célula huésped bacteriana es una célula de *Streptococcus pyogenes*. En otro aspecto preferido, la célula huésped bacteriana es una célula de *Streptococcus uberis*. En otro aspecto preferido, la célula huésped bacteriana es una célula de *Streptococcus equi* subesp. *zooepidemicus*.

45

[0142] La célula huésped bacteriana también puede ser cualquier célula de *Streptomyces*. Células de *Streptomyces* útiles en la práctica de la presente invención incluyen, pero de forma no limitativa, células de *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, y *Streptomyces lividans*.

50

[0143] En un aspecto preferido, la célula huésped bacteriana es una célula de *Streptomyces achromogenes*. En otro aspecto preferido, la célula huésped bacteriana es una célula de *Streptomyces avermitilis*. En otro aspecto preferido, la célula huésped bacteriana es una célula de *Streptomyces coelicolor*. En otro aspecto preferido, la célula huésped bacteriana es una célula de *Streptomyces griseus*. En otro aspecto preferido, la célula huésped bacteriana es una célula de *Streptomyces lividans*.

55

[0144] La introducción de ADN en una célula de *Bacillus* puede, por ejemplo, efectuarse mediante transformación de protoplasto (véase, por ejemplo, Chang y Cohen, 1979, Molecular General Genetics 168: 111- 115), usando células competentes (véase, por ejemplo, Young y Spizizen, 1961, Journal of Bacteriology 81: 823- 829, o Dubnau y Davidoff-Abelson, 1971, Journal of Molecular Biology 56: 209-221), mediante electroporación (véase, por ejemplo, Shigekawa y Dower, 1988 Biotechniques 6: 742-751), o mediante conjugación (véase, por ejemplo, Koehler y Thorne, 1987, Journal of Bacteriology 169: 5271-5278). La introducción de ADN en una célula *E. coli* puede, por ejemplo, efectuarse mediante

60

transformación de protoplasto (véase, por ejemplo, Hanahan, 1983, J. Mol. Biol. 166: 557-580) o electroporación (véase, por ejemplo, Dower et al., 1988, Nucleic Acids Res. 16: 6127-6145). La introducción de ADN en una célula de *Streptomyces* puede, por ejemplo, efectuarse mediante transformación del protoplasto y electroporación (véase, por ejemplo, Gong et al., 2004, Folia Microbiol. (Praga) 49: 399-405), mediante conjugación (véase, por ejemplo, Mazodier et al., 1989, J. Bacteriol. 171: 3583-3585), o mediante transducción (véase, por ejemplo, Burke et al., 2001, Proc. Natl. Acad. Sci. EE.UU. 98: 6289-6294). La introducción de ADN en una célula de *Pseudomonas* puede, por ejemplo, efectuarse mediante electroporación (véase, por ejemplo, Choi et al., 2006, J. Microbiol. Methods 64: 391-397) o mediante conjugación (véase, por ejemplo, Pinedo y Smets, 2005, Appl. Environ. Microbiol. 71: 51-57). La introducción de ADN en una célula de *Streptococcus* puede, por ejemplo, efectuarse mediante competencia natural (véase, por ejemplo, Perry y Kuramitsu, 1981, Infect. Immun. 32: 1295-1297), mediante transformación de protoplasto (véase, por ejemplo, Catt y Jollick, 1991, Microbios. 68: 189-207), mediante electroporación (véase, por ejemplo, Buckley et al., 1999, Appl. Environ. Microbiol. 65: 3800-3804) o mediante conjugación (véase, por ejemplo, Clewell, 1981, Microbiol. Rev. 45: 409-436). No obstante, se puede usar cualquier método conocido en la técnica para la introducción de ADN en una célula huésped.

[0145] La célula huésped también puede ser un eucariota, tal como un mamífero, insecto, planta, o célula fúngica.

[0146] La célula huésped puede ser una célula fúngica. "Hongo" como se utiliza en este caso incluye la fila Ascomycota, Basidiomycota, Chytridiomycota, y Zygomycota (como viene definido por Hawksworth et al., en Ainsworth and Bisby's Dictionary of The Fungi, 8ª edición, 1995, CAB International, University Press, Cambridge, Reino Unido) al igual que la Oomycota (como viene citado en Hawksworth et al., 1995, supra, página 171) y todos los hongos mitospóricos (Hawksworth et al., 1995, supra).

[0147] La célula huésped fúngica puede ser una célula de levadura. "Levadura" como se utiliza por la presente incluye levadura ascoesporógena (Endomycetales), levadura basidioesporógena, y levadura perteneciente a los Fungi Imperfecti (Blastomycetes).

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

[0148] La célula huésped de levadura puede ser una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia*.

[0149] La célula huésped de levadura puede ser una célula de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, o *Saccharomyces oviformis*. En otro aspecto más preferido, la célula huésped de levadura es una célula de *Kluyveromyces lactis*. En otro aspecto más preferido, la célula huésped de levadura es una célula de *Yarrowia lipolytica*.

[0150] La célula huésped fúngica puede ser una célula fúngica filamentosa. "Hongo filamentoso" incluye todas las formas filamentosas de la subdivisión Eumycota y Oomycota (como viene definido por Hawksworth et al., 1995, supra). Los hongos filamentosos se caracterizan generalmente por una pared micelial compuesta de quitina, celulosa, glucano, quitosano, manano, y otros polisacáridos complejos. El crecimiento vegetativo es por alargamiento hifal y el catabolismo de carbono es estrictamente aeróbico. En cambio, el crecimiento vegetativo por levaduras tales como *Saccharomyces cerevisiae* se da mediante injerto de un talo unicelular y el catabolismo de carbono puede ser fermentativo.

[0151] La célula huésped fúngica filamentosa puede ser una célula de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Chrysosporium*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, o *Trichoderma*.

[0152] La célula huésped fúngica filamentosa puede ser una célula de *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* o *Aspergillus oryzae*. En otro aspecto más preferido, la célula huésped fúngica filamentosa es una célula de *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochromum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, o *Fusarium venenatum*. En otro aspecto más preferido, la célula huésped fúngica filamentosa es una célula de *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium tropicum*, *Chrysosporium merdarium*, *Chrysosporium inops*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium zonatum*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola*

insolens, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.

5

[0153] Se pueden transformar células fúngicas mediante un proceso que implica la formación de protoplastos, la transformación de los protoplastos, y la regeneración de la pared celular de una manera conocida *per se*. Se describen procedimientos adecuados para la transformación de células huésped de *Aspergillus* y *Trichoderma* en el documento EP 238 023 y en Yelton et al., 1984, Proceedings of the National Academy of Sciences 81: 1470-1474. Los métodos adecuados para transformar especies de *Fusarium* son descritos por Malardier et al., 1989, Gene 78: 147-156, y el documento WO 96/00787. La levadura se puede transformar utilizando los procedimientos descritos por Becker y Guarente, en Abelson, J.N. y Simon, M.I., editores, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, volumen 194, págs 182-187, Academic Press, Inc., Nueva York; Ito et al., 1983, Journal of Bacteriology 153: 163; y Hinnen et al., 1978, Proceedings of the National Academy of Sciences USA 75: 1920.

10

15

Métodos de producción

[0154] La presente invención también se refiere a métodos de producción de un polipéptido de la presente invención, que comprende: (a) cultivo de una célula de *Thermoascus crustaceus*, que en su forma de tipo natural produce el polipéptido, bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido. En un aspecto más preferido, la célula es *Thermoascus crustaceus* CBS 181.67.

20

[0155] La presente invención también se refiere a métodos de producción de un polipéptido de la presente invención, que comprende: (a) cultivo de una célula huésped recombinante, como se describe en este caso, bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.

25

[0156] En los métodos de producción de la presente invención, las células se cultivan en un medio nutritivo adecuado para la producción del polipéptido utilizando métodos bien conocidos en la técnica. Por ejemplo, la célula se puede cultivar por cultivo en matraz de agitación, y fermentación a pequeña escala o a gran escala (incluyendo fermentaciones continuas, en lote, en lote alimentado, o en estado sólido) en el laboratorio o fermentadores industriales realizadas en un medio adecuado y bajo condiciones que permiten que el polipéptido sea expresado y/o aislado. El cultivo tiene lugar en un medio nutritivo adecuado que comprende fuentes de carbono y nitrógeno y sales inorgánicas, utilizando procedimientos conocidos en la técnica. Hay disponibles medios adecuados de proveedores comerciales o se pueden preparar según composiciones publicadas (por ejemplo, en catálogos de la American Type Culture Collection). Si el polipéptido se secretado en el medio nutritivo, el polipéptido se puede recuperar directamente del medio. Si el polipéptido no es secretado en el medio, puede ser recuperado de lisatos de célula.

30

35

[0157] Los polipéptidos se pueden detectar usando métodos conocidos en la técnica que son específicos para los polipéptidos. Estos métodos de detección pueden incluir el uso de anticuerpos específicos, la formación de un producto enzimático, o la desaparición de un sustrato enzimático. Por ejemplo, se puede utilizar un ensayo enzimático para determinar la actividad del polipéptido como se describe en este caso.

40

[0158] El polipéptido resultante se puede recuperar usando métodos conocidos en la técnica. Por ejemplo, el polipéptido se puede recuperar del medio nutritivo mediante procedimientos convencionales que incluyen, pero no de forma limitante a, centrifugado, filtración, extracción, secado por pulverización, evaporación, o precipitación.

45

[0159] Los polipéptidos de la presente invención se pueden purificar mediante diferentes procedimientos conocidos en la técnica que incluyen, pero no se limitan a, cromatografía (por ejemplo, de intercambio iónico, de afinidad, de cromatoenfoco hidrofóbico, y de exclusión de tamaño), procedimientos electroforéticos (por ejemplo, isoelectroenfoco preparativo), solubilidad diferencial (por ejemplo, precipitación de sulfato amónico), SDS-PAGE, o extracción (véase, por ejemplo, Protein Purification, J.-C. Janson y Lars Riden, editores, VCH Publishers, Nueva York, 1989) para obtener polipéptidos sustancialmente puros.

50

Composiciones

55

[0160] También se describen composiciones que comprenden un polipéptido de la presente invención. Preferiblemente, las composiciones están enriquecidas con tal polipéptido. El término "enriquecido" indica que la actividad de aumento celulolítico de la composición ha sido aumentado, por ejemplo, con un factor de enriquecimiento de al menos 1.1.

60

[0161] La composición puede comprender un polipéptido de la presente invención como el mayor componente enzimático, por ejemplo, una composición monocomponente. Alternativamente, la composición puede comprender

- múltiples actividades enzimáticas, tales como una aminopeptidasa, amilasa, carbohidrasa, carboxipeptidasa, catalasa, celulasa, quitinasa, cutinasa, ciclodextrin glicosiltransferasa, desoxiribonucleasa, esterasa, alfa-galactosidasa, beta-galactosidasa, glucoamilasa, alfa-glucosidasa, beta-glucosidasa, haloperoxidasa, invertasa, lacasa, lipasa, manosidasa, oxidasa, enzima pectinolítica, peptidoglutaminasa, peroxidasa, fitasa, polifenoloxidasa, enzima proteolítica, ribonucleasa, transglutaminasa, o xilanasas. La(s) enzima(s) adicional(es) pueden ser producidas, por ejemplo, mediante un microorganismo del género *Aspergillus*, preferiblemente *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, o *Aspergillus oryzae*; *Fusarium*, preferiblemente *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sulphureum*, *Fusarium toruloseum*, *Fusarium trichothecioides*, o *Fusarium venenatum*; *Humicola*, preferiblemente *Humicola insolens* o *Humicola lanuginosa*; o *Trichoderma*, preferiblemente *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.
- 5
- 10
- 15 [0162] Las composiciones de polipéptido se pueden preparar conforme a métodos conocidos en la técnica y pueden ser en forma de un líquido o una composición seca. Por ejemplo, la composición de polipéptido puede ser en forma de granulado o microgranulado. El polipéptido por incluir en la composición se puede estabilizar conforme a métodos conocidos en la técnica.
- 20 [0163] A continuación se dan ejemplos de usos preferidos de las composiciones de polipéptido de la invención. La dosificación de la composición de polipéptido de la invención y otras condiciones bajo las que se usa la composición se pueden determinar en base a métodos conocidos en la técnica.
- 25 **Procesamiento de material celulósico**
- 30 [0164] La presente invención también se refiere a métodos para la degradación o conversión de un material celulósico, que comprende: tratar el material celulósico con una composición enzimática en presencia de un polipéptido de la presente invención con actividad de aumento celulolítico. En un aspecto preferido, el método comprende además la recuperación del material celulósico degradado o convertido.
- 35 [0165] La presente invención también se refiere a métodos de producción de un producto de fermentación, que comprenden: (a) sacarificación de un material celulósico con una composición enzimática en presencia de un polipéptido de la presente invención con actividad de aumento celulolítico; (b) fermentación del material celulósico sacarificado con uno o más microorganismos de fermentación (diferentes) para producir el producto de fermentación; y (c) recuperación del producto de fermentación de la fermentación.
- [0166] La presente invención también se refiere a métodos de fermentación de un material celulósico, que comprende: fermentación del material celulósico con uno o varios microorganismos de fermentación (diferentes), donde el material celulósico es sacarificado con una composición enzimática en presencia de un polipéptido de la presente invención con actividad de aumento celulolítico. En un aspecto preferido, la fermentación del material celulósico produce un producto de fermentación. En otro aspecto preferido, el método comprende además la recuperación del producto de fermentación de la fermentación.
- 40
- 45 [0167] El polipéptido con actividad de aumento celulolítico puede encontrarse en forma de caldo de fermentación crudo con o sin eliminar las células o en forma de una preparación enzimática semi-purificada o purificada o la composición puede comprender una célula huésped de la presente invención como fuente del polipéptido con actividad de aumento celulolítico en un proceso de fermentación con la biomasa.
- 50 [0168] Los métodos de la presente invención pueden utilizarse para sacarificar un material celulósico en azúcares fermentables y convertir los azúcares fermentables en muchas sustancias útiles, por ejemplo, productos químicos y combustibles. La producción de un producto de fermentación deseado a partir de material celulósico implica típicamente pretratamiento, hidrólisis enzimática (sacarificación), y fermentación.
- 55 [0169] El procesamiento de material celulósico según la presente invención se puede realizar usando procesos convencionales en la técnica. Además, los métodos de la presente invención se pueden implementar utilizando cualquier equipo de tratamiento de biomasa convencional configurado para operar conforme a la invención.
- 60 [0170] La hidrólisis (sacarificación) y fermentación, separadas o simultáneas, incluyen, pero de forma no limitativa, hidrólisis y fermentación separadas (SHF por sus siglas en inglés); sacarificación y fermentación simultáneas (SSF por sus siglas en inglés); sacarificación y cofermentación simultáneas (SSCF por sus siglas en inglés); hidrólisis y fermentación híbridas (HHF por sus siglas en inglés); hidrólisis y co-fermentación separadas (SHCF por sus siglas en

inglés), hidrólisis y fermentación híbridas (HHCF por sus siglas en inglés), y conversión microbiana directa (DMC por sus siglas en inglés). La SHF usa pasos de proceso separados para primero hidrolizar enzimáticamente lignocelulosa en azúcares fermentables, por ejemplo, glucosa, celobiosa, celotriosa, y azúcares de pentosa, y luego fermentar los azúcares fermentables en etanol. En la SSF, la hidrólisis enzimática de lignocelulosa y la fermentación de azúcares en etanol se combinan en un paso (Philippidis, G. P., 1996, Cellulose bioconversion technology, en Handbook on Bioethanol: Production and Utilization, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212). La SSCF implica la cofermentación de múltiples azúcares (Sheehan, J., y Himmel, M., 1999, Enzymes, energy and the environment: A strategic perspective on the U.S. Department of Energy's research and development activities for bioethanol, Biotechnol. Prog. 15: 817-827). La HHF implica un paso de hidrólisis separada, y además un paso de sacarificación e hidrólisis simultáneas, que pueden efectuarse en el mismo reactor. Los pasos en un proceso HHF puede llevarse a cabo a diferentes temperaturas, es decir, la sacarificación enzimática a alta temperatura seguida de la SSF a una temperatura inferior que la tensión de fermentación pueda tolerar. La DMC combina los tres procesos (producción enzimática, hidrólisis de lignocelulosa, y fermentación) en uno o más pasos donde se utiliza el mismo organismo para producir las enzimas para la conversión de la lignocelulosa en azúcares fermentables y para convertir los azúcares fermentables en un producto final (Lynd, L. R., Weimer, P. J., van Zyl, W. H., y Pretorius, I. S., 2002, Microbial cellulose utilization: Fundamentals and biotechnology, Microbiol. Mol. Biol. Reviews 66: 506-577). Se entiende por la presente que cualquier método conocido en la técnica que comprende pretratamiento, hidrólisis enzimática (sacarificación), fermentación, o una combinación de las mismas se puede usar en la práctica de los métodos de la presente invención.

[0171] Un equipo convencional puede incluir un reactor agitado de lote alimentado, un reactor agitado de lote, un reactor agitado de flujo continuo con ultrafiltración, y/o un reactor de columna de flujo de pistón continuo (Fernanda de Castilhos Corazza, Flávio Faria de Moraes, Gisella Maria Zanin e Ivo Neitzel, 2003, Optimal control in fed-batch reactor for the cellobiose hydrolysis, Acta Scientiarum. Technology 25: 33-38; Gusakov, A. V., y Sinitsyn, A. P., 1985, Kinetics of the enzymatic hydrolysis of cellulose: 1. A mathematical model for a batch reactor process, Enz. Microb. Technol. 7: 346-352), un reactor de fricción (Ryu, S. K., y Lee, J. M., 1983, Bioconversion of waste cellulose by using an attrition bioreactor, Biotechnol. Bioeng. 25: 53-65), o un reactor con agitado intensivo inducido por un campo electromagnético (Gusakov, A. V., Sinitsyn, A. P., Davydkin, I. Y., Davydkin, V. Y., Protas, O. V., 1996, Enhancement of enzymatic cellulose hydrolysis using a novel type of bioreactor with intensive stirring induced by electromagnetic field, Appl. Biochem. Biotechnol. 56: 141-153). Otros tipos de reactor adicionales incluyen: reactores de lecho fluidificado, manto de flujo ascendente, inmovilizado, y de tipo extrusor para la hidrólisis y/o fermentación.

[0172] Pretratamiento. En la práctica de los métodos de la presente invención, cualquier proceso de pretratamiento conocido en la técnica puede utilizarse para interrumpir los componentes de pared celular vegetal (Chandra et al., 2007, Substrate pretreatment: The key to effective enzymatic hydrolysis of lignocellulosics? Adv. Biochem. Engin./Biotechnol. 108: 67-93; Galbe y Zacchi, 2007, Pretreatment of lignocellulosic materials for efficient bioethanol production, Adv. Biochem. Engin. / Biotechnol. 108: 41-65; Hendriks y Zeeman, 2009, Pretreatments to enhance the digestibility of lignocellulosic biomass, Bioresource Technol. 100: 10-18; Mosier et al., 2005, Features of promising technologies for pretreatment of lignocellulosic biomass, Bioresource Technol. 96: 673-686; Taherzadeh y Karimi, 2008, Pretreatment of lignocellulosic wastes to improve ethanol and biogas production: A review, Int. J. of Mol. Sci. 9: 1621-1651; Yang y Wyman, 2008, Pretreatment: the key to unlocking low-cost cellulosic ethanol, Biofuels Bioproducts and Biorefining-Biofpr. 2: 26-40).

[0173] El material celulósico también puede ser sometido a reducción de tamaño de partícula, pre-remojo, humidificación, lavado, o preparación antes del pretratamiento usando métodos conocidos en la técnica. Los pretratamientos convencionales incluyen, pero de forma no limitativa, pretratamiento con vapor (con o sin explosión), pretratamiento con ácido diluido, pretratamiento con agua caliente, pretratamiento alcalino, pretratamiento con cal, oxidación en húmedo, explosión en húmedo, explosión en fibra de amoníaco, pretratamiento organosolv, y pretratamiento biológico. Pretratamientos adicionales incluyen pretratamientos con filtración de amoníaco, ultrasonido, electroporación, microondas, CO₂ supercrítico, H₂O supercrítico, ozono, y con irradiación gamma.

[0174] El material celulósico se puede pretratar antes de la hidrólisis y/o fermentación. El pretratamiento es realizado preferiblemente antes de la hidrólisis. Alternativamente, el pretratamiento puede llevarse a cabo simultáneamente con la hidrólisis, así como simultáneamente con el tratamiento del material celulósico con una composición enzimática de la presente invención para liberar azúcares fermentables, tales como glucosa, xilosa, y/o celobiosa. En la mayoría de los casos el paso del pretratamiento mismo produce alguna conversión de biomasa para azúcares fermentables (incluso en ausencia de enzimas).

[0175] Pretratamiento con vapor. En el pretratamiento con vapor, el material celulósico se calienta para alterar los componentes de la pared celular vegetal, incluyendo lignina, hemicelulosa, y celulosa para hacer la celulosa y otras fracciones, por ejemplo, hemicelulosa, accesibles a enzimas. El material de lignocelulosa pasa a o a través de un recipiente de reacción donde se inyecta vapor para aumentar la temperatura hasta la temperatura y presión necesarias y

ES 2 551 141 T3

- se retiene en ellas durante el tiempo de reacción deseado. El pretratamiento con vapor se realiza preferiblemente a 140-230°C, de forma más preferible 160-200°C, y de la forma más preferible 170-190°C, donde el rango de temperatura óptima depende del cualquier adición de un catalizador químico. El periodo de permanencia para el pretratamiento con vapor es preferiblemente 1-15 minutos, de forma más preferible 3-12 minutos, y de la forma más preferible 4-10 minutos, donde el periodo de permanencia óptimo depende del rango de temperatura y cualquier adición de un catalizador químico. El pretratamiento con vapor tiene en cuenta cargas de sólidos relativamente altas, de modo que el material celulósico generalmente es solo húmedo durante el pretratamiento. El pretratamiento con vapor se combina frecuentemente con una descarga explosiva del material después del pretratamiento, lo que se conoce como explosión de vapor, que un sometimiento rápido intermitente a la presión atmosférica y un flujo turbulento del material para aumentar el área de superficie accesible por fragmentación (Duff y Murray, 1996, *Bioresource Technology* 855: 1-33; Galbe y Zacchi, 2002, *Appl. Microbiol. Biotechnol.* 59: 618-628; solicitud de patente estadounidense nº 20020164730). Durante el pretratamiento con vapor, los grupos acetilo de la hemicelulosa se escinden y el ácido resultante autocataliza la hidrólisis parcial de la hemicelulosa en monosacáridos y oligosacáridos. La lignina se elimina en sólo cierta medida.
- 5
- 10
- 15 [0176] Frecuentemente se añade un catalizador tal como H₂SO₄ o SO₂ (típicamente 0,3 a 3% p/p) antes del pretratamiento con vapor, que reduce el tiempo y la temperatura, aumenta la recuperación, y mejora la hidrólisis enzimática (Ballesteros et al., 2006, *Appl. Biochem. Biotechnol.* 129-132: 496- 508; Varga et al., 2004, *Appl. Biochem. Biotechnol.* 113-116: 509-523; Sassner et al., 2006, *Enzyme Microb. Technol.* 39: 756-762).
- 20 [0177] Pretratamiento químico: el término "tratamiento químico" se refiere a cualquier pretratamiento químico que promueva la separación y/o liberación de celulosa, hemicelulosa, y/o lignina. Ejemplos de procesos de pretratamiento químico adecuados incluyen, por ejemplo, pretratamiento con ácido diluido, pretratamiento con cal, oxidación en húmedo, explosión por congelación/fibra de amoníaco (AFEX por sus siglas en inglés), filtración de amoníaco (APR por sus siglas en inglés), y pretratamientos organosolv.
- 25
- 30 [0178] En el pretratamiento con ácido diluido, el material celulósico es mezclado con ácido diluido, típicamente H₂SO₄, y agua para formar un lodo, calentado al vapor hasta la temperatura deseada, y después de un periodo de permanencia es sometido rápidamente a presión atmosférica. El pretratamiento con ácido diluido se puede realizar con diferentes diseños de reactor, por ejemplo, reactores de flujo de pistón, reactores de contracorriente, o reactores de lecho de encogimiento de contracorriente continua (Duff y Murray, 1996, *supra*; Schell et al., 2004, *Bioresource Technol.* 91: 179-188; Lee et al., 1999, *Adv. Biochem. Eng. Biotechnol.* 65: 93-115).
- 35 [0179] También pueden usarse diferentes métodos de pretratamiento bajo condiciones alcalinas. Estos pretratamientos alcalinos incluyen, pero de forma no limitativa, pretratamiento con cal, oxidación en húmedo, filtración de amoníaco (APR), y explosión por congelación / fibra de amoníaco (AFEX).
- 40 [0180] El pretratamiento con cal se realiza con carbonato cálcico, hidróxido sódico, o amoníaco a bajas temperaturas de 85-150°C y tiempos de estancia desde 1 hora a varios días (Wyman et al., 2005, *Bioresource Technol.* 96: 1959-1966; Mosier et al., 2005, *Bioresource Technol.* 96: 673-686). Los documentos WO 2006/110891, WO 2006/11899, WO 2006/11900, y WO 2006/110901 divulgan métodos de pretratamiento que usan amoníaco.
- 45 [0181] La oxidación en húmedo es un pretratamiento térmico realizado típicamente a 180-200°C durante 5-15 minutos con la adición de un agente oxidante tal como peróxido de hidrógeno o sobrepresión de oxígeno (Schmidt y Thomsen, 1998, *Bioresource Technol.* 64: 139-151; Palonen et al., 2004, *Appl. Biochem. Biotechnol.* 117: 1-17; Varga et al., 2004, *Biotechnol. Bioeng.* 88: 567-574; Martin et al., 2006, *J. Chem. Technol. Biotechnol.* 81: 1669-1677). El pretratamiento se realiza a preferiblemente 1-40% sustancia seca, de forma más preferible 2-30% sustancia seca, y de la forma más preferible 5-20% sustancia seca, y frecuentemente el pH inicial es aumentado por la adición de álcalis tal como carbonato de sodio.
- 50 [0182] Una modificación del método de pretratamiento de oxidación en húmedo, conocido como explosión en húmedo (combinación de oxidación en húmedo y explosión de vapor), puede tratar sustancia seca hasta 30%. En la explosión en húmedo, el agente oxidante es introducido durante el pretratamiento después de un determinado tiempo de estancia. El pretratamiento es luego finalizado mediante rápido sometimiento a presión atmosférica (WO 2006/032282).
- 55 [0183] La explosión de fibra de amoníaco (AFEX) implica el tratamiento de material celulósico con amoníaco líquido o gaseoso a temperaturas moderadas tales como 90-100°C y alta presión tal como 17-20 bares durante 5-10 minutos, donde el contenido de sustancia en seco puede ser tan alto como 60% (Gollapalli et al., 2002, *Appl. Biochem. Biotechnol.* 98: 23-35; Chundawat et al., 2007, *Biotechnol. Bioeng.* 96: 219-231; Alizadeh et al., 2005, *Appl. Biochem. Biotechnol.* 121:1133-1141; Teymouri et al., 2005, *Bioresource Technol.* 96: 2014-2018). El pretratamiento AFEX produce la despolimerización de celulosa e hidrólisis parcial de hemicelulosa. Los complejos de lignina-carbohidrato son escindidos.
- 60

ES 2 551 141 T3

- 5 [0184] Pretratamiento organosolv delignifica el material celulósico mediante extracción usando etanol acuoso (40-60% etanol) a 160-200°C durante 30-60 minutos (Pan et al., 2005, *Biotechnol. Bioeng.* 90: 473-481; Pan et al., 2006, *Biotechnol. Bioeng.* 94: 851-861; Kurabi et al., 2005, *Appl. Biochem. Biotechnol.* 121:219-230). El ácido sulfúrico se añade normalmente como catalizador. En el pretratamiento organosolv, la mayoría de la hemicelulosa es eliminada.
- [0185] Otros ejemplos de métodos de pretratamiento adecuados son descritos por Schell et al., 2003, *Appl. Biochem. y Biotechnol.* Vol. 105-108, p. 69-85, y Mosier et al., 2005, *Bioresource Technology* 96: 673-686, y en la solicitud estadounidense publicada 2002/0164730.
- 10 [0186] En un aspecto, el pretratamiento químico se realiza preferiblemente como un tratamiento ácido, y de forma más preferible como un tratamiento de ácido diluido continuo y/o moderado. El ácido es típicamente ácido sulfúrico, pero también pueden usarse otros ácidos, tales como ácido acético, ácido cítrico, ácido nítrico, ácido fosfórico, ácido tartárico, ácido succínico, cloruro de hidrógeno o mezclas derivadas. El tratamiento con ácido moderado se lleva a cabo en el margen de pH de preferiblemente 1-5, de forma más preferible 1-4, y de la forma más preferible 1-3. En un aspecto, la concentración ácida se encuentra en el rango de preferiblemente 0,01 a 20% en peso de ácido, de forma más preferible de 0,05 a 10% en peso de ácido, incluso de forma más preferible 0,1 a 5 % en peso de ácido, y de la forma más preferible 0,2 a 2,0 % en peso de ácido. El ácido se pone en contacto con el material celulósico y se mantiene a una temperatura en el rango de preferiblemente 160-220°C, y de forma más preferible 165-195°C, durante periodos que varían de segundos a minutos a, por ejemplo, de 1 segundo a 60 minutos.
- 15 [0187] El pretratamiento se puede realizar como un paso de explosión de fibra de amoníaco (paso de pretratamiento AFEX).
- [0188] El pretratamiento puede ocurrir en un lodo acuoso. El material celulósico está presente durante el pretratamiento en cantidades preferiblemente entre 10-80% en peso, de forma más preferible entre 20-70% en peso, y de la forma más preferible entre 30-60% en peso, tal como alrededor de 50% en peso. El material celulósico pretratado puede ser lavado o no lavado utilizando cualquier método conocido en la técnica, por ejemplo, lavado con agua.
- 20 [0189] Pretratamiento mecánico: el término "pretratamiento mecánico" se refiere a tipos varios de trituración o fresado (por ejemplo, molienda en seco, molienda en húmedo, o fresado de bola vibratoria).
- [0190] Pretratamiento físico: el término "pretratamiento físico" se refiere a cualquier pretratamiento que promueve la separación y/o liberación de celulosa, hemicelulosa, y/o lignina de material celulósico. Por ejemplo, el pretratamiento físico puede implicar irradiación (por ejemplo, irradiación de microondas), explosión vaporizante/de vapor, hidrotermólisis, y combinaciones de las mismas.
- 25 [0191] El pretratamiento físico puede implicar alta presión y/o alta temperatura (explosión de vapor). En un aspecto, alta presión significa presión en el rango de preferiblemente aproximadamente 300 a aproximadamente 600 psi, de forma más preferible de aproximadamente 350 a aproximadamente 550 psi, y de la forma más preferible de aproximadamente 400 a aproximadamente 500 psi, tal como alrededor de 450 psi. En otro aspecto, temperaturas altas significan temperaturas de aproximadamente 100 a aproximadamente 300°C, preferiblemente de aproximadamente 140 a aproximadamente 235°C. En un aspecto preferido, el pretratamiento mecánico se realiza en un proceso de lote, sistema de hidrolizador de pistola de vapor que usa alta presión y alta temperatura tal como se ha definido anteriormente, por ejemplo, un hidrolizador Sunds disponible de Sunds Defibrator AB, Suecia.
- 30 [0192] Pretratamiento físico y químico combinado: el material celulósico se puede pretratar tanto físicamente como químicamente. Por ejemplo, el paso del pretratamiento puede implicar tratamiento con ácido diluido o moderado y tratamiento de alta temperatura y/o presión. Los pretratamientos físicos y químicos pueden llevarse a cabo consecutivamente o simultáneamente, como se desee. Se puede incluir también un pretratamiento mecánico.
- 35 [0193] El material celulósico se puede someter a pretratamiento mecánico, químico o físico, o a cualquier combinación de los mismos para promover la separación y/o liberación de celulosa, hemicelulosa y/o lignina.
- [0194] Pretratamiento biológico: el término "pretratamiento biológico" se refiere a cualquier pretratamiento biológico que promueve la separación y/o liberación de celulosa, hemicelulosa, y/o lignina del material celulósico. Las técnicas de pretratamiento biológico pueden implicar la aplicación microorganismos de solubilización de lignina (véase, por ejemplo, Hsu, T.-A., 1996, *Pretreatment of biomass*, en *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212; Ghosh y Singh, 1993, *Physicochemical and biological treatments for enzymatic/microbial conversion of cellulosic biomass*, *Adv. Appl. Microbiol.* 39: 295-333; McMillan, J. D., 1994, *Pretreating lignocellulosic biomass: a review*, en *Enzymatic Conversion of Biomass for Fuels Production*, Himmel, M. E., Baker, J. O., y Overend, R. P., eds., ACS Symposium Series 566, American Chemical Society, Washington, DC,
- 40
- 45
- 50
- 55
- 60

ES 2 551 141 T3

capítulo 15; Gong, C. S., Cao, N. J., Du, J., y Tsao, G. T., 1999, Ethanol production from renewable resources, en *Advances in Biochemical Engineering/Biotechnology*, Scheper, T., ed., Springer-Verlag Berlín Heidelberg, Alemania, 65: 207- 241; Olsson y Hahn-Hagerdal, 1996, Fermentation of lignocellulosic hydrolysates for ethanol production, *Enz. Microb. Tech.* 18: 312-331; y Vallander y Eriksson, 1990, Production of ethanol from lignocellulosic materials: State of the art, *Adv. Biochem. Eng./Biotechnol.* 42: 63-95).

[0195] Sacarificación. En el paso de hidrólisis, también conocido como sacarificación, el material celulósico, por ejemplo, pretratado, es hidrolizado para descomponer celulosa y alternativamente también hemicelulosa en azúcares fermentables, tales como glucosa, celobiosa, xilosa, xilulosa, arabinosa, manosa, galactosa, u oligosacáridos solubles. La hidrólisis se realiza enzimáticamente mediante una composición enzimática en presencia de un polipéptido con actividad de aumento celulolítico de la presente invención. La composición puede además comprender una o varias enzimas hemicelulolíticas (diferentes). Las enzimas de las composiciones también pueden ser añadidas consecutivamente.

[0196] La hidrólisis enzimática se realiza preferiblemente en un entorno acuoso adecuado bajo condiciones que pueden ser fácilmente determinadas por un experto en la técnica. En un aspecto preferido, la hidrólisis se realiza bajo condiciones adecuadas para la actividad de la(s) enzima(s), es decir, óptima para la(s) enzima(s). La hidrólisis puede llevarse a cabo como un proceso de lote alimentado o continuo donde el material celulósico pretratado (sustrato) es alimentado gradualmente hasta obtener, por ejemplo, una solución de la hidrólisis que contiene enzimas.

[0197] La sacarificación se realiza generalmente en los reactores o fermentadores de tanque agitado bajo condiciones de pH, temperatura y mezcla controladas. El tiempo, temperatura y condiciones de pH de proceso adecuadas pueden ser determinados fácilmente por un experto en la técnica. Por ejemplo, la sacarificación puede durar hasta 200 horas, pero se realiza típicamente durante preferiblemente aproximadamente 12 a aproximadamente 168 horas, de forma más preferible de aproximadamente 24 a aproximadamente 120 horas, y de la forma más preferible de aproximadamente 48 a aproximadamente 72 horas. La temperatura se encuentra en el rango de preferiblemente aproximadamente 40°C a aproximadamente 70°C, de forma más preferible aproximadamente 45°C a aproximadamente 65°C, y de forma más preferible aproximadamente 50°C a 60°C, en particular aproximadamente 55°C. El pH se encuentra en el rango de preferiblemente aproximadamente 3 a aproximadamente 9, de forma más preferible aproximadamente 3,5 a aproximadamente 8, de forma más preferible aproximadamente 4 a aproximadamente 7, y de la forma más preferible aproximadamente 4,5 a aproximadamente 6, en particular acerca de pH 5. El contenido en sustancias secas se encuentra en el rango de preferiblemente aproximadamente 1 a aproximadamente 50% en peso, de forma más preferible aproximadamente 5 a aproximadamente 40% en peso, de forma más preferible aproximadamente 10 a aproximadamente 30% en peso, y de la forma más preferible aproximadamente 15 a aproximadamente 25% en peso.

[0198] Además de un polipéptido con actividad de aumento celulolítico de la presente invención, los componentes enzimáticos celulolíticos de la composición son preferiblemente enzimas con actividades de endoglucanasa, celobiohidrolasa y beta-glucosidasa. En un aspecto preferido, la composición enzimática comprende una o varias enzimas celulolíticas (diferentes) seleccionadas del grupo consistente en una celulasa, una endoglucanasa, una celobiohidrolasa y una beta-glucosidasa. En otro aspecto preferido, la preparación enzimática celulolítica comprende además o incluso además una o varias actividades enzimáticas adicionales (diferentes) seleccionadas del grupo consistente en hemicelulasas, esterasas (por ejemplo, lipasas, fosfolipasas, y/o cutinasas), proteasas, lacasas, peroxidadas, o mezclas derivadas. En los métodos de la presente invención, la(s) enzima(s) adicional(es) pueden añadirse antes de o durante la fermentación, incluyendo durante o después de la propagación del (de los) microorganismo(s) de fermentación.

[0199] Las enzimas pueden ser derivadas u obtenidas de cualquier origen adecuado, incluyendo origen bacteriano, fúngico, de levadura, planta o mamífero. El término "obtenido" se refiere por la presente a que la enzima puede haber sido aislada a partir de un organismo que produce naturalmente la enzima como una enzima nativa. El término "obtenido" significa también por la presente que la enzima puede haber sido producida recombinantemente en un organismo huésped utilizando métodos descritos por la presente, donde la enzima producida recombinantemente es o bien nativa o bien foránea al organismo huésped o tiene una secuencia de aminoácidos modificada, por ejemplo, con uno o varios aminoácidos (diferentes) que son eliminados, insertados y/o sustituidos, es decir, una enzima producida recombinantemente que es un mutante y/o un fragmento de una secuencia de aminoácidos nativa o una enzima producida mediante procesos de redistribución de ácido nucleico conocidos en la técnica. Incluidas dentro del significado de enzima nativa se encuentran variantes naturales y dentro del significado de una enzima foránea se encuentran variantes obtenidas recombinantemente, tales como por mutagénesis dirigida al sitio o redistribución.

[0200] Las enzimas usadas pueden encontrarse en cualquier forma adecuada para su uso en los métodos descritos por la presente, tales como un caldo de fermentación crudo con o sin células o polipéptidos sustancialmente puros. La(s) enzima(s) puede ser un polvo seco o granulado, un líquido, un líquido estabilizado, o una(s) enzima(s) protegida(s). Las

ES 2 551 141 T3

preparaciones enzimáticas líquidas pueden, por ejemplo, ser estabilizadas añadiendo estabilizadores tales como un azúcar, un alcohol de azúcar u otro poliol, y/o ácido láctico u otro ácido orgánico según un proceso establecido.

5 [0201] Las cantidades óptimas de las enzimas y polipéptidos con actividad de aumento celulolítico dependen de diferentes factores incluyendo, pero no limitándose a, la mezcla de enzimas celulolíticas componentes, el sustrato celulósico, la concentración de sustrato celulósico, el (los) pretratamiento(s) del sustrato celulósico, temperatura, tiempo, pH, e inclusión de organismo fermentador (por ejemplo, levadura para sacarificación y fermentación simultáneas).

10 [0202] Una cantidad eficaz de enzima(s) celulolítica(s) respecto al material celulósico es aproximadamente 0,5 a aproximadamente 50 mg, preferiblemente alrededor de 0,5 a aproximadamente 40 mg, de forma más preferible alrededor de 0,5 a aproximadamente 25 mg, de forma más preferible alrededor de 0,75 a aproximadamente 20 mg, de forma más preferible alrededor de 0,75 a aproximadamente 15 mg, incluso de forma más preferible alrededor de 0,5 a aproximadamente 10 mg, y de la forma más preferible alrededor de 2,5 a aproximadamente 10 mg por g de material celulósico.

15 [0203] Una cantidad eficaz de polipéptido(s) con actividad de aumento celulolítico respecto al material celulósico es aproximadamente 0,01 a aproximadamente 50,0 mg, preferiblemente aproximadamente 0,01 a aproximadamente 40 mg, de forma más preferible aproximadamente 0,01 a aproximadamente 30 mg, de forma más preferible aproximadamente 0,01 a aproximadamente 20 mg, de forma más preferible aproximadamente 0,01 a aproximadamente 10 mg, de forma más preferible aproximadamente 0,01 a aproximadamente 5 mg, de forma más preferible en alrededor de 0,025 a aproximadamente 1,5 mg, de forma más preferible en alrededor de 0,05 a aproximadamente 1,25 mg, de forma más preferible en alrededor de 0,075 a aproximadamente 1,25 mg, de forma más preferible en alrededor de 0,1 a aproximadamente 1,25 mg, incluso de forma más preferible en alrededor de 0,15 a aproximadamente 1,25 mg, y de la forma más preferible en alrededor de 0,25 a aproximadamente 1,0 mg por g de material celulósico.

20 [0204] Una cantidad efectiva de polipéptido(s) con actividad de aumento celulolítico respecto a enzima(s) celulolítica(s) es aproximadamente 0,005 a aproximadamente 1,0 g, preferiblemente alrededor de 0,01 a aproximadamente 1,0 g, de forma más preferible alrededor de 0,15 a aproximadamente 0,75 g, de forma más preferible alrededor de 0,15 a aproximadamente 0,5 g, de forma más preferible alrededor de 0,1 a aproximadamente 0,5 g, incluso de forma más preferible alrededor de 0,1 a aproximadamente 0,5 g, y de la forma más preferible alrededor de 0,05 a aproximadamente 0,2 g por g de enzima(s) celulolítica(s).

25 [0205] En los métodos de la presente invención, la composición enzimática puede comprender cualquier proteína implicada en el procesamiento de un material con celulosa en glucosa, o hemicelulosa en xilosa, manosa, galactosa, y arabinosa, sus polímeros, o productos derivados de éstos como se describe posteriormente. En un aspecto, la composición enzimática comprende una o varias enzimas (diferentes) seleccionadas del grupo consistente en una endoglucanasa, una celobiohidrolasa, y una beta-glucosidasa. En otro aspecto, la composición enzimática además o incluso además comprende una o varias actividades enzimáticas (diferentes) adicionales para mejorar la degradación del material con celulosa. Las enzimas adicionales preferidas son xilanasas, hemicelulasas, esterases (por ejemplo, lipasas, fosfolipasas, y/o cutinasas), proteasas, lacasas, peroxidases, o mezclas derivadas.

30 [0206] La composición enzimática puede ser una preparación monocomponente, por ejemplo, una endoglucanasa, una preparación con varios componentes, por ejemplo, endoglucanasa, celobiohidrolasa, beta-glucosidasa, o una combinación de preparaciones de proteína multicomponente y monocomponente. Las proteínas celulolíticas pueden tener actividad, es decir, hidrolizan celulosa, bien en el rango de pH ácido, neutral, o alcalino.

35 [0207] Uno o varios componentes (diferentes) de la composición enzimática pueden ser un componente recombinante, es decir, producido mediante clonación de una secuencia de ADN que codifica el único componente y la célula posterior transformada con la secuencia de ADN y expresada en un huésped (véase, por ejemplo, los documentos WO 91/17243 y WO 91/17244). El huésped es preferiblemente un huésped heterólogo (la enzima es foránea al huésped), pero el huésped también puede ser bajo ciertas condiciones un huésped homólogo (la enzima es nativa al huésped). Las proteínas celulolíticas monocomponentes también se pueden preparar mediante la purificación de tal proteína a partir de un caldo de fermentación.

40 [0208] Las enzimas usadas en la presente invención pueden encontrarse en cualquier forma adecuada para su uso en los procesos descritos por la presente, tal como, por ejemplo, un caldo de fermentación crudo con o sin células, un polvo o granulado seco, un líquido, un líquido estabilizado, o una enzima protegida. Las preparaciones enzimáticas líquidas pueden, por ejemplo, ser estabilizadas añadiendo estabilizadores tales como un azúcar, un alcohol de azúcar u otro poliol, y/o ácido láctico u otro ácido orgánico según el proceso establecido.

45 [0209] Un polipéptido con actividad enzimática celulolítica puede ser un polipéptido bacteriano. Por ejemplo, el

ES 2 551 141 T3

- 5 polipéptido puede ser un polipéptido bacteriano Gram positivo tal como un polipéptido de *Bacillus*, *Streptococcus*, *Streptomyces*, *Staphylococcus*, *Enterococcus*, *Lactobacillus*, *Lactococcus*, *Clostridium*, *Geobacillus*, u *Oceanobacillus* con actividad enzimática celulolítica, o un polipéptido bacteriano Gram negativo tal como un polipéptido de *E. coli*, *Pseudomonas*, *Salmonella*, *Campylobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Ilyobacter*, *Neisseria*, o *Ureaplasma* con actividad enzimática celulolítica.
- 10 [0210] El polipéptido puede ser un polipéptido de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, o *Bacillus thuringiensis* con actividad enzimática celulolítica.
- 15 [0211] El polipéptido puede ser un polipéptido de *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, o *Streptococcus equi* subesp. *zooepidemicus* con actividad enzimática celulolítica.
- [0212] El polipéptido puede ser un polipéptido de *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, o *Streptomyces lividans* con actividad enzimática celulolítica.
- 20 [0213] El polipéptido con actividad enzimática celulolítica también puede ser un polipéptido fúngico, y de forma más preferible un polipéptido de levadura tal como un polipéptido de *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia* con actividad de enzima celulolítica; o de forma más preferible un polipéptido fúngico filamentoso tal como un *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Aureobasidium*, *Botryosphaeria*, *Ceriporiopsis*, *Chaetomidium*, *Chrysosporium*, *Claviceps*, *Cochliobolus*, *Coprinopsis*, *Coptotermes*, *Corynascus*, *Cryphonectria*, *Cryptococcus*, *Diplodia*, *Exidia*, *Filibasidium*, *Fusarium*, *Gibberella*, *Holomastigotoides*, *Humicola*, *Irpex*, *Lentinula*, *Leptosphaeria*, *Magnaporthe*, *Melanocarpus*, *Meripilus*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Piromyces*, *Poitrasia*, *Pseudoplectania*, *Pseudotriconympha*, *Rhizomucor*, *Schizophyllum*, *Scytalidium*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolyopcladium*, *Trichoderma*, *Trichophaea*, *Verticillium*, *Volvariella*, o *Xylaria* con actividad enzimática celulolítica.
- 25 [0214] El polipéptido puede ser un polipéptido de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, o *Saccharomyces oviformis* con actividad enzimática celulolítica.
- 30 [0215] El polipéptido puede ser un polipéptido de *Acremonium cellulolyticus*, *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, *Aspergillus oryzae*, *Chrysosporium keratinophilum*, *Chrysosporium lucknowense*, *Chrysosporium tropicum*, *Chrysosporium merdarium*, *Chrysosporium inops*, *Chrysosporium pannicola*, *Chrysosporium queenslandicum*, *Chrysosporium zonatum*, *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, *Fusarium venenatum*, *Humicola grisea*, *Humicola insolens*, *Humicola lanuginosa*, *Irpex lacteus*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium funiculosum*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Thielavia achromatica*, *Thielavia albomyces*, *Thielavia albopilosa*, *Thielavia australeinsis*, *Thielavia fimeti*, *Thielavia microspora*, *Thielavia ovispora*, *Thielavia peruviana*, *Thielavia spededonium*, *Thielavia setosa*, *Thielavia subthermophila*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, *Trichoderma viride*, o *Trichophaea saccata* con actividad enzimática celulolítica.
- 35 [0216] También pueden usarse mutantes químicamente modificados o creados genéticamente de proteína de proteínas celulolíticas.
- 40 [0217] Ejemplos de preparaciones de proteína celulolítica comerciales adecuados para su uso en la presente invención incluyen, por ejemplo, CELLIC™ CTec (Novozymes A/S), CELLUCLAST™ (Novozymes A/S), y NOVOZYM™ 188 (Novozymes A/S). Otras preparaciones disponibles comercialmente que comprenden celulasa que se pueden utilizar incluyen CELLUZYME™, CEREFLO™ y ULTRAFLO™ (Novozymes A/S), LAMINEX™ y SPEZYME™ CP (Genencor Int.), ROHAMENT™ 7069 W (Röhm GmbH), y FIBREZYME® LDI, FIBREZYME® LBR, o VISCOSTAR® 150L (Dyadic International, Inc., Jupiter, FL, EE.UU). Las enzimas de celulasa son añadidas en cantidades efectivas desde aproximadamente 0,001% a aproximadamente 5,0% peso de sólidos, de forma más preferible de aproximadamente 0,025% a aproximadamente 4,0% peso de sólidos, y de la forma más preferible de aproximadamente 0,005% a aproximadamente 2,0% peso de sólidos.
- 45 [0218] Ejemplos de endoglucanasas bacterianas que se pueden usar en los métodos de la presente invención, incluyen,
- 50
- 55
- 60

ES 2 551 141 T3

pero de forma no limitativa, una endoglucanasa de *Acidothermus cellulolyticus* (WO 91/05039m WO 93/15186, patente estadounidense nº 5,275,944, WO 96/02551, patente estadounidense nº 5,536,655, WO 00/70031, WO 05/093050), endoglucanasa III de *Thermobifida fusca* (WO 05/093050), y endoglucanasa V de *Thermobifida fusca* (WO 05/093050).

5 [0219] Ejemplos de endoglucanasas fúngicas que se pueden usar en los métodos de la presente invención, incluyen, pero de forma no limitativa, una endoglucanasa I de *Trichoderma reesei* (Penttila et al., 1986, Gene 45: 253-263; número de acceso de GENBANK™ M15665); endoglucanasa II de *Trichoderma reesei* (Saloheimo, et al., 1988, Gene 63:11-22; número de acceso de GENBANK™ M19373); endoglucanasa III de *Trichoderma reesei* (Okada et al., 1988, Appl. Environ. Microbiol. 64: 555-563; número de acceso de GENBANK™ AB003694); endoglucanasa IV de *Trichoderma reesei* (Saloheimo et al., 1997, Eur. J. Biochem. 249: 584-591; número de acceso de GENBANK™ Y11113); y endoglucanasa V de *Trichoderma reesei* (Saloheimo et al., 1994, Molecular Microbiology 13: 219-228; número de acceso de GENBANK™ Z33381); endoglucanasa de *Aspergillus aculeatus* (Ooi et al., 1990, Nucleic Acids Research 18: 5884); endoglucanasa de *Aspergillus kawachii* (Sakamoto et al., 1995, Current Genetics 27: 435-439); endoglucanasa de *Erwinia carotovora* (Saarilahti et al., 1990, Gene 90: 9-14); endoglucanasa de *Fusarium oxysporum* (número de acceso de GENBANK™ L29381); endoglucanasa de *Humicola grisea* var. *thermoidea* (número de acceso de GENBANK™ AB003107); endoglucanasa de *Melanocarpus albomyces* (número de acceso de GENBANK™ MAL515703); endoglucanasa de *Neurospora crassa* (número de acceso de GENBANK™ XM 324477); endoglucanasa V de *Humicola insolens*; endoglucanasa de *Myceliophthora thermophila* CBS 117.65; endoglucanasa de basidiomiceto CBS 495.95; endoglucanasa de basidiomiceto CBS 494.95; endoglucanasa de *Thielavia terrestris* NRRL 8126 CEL6B; endoglucanasa de *Thielavia terrestris* NRRL 8126 CEL6C; endoglucanasa de *Thielavia terrestris* NRRL 8126 CEL7C; endoglucanasa de *Thielavia terrestris* NRRL 8126 CEL7E; endoglucanasa de *Thielavia terrestris* NRRL 8126 CEL7F; endoglucanasa de *Cladorrhinum foecundissimum* ATCC 62373 CEL7A; y endoglucanasa de cepa de *Trichoderma reesei* número VTT-D 80133 (número de acceso de GENBANK™ M15665).

25 [0220] Ejemplos de celobiohidrolasas útiles en los métodos de la presente invención incluyen, pero de forma no limitativa, celobiohidrolasa I de *Trichoderma reesei*; celobiohidrolasa II de *Trichoderma reesei*; celobiohidrolasa I de *Humicola insolens*, celobiohidrolasa II de *Myceliophthora thermophila*, celobiohidrolasa II de *Thielavia terrestris* (CEL6A), celobiohidrolasa I de *Thermophilum chaetomium*, y celobiohidrolasa II de *Chaetomium thermophilum*.

30 [0221] Ejemplos de beta-glucosidasas útiles en los métodos de la presente invención incluyen, pero de forma no limitativa, beta-glucosidasa de *Aspergillus oryzae*; beta-glucosidasa de *Aspergillus fumigatus*; beta-glucosidasa de *Penicillium brasilianum* IBT 20888; beta-glucosidasa de *Aspergillus niger*; y beta-glucosidasa de *Aspergillus aculeatus*.

35 [0222] El polipéptido de *Aspergillus oryzae* con actividad de beta-glucosidasa puede ser obtenido según el documento WO 2002/095014. El polipéptido de *Aspergillus fumigatus* con actividad de beta-glucosidasa puede ser obtenido según el documento WO 2005/047499. El polipéptido de *Penicillium brasilianum* con actividad de beta-glucosidasa puede ser obtenido según el documento WO 2007/019442. El polipéptido de *Aspergillus niger* con actividad de beta-glucosidasa puede ser obtenido según Dan et al., 2000, J. Biol. Chem. 275: 4973-4980. El polipéptido de *Aspergillus aculeatus* con actividad de beta-glucosidasa puede ser obtenido según Kawaguchi et al., 1996, Gene 173: 287-288.

40 [0223] La beta-glucosidasa puede ser una proteína de fusión. En un aspecto, la beta-glucosidasa es la proteína de fusión BG variante de beta-glucosidasa de *Aspergillus oryzae* o la proteína de fusión de beta-glucosidasa de *Aspergillus oryzae* obtenida según el documento WO 2008/057637.

45 [0224] Otras endoglucanasas, celobiohidrolasas, y beta-glucosidasas son divulgadas en numerosas familias de Glicosil Hidrolasa que utilizan la clasificación según Henrissat B., 1991, A classification of glycosyl hydrolases based on amino-acid sequence similarities, Biochem. J. 280: 309-316, y Henrissat B., y Bairoch A., 1996, Updating the sequence-based classification of glycosyl hydrolases, Biochem. J. 316: 695-696.

50 [0225] Otras enzimas celulolíticas que se pueden usar en la presente invención se describen en los documentos EP 495,257, EP 531,315, EP 531,372, WO 89/09259, WO 94/07998, WO 95/24471, WO 96/11262, WO 96/29397, WO 96/034108, WO 97/14804, WO 98/08940, WO 98/012307, WO 98/13465, WO 98/015619, WO 98/015633, WO 98/028411, WO 99/06574, WO 99/10481, WO 99/025846, WO 99/025847, WO 99/031255, WO 2000/009707, WO 2002/050245, WO 2002/0076792, WO 2002/101078, WO 2003/027306, WO 2003/052054, WO 2003/052055, WO 2003/052056, WO 2003/052057, WO 2003/052118, WO 2004/016760, WO 2004/043980, WO 2004/048592, WO 2005/001065, WO 2005/028636, WO 2005/093050, WO 2005/093073, WO 2006/074005, WO 2006/117432, WO 2007/071818, WO 2007/071820, WO 2008/008070, WO 2008/008793, patente estadounidense nº 4,435,307, patente estadounidense nº 5,457,046, patente estadounidense nº 5,648,263, patente estadounidense nº 5,686,593, patente estadounidense nº 5,691,178, patente estadounidense nº 5,763,254, y patente estadounidense nº 5,776,757.

60 [0226] Ejemplos de preparaciones enzimáticas degradantes de xilano comerciales adecuadas para su uso en la presente

ES 2 551 141 T3

invención incluyen, por ejemplo, SHEARZYME™ (Novozymes A/S), CELLIC™ HTec (Novozymes A/S), VISCOZYME® (Novozymes A/S), ULTRAFLO® (Novozymes A/S), PULPZYME® HC (Novozymes A/S), MULTIFECT® Xylanase (Genencor), ECOPULP® TX-200A (AB Enzymes), HSP 6000 Xylanase (DSM), DEPOL™ 333P (Biocatalysts Limit, Gales, UK), DEPOL™ 740L. (Biocatalysts Limit, Gales, UK), y DEPOL™ 762P (Biocatalysts Limit, Gales, UK).

5

[0227] Ejemplos de xilanasas útiles en los métodos de la presente invención incluyen, pero de forma no limitativa, xilanasas de *Aspergillus aculeatus* (GeneSeqP:AAR63790; WO 94/21785), xilanasas de *Aspergillus fumigatus* (WO 2006/078256), y xilanasas de *Thielavia terrestris* NRRL 8126 (WO 2009/079210).

10

[0228] Ejemplos de beta-xilosidasas útiles en los métodos de la presente invención incluyen, pero de forma no limitativa, beta-xilosidasa de *Trichoderma reesei* (número de registro de UniProtKB/TrEMBL Q92458), *Talaromyces emersonii* (número de registro de SwissProt Q8X212), y *Neurospora crassa* (número de registro SwissProt Q7SOW4).

15

[0229] Ejemplos de acetilxilano esterasesas útiles en los métodos de la presente invención incluyen, pero de forma no limitativa, acetilxilano esterasa de *Hypocrea jecorina* (WO 2005/001036), acetilxilano esterasa de *Neurospora crassa* (número de registro de UniProt q7s259), acetilxilano esterasa de *Thielavia terrestris* NRRL 8126 (WO 2009/042846), acetilxilano esterasa de *Chaetomium globosum* (número de registro de Uniprot Q2GWX4), acetilxilano esterasa de *Chaetomium gracile* (número de registro de GeneSeqP AAB82124), acetilxilano esterasa de *Phaeosphaeria nodorum* (número de registro de Uniprot Q0UHJ1), y acetilxilano esterasa de *Humicola insolens* DSM 1800 (WO 2009/073709).

20

[0230] Ejemplos de esterasesas de ácido ferúlico útiles en los métodos de la presente invención incluyen, pero de forma no limitativa, feruloil esterasa de *Humicola insolens* DSM 1800 (WO 2009/076122), feruloil esterasa de *Neurospora crassa* (número de registro de UniProt Q9HGR3), y feruloil esterasa de *Neosartorya fischeri* (número de registro de UniProt A1D9T4).

25

[0231] Ejemplos de arabinofuranosidasas útiles en los métodos de la presente invención incluyen, pero de forma no limitativa, arabinofuranosidasa de *Humicola insolens* DSM 1800 (WO 2009/073383) y arabinofuranosidasa de *Aspergillus niger* (número de registro de GeneSeqP AAR94170).

30

[0232] Ejemplos de alfa-glucuronidasas útiles en los métodos de la presente invención incluyen, pero de forma no limitativa, alfa-glucuronidasa de *Aspergillus clavatus* (número de registro de UniProt alcc12), alfa-glucuronidasa de *Trichoderma reesei* (número de registro de Uniprot Q99024), alfa-glucuronidasa de *Talaromyces emersonii* (número de registro de UniProt Q8X211), alfa-glucuronidasa de *Aspergillus niger* (número de registro de Uniprot Q96WX9), alfa-glucuronidasa de *Aspergillus terreus* (SwissProt número de registro Q0CJP9), y alfa-glucuronidasa de *Aspergillus fumigatus* (SwissProt número de registro Q4WW45).

35

[0233] Las enzimas celulolíticas y proteínas usadas en los métodos de la presente invención se pueden producir mediante fermentación de las cepas microbianas mencionadas anteriormente en un medio nutritivo que contiene fuentes de nitrógeno y carbono adecuadas y sales inorgánicas, usando procedimientos conocidos en la técnica (véase, por ejemplo, Bennett, J.W. y LaSure, L. (eds.), *More Gene Manipulations in Fungi*, Academic Press, CA, 1991). Hay disponibles medios adecuados de proveedores comerciales o se pueden preparar según composiciones publicadas (por ejemplo, en catálogos de la American Type Culture Collection). Los rangos de temperatura y otras condiciones adecuadas para el crecimiento y producción de enzima celulolítica se conocen en la técnica (véase, por ejemplo, Bailey, J.E., y Ollis, D.F., *Biochemical Engineering Fundamentals*, McGraw-Hill Book Company, NY, 1986).

45

[0234] La fermentación puede ser cualquier método de cultivo de una célula que da como resultado la expresión o aislamiento de una enzima celulolítica. La fermentación puede, por lo tanto, entenderse como que comprende un cultivo en matraz de agitación, o fermentación a pequeña o gran escala (incluyendo fermentaciones de lote continuo, de lote alimentado, o fermentaciones de estado sólido) en el laboratorio o fermentadores industriales realizadas en un medio adecuado y bajo condiciones que permiten que la enzima celulolítica sea expresada o aislada. Las enzimas celulolíticas resultantes producidas por los métodos anteriormente descritos se pueden recuperar del medio de fermentación y purificar mediante procedimientos convencionales.

50

[0235] Fermentación. Los azúcares fermentables obtenidos del material celulósico pretratado e hidrolizado pueden ser fermentados por uno o varios microorganismos de fermentación (diferentes) capaces de fermentar los azúcares directa o indirectamente en un producto de fermentación deseado. "Fermentación" o "proceso de fermentación" se refiere a cualquier proceso de fermentación o cualquier proceso que comprende un paso de fermentación. Los procesos de fermentación también incluyen procesos de fermentación usados en la industria del alcohol consumible (por ejemplo, cerveza y vino), industria láctea (por ejemplo, productos lácteos fermentados), industria del cuero, e industria del tabaco. Las condiciones de fermentación dependen del producto de fermentación deseado y el organismo fermentador y pueden ser determinadas fácilmente por un experto en la técnica.

55

60

- 5 [0236] En el paso de fermentación, los azúcares, liberados del material celulósico como resultado de los pasos de pretratamiento y de hidrólisis enzimática, son fermentados dando un producto, por ejemplo, etanol, mediante un organismo fermentador, tal como levadura. La hidrólisis (sacarificación) y fermentación pueden ser separadas o simultáneas. Tales métodos incluyen, pero de forma no limitativa, hidrólisis y fermentación separadas (SHF); sacarificación y fermentación simultáneas (SSF); sacarificación y cofermentación simultáneas (SSCF); hidrólisis y fermentación híbridas (HHF); SHCF (hidrólisis y co-fermentación separadas), HHCF (hidrólisis y fermentación híbridas), y conversión microbiana directa (DMC).
- 10 [0237] Cualquier material celulósico hidrolizado adecuado se puede usar en el paso de fermentación en la práctica de la presente invención. El material se selecciona generalmente en base al producto de fermentación deseado, es decir, la sustancia por obtener de la fermentación, y el proceso empleado, como es bien conocido en la técnica. Ejemplos de sustratos adecuados para su uso en los métodos de la presente invención, incluyen materiales celulósicos, tales como residuos de madera o de planta o azúcares de bajo peso molecular DP1-3 obtenidos de material celulósico procesado que pueden ser metabolizados por el microorganismo fermentador, y que pueden ser suministrados mediante adición directa al medio de fermentación.
- 15 [0238] El término "medio de fermentación" se entiende aquí que se refiere a un medio antes de que el (los) microorganismo(s) de fermentación sea(n) añadidos, tal como un medio resultante de un proceso de sacarificación, al igual que un medio usado en un proceso de sacarificación y fermentación simultáneas (SSF).
- 20 [0239] "Microorganismo fermentador" se refiere a cualquier microorganismo, incluyendo organismos bacterianos y fúngicos, adecuados para su uso en un proceso de fermentación deseado para producir un producto de fermentación. El organismo fermentador puede ser organismos de fermentación C₆ y/o C₅, o una combinación de los mismos. Tanto el organismo de fermentación C₆ como el C₅ son bien conocidos en la técnica. Los microorganismos de fermentación adecuados son capaces de fermentar, es decir, convertir, azúcares, tales como glucosa, xilosa, xilulosa, arabinosa, maltosa, manosa, galactosa, u oligosacáridos, directa o indirectamente en el producto de fermentación deseado.
- 25 [0240] Ejemplos de organismos de fermentación bacterianos y fúngicos productores de etanol son descritos por Lin et al., 2006, Appl. Microbiol. Biotechnol. 69: 627-642.
- 30 [0241] Ejemplos de microorganismos de fermentación que pueden fermentar azúcares C₆ incluyen organismos bacterianos y fúngicos, tales como levadura. La levadura preferida incluye cepas de *Saccharomyces* spp., preferiblemente *Saccharomyces cerevisiae*.
- 35 [0242] Ejemplos de organismos de fermentación que pueden fermentar azúcares C₅ incluyen organismos bacterianos y fúngicos, tales como levadura. La levadura de fermentación C₅ preferida incluye cepas de *Pichia*, preferiblemente *Pichia stipitis*, tal como *Pichia stipitis* CBS 5773; cepas de *Candida*, preferiblemente *Candida boidinii*, *Candida brassicae*, *Candida sheatae*, *Candida diddensii*, *Candida pseudotropicalis*, o *Candida utilis*.
- 40 [0243] Otros organismos de fermentación incluyen cepas de *Zymomonas*, tales como *Zymomonas mobilis*; *Hansenula*, tales como *Hansenula anomala*; *Klyveromyces*, tales como *K. fragilis*; *Schizosaccharomyces*, tal como *S. pombe*; y *E. coli*, especialmente cepas de *E. coli* que han sido genéticamente modificadas para mejorar el rendimiento de etanol.
- 45 [0244] La levadura puede ser una *Saccharomyces* spp. La levadura puede ser *Saccharomyces cerevisiae*. La levadura puede ser *Saccharomyces distaticus*. La levadura puede ser *Saccharomyces uvarum*. La levadura puede ser una *Kluyveromyces*. La levadura puede ser *Kluyveromyces marxianus*. La levadura puede ser *Kluyveromyces fragilis*. La levadura puede ser una *Candida*. La levadura puede ser *Candida boidinii*. La levadura puede ser *Candida brassicae*. La levadura puede ser *Candida diddensii*. La levadura puede ser *Candida pseudotropicalis*. La levadura puede ser *Candida utilis*. La levadura puede ser una *Clavispora*. La levadura puede ser *Clavispora lusitaniae*. La levadura puede ser *Clavispora opuntiae*. La levadura puede ser una *Pachysolen*. La levadura puede ser *Pachysolen tannophilus*. La levadura puede ser una *Pichia*. La levadura puede ser una *Pichia stipitis*. La levadura puede ser una *Bretannomyces*. La levadura puede ser *Bretannomyces clausenii* (Philippidis, G. P., 1996, Cellulose bioconversion technology, en Handbook on Bioethanol: Production and Utilization, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212).
- 50 [0245] Bacterias que pueden fermentar eficazmente hexosa y pentosa en etanol incluyen, por ejemplo, *Zymomonas mobilis* y *Clostridium thermocellum* (Philippidis, 1996, supra).
- 55 [0246] La bacteria puede ser una *Zymomonas*. La bacteria puede ser *Zymomonas mobilis*. La bacteria puede ser una *Clostridium*. La bacteria puede ser *Clostridium thermocellum*.
- 60

ES 2 551 141 T3

- 5 [0247] La levadura disponible comercialmente adecuada para la producción de etanol incluye, por ejemplo, levadura ETHANOL RED™ (disponible de Fermentis/Lesaffre, EE.UU), FALI™ (disponible de Fleischmann's Yeast, EE.UU), levadura fresca SUPERSTART™ y THERMOSACC™ (disponible de Ethanol Technology, WI, EE.UU), BIOFERM™ AFT y XR (disponibles de NABC - North American Bioproducts Corporation, GA, EE.UU), GERT STRAND™ (disponible de GERT Strand AB Suecia), y FERMIOL™ (disponible de DSM Specialties).
- 10 [0248] El microorganismo fermentador puede ser modificado genéticamente para proporcionar la capacidad de fermentar azúcares de pentosa, tal como microorganismos que utilizan xilosa, que utilizan arabinosa, y que co-utilizan xilosa y arabinosa.
- 15 [0249] La clonación de genes heterólogos en varios microorganismos de fermentación ha llevado a la construcción de organismos capaces de convertir hexosas y pentosas en etanol (cofermentación) (Chen y Ho, 1993, Cloning and improving the expression of *Pichia stipitis* xylose reductase gene in *Saccharomyces cerevisiae*, Appl. Biochem. Biotechnol. 39-40: 135-147; Ho et al., 1998, Genetically engineered *Saccharomyces* yeast capable of effectively cofermenting glucose and xylose, Appl. Environ. Microbiol. 64: 1852-1859; Kotter y Ciriacy, 1993, Xylose fermentation by *Saccharomyces cerevisiae*, Appl. Microbiol. Biotechnol. 38: 776-783; Walfridsson et al., 1995, Xylose-metabolizing *Saccharomyces cerevisiae* strains overexpressing the TKL1 and TAL1 genes encoding the pentose phosphate pathway enzymes transketolase and transaldolase, Appl. Environ. Microbiol. 61: 4184-4190; Kuyper et al., 2004, Minimal metabolic engineering of *Saccharomyces cerevisiae* for efficient anaerobic xylose fermentation: a proof of principle, FEMS Yeast Research 4: 655-664; Beall et al., 1991, Parametric studies of ethanol production from xylose and other sugars by recombinant *Escherichia coli*, Biotech. Bioeng. 38: 296-303; Ingram et al., 1998, Metabolic engineering of bacteria for ethanol production, Biotechnol. Bioeng. 58: 204-214; Zhang et al., 1995, Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*, Science 267: 240-243; Deanda et al., 1996, Development of an arabinose-fermenting *Zymomonas mobilis* strain by metabolic pathway engineering, Appl. Environ. Microbiol. 62: 4465-4470).
- 25 [0250] El microorganismo fermentador genéticamente modificado puede ser *Saccharomyces cerevisiae*. El microorganismo fermentador genéticamente modificado puede ser *Zymomonas mobilis*. El microorganismo fermentador genéticamente modificado puede ser *Escherichia coli*. El microorganismo fermentador genéticamente modificado puede ser *Klebsiella oxytoca*.
- 30 [0251] Es bien conocido en la técnica que los organismos descritos anteriormente también pueden usarse para producir otras sustancias, como se describe en este caso.
- 35 [0252] El microorganismo fermentador se añade típicamente a la lignocelulosa o al hidrolizado degradados y la fermentación se realiza durante aproximadamente 8 a aproximadamente 96 horas, tal como aproximadamente 24 a aproximadamente 60 horas. La temperatura se encuentra típicamente entre aproximadamente 26°C a aproximadamente 60°C, en particular aproximadamente 32°C o 50°C, y en alrededor de pH 3 a aproximadamente pH 8, tal como alrededor de pH 4-5, 6, o 7.
- 40 [0253] La levadura y/o otro microorganismo se aplican a la lignocelulosa o hidrolizado degradados y la fermentación se realiza durante aproximadamente 12 a aproximadamente 96 horas, tal como típicamente 24-60 horas. En un aspecto preferido, la temperatura se encuentra preferiblemente entre aproximadamente 20°C a aproximadamente 60°C, de forma más preferible aproximadamente 25°C a aproximadamente 50°C, y de la forma más preferible aproximadamente 32°C a aproximadamente 50°C, en particular aproximadamente 32°C o 50°C, y el pH se encuentra generalmente de alrededor de pH 3 a alrededor de pH 7, preferiblemente alrededor de pH 4-7. No obstante, algún, por ejemplo, organismo de fermentación bacteriano tiene la temperatura de fermentación óptima más alta. La levadura u otro microorganismo se aplica preferiblemente en cantidades de aproximadamente 10^5 a 10^{12} , preferiblemente de aproximadamente 10^7 a 10^{10} , especialmente aproximadamente 2×10^8 concentración de células viables por ml de caldo de fermentación. Se puede encontrar más guía respecto al uso de levadura para la fermentación en, por ejemplo, "The Alcohol Textbook" (Editores K. Jacques, T.P. Lyons y D.R. Kelsall, Nottingham University Press, Reino Unido 1999).
- 45 [0254] El proceso más ampliamente usado en la técnica es el proceso de sacarificación y fermentación simultáneas (SSF) donde no hay fase de retención para la sacarificación, lo que significa que la levadura y enzima son añadidas juntas.
- 50 [0255] Para producción de etanol, después de la fermentación el lodo fermentado se destila para extraer el etanol. Se puede usar el etanol obtenido según los métodos de la invención como, por ejemplo, etanol de combustible, etanol potable, es decir, licores neutrales potables, o etanol industrial.
- 55 [0256] Se puede usar un estimulador de fermentación en combinación con cualquiera de los procesos enzimáticos
- 60

ES 2 551 141 T3

descritos por la presente para mejorar más el proceso de fermentación, y en particular, el rendimiento del microorganismo fermentador, tal como, mejora del índice y rendimiento de etanol. Un "estimulador de fermentación" se refiere a estimuladores para el crecimiento de los microorganismos de fermentación, en particular, levadura. Los estimuladores de fermentación preferidos para el crecimiento incluyen vitaminas y minerales. Ejemplos de vitaminas incluyen multivitaminas, biotina, pantotenato, ácido nicotínico, meso-inositol, tiamina, piridoxina, ácido para-aminobenzoico, ácido fólico, riboflavina, y vitaminas A, B, C, D, y E. Véase, por ejemplo, Alfenore et al., Improving ethanol production and viability of *Saccharomyces cerevisiae* by a vitamin feeding strategy during fed-batch process, Springer-Verlag (2002). Ejemplos de minerales incluyen minerales y sales minerales que pueden suministrar nutrientes que comprenden P, K, Mg, S, Ca, Fe, Zn, Mn, y Cu.

[0257] Productos de fermentación: un producto de fermentación puede ser cualquier sustancia derivada de la fermentación. El producto de fermentación puede ser, sin límite, un alcohol (por ejemplo, arabinol, butanol, etanol, glicerol, metanol, 1,3-propanodiol, sorbitol, y xilitol); un ácido orgánico (por ejemplo, ácido acético, ácido acetónico, ácido adípico, ácido ascórbico, ácido cítrico, ácido 2,5-diketo-D-gluconico, ácido fórmico, ácido fumárico, ácido glucárico, ácido glucónico, ácido glucurónico, ácido glutárico, ácido 3-hidroxiopropiónico, ácido itacónico, ácido láctico, ácido málico, ácido malónico, ácido oxálico, ácido propiónico, ácido succínico, y ácido xilónico); una cetona (por ejemplo, acetona); un aminoácido (por ejemplo, ácido aspártico, ácido glutámico, glicina, lisina, serina, y treonina); y un gas (por ejemplo, metano, hidrógeno (H₂), dióxido de carbono (CO₂), y monóxido de carbono (CO)). El producto de fermentación puede ser también proteína como producto de valor alto.

[0258] El producto de fermentación puede ser un alcohol. Se entiende que el término "alcohol" incluye una sustancia que contiene una o varias fracciones de hidroxilo. El alcohol puede ser arabinol. El alcohol puede ser butanol. El alcohol puede ser etanol. El alcohol puede ser glicerol. El alcohol puede ser metanol. El alcohol puede ser 1,3-propanodiol. El alcohol puede ser sorbitol. El alcohol puede ser xilitol. Véase, por ejemplo, Gong, C. S., Cao, N. J., Du, J., y Tsao, G. T., 1999, Ethanol production from renewable resources, en *Advances in Biochemical Engineering/Biotechnology*, Scheper, T., ed., Springer-Verlag Berlín Heidelberg, Alemania, 65: 207- 241; Silveira, M. M., y Jonas, R., 2002, The biotechnological production of sorbitol, *Appl. Microbiol. Biotechnol.* 59: 400-408; Nigam, P., y Singh, D., 1995, Processes for fermentative production of xylitol - a sugar substitute, *Process Biochemistry* 30 (2): 117-124; Ezeji, T. C., Qureshi, N. y Blaschek, H. P., 2003, Production of acetone, butanol y and ethanol by *Clostridium beijerinckii* BA101 and in situ recovery by gas stripping, *World Journal of Microbiology and Biotechnology* 19 (6): 595-603.

[0259] El producto de fermentación puede ser un ácido orgánico. El ácido orgánico puede ser ácido acético. El ácido orgánico puede ser ácido acetónico. El ácido orgánico puede ser ácido adípico. El ácido orgánico puede ser ácido ascórbico. El ácido orgánico puede ser ácido cítrico. El ácido orgánico puede ser ácido 2,5-diketo-D-gluconico. El ácido orgánico puede ser ácido fórmico. El ácido orgánico puede ser ácido fumárico. El ácido orgánico puede ser ácido glucárico. El ácido orgánico puede ser ácido glucurónico. El ácido orgánico puede ser ácido glutárico. El ácido orgánico puede ser ácido 3-hidroxiopropiónico. El ácido orgánico puede ser ácido itacónico. El ácido orgánico puede ser ácido láctico. El ácido orgánico puede ser ácido málico. El ácido orgánico puede ser ácido malónico. El ácido orgánico puede ser ácido oxálico. El ácido orgánico puede ser ácido propiónico. El ácido orgánico puede ser ácido succínico. El ácido orgánico puede ser ácido xilónico. Véase, por ejemplo, Chen, R., y Lee Y. Y., 1997, Membrane-mediated extractive fermentation for lactic acid production from cellulosic biomass, *Appl. Biochem. Biotechnol.* 63-65: 435-448.

[0260] El producto de fermentación también puede ser una cetona. Se entiende que el término "cetona" abarca una sustancia que contiene una o varias fracciones de cetona. La cetona puede ser acetona. Véase, por ejemplo, Qureshi y Blaschek, 2003, *supra*.

[0261] El producto de fermentación puede ser un aminoácido. El ácido orgánico puede ser ácido aspártico. El aminoácido puede ser ácido glutámico. El aminoácido puede ser glicina. El aminoácido puede ser lisina. El aminoácido puede ser serina. El aminoácido puede ser treonina. Véase, por ejemplo, Ricardo, A., y Margaritis, A., 2004, Empirical modeling of batch fermentation kinetics for poly(glutamic acid) production and other microbial biopolymers, *Biotechnology and Bioengineering* 87 (4): 501-515.

[0262] El producto de fermentación puede ser un gas. El gas puede ser metano. El gas puede ser H₂. El gas puede ser CO₂. El gas puede ser CO. Véase, por ejemplo, Kataoka, N., A. Miya, y K. Kiriya, 1997, Studies on hydrogen production by continuous culture system of hydrogen-producing anaerobic bacteria, *Waster Science and Technology* 36 (6-7): 41-47; y Gunaseelan V.N. en *Biomass and Bioenergy*, Vol. 13 (1-2), págs. 83-114, 1997, Anaerobic digestion of biomass for methane production: A review.

[0263] Recuperación. El (los) producto(s) de fermentación puede(n) ser opcionalmente recuperada(s) del medio de fermentación utilizando cualquier método conocido en la técnica incluyendo, pero no limitándose a, cromatografía,

procedimientos electroforéticos, solubilidad diferencial, destilación, o extracción. Por ejemplo, el alcohol es separado del material celulósico fermentado y purificado por métodos convencionales de destilación. Puede obtenerse etanol con una pureza de hasta aproximadamente 96 vol.%, que se puede usar como, por ejemplo, etanol para combustible, etanol potable, es decir, licores neutrales potables, o etanol industrial.

5

Composiciones detergentes

[0264] Los polipéptidos con actividad de aumento celulolítico de la presente invención se pueden añadir a y así volverse componentes de una composición de detergente.

10

[0265] La composición detergente de la presente invención puede ser formulada, por ejemplo, como una composición de detergente de lavado a mano o a máquina que incluye una composición de aditivo de lavandería adecuada para el pretratamiento de tejidos manchados y una composición suavizante de tejido añadida de enjuague, o ser formulada como una composición detergente para su uso en operaciones de limpieza de superficies duras de casa en general, o ser formulada para operaciones de lavado de vajilla a mano o a máquina.

15

[0266] En un aspecto específico, la presente invención proporciona un aditivo de detergente que comprende un polipéptido de la invención. El aditivo de detergente al igual que la composición de detergente puede comprender una o varias enzimas (diferentes) tales como una proteasa, lipasa, cutinasa, una amilasa, carbohidrasa, celulasa, pectinasa, mananasa, arabinasa, galactanasa, xilanasas, oxidasas, por ejemplo, una lacasa, y/o peroxidasa.

20

[0267] En general las propiedades de la(s) enzima(s) seleccionada(s) deberían ser compatibles con el detergente seleccionado, (es decir, pH óptimo, compatibilidad con otros ingredientes enzimáticos y no enzimáticos, etc.), y la(s) enzima(s) debería(n) estar presente(s) en cantidades eficaces.

25

[0268] Celulasas: celulasas adecuadas incluyen aquellas de origen bacteriano o fúngico. Se incluyen mutantes químicamente modificados o creados genéticamente de proteína. Celulasas adecuadas incluyen celulasas de los géneros *Bacillus*, *Pseudomonas*, *Humicola*, *Fusarium*, *Thielavia*, *Acremonium*, por ejemplo, las celulasas fúngicas producidas a partir de *Humicola insolens*, *Myceliophthora thermophila* y *Fusarium oxysporum* descritas en los documentos US 4,435,307, US 5,648,263, US 5,691,178, US 5,776,757 y WO 89/09259.

30

[0269] Celulasas especialmente adecuadas son las celulasas alcalinas o neutrales con beneficios de cuidado del color. Ejemplos de tales celulasas son las celulasas descritas en los documentos EP 0 495 257, EP 0 531 372, WO 96/11262, WO 96/29397, WO 98/08940. Otros ejemplos son variantes de celulasa como las descritas en los documentos WO 94/07998, EP 0 531 315, US 5,457,046, US 5,686,593, US 5,763,254, WO 95/24471, WO 98/12307 y PCT/DK98/00299.

35

[0270] Las celulasas disponibles comercialmente incluyen CELLUZYME™, y CAREZYME™ (Novozymes A/S), CLAZINASE™, y PURADAX HA™ (Genencor International Inc.), y KAC-500(B)™ (Kao Corporation).

40

[0271] Proteasas: proteasas adecuadas incluyen aquellas de origen animal, vegetal o microbiano. Se prefiere el origen microbiano. Se incluyen mutantes químicamente modificados o creados genéticamente de proteína. La proteasa puede ser una serina proteasa o una metaloproteasa, preferiblemente una proteasa microbiana alcalina o una proteasa de tipo tripsina. Ejemplos de proteasas alcalinas son subtilisinas, especialmente aquellas derivadas de *Bacillus*, por ejemplo, subtilisina Novo, subtilisina Carlsberg, subtilisina 309, subtilisina 147 y subtilisina 168 (descrita en WO 89/06279). Ejemplos de proteasas de tipo tripsina son tripsina (por ejemplo, de origen porcino o bovino) y la proteasa de *Fusarium* descrita en los documentos WO 89/06270 y WO 94/25583.

45

[0272] Ejemplos de proteasas útiles son las variantes descritas en los documentos WO 92/19729, WO 98/20115, WO 98/20116, y WO 98/34946, especialmente las variantes con sustituciones en una o varias de las siguientes posiciones: 27, 36, 57, 76, 87, 97, 101, 104, 120, 123, 167, 170, 194, 206, 218, 222, 224, 235, y 274.

50

[0273] Enzimas proteásicas disponibles comercialmente preferidas incluyen ALCALASE™, SAVINASE™, PRIMASE™, DURALASE™, ESPERASE™, y KANNASE™ (Novozymes A/S), MAXATASE™, MAXACAL™, MAXAPEM™, PROPERASE™, PURAFECT™, PURAFECT OXP™, FN2™, y FN3™ (Genencor International Inc.).

55

[0274] Lipasas: lipasas adecuadas incluyen aquellas de origen bacteriano o fúngico. Se incluyen mutantes químicamente modificados o manipulados genéticamente de proteína. Ejemplos de lipasas útiles incluyen lipasas de *Humicola* (sinónimo *Thermomyces*), por ejemplo, de *H. lanuginosa* (*T. lanuginosus*) como se describe en el EP 258 068 y EP 305 216 o de *H. insolens* como se describe en el documento WO 96/13580, lipasa de *Pseudomonas*, por ejemplo, de *P. alcaligenes* o *P. pseudoalcaligenes* (EP 218 272), *P. cepacia* (EP 331 376), *P. stutzeri* (GB 1,372,034), *P. fluorescens*, *Pseudomonas sp. cepa* SD 705 (WO 95/06720 y WO 96/27002), *P. wisconsinensis* (WO 96/12012), una lipasa de

60

ES 2 551 141 T3

Bacillus, por ejemplo, de *B. subtilis* (Dartois et al., 1993, Biochemica et Biophysica Acta, 1131: 253-360), *B. stearothermophilus* (JP 64/744992) o *B. pumilus* (WO 91/16422).

5 [0275] Otros ejemplos son variantes de lipasa tales como las descritos en los documentos WO 92/05249, WO 94/01541, EP 407 225, EP 260 105, WO 95/35381, WO 96/00292, WO 95/30744, WO 94/25578, WO 95/14783, WO 95/22615, WO 97/04079 y WO 97/07202.

10 [0276] Las enzimas de lipasa disponibles comercialmente preferidas incluyen LIPOLASE™ y LIPOLASE ULTRA™ (Novozymes A/S).

15 [0277] Amilasas: las amilasas adecuadas (α y/o β) incluyen aquellas de origen bacteriano o fúngico. Se incluyen mutantes químicamente modificados o creados genéticamente de proteína. Las amilasas incluyen, por ejemplo, α -amilasas obtenidas a partir de *Bacillus*, por ejemplo, una cepa especial de *Bacillus licheniformis*, descrita con más detalle en GB 1,296,839.

[0278] Ejemplos de amilasas útiles son las variantes descritas en los documentos WO 94/02597, WO 94/18314, WO 96/23873, y WO 97/43424, especialmente las variantes con sustituciones en una o varias de las siguientes posiciones: 15, 23, 105, 106, 124, 128, 133, 154, 156, 181, 188, 190, 197, 202, 208, 209, 243, 264, 304, 305, 391, 408, y 444.

20 [0279] Las amilasas disponibles comercialmente son DURAMYL™, TERMAMYL™, FUNGAMYL™ y BAN™ (Novozymes A/S), RAPIDASE™ y PURASTAR™ (de Genencor International Inc.).

25 [0280] Peroxidasas/oxidases: peroxidasas/oxidases adecuadas incluyen aquellas de origen vegetal, bacteriano o fúngico. Se incluyen mutantes químicamente modificados o creados genéticamente de proteína. Ejemplos de peroxidasas útiles incluyen peroxidasas de *Coprinus*, por ejemplo, de *C. cinereus*, y variantes de las mismas como aquellas descritas en los documentos WO 93/24618, WO 95/10602, y WO 98/15257.

[0281] Las peroxidasas disponibles comercialmente incluyen GUARDZYME™ (Novozymes A/S).

30 [0282] La(s) enzima(s) detergente(s) se puede incluir en una composición detergente añadiendo aditivos separados que contienen una o varias enzimas (diferentes), o añadiendo un aditivo combinado que comprende todas estas enzimas. Un aditivo detergente de la invención, es decir, un aditivo separado o un aditivo combinado, puede ser formulado, por ejemplo, como un granulado, líquido, lodo, etc. Las formulaciones de aditivo detergente preferidas son granulados, en particular granulados no polvorientos, líquidos, en particular líquidos estabilizados, o lodos.

35 [0283] Se pueden producir granulados no polvorientos, por ejemplo, como se describe en los documentos US 4,106,991 y 4,661,452 y pueden estar recubiertos opcionalmente mediante métodos conocidos en la técnica. Ejemplos de materiales de recubrimiento ceroso son productos de óxido de (poli)etileno (polietilenglicol, PEG) con pesos molares medios de 1000 a 20000; nonilfenoles etoxilados con de 16 a 50 unidades de óxido de etileno; alcoholes grasos etoxilados donde el alcohol contiene de 12 a 20 átomos de carbono y en el cual hay de 15 a 80 unidades de óxido de etileno; alcoholes grasos; ácidos grasos; y mono- y di- y triglicéridos de ácidos grasos. Ejemplos de materiales de recubrimiento que forman películas adecuados para la aplicación mediante técnicas de lecho fluidificado se dan en el documento GB 1483591. Las preparaciones enzimáticas líquidas pueden, por ejemplo, ser estabilizadas añadiendo un poliol tal como propilenglicol, un azúcar o alcohol de azúcar, ácido láctico o ácido bórico según métodos establecidos.
45 Las enzimas protegidas se pueden preparar según el método descrito en el documento EP 238,216.

[0284] La composición detergente de la invención puede estar en cualquier forma conveniente, por ejemplo, una barra, una tableta, un polvo, un gránulo, una pasta o un líquido. Un detergente líquido puede ser acuoso, típicamente contiene hasta el 70% agua y 0-30% solvente orgánico, o no acuoso.

50 [0285] La composición detergente comprende uno o varios surfactantes (diferentes), que pueden ser no iónicos incluyendo semipolar y/o aniónicos y/o catiónicos y/o zwitteriónicos. Los surfactantes están presentes típicamente a un nivel de 0,1% a 60% en peso.

55 [0286] Cuando es incluido en ella el detergente normalmente contendrá aproximadamente 1% a aproximadamente 40% de un tensioactivo aniónico tal como alquilbencenosulfonato lineal, alfa-olefinsulfonato, alquil sulfato (sulfato de alcohol graso), etoxisulfato alcohólico, alcanosulfonato secundario, éster metílico de ácido graso alfa-sulfo, ácido alquil o alquenilsuccínico, o jabón.

60 [0287] Cuando el detergente es incluido en él normalmente contendrá aproximadamente del 0,2% a aproximadamente el 40% de un tensioactivo no iónico tal como alcohol etoxilato, nonilfenol etoxilato, alquilpoliglicósido, óxido de

ES 2 551 141 T3

alquildimetilamina, monoetanolamida de ácido graso etoxilado, monoetanolamida de ácido graso, polihidroxi alquil amida de ácido graso, o derivados de n-acilo n-alquilo glucosamina ("glucamidas").

5 [0288] El detergente puede contener 0-65% de un constructor o agente complejante de detergente tal como zeolita, difosfato, trifosfato, fosfonato, carbonato, citrato, ácido nitrilotriacético, ácido etilendiaminatetraacético, ácido dietileno-triaminopentaacético, ácido alquil- o alqueniilsuccínico, silicatos solubles, o silicatos estratificados (por ejemplo, SKS-6 de Hoechst).

10 [0289] El detergente puede comprender uno o varios polímeros (diferentes). Ejemplos son carboximetilcelulosa, poli(vinilpirrolidona), poli(etilenglicol), alcohol (poli)vinílico, óxido (poli)vinilpiridin-N, (poli)vinilimidazol, policarboxilatos tales como poliacrilatos, copolímeros de ácido maléico/acrílico, y copolímeros de ácido metacrilato/acrílico de lauril.

15 [0290] El detergente puede contener un sistema blanqueante que puede comprender una fuente de H₂O₂ tal como perborato o percarbonato que se puede combinar con un activador blanqueante de formación de perácido tal como tetraacetileno-diamina o nonanoiloxibencenosulfonato. Alternativamente, el sistema blanqueante puede comprender peroxiácidos de, por ejemplo, tipo amida, imida, o sulfona.

20 [0291] La(s) enzima(s) de la composición detergente de la invención se pueden estabilizar usando agentes estabilizantes convencionales, por ejemplo, un poliol tal como propilenglicol o glicerol, un azúcar o alcohol de azúcar, ácido láctico, ácido bórico, o un derivado de ácido bórico, por ejemplo, un éster de borato aromático, o un derivado de ácido fenil borónico tal como ácido borónico de 4-formilfenilo, y la composición se puede formular como se describe en, por ejemplo, los documentos WO 92/19709 y WO 92/19708.

25 [0292] El detergente también puede contener otros ingredientes detergentes convencionales tal como, por ejemplo, acondicionadores de tejidos incluyendo arcillas, reforzadores de espuma, supresores de espuma, agentes anticorrosivos, agentes de suspensión de suciedad, agentes antiredeposición de suciedad, colorantes, bactericidas, blanqueadores ópticos, hidrotropos, inhibidores de decoloración, o perfumes.

30 [0293] En las composiciones detergentes, cualquier enzima se puede añadir en una cantidad que corresponde con 0,01-100 mg de proteína enzimática por litro de solución de lavado, preferiblemente 0,05-5 mg de proteína enzimática por litro de solución de lavado, en particular 0,1-1 mg de proteína enzimática por litro de solución de lavado.

35 [0294] En las composiciones detergentes, un polipéptido de la presente invención con actividad de aumento celulolítico se puede adicionar en una cantidad que corresponde con 0,001-100 mg de proteína, preferiblemente 0,005-50 mg de proteína, de forma más preferible 0,01-25 mg de proteína, incluso de forma más preferible 0,05- 10 mg de proteína, de la forma más preferible 0,05-5 mg de proteína, e incluso de la forma más preferible 0,01-1 mg de proteína por litro de solución de lavado.

40 [0295] Un polipéptido de la invención con actividad de aumento celulolítico también se puede incorporar en las formulaciones detergentes descritas en el documento WO 97/07202.

Péptido señal

45 [0296] Describas son un polinucleótido aislado que codifica un péptido señal que comprende o consiste en los aminoácidos 1 a 22 de la SEC ID nº 2, los aminoácidos de 1 a 18 de la SEC ID nº 4, o los aminoácidos de 1 a 23 de la SEC ID nº 6. Hay descritos también constructos de ácidos nucleicos que comprenden un gen que codifica una proteína, donde el gen está enlazado operativamente a tal polinucleótido que codifica un péptido señal que comprende o consiste en los aminoácidos de 1 a 22 de la SEC ID nº 2, los aminoácidos de 1 a 18 de la SEC ID nº 4, o los aminoácidos de 1 a 23 de la SEC ID nº 6, donde el gen es foráneo al polinucleótido.

50 [0297] La secuencia de polinucleótidos puede comprender o consistir en los nucleótidos de 1 a 66 de la SEC ID nº 1, los nucleótidos de 1 a 54 de la SEC ID nº 3, o los nucleótidos de 1 a 69 de la SEC ID nº 5.

55 [0298] Descritos hay vectores de expresión recombinantes y células huésped recombinantes que comprenden tales constructos de ácidos nucleicos.

[0299] La presente invención está descrita posteriormente por los siguientes ejemplos que no deberían ser interpretados como limitación del ámbito de la invención.

60 **Ejemplos**

ES 2 551 141 T3

Cepas

5 [0300] La cepa NN051486 de *Thermoascus crustaceus* (CBS 181,67 o ATCC 16462) fue usada como una fuente de polipéptidos con actividad de aumento celulolítico GH61A, GH61 B, y GH61C. La cepa HowB101 de *Aspergillus oryzae* (WO 95/35385) fue usada como huésped para la expresión de recombinación de los polipéptidos GH61 de *Thermoascus crustaceus* con actividad de aumento celulolítico.

Medios

10 [0301] Las placas PDA estaban compuestas de 39 gramos de ágar de dextrosa de patata y agua desionizada a 1 litro.

15 [0302] Las placas de medio mínimo estaban compuestas por 6 g de NaNO₃, 0,52 g de KCl, 1,52 g de KH₂PO₄, 1 ml de solución de oligoelementos COVE, 20 g de ágar noble, 20 ml de 50% glucosa, 2,5 ml de MgSO₄·7H₂O, 20 ml de una solución de biotina al 0,02%, y agua desionizada a 1 litro.

[0303] La solución de oligoelementos COVE estaba compuesta por 0,04 g de Na₂B₄O₇·10H₂O, 0,4 g de CuSO₄·5H₂O, 1,2 g de FeSO₄·7H₂O, 0,7 g de MnSO₄·H₂O, 0,8 g de Na₂MoO₂·2H₂O, 10 g de ZnSO₄·7H₂O, y agua desionizada a 1 litro.

20 [0304] El medio YPM estaba compuesto por 10 g de extracto de levadura, 10 g de bacto-peptona, 20 g de maltosa, y agua desionizada a 1 litro.

25 [0305] El medio NNCYP-PCS estaba compuesto por 5,0 g de NaNO₃, 3,0 g de NH₄Cl, 2,0 g de MES, 2,5 g de ácido cítrico, 0,2 g de CaCl₂·2H₂O, 1,0 g de bacto-peptona, 5,0 g de extracto de levadura, 0,2 g de MgSO₄·7H₂O, 4,0 g de K₂HPO₄, 1,0 ml de solución de oligoelementos COVE, 2,5 g de glucosa, 25,0 g de rastrojos de maíz pretratados (PCS), y agua desionizada a 1 litro.

30 [0306] La solución de oligoelementos COVE estaba compuesta por 0,04 g de Na₂B₄O₇·10H₂O, 0,4 g de CuSO₄·5H₂O, 1,2 g de FeSO₄·7H₂O, 0,7 g de MnSO₄·H₂O, 0,8 g de Na₂MoO₂·2H₂O, 10 g de ZnSO₄·7H₂O, y agua desionizada a 1 litro.

[0307] Las placas LB estaban compuestas de 10 g de triptona, 5 g de extracto de levadura, 10 g de cloruro sódico, 15 g de agar, y agua desionizada a 1 litro.

35 Ejemplo 1: preparación de ADN y ARN genómico de *Thermoascus crustaceus*

Cepa CBS 181.67

40 [0308] La cepa CBS 181.67 de *Thermoascus crustaceus* fue inoculada sobre una placa PDA e incubada durante 3-4 días a 45°C en la oscuridad. Se inocularon diferentes tapones de micelios de PDA (potato dextrose agar) en 500 ml de frascos de agitación que contienen 100 ml de medio NNCYP-PCS. Los matraces fueron incubados durante 6 días a 45°C con agitación a 160 r.p.m. Los micelios fueron recogidos al día 3, día 4, día 5, y día 6. Luego los micelios de cada día fueron combinados y congelados en nitrógeno líquido, y luego almacenados en un congelador a -80°C hasta su uso.

45 [0309] Se extrajo ADN genómico utilizando un equipo DNEASY® Plant Mini Kit (QIAGEN Inc., Valencia, CA, EE.UU). El ARN total fue aislado usando un equipo RNEASY® Mini Kit (QIAGEN Inc., Valencia, CA, EE.UU). El ADNc fue sintetizado siguiendo las instrucciones del sistema de amplificación rápida 3' Rapid Amplification of cDNA End System (3' RACE por sus siglas en inglés) (Invitrogen Corporation, Carlsbad, CA, EE.UU).

50 Ejemplo 2: clonación de una cepa CBS 181.67 GH61A de *Thermoascus crustaceus*

Gen de polipéptido

55 [0310] Cuatro cebadores degenerados mostrados a continuación fueron diseñados en base a regiones conservadas de secuencias GH61 conocidas.

GH61A scF1:

5'-GCNACNGAYCTNGGNTTTG-3' (SEC ID nº 7)

GH61A scF2:

60 5'-GCNACNGAYCTNGGNTTCG-3' (SEC ID nº 8)

ES 2 551 141 T3

GH61A scF3:

5'-GCNACNGAYTTRGGNTTYG-3' (SEC ID nº 9)

5

GH61A scR1:

5'-CAYTGNGGRTARTTYTGNGC-3' (SEC ID nº 10)

10 [0311] Se realizó la PCR usando una combinación de cebadores hacia adelante GH61A scF1, GH61 scF2, y GH61A scF3, y cebador inverso GH61A scR1 y ADNc como modelo. La reacción de amplificación fue compuesta por 5 µl de 10X tampón PCR (Invitrogen Corporation, Carlsbad, CA, EE.UU), 2 µl de 25 mM MgCl₂, 1 µl de 10 mM dNTP, 1 µl de 100 µM cebador directo, 1 µl de 100 µM cebador inverso, 2 µl de ADNc, 0,5 µl de Taq DNA polymerase High Fidelity (Invitrogen Corporation, Carlsbad, CA, EE.UU), y 37,5 µl de H₂O. La amplificación fue realizada utilizando un ciclador térmico Peltier (Bio-Rad Laboratories, Inc. Hércules, CA, EE.UU) programado para desnaturalizar a 94°C durante 2 minutos; 30 ciclos cada uno a 94°C durante 40 segundos, 50°C durante 40 segundos, y 72°C durante 1 minuto; y una extensión final a 72°C durante 10 minutos.

20 [0312] Un producto PCR de aproximadamente 500 pares de bases fue detectado por 1% electroforesis en gel de agarosa usando búfer TBE. El fragmento de PCR fue escindido del gel y purificado utilizando ILLUSTRATE® GFX® PCR DNA y un equipo Gel Band Purification Kit (GE Healthcare, Buckinghamshire, UK) según las instrucciones del fabricante, directamente ordenadas, y se confirmó que era un gen parcial GH61A por explosión. En base a esta secuencia parcial, se diseñaron cebadores nuevos mostrados a continuación para clonar los terminales 5' y 3' utilizando un equipo Genome Walking Kit (Takara Bio Inc., Otsu, Shiga, Japón):

25

61ASPR1:

5'-TGCAAGGAGCAAGGTAGTTGA-3' (SEC ID nº 11)

30 61ASPR2:

5'-GAGTCCATTCCAGCTTGACGGT-3' (SEC ID nº 12)

61ASPF1:

35

5'-TCAGACAATCTGATAGCGGC-3' (SEC ID nº 13)

61ASPF2:

40 5'-ATCCCAACCACAACCTGCACCT-3' (SEC ID nº: 14)

45 [0313] Para la clonación del terminal 5' y del terminal 3', las amplificaciones primarias fueron compuestas de 2 µl de ADN genómico como modelo, 2,5 mM cada una de dATP, dTTP, dGTP, y dCTP, 100 pmol de AP2 (proporcionado por el equipo Genome Walking Kit) y 10 pmol del cebador 61ASPR1 para la clonación del terminal 5' o 100 pmol de AP3 (proporcionados por el equipo Genome Walking Kit), y 10 pmol de cebador 61ASPF1 para la clonación del terminal 3', 5 µl de 10X LA PCR Buffer II (proporcionada por el equipo Genome Walking Kit), y 2,5 unidades de TakaRa LATaq DNA polymerase (proporcionadas por el equipo Genome Walking Kit) en un volumen final de 50 µl. Las amplificaciones fueron realizadas utilizando un ciclador térmico de Peltier programado para pre-desnaturalizar a 94°C durante 1 minuto y 98°C durante 1 minuto; cinco ciclos cada uno a una temperatura de desnaturalización de 94°C durante 30 segundos; recocimiento a 60°C durante 1 minuto y alargamiento a 72°C durante 2 minutos; 1 ciclo de desnaturalización a 94°C durante 30 segundos; recocido a 25°C durante 3 minutos y alargamiento a 72°C durante 2 minutos; quince repetidores de 2 ciclos a 94°C durante 30 segundos, 62°C durante 1 minuto, y 72°C durante 2 minutos; seguidos de 1 ciclo a 94°C durante 30 segundos, 44°C durante 1 minuto, y 72°C durante 2 minutos; y una extensión final a 72°C durante 10 minutos. El bloque de calor luego pasó a un ciclo de mojado a 4°C.

55

60 [0314] Las amplificaciones secundarias estaban compuestas por 2 µl de 20X producto PCR primario diluido como modelos, 2,5 mM cada uno de dATP, dTTP, dGTP, y dCTP, 100 pmol de AP2, y 10 pmol de cebador 61ASPR2 para la clonación del terminal 5' o 100 pmol de AP3 y 10 pmol de cebador 61ASPF2 para la clonación del terminal 3', 5 µl de 10X LA PCR Buffer II, y 2,5 unidades de TakaRa LA Taq DNA polymerase en un volumen final de 50 µl. Las amplificaciones fueron realizadas utilizando un ciclador térmico de Peltier programado para quince repetidores de 2 ciclos de 94°C durante 30 segundos; 62°C durante 1 minuto; 72°C durante 2 minutos; seguidas de 1 ciclo a 94°C durante

ES 2 551 141 T3

30 segundos, 44°C durante 1 minuto, y 72°C durante 2 minutos; y una extensión final a 72°C durante 10 minutos. El bloque de calor luego pasó a un ciclo mojado a 4°C.

5 [0315] Los productos PCR del terminal 5' y 3' PCR fueron recuperados y secuenciados. Fueron identificados como el terminal 5' y terminal 3' del gen de polipéptido GH61A. Luego se ensamblaron las tres secuencias con el gen parcial, terminal 5', y terminal 3' para generar el GH61A en toda su longitud.

10 [0316] El gen en toda su longitud obtenido mostró que la secuencia contiene una región de codificación de 871 nucleótidos incluyendo 1 intrón y codón de terminación, y codifica 251 aminoácidos con un péptido señal predicho de 22 aminoácidos.

Ejemplo 3: clonación del gen de polipéptido GH61A de *Thermoascus crustaceus* CBS 181.67 a partir del ADN genómico y construcción de un vector de expresión de *Aspergillus oryzae*

15 [0317] En base a la secuencia genética GH61A de *Thermoascus crustaceus* en toda su longitud, se designaron cebadores oligonucleótidos, mostrados a continuación, para amplificar el gen GH61A de ADN genómico de *Thermoascus crustaceus* CBS 181,67. Un equipo IN-FUSION® CF Dry-Down Cloning Kit (Clontech Laboratories, Inc., Mountain View, CA, EE.UU) fue usado para clonar el fragmento directamente en el vector de expresión pPFJO355, sin la necesidad para digestión y ligamiento de restricción.

20 Cebador de sentido:

5'-ACACAACCTGGGGATCCACCATGGCCTTTTCCCAGATAATGGCTA-3' (SEC ID nº 15)

Cebador antisentido:

25

5'-GTCACCCTCTAGATCTGGATCGCAGGAGCGTTCAGA-3' (SEC ID nº 16)

Las letras en negrita representan la secuencia codificante para el cebador de sentido y la secuencia contracorriente del codón de terminación para el cebador antisentido.

30 La secuencia restante es homóloga a los sitios de inserción de pPFJO355.

[0318] El vector de expresión pPFJO355 contiene el promotor de TAKA amilasa de *Aspergillus oryzae*, elementos de terminador de glucoamilasa de *Aspergillus niger*, secuencias derivadas de pUC19 para la selección y propagación en el *E. coli*, y un gen *pyrG*, que codifica una decarboxilasa de orotidina de *Aspergillus nidulans* para la selección de un transformante de una cepa mutante en *pyrG* de *Aspergillus*.

35 [0319] Veinte picomoles de cada uno de los cebadores anteriores fueron usados en una reacción por PCR compuesta por ADN genómico de *Thermoascus crustaceus*, 10 µl de búfer 5X GC (Finnzymes Oy, Espoo, Finlandia), 1,5 µl de DMSO, 2 µl de 2,5 mM cada uno de dATP, dTTP dGTP, y dCTP, y 1 unidad de ADN polimerasa PHUSION™ High-Fidelity DNA Polymerase (Finnzymes Oy, Espoo, Finlandia) en un volumen final de 50 µl. La amplificación fue realizada utilizando un ciclador térmico de Peltier programado para desnaturalizar a 98°C durante 1 minuto, 5 ciclos de desnaturalización a 98°C durante 15 segundos, recocimiento a 70°C durante 30 segundos, con una reducción de 1 °C por ciclo, y alargamiento a 72°C durante 30 segundos, 25 ciclos cada uno a 98°C durante 15 segundos y 72°C durante 90 segundos; y una extensión final a 72°C durante 10 minutos. El bloque de calor luego fue llevado a un ciclo mojado a 4°C.

40 [0320] Los productos reactivos fueron aislados por 1,0% electroforesis en gel de agarosa usando 90 mM tris-borato y 1 mM búfer EDTA (TBE por sus siglas en inglés) donde una banda de producto de aproximadamente 1,0 kb fue escindida del gel, y purificada utilizando un ILLUSTRATE® GFX® PCR DNA y un equipo Gel Band Purification Kit según las instrucciones del fabricante.

45 [0321] El plásmido pPFJO355 fue digerido con Bam HI y Bgl II, aislado por 1,0% electroforesis en gel de agarosa usando búfer TBE, y purificado usando un ILLUSTRATE® GFX® PCR DNA y un equipo Gel Band Purification Kit según las instrucciones del fabricante.

50 [0322] El fragmento del gen y el vector digerido fueron ligados entre sí utilizando un equipo de clonación IN-FUSION® CF Dry-Down PCR Cloning Kit dando como resultado pGH61a51486 (figura 4) donde la transcripción del gen GH61A de *Thermoascus crustaceus* estaba bajo el control del promotor de TAKA alfa amilasa de *Aspergillus oryzae*. En resumen, 30 ng de pPFJO355 digeridos con *Bam* HI y *Bgl* II y 50 ng del producto PCR purificado del gen GH61A de *Thermoascus crustaceus* fueron añadidos a un frasco de reacción y resuspendidos en un volumen final de 10 µl con agua desionizada. La reacción fue incubada a 37°C durante 15 minutos y luego a 50°C durante 15 minutos. Tres µl de la reacción fueron

55

60

ES 2 551 141 T3

usados para transformar células competentes TOP10 de *E. coli* (TIANGEN Biotech (Pekín) Co. Ltd., Pekín, China) según las instrucciones del fabricante. La transformación fue extendida en placas LB suplementada con 100 µg de ampicilina por ml e incubadas a 37°C durante 1 día. Un transformante de *E. coli* que contenía un plásmido designado pGH61a51486 fue detectado por la colonia PCR y se preparó ADN plásmido utilizando un equipo QIAprep Spin Miniprep Kit (QIAGEN Inc., Valencia, CA, EE.UU.). El inserto de gen GH61A de *Thermoascus crustaceus* en pGH61a51486 fue confirmado mediante secuenciación del ADN utilizando un analizador 3730XL DNA Analyzer (Applied Biosystems Inc, Foster City, CA, EE.UU).

[0323] El mismo fragmento de PCR fue clonado en vector pGEM-T (Promega Corporation, Madison, WI, EE.UU) utilizando un sistema pGEM-T Vector System, para generar pGEM-T-GH61a51486. El inserto de gen GH61A de *Thermoascus crustaceus* en pGEM-T-GH61a51486 fue confirmado mediante secuenciación del ADN utilizando un analizador 3730XL DNA Analyzer. La cepa T-51486A de *E. coli*, designada NN059126, que contiene pGEM-T-GH61a51486, fue depositada en el Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM), Mascheroder Braunschweig, Alemania, el 10 de junio de 2009, y se le asignó el número de registro DSM 22656.

Ejemplo 4: caracterización de la secuencia genómica de *Thermoascus crustaceus* que codifica un polipéptido GH61A con actividad de aumento celulolítico

[0324] Los datos de secuencia de nucleótidos fueron escrutinados para observar la calidad y todas las secuencias fueron comparadas entre sí con ayuda del software PHRED/PHRAP (Universidad de Washington, Seattle, WA, EE.UU.).

[0325] La secuencia de nucleótidos (SEC ID nº 1) y la secuencia de aminoácidos deducida (SEC ID nº 2) del gen GH61A de *Thermoascus crustaceus* se muestra en las figuras 1A y 1B. La secuencia codificante es de 871 pares de bases incluyendo el codón de terminación y está interrumpida por un intrón de 115 pares de bases (nucleótidos 105-219). La proteína predicha codificada es de 251 aminoácidos. El % de contenido de G+C de la secuencia codificante en toda su longitud y de la secuencia codificante madura es 50,23% y 52,55%, respectivamente. Usando el programa de software SignalP (Nielsen et al., 1997, Protein Engineering 10: 1-6), se predijo un péptido señal de 22 residuos. La proteína madura predicha contiene 229 aminoácidos con una masa molecular predicha de 26,35 kDa.

[0326] Un alineamiento global por pares comparativo de secuencias de aminoácidos se determinó utilizando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, J. Mol. Biol. 48: 443-453) como viene implementado en el programa Needle de EMBOSS con penalización por apertura de gap de 10, penalización por extensión de gap de 0,5, y la matriz EBLOSUM62. La alineación mostró que la secuencia de aminoácidos deducida del polipéptido maduro del gen de polipéptido GH61A de *Thermoascus crustaceus* compartía el 86% de identidad (excluyendo los gaps) a la secuencia de aminoácidos deducida de un gen GH61A de *Thermoascus aurantiacus* (número de registro GeneSeq AEC05922 o AUP68836).

Ejemplo 5: expresión del gen de polipéptido GH61A de *Thermoascus crustaceus* CBS 181.67 en *Aspergillus oryzae*

[0327] Los protoplastos HowB101 de *Aspergillus oryzae* fueron preparados según el método de Christensen et al., 1988, Bio/Technology 6: 1419-1422 y transformados con 3 µg de pGH61a51486. La transformación dio como resultado aproximadamente 50 transformantes. Doce transformantes fueron aislados en placas de medio mínimo individuales.

[0328] Cuatro transformantes fueron inoculados separadamente en 3 ml de medio YPM en una placa de 24 pocillos e incubados a 30°C con agitación a 150 r.p.m. Después de 3 días de incubación, 20 µl de sobrenadante de cada cultivo fueron analizados utilizando un gel NUPAGE® NOVEX® 4-12% Bis-Tris Gel con ácido 2-(N-morfolino)etanosulfónico (MES por sus siglas en inglés) (Invitrogen Corporation, Carlsbad, CA, EE.UU.) según las instrucciones del fabricante. El gel resultante fue teñido con INSTANT BLUE™ (Expedeon Ltd., Babraham Cambridge, UK). Los perfiles SDS-PAGE de los cultivos mostraron que la mayoría de los transformantes tenían una gran banda de aproximadamente 30 kDa.

La cepa de expresión fue llamada *Aspergillus oryzae* EXP03151. Un cultivo oblicuo de *Aspergillus oryzae* EXP03151 fue lavado con 10 ml de medio YPM e inoculado en un matraz de 2 litros que contiene 400 ml de medio YPM para generar caldo para la caracterización de la enzima. El cultivo fue cosechado en el día 3 y filtrado utilizando una membrana DURAPORE® de 0,45 µm (Millipore, Bedford, MA, EE.UU.).

Ejemplo 6: preparación de micelios de *Thermoascus crustaceus* CBS 181.67 para la extracción total de ARN

[0329] La cepa de *Thermoascus crustaceus* fue inoculada sobre una placa PDA e incubada durante 4 días a 45°C en la oscuridad. Diferentes tapones de PDA micelial fueron inoculados en frascos de agitación de 500 ml que contienen 100 ml de medio NNCYP-PCS. Los matraces fueron incubados durante 6 días a 45°C con agitación a 160 r.p.m. Los micelios fueron recogidos en el día 4, día 5, y día 6. Luego los micelios de cada día fueron combinados y congelados en nitrógeno líquido, y luego almacenados en un congelador a -80°C hasta su uso.

ES 2 551 141 T3

Ejemplo 7: preparación de ARN de *Thermoascus crustaceus* CBS 181.67

5 [0330] Los micelios congelados obtenidos en el ejemplo 6 fueron transferidos a un mortero y mano de mortero preenfriado con nitrógeno líquido y molidos hasta obtener un polvo fino. El ARN total fue obtenido a partir de los micelios en polvo mediante extracción con reactivo TRIZOL® (Invitrogen, Carlsbad, CA, EE.UU) y purificado utilizando un equipo RNEASY® Mini Kit según el protocolo del fabricante. Cincuenta microgramos de ARN total fueron sometidos a secuenciación como se describe en el ejemplo 8.

10 Ejemplo 8: ensamblaje y extracción de secuencia

15 [0331] El ARN total enriquecido para secuencias poliA con el protocolo mRNASeq fue ordenado utilizando un sistema ILLUMINA® GA2 System (Illumina, Inc., San Diego, CA, EE.UU). Las lecturas raw de 36 pares de bases fueron ensambladas con un ensamblador doméstico. Las secuencias ensambladas fueron analizadas utilizando métodos de bioinformática estándar para la predicción de hallazgo y funcional de genes. Se usó ESTscan 2.0 para predicción de gen. Se usaron NCBI blastall versión 2.2.10 y HMMER versión 2.1.1 para predecir la función basada en homología estructural. Los candidatos de la familia GH61 fueron identificados directamente mediante análisis de los resultados de Blast.

20 Ejemplo 9: clonación del gen de polipéptido GH61A de *Thermoascus crustaceus* a partir de ADN genómico y construcción de un vector de expresión de *Aspergillus oryzae*

25 [0332] En base a la información de secuenciación ILLUMINA®, los cebadores oligonucleótidos, mostrados a continuación, fueron diseñados para amplificar el gen de polipéptido GH61 B a partir de ADN genómico de *Thermoascus crustaceus* CBS 181.67. Se usó un equipo IN-FUSION® CF Dry-Down Cloning Kit para clonar el fragmento directamente en el vector de expresión pPFJO355, sin necesidad de digestión y ligamiento de restricción. Cebador de sentido:

30 5'-ACACAACCTGGGGATCCACCATGTCATTCTCGAAGATACTTGCTA-3' (SEC ID nº 17)

Cebador antisentido:

5'-GTCACCCTCTAGATCTCATCTCGTCTTTTCGTATCAGTGA-3' (SEC ID nº 18)

35 Las letras en negrita representan la secuencia codificante. La secuencia restante es homóloga a los sitios de inserción de pPFJO355.

40 [0333] Diez picomoles de cada uno de los cebadores anteriores fueron usados en una reacción PCR compuesta de ADN genómico de *Thermoascus crustaceus* CBS 181.67, 10 µl de 5X GC búfer, 1,5 µl de DMSO, 2 µl de 2,5 mM de cada uno de dATP, dTTP dGTP, y dCTP, y 0,6 unidades de ADN polimerasa PHUSION™ High-Fidelity DNA Polymerase en un volumen final de 50 µl. La amplificación fue realizada utilizando un ciclador térmico de Peltier programado para desnaturalizar a 98°C durante 1 minuto, 5 ciclos de desnaturalización a 98°C durante 15 segundos, recocimiento a 58°C durante 30 segundos, con un 1°C aumento por ciclo, y alargamiento a 72°C durante 60 segundos; 25 ciclos cada uno a 98°C durante 15 segundos, 65°C durante 30 segundos y 72°C durante 60 segundos, y una extensión final a 72°C durante 10 minutos. El bloque de calor luego fue llevado a un ciclo mojado a 4° C.

45 [0334] Los productos reactivos fueron aislados por 1,0% electroforesis en gel de agarosa usando búfer TBE donde una banda de producto de aproximadamente 1,2 kb fue escindida del gel, y purificada utilizando un equipo ILLUSTRATE® GFX® PCR DNA y un equipo Gel Band Purification Kit según las instrucciones del fabricante.

50 [0335] El plásmido pPFJO355 fue digerido con *Bam* HI y *Bgl* II, aislado por 1,0% electroforesis en gel de agarosa usando búfer TBE, y purificado utilizando ILLUSTRATE® GFX® PCR DNA y un equipo Gel Band Purification Kit según las instrucciones del fabricante.

55 [0336] El fragmento de gen y el vector digerido fueron ligados entre sí utilizando un equipo IN-FUSION® CF Dry-Down PCR Cloning Kit dando como resultado pGH61D14YF (figura 5) donde la transcripción del gen de polipéptido GH61B de *Thermoascus crustaceus* estaba bajo el control de un promotor del gen alfa-amilasa de *Aspergillus oryzae*. En resumen, 30 ng de pPFJO355 digeridos con *Bam* HI y *Bgl* II y 240 ng del producto PCR purificado del gen GH61B de *Thermoascus crustaceus* fueron añadidos a un frasco de reacción y resuspendidos en un volumen final de 10 µl con agua desionizada. La reacción fue incubada a 37°C durante 15 minutos y luego a 50°C durante 15 minutos. Tres µl de la reacción fueron usados para transformar células competentes TOP10 de *E. coli* según las instrucciones del fabricante.

60

ES 2 551 141 T3

La transformación fue extendida en placas LB suplementadas con 100 µg de ampicilina por ml e incubadas a 37°C durante 1 día. Un transformante de *E. coli* con un plásmido designado pGH61 D14YF fue detectado por la colonia PCR y se preparó ADN plásmido utilizando un equipo QIAprep Spin Miniprep Kit. El inserto de gen de polipéptido GH61B de *Thermoascus crustaceus* en pGH61 D14YF fue confirmado mediante secuenciación del ADN utilizando un analizador 3730XL DNA Analyzer.

[0337] El mismo fragmento de PCR fue clonado en vector pGEM-T utilizando un sistema pGEM-T Vector System, para generar pGEM-T-GH61 D14YF. El gen GH61B de *Thermoascus crustaceus* insertado en pGEM-T-GH61 D14YF fue confirmado mediante secuenciación del ADN que utiliza un analizador 3730XL DNA Analyzer. La cepa T-51486B de *E. coli*, designada NN059120, que contiene pGEM-T-GH61 DYF, fue depositada en el Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Braunschweig, Alemania, el 10 de junio de 2009, y se le asignó el número de registro DSM 22654.

Ejemplo 10: caracterización de la secuencia genómica de *Thermoascus crustaceus* que codifica un polipéptido GH61B con actividad de aumento celulolítico

[0338] Los datos de secuencia de nucleótidos fueron escrutinados para observar la calidad y todas las secuencias fueron comparadas entre sí con ayuda del software PHRED/PHRAP (Universidad de Washington, Seattle, WA, EE.UU.).

[0339] La secuencia de nucleótidos (SEC ID nº 3) y la secuencia de aminoácidos deducida (SEC ID nº 4) del gen GH61B de *Thermoascus crustaceus* se muestran en las figuras 2A e 2B. El fragmento genómico codifica un polipéptido de 349 aminoácidos, interrumpido por 1 intrón de 51 pares de bases (nucleótidos 102-153). El contenido en % de G+C de la secuencia codificante en toda su longitud y de la secuencia codificante madura son 52,86% y 53,58%, respectivamente. Usando el programa de software SignalP (Nielsen et al., 1997, *supra*), se predijo un péptido señal de 18 residuos. La proteína madura predicha contiene 229 aminoácidos con una masa molecular predicha de 26,35 kDa.

[0340] Un alineamiento global por pares comparativo de secuencias de aminoácidos fue determinado utilizando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, *supra*) como viene implementado en el programa Needle de EMBOSS con penalización por apertura de gap de 10, penalización por extensión de gap de 0,5, y la matriz EBLOSUM62. La alineación mostró que la secuencia de aminoácidos deducida del polipéptido maduro del gen de polipéptido GH61B de *Thermoascus crustaceus* compartía el 75% de identidad (excluyendo los gaps) a la secuencia de aminoácidos deducida de un gen de polipéptido GH61A de *Thermoascus aurantiacus* (número de registro GeneSeq AUP68836).

Ejemplo 11: expresión del gen de polipéptido GH61B de *Thermoascus crustaceus* en un hésped *Aspergillus oryzae*

[0341] Los protoplastos HowB101 de *Aspergillus oryzae* fueron preparados según el método de Christensen et al., 1988, *supra*, y transformados con 3 µg de pGH61 D14YF. La transformación dio como resultado aproximadamente 50 transformantes. Treinta transformantes fueron aislados en placas de medio mínimo individuales.

[0342] Los treinta transformantes fueron inoculados separadamente en 3 ml de medio YPM en una placa de 24 pocillos e incubados a 30°C, 150 r.p.m. Después de 3 días de incubación, se analizaron 20 µl de sobrenadante de cada cultivo utilizando un NUPAGE® NOVEX® 4-12% Bis-Tris Gel con MES según las instrucciones del fabricante. El gel resultante fue teñido con INSTANT BLUE™. Los perfiles SDS-PAGE perfiles de los cultivos mostraron que la mayoría de los transformantes tenían una gran banda de aproximadamente 55 kDa. La cepa de expresión fue designada *Aspergillus oryzae* EXP03090.

[0343] Un cultivo oblicuo de *Aspergillus oryzae* EXP03090 con 10 ml de YPM e inoculado en un matraz de 2 litros que contiene 400 ml de medio YPM para generar caldo para caracterización.

El cultivo fue cosechado en el día 3 y filtrado utilizando una membrana DURAPORE® Membrane de 0,45 µm.

Ejemplo 12: clonación del gen de polipéptido GH61C de *Thermoascus crustaceus* a partir de ADN genómico y construcción de un vector de expresión de *Aspergillus oryzae*

[0344] En base a la información de secuenciación de ILLUMINA® en el ejemplo 8, los cebadores oligonucleótidos, mostrados a continuación, fueron diseñados para amplificar el gen de polipéptido GH61C a partir de ADN genómico de *Thermoascus crustaceus* CBS 181.67. Se usó un equipo IN-FUSION™ CF Dry-down Cloning Kit para clonar el fragmento directamente en el vector de expresión pPFJO355, sin necesidad de digestión y ligamiento de restricción. Cebador de sentido:

5'-ACACAACCTGGGGATCCACCATGTTGTCAATTCACCCACCAAGTCA-3' (SEC ID nº: 19)

ES 2 551 141 T3

Cebador antisentido:

5'- GTCACCCTCTAGATCT**ACACCCTCATGCACCTCTCCTTCTAA** -3' (SEC ID nº: 20)

5

Las letras en negrita representan secuencia codificante. La secuencia restante es homóloga a los sitios de inserción de pPFJO355.

[0345] Diez picomoles de cada uno de los cebadores anteriores fueron usados en una reacción por PCR compuesta por ADN genómico de *Thermoascus crustaceus* CBS 181.67, 10 µl de 5X búfer GC, 1,5 µl de DMSO, 2,5 mM de cada uno de dATP, dTTP, dGTP, y dCTP, y 0,6 unidades de ADN polimerasa PHUSION™ High-Fidelity DNA Polymerase, en un volumen final de 50 µl. La amplificación fue realizada utilizando un ciclador térmico de Peltier programado para desnaturalizar a 98°C durante 1 minuto; 5 ciclos de desnaturalización a 98°C durante 15 segundos, recocimiento a 58°C durante 30 segundos, con un aumento de 1°C por ciclo y alargamiento a 72°C durante 60 segundos; 25 ciclos cada uno a 98°C durante 15 segundos, 65°C durante 30 segundos, y 72°C durante 60 segundos; y una extensión final a 72°C durante 10 minutos. El bloque de calor luego fue llevado a un ciclo mojado a 4°C.

10
15

[0346] Los productos reactivos fueron aislados por 1,0% electroforesis en gel de agarosa usando un búfer TBE donde una banda de producto de aproximadamente 1,5 kb fue escindida del gel, y purificada utilizando un ILLUSTRATE® GFX® PCR ADN y un Gel Band Purification Kit según las instrucciones del fabricante.

20

[0347] El plásmido pPFJO355 fue digerido con *Bam* I y *Bgl* II, aislado por 1,0% electroforesis en gel de agarosa utilizando búfer TBE, y purificado utilizando un ILLUSTRATE® GFX® PCR ADN y un equipo Gel Band Purification Kit según las instrucciones del fabricante.

25

[0348] El fragmento de gen y el vector digerido fueron ligados entre sí utilizando un equipo IN-FUSION™ CF Dry-down PCR Cloning Kit dando como resultado pGH61D14YH (figura 6) donde la transcripción del gen de polipéptido GH61C estaba bajo el control del promotor de TAKA alfa-amilasa de *Aspergillus oryzae*. En resumen, 30 ng de pPFJO355 digeridos con *Bam* I y *Bgl* II, y 300 ng del producto de PCR purificado del gen de polipéptido GH61C de *Thermoascus crustaceus* fueron añadidos a un frasco de reacción y resuspendidos en un volumen final de 10 µl con adición de agua desionizada. La reacción fue incubada a 37°C durante 15 minutos y luego a 50°C durante 15 minutos. Tres µl de la reacción fueron usados para transformar células competentes TOP10 de *E. coli* (TIANGEN Biotech Co. Ltd., Pekín, China). Un transformante de *E. coli* que contiene pGH61D14YH fue detectado por la colonia PCR y se preparó ADN plásmido utilizando un equipo QIAprep Spin Miniprep Kit. El inserto de gen GH61C de *Thermoascus crustaceus* en pGH61 D14YH fue confirmado mediante secuenciación del ADN que utiliza un analizador 3730XL DNA Analyzer.

30
35

[0349] El mismo fragmento de PCR fue clonado en vector pGEM-T (Promega Corporation, Madison, WI, EE.UU) utilizando un sistema pGEM-T Vector System, para generar pGEM-T-GH61 D14YH. El inserto de gen de polipéptido GH61C de *Thermoascus crustaceus* en pGEM-T-GH61D14YH fue confirmado mediante secuenciación del ADN que utiliza un analizador 3730XL DNA Analyzer. La cepa T-51486C de *E. coli*, designada NN059127, que contiene pGEM-T-GH61D14YH, fue depositada en el Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Braunschweig, Alemania, el 10 de junio de 2009, y se le asignó el número de registro DSM 22657.

40

Ejemplo 13: caracterización del gen de polipéptido GH61C de *Thermoascus crustaceus*

45

[0350] Los datos de secuencia de nucleótidos fueron escrutinados para observar la calidad y todas las secuencias fueron comparadas entre sí con ayuda del software PHRED/PHRAP (Universidad de Washington, Seattle, WA, EE.UU.).

[0351] La secuencia de nucleótidos (SEC ID nº 3) y secuencia de aminoácidos deducida (SEC ID nº 4) del gen GH61C de *Thermoascus crustaceus* se muestran en las figuras 3A e 3B. El fragmento genómico codifica un polipéptido de 436 aminoácidos, interrumpido por 2 intrones de 80 pares de bases (nucleótidos 200-279) y 95 pares de bases (nucleótidos 1134-1228). El contenido % de G+C de la secuencia codificante en toda su longitud y de la secuencia codificante madura es 53,16 % y 55,61%, respectivamente. Usando el programa de software SignalP (Nielsen et al., 1997, *supra*), se predijo un péptido señal de 23 residuos. La proteína madura predicha contiene 413 aminoácidos con una masa molecular predicha de 45,08 kDa.

50

55

[0352] Un alineamiento global por pares comparativo de secuencias de aminoácidos fue determinado utilizando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, *supra*) como viene implementado en el programa Needle de EMBOSS con penalización por apertura de gap de 10, penalización por extensión de gap de 0,5, y la matriz EBLOSUM62. La alineación mostró que la secuencia de aminoácidos deducida del polipéptido maduro del gen de polipéptido GH61B de *Thermoascus crustaceus* comparte el 75% de identidad (excluyendo los gaps) a la secuencia de

60

ES 2 551 141 T3

aminoácidos deducida de un gen GH61A de *Thermoascus aurantiacus* (número de registro GeneSeq AUP68836).

Ejemplo 14: hidrólisis de rastrojos de maíz pretratados por polipéptidos GH61A y GH61B de *Thermoascus crustaceus* con actividad de aumento celulolítico

5

[0353] Los caldos de cultivo para los polipéptidos GH61A y GH61B de *T. crustaceus* con actividad de aumento celulolítico fueron preparados como se describe en el ejemplo 5 y ejemplo 11, respectivamente, y cada uno fue concentrado aproximadamente en 20 veces utilizando un dispositivo de ultrafiltración Amicon (Millipore, Bedford, MA, EE.UU, 10 kDa membrana de polietersulfona, 40 psi, 4°C). La concentración de proteína fue estimada mediante densitometría a continuación del SDS-PAGE y tinción con Coomassie blue. Los rastrojos de maíz fueron pretratados y preparados como un sustrato de ensayo como se describe en el documento WO 2005/074647 para generar rastrojos de maíz pretratados (PCS por sus siglas en inglés). La mezcla de celulosa de base usada para ensayar la actividad de aumento fue preparada a partir de la cepa SMA135 de *Trichoderma reesei* (WO 2008/057637).

10

15

[0354] La hidrólisis de PCS fue llevada a cabo usando placas de pocillos de 1,6 ml de profundidad (Axygen, Santa Clara, CA, EE.UU) utilizando un volumen de reacción total de 1,0 ml y una concentración PCS de 50 mg/ml en 1 mM sulfato de manganeso - 50 mM acetato sódico a pH 5,0. El polipéptido GH61B de *Thermoascus crustaceus* fue añadido separadamente a la mezcla de celulosa de base a concentraciones que varían del 0 al 100% de la concentración de proteína de la mezcla de celulosa de base. La incubación se realizó a 50°C durante 72 horas. Se realizaron ensayos por triplicado. Las partes alicuotas fueron centrifugadas, y el líquido sobrenadante fue filtrado por centrifugado (MULTISCREEN® HV 0,45 µm, Millipore, Billerica, MA, EE.UU) a 3000 r.p.m. durante 10 minutos utilizando un centrifugador de placa (SORVALL® RT7, Thermo Fisher Scientific, Waltham, MA, EE.UU.). Cuando no fueron usadas inmediatamente, las partes alicuotas filtradas del hidrolizado fueron congeladas a -20°C. Las concentraciones de azúcar de las muestras diluidas en 0,005 M H₂SO₄ con 0,05% p/p ácido benzoico fueron medidas después de la elución por 0,005 M H₂SO₄ con 0,05% p/p ácido benzoico a una velocidad de flujo de 0,6 ml/minutos de una columna AMINEX® HPX-87H de 4,6 x 250 mm (Bio-Rad Laboratories, Inc., Hércules, CA, EE.UU) a 65°C con cantidad mediante integración de señales de glucosa y celobiosa a partir de detección de índice de refracción (CHEMSTATION®, AGILENT® 1100 HPLC, Agilent Technologies, Santa Clara, CA, EE.UU) calibrado mediante muestras de azúcar puro (Absolute Standards Inc., Hamden, CT, EE.UU). Los equivalentes resultantes fueron usados para calcular el porcentaje de conversión de celulosa para cada reacción. El grado de conversión de celulosa en glucosa más azúcares de celobiosa (conversión; %) fue calculado utilizando la siguiente ecuación:

20

25

30

Conversión (%) = (glucosa + celobiosa x 1,053) (mg/ml) x 100 x 162 / (Celulosa (mg/ml) x 180) = (glucosa + celobiosa x 1,053) (mg/ml) x 100 / (Celulosa (mg/ml) x 1,111)

35

En esta ecuación el factor 1,111 refleja el aumento de peso en la conversión de celulosa en glucosa, y el factor 1,053 refleja el aumento de peso en la conversión de celobiosa en glucosa. La celulosa en PCS fue determinada mediante una digestión de PCS de límite para liberar glucosa y celobiosa.

40

[0355] El resultado de añadir cantidades en aumento del polipéptido GH61A de *Thermoascus crustaceus* a la mezcla de celulosa de base se muestra en la figura 7. La adición del polipéptido GH61A de *Thermoascus crustaceus* proporcionó un factor de estimulación de 1,37 a un 100% de nivel de adición.

45

[0356] El resultado de añadir cantidades en aumento del polipéptido GH61B del *Thermoascus crustaceus* a la mezcla de celulosa de base se muestra en la figura 8. La adición del polipéptido GH61B de *Thermoascus crustaceus* proporcionó un factor de estimulación de 1,29 a un 100% de nivel de adición.

Depósito de material biológico

50

[0357] El siguiente material biológico ha sido depositado según las condiciones del Tratado de Budapest con el Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ), Mascheroder Weg 1 B, D-38124 Braunschweig, Alemania, y asignaron el siguiente número de acceso:

Depósito	Número de acceso	Fecha de depósito
<i>E. coli</i> (NN059126)	DSMZ 22656	10 de junio de 2009
<i>E. coli</i> (NN059120)	DSMZ 22654	10 de junio de 2009
<i>E. coli</i> (NN059127)	DSMZ 22657	10 de junio de 2009

55

[0358] Las cepas han sido depositadas bajo condiciones que aseguran que el acceso a los cultivos estará disponible durante la pendencia de esta solicitud de patente para un determinado por leyes de patente extranjera a ser autorizadas a ella. Los depósitos representan cultivos sustancialmente puro de las tensiones depositadas. Los depósitos están

ES 2 551 141 T3

disponibles según sea necesario por leyes de patente extranjera en países donde duplicados de la aplicación de sujeto, o su descendiente son solicitados. No obstante, se debe entender que la disponibilidad de un depósito no constituye una licencia para practicar la presente invención en dérogas de derecho resultante de las patentes ha concedido por acción gubernativa.

5

Listado de secuencias

[0359]

10 <110> Novozymes, Inc.

Novozymes A/S

Tang, Lan

Liu, Ye

Duan, Junxin

15 Kramer, Randall

Wu, Wenping

<120> Polipéptidos con actividad de aumento celulolítico y polinucleótidos que codifican la misma

20 <130> 11667-WO-PCT

<150> US 61/247,250

<151> 2009-09-30

25 <160> 20

<170> Versión de PatentIn 3.5

<210> 1

30 <211> 871

<212> ADN

<213> Thermoascus crustaceus

<400> 1

ES 2 551 141 T3

```

atggcctttt cccagataat ggctattacc ggcgtttttc ttgcctctgc ttcctgggtg      60
gctggccatg gctttgttca gaatatcgtg attgatggta aaaggtacct aactacctac      120
cttactatct gatgtcattt acaagaaagg gcacagacac aagcggcaaa aaaaagaaag      180
aaagaaagaa agaaagaaag ctgacaaaaa ttcaacaagt tatggcgggt acatcgtgaa      240
ccaatatcca tacatgtcag atcctccgga ggtcgtcggc tggctacca ccgcaaccga      300
cctcggattc gtggacggta ccgatacca aggacctgat atcatctgcc acaggggcgc      360
caagcctgca gccctgactg ccaagtggc cgccggagga accgtcaagc tggaatggac      420
tccatggcct gattctcacc acggcccggg gatcaactac cttgctcctt gcaacgggta      480
ctgttccacc gtggacaaga cccaattgaa attcttcaag atcgcccagg cgggtctcat      540
cgatgacaac agtcctcctg gtatctgggc ctcaacaat ctgatagcgg ccaacaacag      600
ctggactgtc accatcccaa ccacaactgc acctggaaac tatgttctaa ggcatgagat      660
cattgctctc cactcagctg ggaacaagga tggcgcgag aactatcccc agtgcacaa      720
cctgaaggtc actgaaaatg gttctggcaa tcctcctgct ggtgctcttg gaacggcact      780
ctacaaggat acagatccgg gaattctgat caatatctac cagaaacttt ccagctatgt      840
tattcctggg cctgctttgt aactggta g      871

```

<210> 2

<211> 251

5 <212> PRT

<213> Thermoascus crustaceus

<400> 2

ES 2 551 141 T3

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

Ala Ser Leu Val Ala Gly His Gly Phe Val Gln Asn Ile Val Ile Asp
 20 25 30

Gly Lys Ser Tyr Gly Gly Tyr Ile Val Asn Gln Tyr Pro Tyr Met Ser
 35 40 45

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

Phe Val Asp Gly Thr Gly Tyr Gln Gly Pro Asp Ile Ile Cys His Arg
 65 70 75 80

Gly Ala Lys Pro Ala Ala Leu Thr Ala Gln Val Ala Ala Gly Gly Thr
 85 90 95

Val Lys Leu Glu Trp Thr Pro Trp Pro Asp Ser His His Gly Pro Val
 100 105 110

Ile Asn Tyr Leu Ala Pro Cys Asn Gly Asp Cys Ser Thr Val Asp Lys
 115 120 125

Thr Gln Leu Lys Phe Phe Lys Ile Ala Gln Ala Gly Leu Ile Asp Asp
 130 135 140

Asn Ser Pro Pro Gly Ile Trp Ala Ser Asp Asn Leu Ile Ala Ala Asn
 145 150 155 160

Asn Ser Trp Thr Val Thr Ile Pro Thr Thr Thr Ala Pro Gly Asn Tyr
 165 170 175

Val Leu Arg His Glu Ile Ile Ala Leu His Ser Ala Gly Asn Lys Asp
 180 185 190

Gly Ala Gln Asn Tyr Pro Gln Cys Ile Asn Leu Lys Val Thr Gly Asn
 195 200 205

Gly Ser Gly Asn Pro Pro Ala Gly Ala Leu Gly Thr Ala Leu Tyr Lys
 210 215 220

Asp Thr Asp Pro Gly Ile Leu Ile Asn Ile Tyr Gln Lys Leu Ser Ser
 225 230 235 240

Tyr Val Ile Pro Gly Pro Ala Leu Tyr Thr Gly
 245 250

<210> 3

ES 2 551 141 T3

<211> 1102
 <212> ADN
 <213> Thermoascus crustaceus

5 <400> 3

atgtcattct cgaagatact tgctatcgct ggggccatta cctacgcatac ttcagctgcc	60
gctcatgggt atgtccaggg aattgtgtc gatggcagct agtatgtcac tctggatgga	120
accttcagca cgtactgtac taacaatcag cagctacggg ggatatatgg tgaccaata	180
tccctacacc gctcaacctc cggaactcat cgcctggtcc actaaagcaa ccgatcttg	240
gtttgtggac ggcagtggct atacttctcc tgatatcacc tgccataaag gtgctgagcc	300
tggtgcccag agcgccaaag tggcagctgg agggaccggt gagctgcagt ggacggcatg	360
gcccaggtct cacaagggcc cagttattga ctacctcgcc gcctgcgacg gggactgctc	420
atctgttgat aagactgcac taaagtctt taagattgac gagagtggc tgattgacgg	480
caacggtgct ggaacatggg cctctgatac gttgatcaaa aataacaaca gctggactgt	540
caccatccca agcacaattg cttccggaaa ctacgtacta agacacgaaa taattgcgct	600
ccattctgcc ggaacaaaag atggtgctca gaactatccc cagtgtatca acctcgaggt	660
cactggtagt ggcaccgaaa accctgctgg cactctogga acagcgcttt acacagacac	720
tgatcctggc cttctggca acatctacca gggctctgccc aactattcaa tcctggtcc	780
tgctctgtat agcggcaaca gtgataacgc tggttcctc aaccctacca ccacgcccgc	840
aattcagaat gctgctgctg ctccctccac ttccacagca tctgtgtca ctgattcttc	900
gtcagccacc cagactgcta gtgtcgccgc cagactcca gcctccactt cggctgttac	960
agcctcacca gctcccgata ctggaagcga cgtaacccaaa tatctggatt cgatgagctc	1020
ggatgaggtc ctcaccctgg tgcgcgggac cctgtcttgg ctggtttcta acaagaaaca	1080
tgcgcgggat ctttctcact ga	1102

10 <210> 4
 <211> 349
 <212> PRT
 <213> Thermoascus crustaceus

<400> 4

ES 2 551 141 T3

Met Ser Phe Ser Lys Ile Leu Ala Ile Ala Gly Ala Ile Thr Tyr Ala
1 5 10 15

Ser Ser Ala Ala Ala His Gly Tyr Val Gln Gly Ile Val Val Asp Gly
20 25 30

Ser Tyr Tyr Gly Gly Tyr Met Val Thr Gln Tyr Pro Tyr Thr Ala Gln
35 40 45

Pro Pro Glu Leu Ile Ala Trp Ser Thr Lys Ala Thr Asp Leu Gly Phe

ES 2 551 141 T3

50						55										60
Val	Asp	Gly	Ser	Gly	Tyr	Thr	Ser	Pro	Asp	Ile	Ile	Cys	His	Lys	Gly	
65					70					75					80	
Ala	Glu	Pro	Gly	Ala	Gln	Ser	Ala	Lys	Val	Ala	Ala	Gly	Gly	Thr	Val	
				85					90					95		
Glu	Leu	Gln	Trp	Thr	Ala	Trp	Pro	Glu	Ser	His	Lys	Gly	Pro	Val	Ile	
			100					105					110			
Asp	Tyr	Leu	Ala	Ala	Cys	Asp	Gly	Asp	Cys	Ser	Ser	Val	Asp	Lys	Thr	
		115					120					125				
Ala	Leu	Lys	Phe	Phe	Lys	Ile	Asp	Glu	Ser	Gly	Leu	Ile	Asp	Gly	Asn	
	130					135					140					
Gly	Ala	Gly	Thr	Trp	Ala	Ser	Asp	Thr	Leu	Ile	Lys	Asn	Asn	Asn	Ser	
145					150					155					160	
Trp	Thr	Val	Thr	Ile	Pro	Ser	Thr	Ile	Ala	Ser	Gly	Asn	Tyr	Val	Leu	
				165					170					175		
Arg	His	Glu	Ile	Ile	Ala	Leu	His	Ser	Ala	Gly	Asn	Lys	Asp	Gly	Ala	
			180					185					190			
Gln	Asn	Tyr	Pro	Gln	Cys	Ile	Asn	Leu	Glu	Val	Thr	Gly	Ser	Gly	Thr	
		195					200					205				
Glu	Asn	Pro	Ala	Gly	Thr	Leu	Gly	Thr	Ala	Leu	Tyr	Thr	Asp	Thr	Asp	
	210					215					220					
Pro	Gly	Leu	Leu	Val	Asn	Ile	Tyr	Gln	Gly	Leu	Ser	Asn	Tyr	Ser	Ile	
225					230					235					240	
Pro	Gly	Pro	Ala	Leu	Tyr	Ser	Gly	Asn	Ser	Asp	Asn	Ala	Gly	Ser	Leu	
				245					250					255		
Asn	Pro	Thr	Thr	Thr	Pro	Ser	Ile	Gln	Asn	Ala	Ala	Ala	Ala	Pro	Ser	
			260					265					270			
Thr	Ser	Thr	Ala	Ser	Val	Val	Thr	Asp	Ser	Ser	Ser	Ala	Thr	Gln	Thr	
		275					280					285				
Ala	Ser	Val	Ala	Ala	Thr	Thr	Pro	Ala	Ser	Thr	Ser	Ala	Val	Thr	Ala	
	290					295					300					
Ser	Pro	Ala	Pro	Asp	Thr	Gly	Ser	Asp	Val	Thr	Lys	Tyr	Leu	Asp	Ser	
305					310					315					320	

ES 2 551 141 T3

Met Ser Ser Asp Glu Val Leu Thr Leu Val Arg Gly Thr Leu Ser Trp
325 330 335

Leu Val Ser Asn Lys Lys His Ala Arg Asp Leu Ser His
340 345

<210> 5

<211> 1493

5 <212> ADN

<213> Thermoascus crustaceus

<400> 5

ES 2 551 141 T3

```

atggtgcat tcattccac caagtcagct gcgctgacga ctcttctact tcttggaca    60
gctcatgctc aactttgat gaccaccatg tttgtggacg gcgtcaacca gggagatggt    120
gtctgcattc gcatgaacaa tgacggcgga actgccaata cctatatcca gcctatcacg    180
agcaaggata tcgctcgcg taagtaccca gatgtcatca tactctgcca taacatccgt    240
catatctact agaatcggag caatgttaag tatttccagg catccaaggc gaaatcggcg    300
cctcccagat ctgccagtc aaggcatctt ccaccctaac cttccaattc cgcgagcaac    360
ccaacaaccc aaactcctcc cctctcgatc catcgcaaaa aggccccgcc gcggtgtacc    420
tgaaaaaggt cgactccgcc atcgcgagca acaacgccgc cggagacagc tggttcaaga    480
tctgggagtc cgtctacgac gaggccacgg gcaaatgggg cacgaccaag atgatcgaga    540
acaacgggca catctccgct aagggtgccc atgatatcga ggggtggttac tatcttgccc    600
ggacggagct gctggcgcta cattctgctg atcaggggga tccgcagttc tatggttggt    660
gtgctgagct gtttatcgat tcggatggga cggcgaaacc gccactggt tctattggag    720
aggggacgta cgatctgagc atgcctgcca tgacgtataa tatctgggag acaccgttg    780
ctctgccgta tccgatgat gggcctcctg tctatacgcc tggtctggt tctggatcag    840
tccgtgcgac gagctcttct gctgtcccta ctgcaaccga atcctctttt gtagaggaaa    900
gagcaaacc cgtcacggca aacagtgttt attctgcaag gggcaaattc aaaacctgga    960
ttgataaact gtcattggcg ggaaggtcc gtgagaacgt cagacaagcc gcgggaagaa   1020
gaagcactct cgtccagact gtgggtctaa agccaaaagg ctgcatcttc gtcaatggaa   1080
actggtgctg cttcgagggt cccgactaca acgatgcgga gagctgctgg gctgtatggt   1140
cccctcctta gcctcttaca tccctaagta ctacatttga aaacaacaaa aagaaatgta   1200
tatactaact acgtacgctc tactctaggc ctccgacaac tgctggaaac agtccgacgc   1260
ctgctggaac aagaccaaac ccacgggcta caataactgc cagatctggc aggacaagaa   1320
atgcaaggtc atccaggatt cctgtagcgg acccaaccg catggaccac cgaataaggg   1380
caaggatttg actccggagt ggccgccact gaagggtctg atggatacgt tctccaagcg   1440
tactatcggg taccgcgatt ggattgtag aaggagaggt gcatgagggt gta           1493

```

<210> 6
 <211> 436
 <212> PRT
 <213> Thermoascus crustaceus

 <400> 6

5

ES 2 551 141 T3

Met Leu Ser Phe Ile Pro Thr Lys Ser Ala Ala Leu Thr Thr Leu Leu
 1 5 10 15

Leu Leu Gly Thr Ala His Ala His Thr Leu Met Thr Thr Met Phe Val
 20 25 30

Asp Gly Val Asn Gln Gly Asp Gly Val Cys Ile Arg Met Asn Asn Asp
 35 40 45

Gly Gly Thr Ala Asn Thr Tyr Ile Gln Pro Ile Thr Ser Lys Asp Ile
 50 55 60

Ala Cys Gly Ile Gln Gly Glu Ile Gly Ala Ser Arg Val Cys Pro Val
 65 70 75 80

Lys Ala Ser Ser Thr Leu Thr Phe Gln Phe Arg Glu Gln Pro Asn Asn
 85 90 95

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

Tyr Leu Lys Lys Val Asp Ser Ala Ile Ala Ser Asn Asn Ala Ala Gly
 115 120 125

Asp Ser Trp Phe Lys Ile Trp Glu Ser Val Tyr Asp Glu Ser Thr Gly
 130 135 140

Lys Trp Gly Thr Thr Lys Met Ile Glu Asn Asn Gly His Ile Ser Val
 145 150 155 160

Lys Val Pro Asp Asp Ile Glu Gly Gly Tyr Tyr Leu Ala Arg Thr Glu
 165 170 175

Leu Leu Ala Leu His Ser Ala Asp Gln Gly Asp Pro Gln Phe Tyr Val
 180 185 190

Gly Cys Ala Gln Leu Phe Ile Asp Ser Asp Gly Thr Ala Lys Pro Pro
 195 200 205

Thr Val Ser Ile Gly Glu Gly Thr Tyr Asp Leu Ser Met Pro Ala Met
 210 215 220

ES 2 551 141 T3

Thr Tyr Asn Ile Trp Glu Thr Pro Leu Ala Leu Pro Tyr Pro Met Tyr
225 230 235 240

Gly Pro Pro Val Tyr Thr Pro Gly Ser Gly Ser Gly Ser Val Arg Ala
245 250 255

Thr Ser Ser Ser Ala Val Pro Thr Ala Thr Glu Ser Ser Phe Val Glu
260 265 270

Glu Arg Ala Asn Pro Val Thr Ala Asn Ser Val Tyr Ser Ala Arg Gly
275 280 285

Lys Phe Lys Thr Trp Ile Asp Lys Leu Ser Trp Arg Gly Lys Val Arg
290 295 300

Glu Asn Val Arg Gln Ala Ala Gly Arg Arg Ser Thr Leu Val Gln Thr
305 310 315 320

Val Gly Leu Lys Pro Lys Gly Cys Ile Phe Val Asn Gly Asn Trp Cys
325 330 335

Gly Phe Glu Val Pro Asp Tyr Asn Asp Ala Glu Ser Cys Trp Ala Ala
340 345 350

Ser Asp Asn Cys Trp Lys Gln Ser Asp Ala Cys Trp Asn Lys Thr Gln
355 360 365

Pro Thr Gly Tyr Asn Asn Cys Gln Ile Trp Gln Asp Lys Lys Cys Lys
370 375 380

Val Ile Gln Asp Ser Cys Ser Gly Pro Asn Pro His Gly Pro Pro Asn
385 390 395 400

Lys Gly Lys Asp Leu Thr Pro Glu Trp Pro Pro Leu Lys Gly Ser Met
405 410 415

Asp Thr Phe Ser Lys Arg Thr Ile Gly Tyr Arg Asp Trp Ile Val Arg
420 425 430

Arg Arg Gly Ala
435

<210> 7
<211> 19
<212> ADN

ES 2 551 141 T3

<213> Thermoascus crustaceus

<220>
<221> característica diferente
5 <222> (3)..(3)
<223> N=A,C,G, o T

<220>
<221> característica diferente
10 <222> (6)..(6)
<223> N=A,C,G, o T

<220>
<221> característica diferente
15 <222> (12)..(12)
<223> N=A,C,G, o T

<220>
<221> diferente característica
20 <222> (15)..(15)
<223> N=A,C,G, o T

<400> 7
gcnacngayc tnggnttg 19
25

<210> 8
<211> 19
<212> ADN
<213> Thermoascus crustaceus
30

<220>
<221> característica diferente
<222> (3). (3)
<223> N-A,C,G, o T
35

<220>
<221> característica diferente
<222> (6). (6)
<223> N-A,C,G, o T
40

<220>
<221> característica diferente
<222> (12) .(12)
<223> N-A,C,G, o T
45

<220>
<221> característica diferente
<222> (15)..(15)
<223> N-A,C,G, o T
50

<400> 8
gcnacngayc tnggnttcg 19

<210> 9
<211> 19
<212> ADN
<213> Thermoascus crustaceus
55

<220>
<221> característica diferente
60 <222> (3)..(3)

ES 2 551 141 T3

<223> N=A,C,G, o T

<220>
<221> característica diferente
5 <222> (6) .. (6)
<223> N=A,C,G, o T

<220>
<221> característica diferente
10 <222> (15)..(15)
<223> N=A,C,G, o T

<400> 9
gcnacngayt trggnttyg 19

15 <210> 10
<211> 20
<212> ADN
<213> Thermoascus crustaceus

20 <220>
<221> característica diferente
<222> (6)..(6)
<223> N=A,C,G, o T

25 <220>
<221> característica diferente
<222> (18)..(18)
<223> N=A,C,G, o T

30 <400> 10
caytgnggrt arttytgnc 20

35 <210> 11
<211> 21
<212> ADN
<213> Thermoascus crustaceus

40 <400> 11
tgcaaggagc aaggtagttg a 21

45 <210> 12
<211> 22
<212> ADN
<213> Thermoascus crustaceus

<400> 12
gagtcattc cagcttgacg gt 22

50 <210> 13
<211> 20
<212> ADN
<213> Thermoascus crustaceus

55 <400> 13
tcagacaatc tgatagcggc 20

60 <210> 14
<211> 21
<212> ADN
<213> Thermoascus crustaceus

ES 2 551 141 T3

<400> 14
atcccaacca caactgcacc t 21

5 <210> 15
<211> 44
<212> ADN
<213> Thermoascus crustaceus

10 <400> 15
acacaactgg ggatccacca tggcctttc ccagataatg gcta 44

<210> 16
<211> 36

15 <212> ADN
<213> Thermoascus crustaceus

<400> 16
gtcaccctct agatctggat cgcaggagcg ttcaga 36

20 <210> 17
<211> 44
<212> ADN
<213> Thermoascus crustaceus

25 <400> 17
acacaactgg ggatccacca tgtcattctc gaagatactt gcta 44

<210> 18
<211> 39
<212> ADN
<213> Thermoascus crustaceus

30 <400> 18
gtcaccctct agatctcctc tcgtctttcg tatcagtga 39

<210> 19
<211> 46
<212> ADN

40 <213> Thermoascus crustaceus

<400> 19
acacaactgg ggatccacca tgggtcatt cattcccacc aagtca 46

45 <210> 20
<211> 42
<212> ADN
<213> Thermoascus crustaceus

50 <400> 20
gtcaccctct agatctacac cctcatgcac ctctctct aa 42

REIVINDICACIONES

1. Polipéptido aislado con actividad de aumento celulolítico, seleccionado del grupo consistente en:
- 5 (a) un polipéptido que comprende una secuencia de aminoácidos con al menos el 90% de identidad al polipéptido maduro mostrado como los aminoácidos 23-251 de la SEC ID nº 2;
(b) un polipéptido codificado por un polinucleótido que comprende una secuencia de nucleótidos con al menos el 90% de identidad a la secuencia codificante de polipéptido maduro mostrada como los nucleótidos 67-868 de la SEC ID nº 1.
- 10 2. Polipéptido según la reivindicación 1, que comprende o consiste en el polipéptido maduro de la SEC ID nº 2.
3. Polipéptido según la reivindicación 1, que es codificado por el polinucleótido contenido en plásmido pGEM-T-GH61a51486 que es contenido en *E. coli* DSM 22656.
- 15 4. Polinucleótido aislado que comprende una secuencia de nucleótidos que codifica el polipéptido según la reivindicación 1.
5. Método para producir el polipéptido según la reivindicación 1, que comprende: (a) cultivo de una célula de *Thermoascus crustaceus*, que en su forma de tipo natural produce el polipéptido, bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.
- 20
6. Método para la producción del polipéptido con actividad de aumento celulolítico, que comprende: (a) cultivo de una célula huésped recombinante que comprende el polinucleótido de la reivindicación 4 bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.
- 25
7. Composición detergente que comprende el polipéptido según la reivindicación 1 y un tensioactivo.
8. Método para la degradación o conversión de un material celulósico, que comprende: tratar el material celulósico con una composición enzimática en la presencia del polipéptido según la reivindicación 1 con actividad de aumento celulolítico.
- 30
9. Método según la reivindicación 8, que comprende además la recuperación del material celulósico degradado.
10. Método para la producción de un producto de fermentación, que comprende:
- 35 (a) sacarificación de un material celulósico con una composición enzimática en presencia del polipéptido según la reivindicación 1 con actividad de aumento celulolítico;
(b) fermentación del material celulósico sacarificado con uno o varios microorganismos de fermentación (diferentes) para la producción del producto de fermentación; y
(c) recuperación del producto de fermentación de la fermentación.
- 40
11. Método de fermentación de un material celulósico, que comprende: fermentación del material celulósico con uno o varios microorganismos de fermentación (diferentes), donde el material celulósico es sacarificado con una composición enzimática en presencia de un polipéptido según la reivindicación 1 con actividad de aumento celulolítico.
- 45
12. Método según la reivindicación 11, donde la fermentación del material celulósico produce un producto de fermentación.
13. Método según la reivindicación 12, que comprende además la recuperación del producto de fermentación de la fermentación.
- 50

ES 2 551 141 T3

```

      M A F S Q I M A I T G V F L A S A S L V
1   ATGGCCTTTT CCCAGATAAT GGCTATTACC GGCCTTTTTC TTGCCTCTGC TTCCTGGTG
      A G H G F V Q N I V I D G K S
61  GCTGGCCATG GCTTTGTTCA GAATATCGTG ATTGATGGTA AAAGGTACCT AACTACCTAC
121 CTTACTATCT GATGTCATT ACAAGAAAGG GCACAGACAC AAGCGGCAAA AAAAAGAAAG
      Y G G Y I V N
181 AAAGAAAGAA AGAAAGAAAG CTGACAAAAA TTCAACAAGT TATGGCGGGT ACATCGTGAA
      · Q Y P Y M S D P P E V V G W S T T A T D ·
241 CCAATATCCA TACATGTCAG ATCCTCCGGA GGTCGTCCGC TGGTCTACCA CCGCAACCGA
      · L G F V D G T G Y Q G P D I I C H R G A ·
301 CCTCGGATTC GTGGACGGTA CCGGATACCA AGGACCTGAT ATCATCTGCC ACAGGGGCGC
      · K P A A L T A Q V A A G G T V K L E W T ·
361 CAAGCCTGCA GCCCTGACTG CCAAGTGGC CGCCGGAGGA ACCGTCAAGC TGAATGGAC
      · P W P D S H H G P V I N Y L A P C N G D ·
421 TCCATGGCCT GATTCTCACC ACGGCCCGGT GATCAACTAC CTTGCTCCTT GCAACGGTGA
      · C S T V D K T Q L K F F K I A Q A G L I ·
481 CTGTTCCACC GTGGACAAGA CCAATTGAA ATTCTTCAAG ATCGCCCAGG CCGGTCTCAT
      · D D N S P P G I W A S D N L I A A A N N S ·
541 CGATGACAAC AGTCCTCCTG GTATCTGGGC CTCAGACAAT CTGATAGCGG CCAACAACAG
      · W T V T I P T T T A P G N Y V L R H E I ·
601 CTGGACTGTC ACCATCCCAA CCACAAGTGC ACCTGGAAAC TATGTTCTAA GGCATGAGAT
      · I A L H S A G N K D G A Q N Y P Q C I N ·
661 CATTGCTCTC CACTCAGCTG GGAACAAGGA TGGTGCCGAG AACTATCCCC AGTGCATCAA
      · L K V T G N G S G N P P A G A L G T A L ·
721 CCTGAAGGTC ACTGGAAATG GTTCTGGCAA TCCTCCTGCT GGTGCTCTTG GAACGGCACT
      · Y K D T D P G I L I N I Y Q K L S S Y V ·
781 CTACAAGGAT ACAGATCCGG GAATTCTGAT CAATATCTAC CAGAACTTT CCAGCTATGT
      · I P G P A L Y T G *
841 TATTCTGGT CCTGCTTGT AACTGGTTA G

```

FIG. 1

ES 2 551 141 T3

```

      M S F S K I L A I A G A I T Y A S S A A
1  ATGTCATTCT CGAAGATACT TGCTATCGCT GGGGCCATTA CCTACGCATC TTCAGCTGCC
      A H G Y V Q G I V V D G S Y
61  GCTCATGGTT ATGTCCAGGG AATTGTTGTC GATGGCAGCT AGTATGTCAC TCTGGATGGA
      Y G G Y M V T Q Y
121  ACCTTCAGCA CGTACTGTAC TAACAATCAG CAGCTACGGG GGATATATGG TGACCCAATA
      P Y T A Q P P E L I A W S T K A T D L G
181  TCCCTACACC GCTCAACCTC CGGAATCAT CGCCTGGTCC ACTAAAGCAA CCGATCTTGG
      F V D G S G Y T S P D I I C H K G A E P
241  GTTTGTGGAC GGCAGTGGCT ATACTTCTCC TGATATCATC TGCCATAAGG GTGCTGAGCC
      G A Q S A K V A A G G T V E L Q W T A W
301  TGGTGGCCAG AGCGCCAAAG TGGCAGCTGG AGGGACCGTT GAGCTGCAGT GGACGGCATG
      P E S H K G P V I D Y L A A C D G D C S
361  GCCCGAGTCT CACAAGGGCC CAGTTATTGA CTACCTCGCC GCCTGCGACG GGGACTGCTC
      S V D K T A L K F F K I D E S G L I D G
421  ATCTGTTGAT AAGACTGCAC TAAAGTTCCT TAAGATTGAC GAGAGTGGTC TGATTGACGG
      N G A G T W A S D T L I K N N N S W T V
481  CAACGGTGCT GGAACATGGG CCTCTGATAC GTTGATCAAA AATAACAACA GCTGGACTGT
      T I P S T I A S G N Y V L R H E I I A L
541  CACCATCCCA AGCACAATTG CTTCCGGAAA CTACGTAATA AGACACGAAA TAATTGCGCT
      H S A G N K D G A Q N Y P Q C I N L E V
601  CCATTCTGCC GGAAACAAAG ATGGTGCTCA GAACTATCCC CAGTGTATCA ACCTCGAGGT
      T G S G T E N P A G T L G T A L Y T D T
661  CACTGGTAGT GGCACCGAAA ACCCTGCTGG CACTCTCGGA ACAGCGCTTT ACACAGACAC
      D P G L L V N I Y Q G L S N Y S I P G P
721  TGATCCTGGC CTTCTGGTCA ACATCTACCA GGGTCTGTCC AACTATTCAA TCCCTGGTCC
      A L Y S G N S D N A G S L N P T T T P S
781  TGCTCTGTAT AGCGGCAACA GTGATAACGC TGGTTCCCTC AACCTTACCA CCACGCCGTC
      I Q N A A A A P S T S T A S V V T D S S
841  AATTCAGAAT GCTGCTGCTG CTCCTCCAC TTCCACAGCA TCTGTTGTCA CTGATTCTTC
      S A T Q T A S V A A T T P A S T S A V T
901  GTCAGCCACC CAGACTGCTA GTGTCGCCGC CACGACTCCA GCCTCCACTT CCGCTGTTAC
      A S P A P D T G S D V T K Y L D S M S S
961  AGCCTACCA GCTCCCGATA CTGGAAGCGA CGTAACCAAA TATCTGGATT CGATGAGCTC
      D E V L T L V R G T L S W L V S N K K H
1021  GGATGAGGTC CTCACCCTGG TGCGCGGGAC CCTGTCTTGG CTGGTTTCTA ACAAGAAACA
      A R D L S H *
1081  TGCGCGGGAT CTTTCTCACT GA

```

FIG. 2

ES 2 551 141 T3

M L S F I P T K S A A L T T L L L L G T
 1 ATGTTGTCAT TCATTCCCAC CAAGTCAGCT GCGCTGACGA CTCTTCTACT TCTTGGAACA
 A H A H T L M T T M F V D G V N Q G D G
 61 GCTCATGCTC ACACTTTGAT GACCACCATG TTTGTGGACG CCGTCAACCA GGGAGATGGT
 V C I R M N N D G G T A N T Y I Q P I T
 121 GTCTGCATTC GCATGAACAA TGACGGCGGA ACTGCCAATA CCTATATCCA GCCTATCAGC
 S K D I A C
 181 AGCAAGGATA TCGCCTGCGG TAAGTACCCA GATGTCATCA TACTCTGCCA TAACATCCGT
 G I Q G E I G A ·
 241 CATATCTACT AGAATCGGAG CAATGTTAAG TATTCCAGG CATCCAAGGC GAAATCGGGC
 · S R V C P V K A S S T L T F Q F R E Q P ·
 301 CCTCCCAGT CTGCCAGTC AAGGCATCTT CCACCTAAC CTTCCAATTC CGCGAGCAAC
 · N N P N S S P L D P S H K G P A A V Y L ·
 361 CCAACAACCC AAATCCTCC CCTCTCGATC CATCGCACAA AGGCCCCGCC GCGGTGTACC
 · K K V D S A I A S N N A A G D S W F K I ·
 421 TGAAAAGGT CGACTCCGCC ATCGGAGCA ACAACGCCGC CGGAGACAGC TGGTCAAGA
 · W E S V Y D E S T G K W G T T K M I E N ·
 481 TCTGGGAGC CGTCTAGCAG GAGTCCACGG GCAAATGGGG CACGACCAAG ATGATCGAGA
 · N G H I S V K V P D D I E G G Y Y L A R ·
 541 ACAACGGGCA CATCTCCGTC AAGGTGCCCG ATGATATCGA GGGTGGTTAC TATCTGCCC
 · T E L L A L H S A D Q G D P Q F Y V G C ·
 601 GGACGGAGCT GCTGGCGCTA CATTCTGCGG ATCAGGGGGA TCCGCAGTTC TATGTGGCT
 · A Q L F I D S D G T A K P P T V S I G E ·
 661 GTGCGCAGCT GTTTATCGAT TCGGATGGGA CGGCGAAACC GCCCACTGTT TCTATTGGAG
 · G T Y D L S M P A M T Y N I W E T P L A ·
 721 AGGGGACGTA CGATCTGAGC ATGCCTGCCA TGACGTATAA TATCTGGGAG ACACCGTTGG
 · L P Y P M Y G P P V Y T P G S G S G S V ·
 781 CTCTGCCGTA TCCGATGTAT GGGCCTCCTG TCTATACGCC TGGCTCTGGT TCTGGATCAG
 · R A T S S S A V P T A T E S S F V E E R ·
 841 TCCGTGCGAC GAGCTCTTCT GCTGTCCCTA CTGCAACCGA ATCCTCTTTT GTAGAGGAAA
 · A N P V T A N S V Y S A R G K F K T W I ·
 901 GAGCAAACCC CGTCACGGCA AACAGTGTTT ATTCTGCAAG GGGCAAATTC AAAACCTGGA
 · D K L S W R G K V R E N V R Q A A G R R ·
 961 TTGATAAACT GTCATGGCGC GGAAGGTCC GTGAGAACGT CAGACAAGCC GCGGGAAGAA
 · S T L V Q T V G G L K P K G C I F V N G N ·
 1021 GAAGCACTCT CGTCCAGACT GTGGGTCTAA AGCCAAAAGG CTGCATCTTC GTCATGGAA
 · W C G F E V P D Y N D A E S C W A
 1081 ACTGGTGGCG CTTTCGAGTT CCCGACTACA ACGATGCGGA GAGCTGCTGG GCTGTATGTT
 1141 CCCCTCCTTA GCCTCTTACA TCCCTAAGTA CTACATTGTA AAACAACAAA AAGAAATGTA
 A S D N C W K Q S D A ·
 1201 TATACTAACT ACGTACGCTC TACTCTAGGC CTCCGACAAC TGCTGGAAAC AGTCCGACGC
 · C W N K T Q P T G Y N N C Q I W Q D K K ·
 1261 CTGCTGGAAC AAGACCCAAC CCACGGGCTA CAATAACTGC CAGATCTGGC AGGACAAGAA
 · C K V I Q D S C S G P N P H G P P N K G ·
 1321 ATGCAAGGTC ATCCAGGATT CCTGTAGCGG ACCCAACCCG CATGGACCAC CGAATAAGGG
 · K D L T P E W P P L K G S M D T F S K R ·
 1381 CAAGGATTTG ACTCCGAGT GCCCGCACT GAAGGGCTCG ATGGATACGT TCTCCAAGCG
 · T I G Y R D W I V R R R G A *
 1441 TACTATCGGT TACCGGATT GGATTGTTAG AAGGAGAGGT GCATGAGGGT GTA

FIG. 3

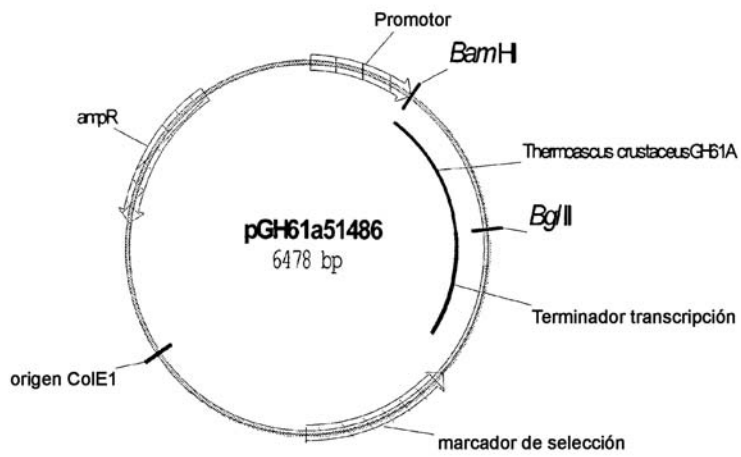


FIG. 4

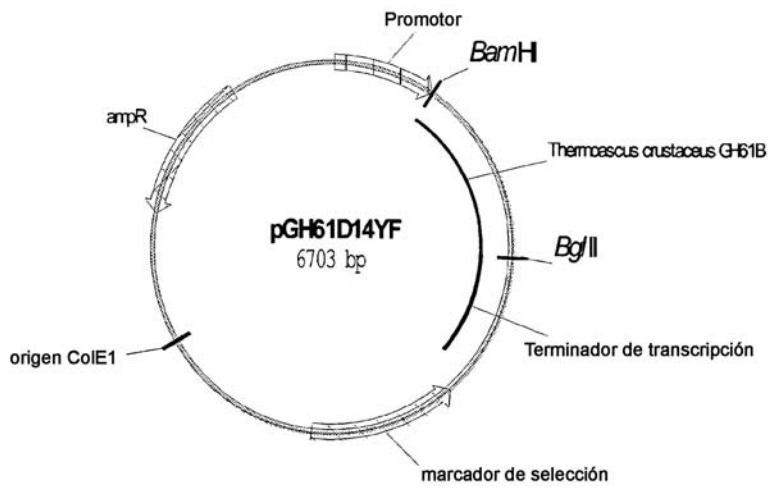


FIG. 5

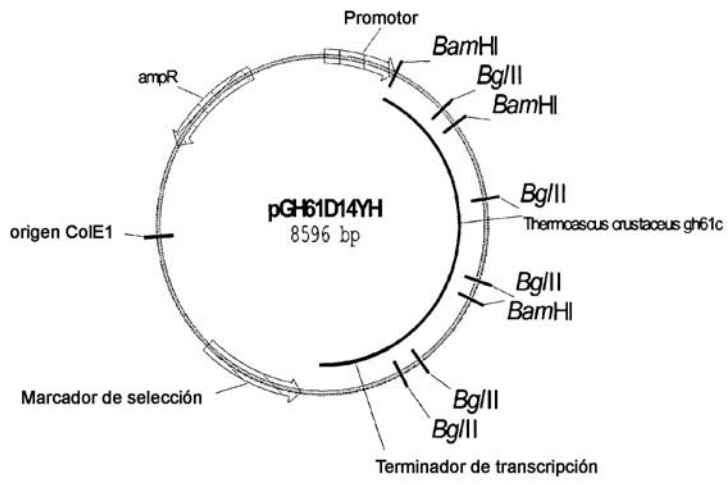


FIG. 6

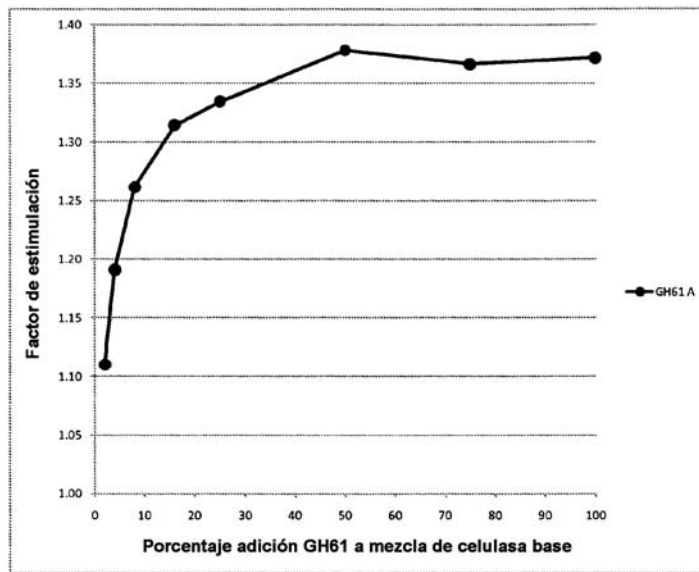


FIG. 7

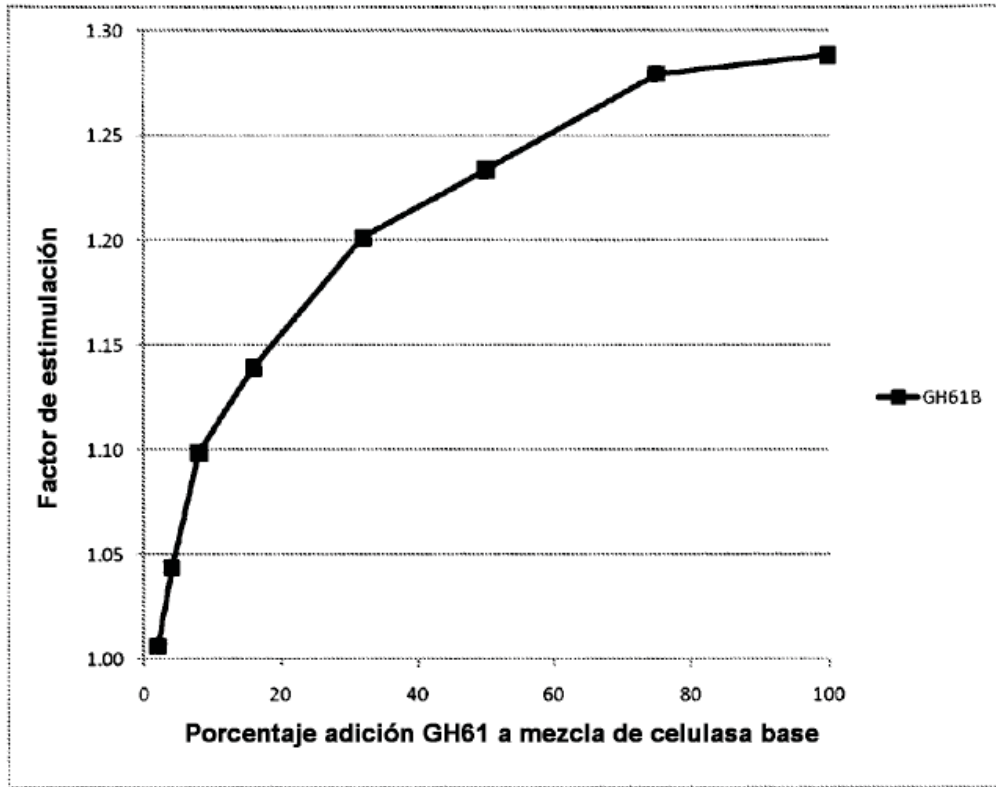


FIG. 8