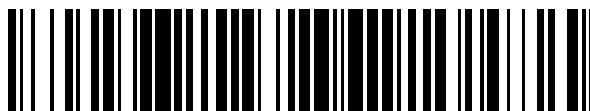


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 368 608**

51 Int. Cl.:

C12N 9/42 (2006.01)

C12N 15/56 (2006.01)

C12N 15/11 (2006.01)

C12N 5/10 (2006.01)

A01H 5/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07758275 .7**

96 Fecha de presentación: **09.03.2007**

97 Número de publicación de la solicitud: **1999257**

97 Fecha de publicación de la solicitud: **10.12.2008**

54 Título: **POLIPÉPTIDOS CON ACTIVIDAD ENDOGLUCANASA Y POLINUCLEÓTIDOS CODIFICANDO LOS POLIPÉPTIDOS.**

30 Prioridad:
20.03.2006 US 784088 P

45 Fecha de publicación de la mención BOPI:
18.11.2011

45 Fecha de la publicación del folleto de la patente:
18.11.2011

73 Titular/es:
**NOVOZYMES, INC.
1445 DREW AVENUE
DAVIS, CA 95618, US y
NOVOZYMES A/S**

72 Inventor/es:
**HARRIS, Paul;
VLASENKO, Elena;
LASSEN, Soren Flensted y
KROGH, Kristian, Bertel, Romer, Morkeberg**

74 Agente: **Tomas Gil, Tesifonte Enrique**

ES 2 368 608 T3

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

DESCRIPCIÓN

Polipéptidos con actividad endoglucanasa y polinucleótidos codificando los polipéptidos.

Referencia a un listado de secuencias

5 [0001] Esta solicitud contiene un listado de secuencias en soporte de lectura informática. Este soporte de lectura informática está incorporado aquí por referencia.

Referencia a un depósito de material biológico

[0002] Esta solicitud contiene una referencia a depósitos de material biológico que han sido producidos en el Northern Regional Research Center (NRRL) conforme al Tratado de Budapest y se le asignó el número de registro NRRL B-30902.

10 **Antecedentes de la invención****Campo de la invención**

[0003] La presente invención se refiere a polipéptidos aislados que tienen actividad endoglucanasa y a polinucleótidos aislados que codifican los polipéptidos. La invención también se refiere a constructos de ácidos nucleicos, vectores y células huéspedes comprendiendo tanto los polinucleótidos como los métodos para producir y usar los polipéptidos.

15 **Descripción de la técnica relacionada**

[0004] La celulosa es un polímero de la glucosa de azúcar simple unido de manera covalente por enlaces beta-1,4. Muchos microorganismos producen enzimas que hidrolizan betaglucanos. Estas enzimas incluyen endoglucanasas, celobiohidrolasas, y beta-glucosidasas. Las endoglucanasas digieren el polímero de celulosa en lugares aleatorios, abriéndolo para ser atacado por celobiohidrolasas. Las celobiohidrolasas liberan secuencialmente moléculas de celobiosa de las extremidades del polímero de celulosa. La celobiohidrolasa I es una actividad 1,4-beta-D-glucano celobiohidrolasa (E.C. 3,2,1,91) que cataliza la hidrólisis de los enlaces 1,4-beta-D-glucosídicos en celulosa, celotetrosa, o cualquier glucosa beta-1,4- enlazada que contenga polímero, liberando celobiosa de las extremidades reductoras de la cadena. La celobiohidrolasa II es una actividad 1,4-D-glucano celobiohidrolasa (E.C. 3,2,1,91) que cataliza la hidrólisis de los enlaces 1,4-beta-D-glucosídicos en celulosa, celotetrosa, o cualquier glucosa beta 1,4-enlazada que contenga polímero, liberando celobiosa de las extremidades no reductoras de la cadena. La celobiosa es un dímero hidrosoluble beta-1,4-enlazado de glucosa. Las beta-glucosidasas hidrolizan la celobiosa a glucosa.

[0005] La conversión de materias primas celulósicas en etanol tiene las ventajas de la disponibilidad inmediata de cantidades grandes de materia prima, la conveniencia de evitar el quemado o vertido de los materiales, y la limpieza del combustible de etanol. La madera, residuos agrícolas, brotes herbáceos, y desperdicios sólidos municipales han sido considerados como materias primas para la producción de etanol. Estos materiales principalmente consisten en celulosa, hemicelulosa y lignina. Una vez que la celulosa se convierte en glucosa, la glucosa es fácilmente fermentada por levadura en etanol.

[0006] Roy et al., 1990, Journal of General Microbiology 136: 1967-1972, da a conocer la purificación y propiedades de una endoglucanasa extracelular de *Myceliophthora thermophila* ATCC 48104. Chernoglazov et al., 1988, Biokhimiya 53: 475-482, da a conocer el aislamiento, purificación, y especificación de sustrato de una endoglucanasa de *Myceliophthora thermophila*. Kliosov et al., 1988, Cartas de Biotecnología 10: 351-354, da a conocer una endoglucanasa termoestable de *Myceliophthora thermophila*. Guzhova y Loginova, 1987, Prikladnaia Biokhimiya y Mikrobiologiya 23: 820-825, da a conocer enzimas celulolíticas de *Myceliophthora thermophila*. Rabinovich et al., 1986, Bioorganicheskaya Khimiya 12: 1549-1560, da a conocer la purificación y caracterización de una endoglucanasa de *Myceliophthora thermophila*. Svistova et al., 1986, Mikrobiologiya 55: 49-54, da a conocer la regulación de la biosíntesis de celulosa en *Myceliophthora thermophila*. Bhat y Maheshwari, 1987, Applied and Environmental Microbiology 53: 2175-2182, da a conocer la actividad de componentes del sistema de celulosa extracelular de *Myceliophthora thermophila*. Kliosov et al., 1987, Prikladnaya Biokhimiya i Mikrobiologiya 23: 44-50, da a conocer una endoglucanasa termoestable de *Myceliophthora thermophila*. Jorgensen et al., 2003, Enzyme and Microbial Technology 32: 851-861, y Tigesen et al., 2003, Enzyme and Microbial Technology 32: 606-615, dan a conocer enzimas degradantes de celulosa de *Penicillium brasilianum* IBT 20888. WO2005/659084 describe una endoglucanasa *Humicola insolens* con un grado de identidad del 70,9% con la endoglucanasa de la invención, y WO2003/062409 describe una endoglucanasa Cel5A de *Thermoascus aurantiacus* con un grado de identidad del 63% con la endoglucanasa de la invención.

50 [0007] Sería una ventaja para la técnica el identificar endoglucanasas nuevas con propiedades mejoradas, tales como un índice de hidrólisis mejorado, una mejor termoestabilidad, una adsorción reducida a lignina, y la capacidad de hidrolizar componentes no celulósicos de biomasa, tales como la hemicelulosa, además de hidrolizar celulosa. Las endoglucanasas con un amplio intervalo de actividades laterales en hemicelulosa pueden ser especialmente beneficiosas para mejorar el rendimiento de la hidrólisis general de complejos sustratos de biomasa ricos en

hemicelulosa.

[0008] Es un objetivo de esta invención el proporcionar polipéptidos mejorados con actividad de endoglucanasa y polinucleótidos que codifican los polipéptidos.

Resumen de la invención

5 [0009] La presente invención se refiere a polipéptidos aislados con actividad endoglucanasa seleccionados del grupo que consiste en:

(a) un polipéptido que comprende una secuencia de aminoácidos que tiene al menos un 85% de grado de identidad con el polipéptido maduro de la SEC ID n°: 4;

10 (b) un polipéptido que es codificado por un polinucleótido que hibridiza bajo al menos condiciones de astringencia altas con (i) la secuencia codificante de polipéptido maduro de la SEC ID n°: 3, (ii) la secuencia de ADNc contenida en la secuencia codificante de polipéptido maduro de la SEC ID n°: 3 o (iii) una cadena complementaria en toda su longitud de (i) o (ii) donde una alta astringencia se define como hibridación en 5xSSPE y 50% formamida a 42°C y tres lavados de 15 min usando 2x SSC, 0,2% de SDS al menos a 65°C; y

15 (c) un polipéptido que se codifica por un polinucleótido con al menos un 85% de identidad con la secuencia codificante de polipéptido maduro de la SEC ID n°: 3;

y donde la actividad endoglucanasa del polipéptido aislado es de al menos un 70% de la actividad endoglucanasa del polipéptido dado a conocer como SEC ID n°. 4 determinado usando hidrólisis de carboximetilcelulosa y donde 1 unidad es definida como 1,0 micromol de azúcares reductores producidos por minuto a 50°C, pH 4.8.

20 [0010] Esta invención también se refiere a polinucleótidos aislados que codifican polipéptidos con actividad endoglucanasa, seleccionados del grupo que consiste en:

(a) un polinucleótido que codifica un polipéptido que comprende una secuencia de aminoácidos que tiene al menos un grado de identidad de un 85% con el polipéptido maduro de la SEC ID n°: 4;

25 (b) un polinucleótido que hibridiza bajo al menos condiciones de alta astringencia con (i) la secuencia codificante de polipéptido maduro de la SEC ID n°: 3, (ii) la secuencia de ADNc contenida en la secuencia codificante de polipéptido maduro de la SEC ID n°: 3 o (iii) una cadena complementaria de (i) o (ii);

(c) un polinucleótido con al menos un grado de identidad del 85% con la secuencia codificante de polipéptido maduro de la SEC ID n°: 3

30 [0011] En un aspecto preferido, el polipéptido maduro consiste en los aminoácidos de 17 a 389 de la SEC ID n°: 4. En otro aspecto preferido, la secuencia codificante de polipéptido maduro consiste en los nucleótidos 67 a 1185 de la SEC ID n°: 3.

[0012] La presente invención también se refiere a constructos de ácidos nucleicos, vectores de expresión recombinantes, células huéspedes recombinantes que comprenden los polinucleótidos, y a métodos de producción de un polipéptido con actividad endoglucanasa.

35 [0013] La presente invención también se refiere a métodos de uso de los polipéptidos con actividad endoglucanasa en la conversión de celulosa a glucosa y sustancias varias.

[0014] La presente invención también se refiere a plantas que comprenden un polinucleótido aislado codificando tal polipéptido con actividad endoglucanasa.

40 [0015] La presente invención también se refiere a métodos para producir tal polipéptido con actividad de endoglucanasa, que comprende: (a) cultivar una planta transgénica o una célula vegetal que comprenda un polinucleótido codificando tal polipéptido con actividad endoglucanasa bajo las condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.

Breve descripción de las figuras

[0016]

La Figura 1 muestra un mapa de restricción de pCIC161.

45 La Figura 2 muestra la secuencia de ADN genómico y la secuencia deducida de aminoácidos de una endoglucanasa de *Myceliophthora thermophila* CBS 117.65 (SEC ID NOs: 3 y 4, respectivamente).

La Figura 3 muestra un mapa de restricción de pA2C161.

Figura 4 muestra un mapa de restricción de pCIC453.

La Figura 5 muestra un mapa de restricción de pA2C453.

La Figura 6 muestra la secuencia de ADNc y la secuencia deducida de aminoácidos de una endoglucanasa de *basidiomycete* CBS 495.95 (SEC ID NOs: 5 y 6, respectivamente).

La Figura 7 muestra un mapa de restricción de pCIC486.

5 La Figura 8 muestra un mapa de restricción de pA2C486.

La Figura 9 muestra la secuencia de ADNc y la secuencia de aminoácidos deducida de una endoglucanasa de *basidiomycete* CBS 494.95 (SEC ID NOs: 7 y 8, respectivamente).

Las Figuras 10A e 10B muestran la secuencia de ADN genómico y la secuencia de aminoácidos deducida de una endoglucanasa de cepa de *Penicillium brasilianum* IBT 20888 (SEC ID NOs: 9 y 10, respectivamente).

10 La Figura 11 muestra un mapa de restricción de pKBK03.

La Figura 12 muestra un mapa de restricción de pPBCel5C.

La Figura 13 muestra la actividad específica de la endoglucanasa CEL5C de *Penicillium brasilianum* IBT 20888 a valores diferentes de pH y 50°C (n=2).

15 La Figura 14 muestra la actividad específica de la endoglucanasa CEL5C de *Penicillium brasilianum* IBT 20888 a temperaturas diferentes y pH 4.8 (n=2).

La Figura 15 muestra la actividad residual de la endoglucanasa CEL5C de *Penicillium brasilianum* IBT 20888 después de 20 horas de incubación a valores diferentes de pH y 25°C y 50°C (n=2).

La Figura 16 muestra la actividad relativa en PASC (2 mg/ml) como una función de temperatura para *basidiomycete* CBS 494.95 y *basidiomycete* CBS 495.95 a pH 5.0.

20 La Figura 17 muestra la conversión relativa de PASC (2 mg/ml) como una función de temperatura tras 45 horas de hidrólisis con *basidiomycete* CBS 494.95 y *basidiomycete* CBS 495.95 (0,5 mg de proteína por g de PASC) a pH 5.0.

La Figura 18 muestra una comparación de endoglucanasas de *Myceliophthora thermophila* CBS 117.65, *basidiomycete* CBS 494.95, *basidiomycete* CBS 495.95, y *Trichoderma reesei* para la producción de azúcares reductores de beta-glucano (1% p/v) tras 2 horas de reacción de hidrólisis a pH 5.5 y 60°C.

25 La Figura 19 muestra una comparación de endoglucanasas de *Myceliophthora thermophila* CBS 117.65, *basidiomycete* CBS 494.95, *basidiomycete* CBS 495.95, y *Trichoderma reesei* para la producción de azúcares reductores de beta-glucano (1% p/v) tras 2 horas de reacción de hidrólisis a pH 5.5 y 60°C.

Definiciones

30 [0017] **Actividad de endoglucanasa:** el término "actividad endoglucanasa" se define aquí como una endo-1,4-beta-D-glucan 4-glucanohidrolasa (E.C. N°. 3.2.1.4) que cataliza la endohidrólisis de los enlaces 1,4-beta-D-glicosídicos en la celulosa, derivados de la celulosa (tal como carboximetilcelulosa y la hidroxietilcelulosa), liquenina, enlaces beta-1,4 en beta-1,3 glucanos mezclados tales como beta-D-glucanos de cereal o xiloglucanos, y otros materiales vegetales que contengan componentes celulósicos. Para los propósitos de la presente invención, la actividad endoglucanasa se determina usando la hidrólisis de carboximetilcelulosa (CMC) según el procedimiento de Ghose, 1987 Pure and Appl. Chem. 59: 257-268. Una unidad de actividad endoglucanasa se define como 1,0 μmol de azúcares reductores producidos por minuto a 50°C, pH 4.8.

40 [0018] En un aspecto preferido, los polipéptidos de la presente invención con actividad endoglucanasa además tienen actividad enzimática hacia uno o más sustratos seleccionados del grupo que consiste de xilano, xiloglucano, arabinoxilano, 1,4- beta-D-manano, y galactomanano. La actividad de los polipéptidos con actividad endoglucanasa en estos sustratos polisacáridos se determina como porcentaje del sustrato hidrolizado a azúcares reductores después de incubar el sustrato (5 mg por ml) con un polipéptido con actividad endoglucanasa de la presente invención (5 mg de proteína por g de sustrato) durante 24 horas con agitado intermitente a pH 5.0 (50 mM de acetato sódico) y 50°C. Los azúcares reductores en mezclas de hidrólisis se determinan por el ensayo de hidrácido de ácido p-hidroxibenzoico (PHBAH).

45 [0019] En un aspecto más preferido, los polipéptidos de la presente invención con actividad endoglucanasa además tienen actividad enzimática hacia el xilano. En otro aspecto más preferido, los polipéptidos de la presente invención con actividad endoglucanasa además tienen actividad enzimática hacia el xiloglucano. En otro aspecto más preferido, los polipéptidos de la presente invención con actividad endoglucanasa además tienen actividad enzimática hacia el arabinoxilano. En otro aspecto más preferido, los polipéptidos de la presente invención con actividad endoglucanasa además tienen actividad enzimática hacia el 1,4-beta-D-manano. En otro aspecto más preferido, los polipéptidos de la presente invención con actividad endoglucanasa además tienen actividad enzimática hacia el galactomanano. En otro

aspecto más preferido, los polipéptidos de la presente invención con actividad endoglucanasa además tienen actividad enzimática hacia el xilano, xiloglucano, arabinoxilano, 1,4-beta-D-manano y/o galactomanano.

5 [0020] Los polipéptidos de la presente invención tienen al menos el 20%, preferiblemente al menos el 40%, más preferiblemente al menos el 50%, más preferiblemente al menos el 60%, más preferiblemente al menos el 70%, más preferiblemente al menos el 80%, incluso más preferiblemente al menos el 90%, y de la forma más preferible al menos el 95%, e incluso de la forma más preferible aún al menos el 100% de la actividad endoglucanasa de polipéptido maduro de la SEC ID n°: 4.

10 [0021] **Familia 5 de glucósido hidrolasa o familia GH5:** el término "Familia 5 de glucósido hidrolasa" o "Familia GH5" se define aquí como un polipéptido que cae en la familia 5 de glucósido hidrolasa 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.

15 [0022] **Polipéptido aislado:** el término "polipéptido aislado" según se usa en este caso se refiere a un polipéptido que es al menos un 20% puro, preferiblemente al menos un 40% puro, más preferiblemente al menos un 60%, puro incluso más preferiblemente al menos un 80% puro, de la forma más preferible al menos un 90% puro, e incluso de la forma más preferible al menos un 95% puro, según determinada la SDS-PAGE.

20 [0023] **Polipéptido sustancialmente puro:** el término "polipéptido sustancialmente puro" denota en la presente una preparación de polipéptido que contiene como máximo el 10%, preferiblemente como máximo el 8%, más preferiblemente como máximo el 6%, más preferiblemente como máximo el 5%, más preferiblemente como máximo el 4%, más preferiblemente como máximo el 3%, incluso más preferiblemente como máximo el 2%, de la forma más preferible como máximo el 1%, e incluso de la forma más preferible como máximo el 0,5% en peso de otro material de polipéptido con el cual está asociado original o recombinantemente. Por lo tanto se prefiere que el polipéptido sustancialmente puro sea al menos un 92% puro, preferiblemente al menos un 94% puro, más preferiblemente al menos un 95% puro, más preferiblemente al menos un 96% puro, más preferiblemente al menos un 96% puro, más preferiblemente al menos un 97% puro, más preferiblemente al menos un 98%, puro incluso más preferiblemente al menos un 99%, de la forma más preferible al menos un 99,5% puro, e incluso de la forma más preferible un 100% puro en peso del material polipéptido total presente en la preparación.

30 [0024] Los polipéptidos en la presente invención se encuentran preferiblemente en una forma sustancialmente pura. En particular, se prefiere que los polipéptidos estén en "forma esencialmente pura", es decir, que la preparación del polipéptido esté esencialmente libre de otro material de polipéptido con el que esté asociado original o recombinantemente. Esto se puede conseguir, por ejemplo, preparando el polipéptido bien mediante métodos recombinantes conocidos o por métodos de purificación tradicionales.

[0025] En la presente, el término "polipéptido maduro" es sinónimo de los términos "polipéptido aislado" y "polipéptido en forma aislada."

35 [0026] **Polipéptido maduro:** el término "polipéptido maduro" se define en la presente como un polipéptido con actividad endoglucanasa que está en su forma final tras la traducción y cualquier modificación postraduccional, tal como el procesamiento N-terminal, truncamiento C-terminal, glicosilación, etc.

[0027] **Secuencia codificante de polipéptido maduro:** el término "secuencia codificante del polipéptido maduro" se define en la presente como una secuencia de nucleótidos que codifica un polipéptido maduro con actividad endoglucanasa.

40 [0028] **Identidad:** la relación entre dos secuencias de aminoácidos o entre dos secuencias de nucleótidos es descrita por el parámetro "identidad".

45 [0029] Para los objetivos de la presente invención, el grado de identidad entre dos secuencias de aminoácidos se determina usando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, J. Mol. Biol. 48: 443-453) tal y como está implementado en el programa Needle de EMBOSS con una penalización de hueco abierto de 10, penalización de extensión de hueco de 0.5, y el matriz EBLOSUM62. El resultado de Needle etiquetado como "la identidad más larga" se usa como el porcentaje de identidad y se calcula de la siguiente manera:

(Residuos idénticos x 100)/(Longitud de alineamiento – Número de huecos en alineamiento)

50 [0030] Para los objetivos de la presente invención, el grado de identidad entre dos secuencias de nucleótidos se determina usando el algoritmo de Needleman-Wunsch (Needleman y Wunsch, 1970, J. Mol. Biol. 48: 443-453) tal y como está implementado en el programa Needle de EMBOSS con una penalización de hueco abierto de 10, penalización de extensión de hueco de 0.5, y el matriz EDNAFULL. El resultado de Needle etiquetado como "la identidad más larga" se usa como el porcentaje de identidad y se calcula de la siguiente manera:

(Residuos idénticos x 100)/(Longitud de alineamiento – Número de huecos en alineamiento)

[0031] **Secuencia homóloga:** el término "secuencia homóloga" se define en la presente como una proteína predicha que da un valor E (o puntuación de expectativa) inferior a 0,001 en una búsqueda de FASTA (Pearson, W.R., 1999, in Bioinformatics Methods and Protocols, S. Misener and S. A. Krawetz, ed., pp. 185-219) con la endoglucanasa madura de la SEC ID n°: 4, SEC ID n°: 6, SEC ID n°: 8, o SEC ID n°: 10.

[0032] **Fragmento polipeptídico:** el término "fragmento polipeptídico" se define aquí como un polipéptido con uno o más aminoácidos eliminados del amino y/o carboxilo terminal del polipéptido maduro de la SEC ID n°: 4 o una secuencia homóloga del mismo; donde el fragmento tiene actividad endoglucanasa. En un aspecto preferido, un fragmento contiene al menos 295 residuos de aminoácidos, más preferiblemente al menos 315 residuos de aminoácidos, y de la forma más preferible al menos 335 residuos de aminoácidos del polipéptido maduro de la SEC ID n°: 4 o una secuencia homóloga del mismo. **Subsecuencia:** el término "subsecuencia" se define en la presente como una secuencia de nucleótidos con uno o más nucleótidos eliminados del extremo 5' y/o 3' de la secuencia codificante del polipéptido maduro de la SEC ID n°: 3 o una secuencia homóloga de la misma; donde la subsecuencia codifica un fragmento de polipéptido con actividad endoglucanasa. Preferentemente, una subsecuencia contiene al menos 885 nucleótidos, más preferiblemente al menos 945 nucleótidos, y de la forma más preferible al menos 1005 nucleótidos de la secuencia codificante de polipéptido maduro de la SEC ID n°: 3 o una secuencia homóloga del mismo.

[0033] **Variante alélica:** el término "variante alélica" denota aquí cualquiera de dos o más formas alternativas de un gen que ocupan el mismo locus cromosómico. La variación alélica surge naturalmente a través de mutación, y puede resultar en polimorfismo dentro de las poblaciones. Las mutaciones genéticas pueden ser silenciosas (ningún cambio en el polipéptido codificado) o pueden codificar polipéptidos con secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

[0034] **Polinucleótido aislado:** el término "polinucleótido aislado" según se usa en la presente se refiere a un polinucleótido que es al menos un 20% puro, preferiblemente al menos un 40% puro, más preferiblemente al menos un 60% puro, incluso más preferiblemente al menos un 80% puro, de la forma más preferible al menos un 90% puro, e incluso de la forma más preferible al menos un 95% puro, según lo determina la electroforesis de agarosa.

[0035] **Polinucleótido sustancialmente puro:** el término "polinucleótido sustancialmente puro" según se usa en la presente se refiere a una preparación polinucleótida libre de otros nucleótidos indeseados o extraños y en una forma adecuada para uso dentro de los sistemas de producción de proteínas modificadas genéticamente. Así, un polinucleótido sustancialmente puro contiene como máximo el 10%, preferiblemente como máximo el 8%, más preferiblemente como máximo el 6%, más preferiblemente como máximo el 5%, más preferiblemente como máximo el 4%, más preferiblemente como máximo el 3%, incluso más preferiblemente como máximo el 2%, de la forma más preferible como máximo el 1%, e incluso de la forma más preferible como máximo el 0,5% en peso de otro material polinucleótido con el cual está asociado original o recombinantemente. Un polinucleótido sustancialmente puro puede, no obstante, incluir regiones no traducidas 5' y 3' de origen natural, tales como promotores y terminadores. Se prefiere que el polinucleótido sustancialmente puro sea al menos un 90% puro, preferiblemente al menos un 92% puro, más preferiblemente al menos un 94% puro, más preferiblemente al menos un 95% puro, más preferiblemente al menos un 96% puro, más preferiblemente al menos un 97%, puro, incluso más preferiblemente al menos un 98% puro, de la forma más preferible al menos un 99% e incluso de la forma más preferible al menos un 99,5% puro en peso. Los polinucleótidos de la presente invención están preferiblemente en una forma sustancialmente pura. En particular, se prefiere que los polinucleótidos descritos aquí estén en "forma esencialmente pura", es decir, que la preparación polinucleótida esté esencialmente libre de otra materia polinucleótida con la cual está asociada original o recombinantemente. En la presente, el término "polinucleótido sustancialmente puro" es sinónimo de los términos "polinucleótido en forma aislada" y "polinucleótido aislado". Los polinucleótidos pueden ser de origen genómico, ADNc, ARN, semisintético, sintético, o cualquier combinaciones de los mismos.

[0036] **Secuencia codificante:** cuando en la presente se usa el término "secuencia codificante", se refiere a una secuencia de nucleótidos, que directamente especifica la secuencia de aminoácidos de su producto proteínico. Los bordes de la secuencia codificante son determinados generalmente por un marco de lectura abierto, que normalmente se inicia con el codón de iniciación ATG o codones de iniciación alternativos tales como GTG y TTG y extremos con un codón de terminación tales como TAA, TAG, y TGA. La secuencia codificante puede ser un ADN, ADNc, o una secuencia de nucleótidos recombinante.

[0037] **Secuencia codificante de polipéptido maduro:** el término "secuencia codificante de polipéptido maduro" se define en la presente como una secuencia de nucleótidos que codifica un polipéptido maduro con actividad endoglucanasa.

[0038] **ADNc:** el término "ADNc" se define en la presente como una molécula de ADN que se puede preparar por transcripción inversa a partir de una molécula ARNm madura, dividida, obtenida de una célula eucariótica. El ADNc no tiene secuencias de intrón que normalmente están presentes en el ADN genómico correspondiente. La transcripción de

ARN inicial primaria es un precursor del ARNm que se procesa a través de una serie de fases antes de aparecer como ARNm maduro empalmado. Estos pasos incluyen la eliminación de secuencias de intrón mediante un proceso llamado empalme. El ADNc derivado de ARNm carece, por lo tanto, de cualquier secuencia de intrón.

5 [0039] **Constructo de ácidos nucleicos:** el término "constructos de ácidos nucleicos" según se usa en la presente se refiere a una molécula de ácido nucleico, ya sea mono o bicatenaria, que se aísla de un gen de origen natural o que es modificada para contener segmentos de ácidos nucleicos que de otra manera no existirían en la naturaleza. El término "constructo de ácidos nucleicos" es sinónimo del término "casete 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.

10 [0040] **Secuencia de control:** el término "secuencias de control" se define en la presente para incluir todos los componentes que son necesarios o ventajosos para la expresión de un polinucleótido que codifica un polipéptido de la presente invención. Cada secuencia de control puede ser nativa o exógena a la secuencia de nucleótidos que codifica el polipéptido o exógeno o nativo entre sí. Tales secuencias de control incluyen, pero de forma no limitativa, una guía, secuencia de poliadenilación, secuencia de propéptido, promotor, secuencia de péptido señal y terminador de transcripción. Como mínimo, las secuencias de control incluyen un promotor y señales de terminación traduccionales y transcripcionales. Las secuencias de control pueden ser provistas de enlaces con el objetivo de introducir sitios de restricción específicos facilitando la ligadura de las secuencias de control con la región de codificación de la secuencia de nucleótidos que codifica un polipéptido.

20 [0041] **Operativamente enlazado:** el término "operativamente enlazado" denota en la presente una configuración en la que una secuencia de control se coloca en una posición apropiada en relación a la secuencia de codificación de la secuencia polinucleótida de tal manera que la secuencia de control dirige la expresión de la secuencia codificante de un polipéptido.

[0042] **Expresión:** el término "expresión" incluye cualquier fase implicada en la producción del polipéptido incluyendo, pero sin limitarse a ello, transcripción, modificación postranscripcional, traducción, modificación postraduccional y secreción.

25 [0043] **Vector de expresión:** el término "vector de expresión" se define en la presente como una molécula de ADN circular o lineal que comprende un polinucleótido que codifica un polipéptido de la presente invención y que está operativamente enlazado a nucleótidos adicionales que permiten su expresión.

30 [0044] **Célula huésped:** el término "célula huésped", según se utiliza en este caso, incluye cualquier tipo de célula susceptible a transformación, transfección, transducción, y similares con un constructo de ácidos nucleicos o vector de expresión que comprenda un polinucleótido de la presente invención.

[0045] **Modificación:** el término "modificación" significa en la presente cualquier modificación química del polipéptido que consiste en el polipéptido maduro de la SEC ID n°: 4 o una secuencia homóloga de la misma; así como la manipulación genética del ADN que codifica tal polipéptido. La modificación puede ser sustituciones, deleciones y/o inserciones de uno o más aminoácidos al igual que sustituciones de una o más cadenas laterales de aminoácidos.

35 [0046] **Variante Artificial:** cuando usado en la presente, el término "variante artificial" se refiere a un polipéptido con actividad endoglucanasa producida por una organismo que expresa una secuencia de nucleótidos modificada de la secuencia codificante de polipéptido maduro de la SEC ID n°: 3 o una secuencia homóloga de la misma. La secuencia de nucleótidos modificada se obtiene a través de intervención humana mediante la modificación de la secuencia de nucleótidos descrita en la SEC ID n°: 3 o una secuencia homóloga de la misma.

40 **Descripción detallada de la invención**

Polipéptidos con actividad endoglucanasa

45 [0047] En un primer aspecto, la presente invención se refiere a polipéptidos que comprenden una secuencia de aminoácidos que tiene un grado de identidad con el polipéptido maduro de la SEC ID n°: 4, de al menos el 85%, incluso más preferiblemente de al menos el 90%, de la forma más preferible de al menos el 95%, e incluso de la forma más preferible de al menos el 96%, 97%, 98%, o 99%, que tiene actividad endoglucanasa (de ahora en adelante "polipéptidos homólogos"). En un aspecto preferido, los polipéptidos homólogos tienen una secuencia de aminoácidos que difiere por diez aminoácidos, preferiblemente por cinco aminoácidos, más preferiblemente por cuatro aminoácidos, incluso más preferiblemente por tres aminoácidos, de la forma más preferible por dos aminoácidos, e incluso de la forma más preferible por un aminoácido del polipéptido maduro de la SEC ID n°: 4.

50 [0048] Un polipéptido de la presente invención preferiblemente comprende la secuencia de aminoácidos de la SEC ID n°: 4. En un aspecto preferido, un polipéptido comprende la secuencia de aminoácidos de la SEC ID n°: 4. En otro aspecto preferido, un polipéptido comprende el polipéptido maduro de la SEC ID n°: 4. En otro aspecto preferido, un polipéptido comprende aminoácidos 17 a 389 de la SEC ID n°: 4. En otro aspecto preferido, un polipéptido comprende aminoácidos 17 a 389 de la SEC ID n°: 4. En otro aspecto preferido, un polipéptido consiste en la secuencia de aminoácidos de la

SEC ID n°: 4. En otro aspecto preferido, un polipéptido consiste en la secuencia de aminoácidos de la SEC ID n°: 4. En otro aspecto preferido, un polipéptido consiste en el polipéptido maduro de la SEC ID n°: 4. En otro aspecto preferido, un polipéptido consiste en aminoácidos 17 a 389 de la SEC ID n°: 4. En otro aspecto preferido, un polipéptido consiste en aminoácidos 17 a 389 de la SEC ID n°: 4.

5 [0049] En un segundo aspecto, la presente invención se refiere a polipéptidos aislados con actividad endoglucanasa codificados por polinucleótidos que hibridan bajo condiciones de astringencia alta, y de la forma más preferible bajo condiciones de astringencia muy alta con (i) la secuencia codificante de polipéptido maduro de la SEC ID n°: 3, (ii) la secuencia de ADNc contenida en la secuencia codificante de polipéptido maduro de la SEC ID n°: 3, (iii) una subsecuencia de (i) o (ii), o (iv) una cadena complementaria de (i); (ii), o (iii) (J. Sambrook, E.F. Fritsch, y T. Maniatis, 10 1989, Molecular Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, Nueva York). Una subsecuencia de la secuencia codificante de polipéptido maduro de la SEC ID n°: 3, SEC ID n°: 5, SEC ID n°: 7, o la SEC ID n°: 9 contiene un mínimo de 100 nucleótidos contiguos o preferiblemente al menos 200 nucleótidos contiguos. Por otra parte, la subsecuencia puede codificar un fragmento de polipéptido que tiene actividad endoglucanasa. En un aspecto preferido, la secuencia codificante de polipéptido maduro consiste en los nucleótidos 67 a 1185 de la SEC ID n°: 3. La secuencia de nucleótidos de la SEC ID n°: 3 o una subsecuencia de la misma; así como la secuencia de aminoácidos de la SEC ID n°: 4 o un fragmento de la misma pueden ser usados para diseñar una sonda de ácidos nucleicos para identificar y clonar ADN que codifica polipéptidos con actividad endoglucanasa de cepas de géneros o especies diferentes según métodos bien conocidos en la técnica. En particular, tales sondas se pueden usar para la hibridación con el genómico o ADNc del género o especie de interés, siguiendo los procedimientos estándar de transferencia de Southern, para 15 identificar y aislar el gen correspondiente en ello. Tales sondas pueden ser considerablemente más cortas que la secuencia entera, pero deberían ser al menos de 14, preferiblemente al menos de 25, más preferiblemente al menos de 35, y de la forma más preferible al menos de 70 nucleótidos de longitud. No obstante, se prefiere que la sonda de ácidos nucleicos sea al menos de 100 nucleótidos de longitud. Por ejemplo, la sonda de ácidos nucleicos puede ser de al menos 200 nucleótidos, preferiblemente de al menos 300 nucleótidos, más preferiblemente de al menos 400 nucleótidos, o de la forma más preferible de al menos 500 nucleótidos de longitud. Sondas incluso más largas pueden ser utilizadas, p.ej., sondas de ácidos nucleicos que sean de al menos 600 nucleótidos, al menos preferiblemente de al menos 700 nucleótidos, más preferiblemente de al menos 800 nucleótidos, o de la forma más preferible de al menos 900 nucleótidos de longitud. Tanto sondas de ADN como de ARN pueden ser empleadas. Las sondas normalmente están marcadas para la detección del gen correspondiente (p.ej., con ³²P, ³H, ³⁵S, biotina, o avidina). Tales sondas están englobadas en la presente invención.

[0050] Un banco de ADN genómico o de ADNc preparado a partir de tales otros organismos pueden, por lo tanto, ser seleccionado para ADN que hibridiza con las sondas anteriormente descritas y que codifica un polipéptido con actividad endoglucanasa. El ADN genómico u otro ADN de tales otros organismos se pueden separar mediante electroforesis en gel de agarosa o poliacrilamida, u otras técnicas de separación. ADN de los bancos o el ADN separado se pueden transferir e inmovilizar en nitrocelulosa u otro material portador adecuado. Para identificar un clon o ADN que es homólogo con la SEC ID n°: 3 o una subsecuencia de la misma; el material portador preferiblemente es usado en una transferencia de Southern.

[0051] Para los objetivos de la presente invención, la hibridación indica que la secuencia de nucleótidos hibridiza a una sonda de ácidos nucleicos marcada correspondiente a la secuencia codificante de polipéptido maduro de la SEC ID n°: 3; la secuencia de ADNc contenida en la secuencia codificante de polipéptido maduro de la SEC ID n°: 3; su cadena complementaria; o una subsecuencia de la misma; bajo condiciones de astringencia alta. Las moléculas a las que la sonda de ácidos nucleicos hibridiza bajo estas condiciones pueden ser detectadas usando, p.ej., película radiográfica.

[0052] En un aspecto preferido, la sonda de ácidos nucleicos es la secuencia codificante de polipéptido maduro de la SEC ID n°: 3. En otro aspecto preferido, la sonda de ácidos nucleicos consiste en los nucleótidos 67 a 1185 de la SEC ID n°: 3. En otro aspecto preferido, la sonda de ácidos nucleicos es una secuencia polinucleótida que codifica el polipéptido de la SEC ID n°: 4, o una subsecuencia de la misma. En otro aspecto preferido, la sonda de ácidos nucleicos es la SEC ID n°: 3. En otro aspecto preferido, la sonda de ácidos nucleicos es la secuencia polinucleótida contenida en el plásmido pCIC161 que se contiene en *E. coli* NRRL B-30902, donde la secuencia polinucleótida de la misma codifica un polipéptido con actividad endoglucanasa. En otro aspecto preferido, la sonda de ácidos nucleicos es la región de codificación del polipéptido maduro contenido en el plásmido pCIC161 que se contiene en *E. coli* NRRL B-30902.

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

[0054] Para sondas largas de al menos 100 nucleótidos de longitud, el material portador es finalmente lavado tres veces durante 15 minutos cada vez usando 2X SSC, 0,2% de SDS preferiblemente al menos a 45°C (astringencia muy baja),

más preferiblemente al menos a 50°C (astringencia baja), más preferiblemente al menos a 55°C (astringencia media), más preferiblemente al menos a 60°C (astringencia media alta), incluso más preferiblemente al menos a 65°C (astringencia alta), y de la forma más preferible al menos a 70°C (astringencia muy alta).

5 [0055] Para sondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, las condiciones de astringencia se definen como prehibridación, hibridación, y lavado de post-hibridación a aproximadamente 5°C a sobre 10°C por debajo del T_m calculado usando el cálculo según Bolton y McCarthy (1962, Proceedings of the National Academy of Sciences EEUU 48:1390) en 0,9 M NaCl, 0,09 M de Tris-HCl pH 7.6, 6 mM de EDTA, 0,5% de NP-40, 1X solución de Denhardt, 1 mM de pirofosfato de sodio, 1 mM de fosfato monobásico de sodio, 0,1 mM de ATP, y 0,2 mg de ARN de levadura por ml siguiendo procedimientos estándar de transferencia de Southern durante 12 a 24 horas óptimamente.

[0056] Para las sondas cortas de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, el material portador se lava una vez en 6X SCC más 0,1% de SDS durante 15 minutos y dos veces cada durante 15 minutos usando 6X SSC a 5°C a 10°C por debajo de la T_m calculada.

15 [0057] En un tercer aspecto, la presente invención se refiere a polipéptidos aislados codificados por polinucleótidos que comprenden o que consisten de secuencias de nucleótidos que tienen un grado de identidad a la secuencia codificante de polipéptido maduro de la SEC ID n°: 3 de por lo menos el 85%, más preferiblemente al menos el 90%, incluso más preferiblemente al menos el 95%, y de la forma más preferible al menos el 97% de identidad, que codifica un polipéptido activo. En un aspecto preferido, la secuencia codificante de polipéptido maduro consiste en los nucleótidos 67 a 1185 de la SEC ID n°: 3. Ver sección polinucleótida en la presente.

20 [0058] La presente divulgación se refiere a variantes artificiales que comprenden una sustitución, deleción y/o inserción de uno o más aminoácidos del polipéptido maduro de la SEC ID n°: 4 o una secuencia homóloga de la misma. Preferiblemente, los cambios de los aminoácidos son de una naturaleza menor, es decir, sustituciones o inserciones conservadoras de aminoácidos que no afectan significativamente al pliegue y/o actividad de la proteína; deleciones pequeñas, normalmente de 1 a 30 aminoácidos; pequeñas extensiones amino-terminales o carboxi-terminales, tales como un residuo de metionina amino-terminal; un pequeño péptido de enlace de hasta sobre 20-25 residuos; o una pequeña extensión que facilita la purificación cambiando la carga neta u otra función, tal como un tracto de polihistidina, un epítipo antigénico o un dominio de unión.

25 [0059] Ejemplos de sustituciones conservadoras se encuentran dentro del grupo de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrofóbicos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano y tirosina), y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina). Las sustituciones de aminoácidos que generalmente no alteran la actividad específica son conocidas en la técnica y son descritas, por ejemplo, por H. Neurat y R.L. Colina, 1979, In, The Proteins, Academia Press, Nueva York. Los intercambios más frecuentes son Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu, y Asp/Gly.

30 [0060] Además de los 20 aminoácidos estándar, los aminoácidos no estándar (tales como 4-hidroxiprolina, 6-N-metil-lisina, ácido 2-aminoisobutírico, isovalina, y alfa-metil serina) se puede sustituir por residuos de aminoácidos de un polipéptido de tipo salvaje. Un número limitado de aminoácidos no conservadores, aminoácidos que no se codifican por el código genético, y aminoácidos no naturales se pueden sustituir por residuos de aminoácidos. Los "aminoácidos no naturales" han sido modificados después de la síntesis de proteína, y/o tienen una estructura química en su/s cadena/s lateral/es diferente de la de los aminoácidos estándar. Los aminoácidos no naturales pueden ser sintetizados químicamente, y preferiblemente, están comercialmente disponibles, e incluyen ácido piperídico, ácido carboxílico de tiazolidina, dehidroprolina, 3- y 4-metilprolina, y 3,3-dimetilprolina.

35 [0061] Alternativamente, los cambios de aminoácidos son de una naturaleza tal que las propiedades físico químicas de los polipéptidos son alteradas. Por ejemplo, los cambios de los aminoácidos pueden mejorar la termoestabilidad del polipéptido, alterar la especificidad de sustrato, cambiar el pH óptimo, y similares.

40 [0062] Los aminoácidos esenciales en el polipéptido progenitor se pueden identificar según procedimientos conocidos en la técnica, tal como la mutagénesis dirigida o mutagénesis de barrido de alanina (Cunningham and Wells, 1989, Science 244: 1081- 1085). En esta técnica se introducen mutaciones simples de alanina en cada residuo de la molécula, y las moléculas mutantes resultantes se evalúan para detectar actividad biológica (es decir, actividad endoglucanasa) para identificar residuos de aminoácidos que son críticos para la actividad de la molécula. Ver también, Hilton et al., 1996, J. Biol. Chem. 271: 4699-4708. El sitio activo de la enzima u otra interacción biológica también pueden ser determinadas por análisis físico de la estructura, como está determinado por tales técnicas como la resonancia magnética nuclear, cristalografía, difracción electrónica, o marcaje por fotoafinidad, conjuntamente con mutación de aminoácidos de sitio de contacto putativo. Ver, por ejemplo, de Vos et al., 1992, Science 255: 306-312; Smith et al., 1992, J. Mol. Biol. 224: 899-904; Wlodaver et al., 1992, FEBS Lett. 309: 59-64. Las identidades de los aminoácidos esenciales también pueden ser

inferidas a partir de análisis de identidades con polipéptidos que están relacionados con un polipéptido según la invención.

5 [0063] Las sustituciones de aminoácidos múltiples o simples pueden ser realizadas y evaluadas usando métodos conocidos de mutagénesis, recombinación, y/o redistribución, seguidos de un procedimiento de selección pertinente, tal como los descritos por Reidhaar-Olson and Sauer, 1988, Science 241: 53-57; Bowie and Sauer, 1989, Proc. Natl. Acad. Sci. EEUU 86: 2152-2156; WO 95/17413; or WO 95/22625. Otros métodos que pueden ser usados incluyen PCR propenso al error, exposición en fago (e.g., Lowman et al., 1991, Biochem. 30: 10832-10837; Patente estadounidense No. 5,223,409; WO 92/06204), y mutagénesis dirigida (Derbyshire et al., 1986, Gene 46: 145 ; Ner et al., 1988, DNA 7:127).

10 [0064] Los métodos de mutagénesis/redistribución se pueden combinar con métodos de selección automatizados de alto rendimiento para detectar actividad de polipéptidos mutagenizados clonados, expresados por células huéspedes (Ness et al., 1999, Nature Biotechnology 17: 893-896). Las moléculas mutagenizadas de ADN que codifican polipéptidos activos pueden ser recuperadas de las células huéspedes y rápidamente ordenadas usando métodos estándar en la técnica. Estos métodos permiten la rápida determinación de la importancia de residuos de aminoácidos individuales en un polipéptido de interés, y se puede aplicar a polipéptidos de estructura desconocida.

15 [0065] La cantidad total de sustituciones de aminoácido, deleciones y/o inserciones del polipéptido maduro de la SEC ID n°: 4 es de 10, preferiblemente 9, más preferiblemente 8, más preferiblemente 7, más preferiblemente como máximo 6, más preferiblemente 5, más preferiblemente 4, incluso más preferiblemente 3, de la forma más preferible 2, e incluso de la forma más preferible 1.

20 [0066] En un aspecto más preferido, el polipéptido es un polipéptido de *Myceliophthora thermophila*, y de la forma más preferible un polipéptido de *Myceliophthora thermophila* CBS 111.65, p.ej., el polipéptido de la SEC ID n°: 4, o el polipéptido maduro del mismo.

25 [0067] Los polipéptidos de la presente invención también incluyen polipéptidos fusionados o polipéptidos de fusión escindibles en donde otro polipéptido se fusiona al N-terminal o C-terminal del polipéptido o al fragmento del mismo. Un polipéptido fusionado se produce por fusión de una secuencia de nucleótidos (o una parte de la misma) codificando otro polipéptido a una secuencia de nucleótidos (o una parte de la misma) de la presente invención. Las técnicas para producir polipéptidos de fusión son conocidas en la técnica, e incluyen ligar las secuencias de codificación que codifican los polipéptidos de modo que en el marco y que expresión del polipéptido fusionado esté bajo control del mismo promotor(es) y terminador.

30 [0068] Un polipéptido de fusión puede además comprender un sitio de escisión. En la secreción de la proteína de fusión, el sitio se escinde liberando el polipéptido con actividad endoglucanasa de la proteína de fusión.

35 [0069] Ejemplos de sitios de escisión incluyen, pero no se limitan a, un sitio Kex2 que codifica el dipéptido Lys-Arg (Martín et al., 2003, J. Ind. Microbiol. Biotechnol. 3: 568-76; 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-Ly, que es escindido por una enteropeptidasa después de la lisina (Collins-Racie et al., 1995, Biotechnology 13: 982-987); un sitio His-Tyr-Glu o His-Tyr-Asp, que se divide por Genenase I (Carter et al., 1989, Proteins: Structure, Function and Genetics 6: 240-248); un sitio Leu- Val-Pro-Arg-Gly-Ser site, que se escinde mediante trombina después del Arg (Stevens, 2003, Drug Discovery World 4: 35-48); un sitio Glu-Asn-Leu-Tyr-Phe-Gln-Gly, que se escinde mediante proteasa TEV después del 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 3C de rinovirus humano después del Gln (Stevens, 2003, supra).

Polinucleótidos

45 [0070] La presente invención también se refiere a un polinucleótido aislado que comprende o que consiste en una secuencia de nucleótidos que codifica un polipéptido de la presente invención con actividad endoglucanasa.

[0071] En un aspecto preferido, la secuencia de nucleótidos comprende o consiste en la SEC ID n°: 3. En otro aspecto más preferido, la secuencia de nucleótidos comprende o consiste en la secuencia contenida en el plásmido pCIC161 contenido en *E. coli* NRRL B-30902. En otro aspecto preferido, la secuencia de nucleótidos comprende o consiste en la región de codificación de polipéptido maduro de la SEC ID n°: 3. En otro aspecto preferido, la secuencia de nucleótidos comprende o consiste en nucleótidos 67 a 1185 de la SEC ID n°: 3. En otro aspecto más preferido, la secuencia de nucleótidos comprende o consiste en la región de codificación de polipéptido maduro contenida en el plásmido pCIC161 contenido en *E. coli* NRRL B-30902. La presente divulgación también comprende las secuencias de nucleótidos que codifican polipéptidos que comprenden o que consisten en la secuencia de aminoácidos de la SEC ID n°: 4 o el polipéptido maduro de la misma, que difiere de la SEC ID n°: 3 o de la secuencia de codificación del polipéptido maduro del mismo en virtud de la degeneración del código genético. La presente divulgación también se refiere a subsecuencias

de la SEC ID nº: 3 que codifican fragmentos de la SEC ID nº: 4 que tienen actividad endoglucanasa.

[0072] Las técnicas usadas para aislar o clonar un polinucleótido que codifica un polipéptido son conocidas en la técnica e incluyen 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 ser efectuada, p.ej., usando la bien conocida reacción en cadena de polimerasa (PCR) o la selección de anticuerpos de genotecas de expresión para detectar fragmentos de ADN clonados con características estructurales compartidas. Ver, p.ej., Innis et al., 1990, PCR: A Guide to Methods and Application, Academic Press, Nueva York. Otros procedimientos de amplificación de ácidos nucleicos como la reacción en cadena de la ligasa (LCR), la transcripción activada ligada (LAT) y la amplificación basada en secuencias de nucleótidos (NASBA) pueden ser utilizados. Los polinucleótidos pueden clonarse de una cepa de *Myceliophthora thermophila* CBS 117.65, *basidiomycete* CBS 494.95, o *basidiomycete* CBS 495.95, u otro u organismo relacionado y así, por ejemplo, puede ser una variante alélica o de especies de la región de codificación de polipéptido de la secuencia de nucleótidos.

[0073] La presente invención también se refiere a polinucleótidos aislados que comprenden o que consisten en secuencias de nucleótidos que tienen un grado de identidad con la secuencia codificante de polipéptido maduro de la SEC ID nº: 3 de al menos 85%, más preferiblemente de al menos 90%, incluso más preferiblemente de al menos 95%, y de la forma más preferible de al menos 97% identidad, que codifica un polipéptido activo. En un aspecto preferido, la secuencia codificante de polipéptido maduro es nucleótidos 67 a 1185 de la SEC ID nº: 3. 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 las formas del polipéptido que no ocurren naturalmente. Estos polipéptidos puede diferir de alguna forma modificada del polipéptido aislado de su fuente nativa, p.ej., variantes artificiales que difieren en actividad específica, termoestabilidad, pH óptimo, o similares. La secuencia variante puede ser construida basándose en la secuencia de nucleótidos presentada como la región de codificación de polipéptido de la SEC ID nº: 3, p.ej., una subsecuencia de la misma, y/o por introducción de sustituciones de nucleótido que no dan lugar a otra secuencia de aminoácidos del polipéptido codificado por la secuencia de nucleótidos, pero que corresponden al uso de codón del organismo huésped destinado para la producción de la enzima, o por la introducción de sustituciones de nucleótido que pueden dar lugar a una secuencia de aminoácidos diferente. Para una descripción general de sustitución de nucleótido, ver, p.ej., Ford et al., 1991, Protein Expression and Purification 2: 95-107.

[0074] Será evidente para los expertos en la técnica que tales sustituciones pueden ser realizadas fuera de las regiones críticas para la función de la molécula y todavía resultar en 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 sujeto a sustitución, pueden ser identificados según procedimientos conocidos en la técnica, tales como la mutagénesis dirigida o mutagénesis de barrido de alanina (ver, p.ej. Cunningham y Wells, 1989, supra). En esta última técnica, las mutaciones se introducen en cada residuo cargado positivamente en la molécula, y las moléculas mutantes resultantes se evalúan para detectar actividad endoglucanasa para identificar residuos de aminoácidos que son críticos para la actividad de la molécula. Los sitios de interacción de sustrato de enzima también pueden ser determinados por el análisis de la estructura tridimensional como determinan tales técnicas como el análisis por resonancia magnética nuclear, cristalografía o marcaje por fotoafinidad (ver, P.ej. de Vos et al., 1992, supra; Smith et al., 1992, supra; Wlodaven et al., 1992, supra).

[0075] La presente invención también se refiere a polinucleótidos aislados que codifican un polipéptido de la presente invención, que hibridan bajo condiciones de astringencia alta, y de la forma más preferible condiciones de astringencia muy alta con (i) la secuencia codificante de polipéptido maduro de la SEC ID nº: 3, (ii) la secuencia de ADNc contenida en la secuencia codificante de polipéptido maduro de la SEC ID nº: 3, o (iii) una cadena complementaria de (i) o (ii); o variantes alélicas y subsecuencias de las mismas (Sambrook et al., 1989, supra), tal y como se define aquí. En un aspecto preferido, la secuencia codificante de polipéptido maduro de la SEC ID nº: 3 consiste en los nucleótidos 67 a 1185.

[0076] La descripción también se refiere a polinucleótidos aislados obtenidos mediante (a) la hibridación de una población de ADN bajo condiciones de astringencia muy baja, baja, media, media alta, alta, o muy alta con (i) la secuencia codificante de polipéptido maduro de la SEC ID nº: 3, (ii) la secuencia de ADNc contenida en la secuencia codificante de polipéptido maduro de la SEC ID nº: 3, o (iii) una cadena complementaria de (i) o (ii); y (b) el aislamiento del polinucleótido de hibridación, que codifica un polipéptido con actividad endoglucanasa. La secuencia codificante de polipéptido maduro de la SEC ID nº: 3 consiste en los nucleótidos 67 a 1185.

Constructos de ácidos nucleicos

[0077] La presente invención también se refiere a constructos de ácidos nucleicos que comprenden un polinucleótido aislado de la presente invención operativamente enlazado a una o más secuencias de control que dirigen la expresión de la secuencia codificante en una célula huésped adecuada bajo condiciones compatibles con las secuencias de control.

- [0078] Un polinucleótido aislado que codifica un polipéptido de la presente invención puede ser manipulado de varias maneras para proporcionar la expresión del polipéptido. La manipulación de la secuencia del polinucleótido antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar secuencias polinucleótidas utilizando métodos de ADN recombinante son bien conocidas en la técnica.
- 5 [0079] La secuencia de control puede ser una secuencia promotora apropiada, una secuencia de nucleótidos 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 muestran actividad transcripcional en la célula huésped de elección incluyendo, promotores mutantes, truncados, e híbridos, y se pueden obtener de genes que codifican polipéptidos extracelulares o intracelulares homólogos o heterólogos a la célula huésped.
- 10 [0080] Ejemplos de promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente invención, especialmente en una célula huésped bacteriana, son los promotores obtenidos del operón *lac* de *E. coli*, gen de agarasa de *Streptomyces coelicolor* (*dagA*), gen de levansucrasa de *Bacillus subtilis* (*sacB*), gen de alfa-amilasa de *Bacillus licheniformis* (*amyL*), gen de amilasa maltogénica de *Bacillus stearothermophilus* (*amyM*), gen de alfa-amilasa (*amyQ*) de *Bacillus amyloliquefaciens*, gen de penicilinas de *Bacillus licheniformis* (*penP*), genes *xy/A* y *xy/B* de *Bacillus subtilis*, y gen procariótico de beta lactamasa (Villa-Kamaroff et al., 1978, Proceedings of the National Academy of Sciences EEUU 75: 3727-3731), así como el promotor *tac* (DeBoer et al., 1983, Proceedings of the National Academy of Sciences EEUU 80: 21-25). Se describen más promotores en "Useful proteins from recombinant bacteria" en Scientific American, 1980, 242: 74-94; y en Sambrook et al., 1989, supra.
- 15 [0081] Ejemplos de promotores adecuados para dirigir la transcripción de los constructos de ácidos nucleicos de la presente invención en una célula huésped fúngica filamentosa son promotores obtenidos de los genes para TAKA amilasa de *Aspergillus oryzae*, proteinasa aspártica de *Rhizomucor mizei*, alfa-amilasa neutra de *Aspergillus niger*, alfa-amilasa estable en ácido de *Aspergillus niger*, glucoamilasa de *Aspergillus niger* o *Aspergillus awamori* (*glaA*), lipasa de *Rhizomucor miehei*, proteasa alcalina de *Aspergillus oryzae*, Triosa fosfato isomerasa de *Aspergillus oryzae*, acetamidasa de *Aspergillus nidulans*, amiloglucosidasa de *Fusarium venenatum* (WO 00/56900), *Fusarium venenatum* Daria (WO 00/56900), *Fusarium venenatum* Quinn(WO 00/56900), Proteasa 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*, xilanasas I de *Trichoderma reesei*, xilanasas II de *Trichoderma reesei*, beta-xilosidasa de *Trichoderma reesei*, al igual que el promotor NA2-tpi (un híbrido de los promotores de los genes para alfa-amilasa neutra de *Aspergillus niger* y triosa fosfato isomerasa de *Aspergillus oryzae*); y promotores mutantes, truncados, e híbridos de los mismos.
- 20 [0082] En un huésped de levadura, promotores útiles se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO- 1), galactoquinasa de *Saccharomyces cerevisiae* (GAL1), alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH1 ADH2/GAP), triosa fosfato isomerasa de *Saccharomyces cerevisiae* (TPI), metalotioneína de *Saccharomyces cerevisiae* (CUP1), y 3-fosfoglicerato-quinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para células huéspedes de levadura son descritas por Romanos et al., 1992, Yeast 8: 423-488.
- 25 [0083] La secuencia de control puede también ser una secuencia del terminador de la transcripción adecuada, una secuencia reconocida por una célula huésped para terminar la transcripción. La secuencia del terminador está operativamente enlazada al 3' terminal de la secuencia de nucleótidos que codifica el polipéptido. Cualquier terminador que sea funcional en la célula huésped de elección se ser usado en la presente invención.
- 30 [0084] Los terminadores preferidos para células huéspedes filamentosas fúngicas 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 tipo tripsina de *Fusarium oxysporum*.
- 35 [0085] Terminadores preferidos para células huéspedes de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae*, citocromo C (CYC1) de *Saccharomyces cerevisiae*, y gliceraldehído-3-deshidrogenasa fosfato de *Saccharomyces cerevisiae*. Otros terminadores útil para células huéspedes de levadura son descritas por Romanos et al., 1992, supra.
- 40 [0086] La secuencia de control puede también ser una secuencia líder adecuada, una región no traducida de un ARNm que es importante para traducción por la célula huésped. La secuencia líder está operativamente enlazada al extremo 5' de la secuencia de nucleótidos que codifica el polipéptido. Cualquier secuencia líder que sea funcional en la célula huésped de elección puede ser utilizado en la presente invención.
- 45 [0087] Los líderes preferidos para células huéspedes fúngicas filamentosas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae* y Triosa fosfato isomerasa de *Aspergillus nidulans*.
- 50
- 55

- [0088] Los líderes adecuados para células huéspedes de levadura se obtienen de los genes para enolasa (ENO-1) de *Saccharomyces cerevisiae*, fosfoglicerato-quinasa de *Saccharomyces cerevisiae*, factor alfa de *Saccharomyces cerevisiae*, y alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa (ADH2/GAP) de *Saccharomyces cerevisiae*
- 5 [0089] La secuencia de control puede también ser una secuencia de poliadenilación, una secuencia operativamente enlazada al extremo 3' de la secuencia de nucleótidos y que, cuando transcrita, es reconocida por la célula huésped como una señal para añadir residuos de poliadenosina a ARNm transcrito. Cualquier secuencia de poliadenilación que sea funcional en la célula huésped de elección se puede utilizar en la presente invención.
- 10 [0090] Las secuencias de poliadenilación preferidas para células huéspedes fúngicas filamentosas se obtienen a partir 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*.
- [0091] Secuencias de poliadenilación útil para células huéspedes de levadura son descritas por Guo y Sherman, 1995, Molecular Cellular Biology 15:5983-5990.
- 15 [0092] La secuencia de control puede también ser una región de codificación de péptido señal que codifica una secuencia de aminoácidos enlazada al amino término de un polipéptido y dirige el polipéptido codificado en la vía secretora de la célula. El extremo 5' de la secuencia codificante de la secuencia de nucleótidos intrínsecamente puede contener una región de codificación de péptido señal naturalmente enlazado en el marco de lectura de traducción con el segmento de la región de codificación que codifica el polipéptido segregado. Alternativamente, el extremo 5' de la secuencia codificante puede contener una región de codificación de péptido señal que es extranjera a la secuencia
- 20 codificante. La región de codificación de péptido señal extranjera puede ser requerido dónde la secuencia codificante no contener una región de codificación de péptido señal de forma natural. Alternativamente, la región de codificación de péptido señal extranjera puede simplemente reemplazar la región de codificación de péptido señal natural para mejorar la secreción del polipéptido. No obstante, cualquier región de codificación de péptido señal que dirija el polipéptido expresado en la vía secretora de una célula huésped de elección, es decir, segregado en un medio de cultivo, puede ser
- 25 empleada en la presente invención.
- [0093] Las regiones de codificación de péptido señale eficaces para células huéspedes bacterianas son las regiones de codificación de péptido señal obtenidas de los genes para amilasa maltogénica Bacillus de NCIB 11837, alfa-amilasa de *Bacillus stearothermophilus*, subtilisina de *Bacillus licheniformis*, beta-lactamasa de *Bacillus licheniformis*, proteasas neutras (nprT, nprS, nprM) de *Bacillus stearothermophilus*, y *Bacillus subtilis* prsA. Más péptidos señal son descritos por Simonen y Palva, 1993, Microbiological Reviews 57: 109-137.
- 30 [0094] Regiones de codificación de péptido señales eficaz para células huésped fúngicas filamentosas son las regiones de codificación de péptido señal obtenidas 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*.
- 35 [0095] Los péptidos señal útiles para células huéspedes de levadura se obtienen a partir de los genes para factor alfa de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*. Otras regiones de codificación de péptidos señal útiles son descritos por Romanos et al., 1992, supra.
- [0096] El péptido señal comprende o consiste en aminoácidos 1 a 16 de la SEC ID n°: 4. La región de codificación del péptido señal es nucleótidos 19 a 69 de la SEC ID n°: 3.
- 40 [0097] La secuencia de control puede también ser una región de codificación de propéptido que codifica para una secuencia de aminoácidos situada al amino término de un polipéptido. El polipéptido resultante es conocido como una proenzima o propolipéptido (o un zimógeno en algunos casos). Un propéptido es generalmente inactivo y se puede convertir a un polipéptido maduro activo por escisión autocatalítica o catalítica del propéptido del propolipéptido. La región de codificación de propéptido se puede obtener a partir de los genes para proteasa alcalina de *Bacillus subtilis*
- 45 (aprE), proteasa neutra de *Bacillus subtilis* (nprT), factor alfa de *Saccharomyces cerevisiae*, proteinasa aspártica de *Rhizomucor miehei*, y lacasa de *Myceliophthora thermophila* (WO 95/33836).
- [0098] Donde ambos péptido señal y regiones de propéptido están presentes al amino término de un polipéptido, la región de propéptido está situado junto al amino término de un polipéptido y la región de péptido señal está situado junto al amino término de la región de propéptido.
- 50 [0099] Puede también ser deseable el añadir secuencias reguladores que permitan la regulación de la expresión del polipéptido relativamente al crecimiento de la célula huésped. Ejemplos de sistemas reguladores son los que causan que la expresión del gen activada o desactivada en respuesta a un estímulo físico o químico, incluyendo la presencia de un compuesto regulador. Sistemas reguladores en sistemas procarióticos incluyen los sistemas operadores *lac*, *tac*, y *trp*. En levadura, se puede usar el sistema ADH2 o sistema GAL1. En hongos filamentosos, el promotor TAKA de alfa-amilasa, promotor de glucoamilasa de *Aspergillus niger*, y promotor de glucoamilasa de *Aspergillus oryzae* se pueden
- 55

utilizar como secuencias reguladoras. Otros ejemplos de secuencias reguladoras son los que permiten amplificación del gen. En sistemas eucarióticos, estos 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 operativamente enlazado con la secuencia reguladora.

5 Vectores de expresión

[0100] La presente invención también se refiere a vectores de expresión recombinantes que comprenden un polinucleótido de la presente invención, un promotor, y señales de parada traduccionales y transcripcionales. Los diferentes ácidos nucleicos y secuencias de control descritas aquí pueden ser unidos para producir un vector de expresión recombinante que puede incluir uno o más sitios de restricción convenientes para permitir para inserción o sustitución de la secuencia de nucleótidos que codifica el polipéptido en tales sitios. Alternativamente, una secuencia polinucleótida 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 el vector de expresión, la secuencia codificante se localiza en el vector de modo que la secuencia codificante está operativamente enlazado con las secuencias de control apropiadas para la expresión.

[0101] El vector de expresión recombinante puede ser cualquier vector (p. ej. un plásmido o virus) que puede ser convenientemente sometido a procedimientos de ADN recombinante y puede provocar expresión de la secuencia de nucleótidos. La elección del vector típicamente dependerá de la compatibilidad del vector con la célula huésped en la que se va a introducir el vector. Los vectores puede ser plásmidos circulares lineales o cerrados.

[0102] El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es independiente de la replicación cromosómica, p.ej., un plásmido, un elemento extracromosómico, un minicromosoma, o un cromosoma artificial. El vector puede contener cualquier medio para asegurar autorreplicación. Alternativamente, el vector puede ser uno que, cuando se introduce en la célula huésped, se integra en el genoma y se replica junto con el/los cromosoma/s en que ha sido integrada. Además, se puede usar un único vector o plásmido o dos o más vectores o plásmidos que juntos contienen el ADN total a ser introducido en el genoma de la célula huésped, o un transposón.

[0103] Los vectores de la presente invención preferiblemente contienen uno o más marcadores seleccionables que permiten selección fácil células transformadas, modificadas, transducidas, o similares. Un marcador seleccionable es un gen cuyo producto proporciona resistencia biocida o vírica, resistencia a metales pesados, prototrofia a auxótrofos, y similares.

[0104] Ejemplos de marcadores seleccionables bacterianos son los genes del de *Bacillus subtilis* o *Bacillus licheniformis*, o marcadores que confieren resistencia antibiótica tal como ampicilina, canamicina, cloranfenicol, o tetraciclina. Marcadores adecuados para células huéspedes de levadura son ADE2, HIS3, LEU2, LYS2, MET3, TRP1, y URA3. Marcadores seleccionables para el uso en una célula huésped fúngica filamentosa incluyen, pero no se limitan a, *amdS* (acetamidasa), *argB* (ornitina-carbamoiltransferasa, *bar* (fosfonitricina acetiltransferasa, *hph* (higromicina fosfotransferasa, *niaD* (nitrato-reductasa), *pyrG* (orotidina 5'-fosfato decarboxilasa), *sC* (sulfato adeniltransferasa), y *trpC* (antranilato sintasa), al igual que los equivalentes de los mismos. Preferidos para uso en una célula de *Aspergillus* son los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces higroscopicus*.

[0105] Los vectores de la presente invención preferiblemente contienen un elemento/s que permite/n la integración del vector en el genoma huésped de la célula o replicación autónoma del vector en la célula independiente del genoma.

[0106] Para la integración en el genoma de la célula huésped, el vector puede depender 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 de nucleótidos adicionales para dirigir la integración por recombinación homóloga en el genoma de la célula huésped a una/s ubicación/es preciso/s en el/los cromosoma/s. Para aumentar la probabilidad de integración en una ubicación precisa, los elementos integracionales preferiblemente deberían contener un número suficiente de ácidos nucleicos, tal como de 100 a 10.000 pares de bases, preferiblemente de 400 a 10.000 pares de bases, y de la forma más preferible de 800 a 10.000 pares de bases, que tienen un alto grado de identidad con la secuencia objetiva correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos integracionales pueden ser cualquier secuencia que sea homóloga a la secuencia objetivo en el genoma de la célula huésped. Además, los elementos integracionales pueden ser secuencias nucleótidas codificadores o no codificadores. Por otro lado, el vector se puede integrado en el genoma de la célula huésped por recombinación no homóloga.

[0107] Para la replicación autónoma, el vector puede comprender además un origen de replicación permitiendo al vector replicar de manera autónoma en la célula huésped en cuestión. El origen de replicación puede ser cualquier plásmido replicador que medie la replicación autónoma que funciona en una célula. El término "plásmido replicador" u "origen de replicación" se define aquí como una secuencia de nucleótidos que habilita un plásmido o vector para replicar in vivo.

[0108] Ejemplos de orígenes bacterianos de replicación son los orígenes de replicación de plásmidos pBR322, pUC19, pACYC177, and pACYC184 permitiendo replicación en *E. coli*, y pUB110, pE194, pTA1060, y pAMB1 permitiendo replicación en *Bacillus*.

5 [0109] Ejemplos de orígenes de replicación para su uso en una célula huésped de levadura son los orígenes de replicación de 2 micrones, ARS1, ARS4, la combinación de ARS1 y CEN3, y la combinación de ARS4 y CEN6.

[0110] Ejemplos de orígenes de replicación útiles en una célula fungosa 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 realizar según los métodos descritos en WO 00/24883.

10 [0111] Más de una copia de un polinucleótido de la presente invención puede ser insertado en una célula huésped para aumentar la producción del producto genético. Un aumento en el número de copias del polinucleótido se puede obtener 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
15 células en presencia del agente seleccionable apropiado.

[0112] Los procedimientos usados para enlazar los elementos anteriormente descritos para construir de los vectores de expresión recombinantes de la presente invención son conocidos por un experto en la materia (ver, p.ej. Sambrook et al., 1989, supra).

Células huéspedes

20 [0113] La célula huésped puede ser cualquier célula útil en la producción recombinante de un polipéptido de la presente invención, p.ej. un procariota o un eucariota.

[0114] La célula huésped procariótica puede ser cualquier bacteria Gram positiva o una bacteria Gram negativo. 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*, *Campilobacter*, *Helicobacter*, *Flavobacterium*, *Fusobacterium*, *Iliobacter*, *Neisseria*, y *Ureaplasma*.

[0115] La célula huésped bacteriana puede ser cualquier célula de *Bacillus*. Células de bacillus útil en la práctica de la presente invención incluyen, pero no se limitan a, células de *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus firmus*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus pumilus*, *Bacillus stearothermophilus*, *Bacillus subtilis*, y de *Bacillus thuringiensis*.

[0116] En un aspecto preferido, la célula huésped bacteriana es un *Bacillus amyloliquefaciens*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus* o célula de *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*.

[0117] La célula huésped bacteriana puede también ser cualquier célula de *Streptococcus*. Células de *Streptococcus* útiles en la práctica de la presente invención incluyen, pero no se limitan a, *Streptococcus equisimilis*, *Streptococcus pyogenes*, *Streptococcus uberis*, y *Streptococcus equi* subesp. *Zooepidemicus*.

40 [0118] 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 *Streptococcus uberis*. En otro aspecto preferido, la célula huésped bacteriana es una célula *Streptococcus equi* subsp. *Zooepidemicus*.

[0119] La célula huésped bacteriana puede también ser cualquier célula de *Streptomyces*. Células de *Streptomyces* útiles en la práctica de la presente invención incluyen, pero no se limitan a, *Streptomyces achromogenes*, *Streptomyces avermitilis*, *Streptomyces coelicolor*, *Streptomyces griseus*, y *Streptomyces lividans*.

[0120] 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*.

50 [0121] La introducción de ADN en una célula de *Bacillus* puede ser efectuado, p.ej. por transformación por protoplasto (ver, p.ej. Chang and Cohen, 1979, Molecular General Genetics 168: 111-115), usando células competentes (ver, e.g.,

Young and Spizizen, 1961, Journal of Bacteriology 81: 823-829 , or Dubnau and Davidoff-Abelson, 1971, Journal of Molecular Biology 56: 209-221), por electroporación (ver, e.g., Shigekawa and Dower, 1988, Biotechniques 6: 742-751), o por conjugación (see, e.g., Koehler and Thorne, 1987, Journal of Bacteriology 169: 5271-5278). La introducción de ADN en una célula de *E. coli* puede, p.ej. ser efectuado por transformación por protoplasto (ver, p.ej. Hanahan, 1983, J. Mol. Biol. 166: 557-580) o electroporación (ver, p.ej. Dower et al., 1988, Nucleic Acids Res. 16: 6127-6145). La introducción de ADN en una célula de *Streptomyces* puede, p.ej. ser efectuado por transformación por protoplasto y electroporación (ver, p.ej. Gong et al., 2004, Folia Microbiol. (Praha) 49: 399-405), por conjugación (ver, p.ej. Mazodier et al., 1989, J. Bacteriol. 171: 3583-3585), o por transducción (ver, p.ej. Burke et al., 2001, Proc. Natl. Acad. Sci. USA 98:6289-6294). La introducción de ADN en una célula de *Pseudomonas* puede, p.ej. ser efectuado por electroporación (ver, por ejemplo, Choi et al., 2006, J. Microbiol. Methods 64: 391-397) o por conjugación (ver, p.ej. Pinedo y Smets, 2005, Appl. Environ. Microbiol. 71: 51-57). La introducción de ADN en una célula de *Streptococcus* puede, p.ej. ser efectuado por competencia natural (ver, p.ej. Perry and Kuramitsu, 1981, Infect. Immun. 32: 1295-1297), por transformación por protoplasto (ver, p.ej. Catt y Jollick, 1991, Microbios. 68: 189-2070, por electroporación (ver, p.ej. Buckley et al., 1999, Appl. Environ. Microbiol. 65: 3800-3804) o por conjugación (ver, p.ej. Clewell, 1981, Microbiol. Rev. 45: 409-436). Sin embargo, se puede usar cualquier método conocido en la técnica para a introducción de ADN en una célula huésped.

[0122] La célula huésped puede también ser un eucariota, tal como una célula de mamífero, insecto, planta, o fúngica.

[0123] En un aspecto preferido, la célula huésped es una célula fúngica. "Hongo" según se utiliza en este caso incluye la *Phyla Ascomycota, Basidiomycota, Chytridiomycota, y Zygomycota* (como definido por Hawksworth et al., en Ainsworth and Bisby's Dictionary of The Fungi, 8th edition, 1995, CAB International, University Press, Cambridge, UK) así como la Oomycota (como se cita in Hawksworth et al., 1995, supra, page 171) y todos los hongos mitospóricos (Hawksworth et al., 1995, supra).

[0124] En un aspecto más preferido, la célula huésped fúngica es una célula de levadura. "Levadura" según se utiliza en este caso incluye la levadura ascoesporógena (Endomycetales), levadura basidio esporogena, y levadura de los Hongos Imperfecti (Blastomycetes). Ya que la clasificación de la levadura puede cambiar en el futuro, para los objetivos 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 N°. 9, 1980).

[0125] En un aspecto incluso más preferido, la célula huésped de levadura es una célula de *Candida, Hansenula, Kluyveromyces, Pichia, Saccharomyces, Schizosaccharomyces, o Yarrowia*.

[0126] En un aspecto más preferido, la célula huésped de levadura es 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*.

[0127] En otro aspecto más preferido, la célula huésped fúngica es una célula filamentosa fúngica. "Hongos filamentosos" incluye todas las formas filamentosas de la subdivisión Eumycota y Oomycota (como definido por Hawksworth et al., 1995, supra). Los hongos filamentosos generalmente se caracterizan por el hecho de que una pared micelial compuesta por quitina, celulosa, glucano, quitosano, manano, y otros polisacáridos complejos. El crecimiento vegetativo es mediante alargamiento hifal y el catabolismo de carbono es estrictamente aeróbico. En cambio, crecimiento vegetativo por levaduras tal como el *Saccharomyces cerevisiae* es mediante injerto de un talo unicelular y el catabolismo de carbono puede ser fermentativo.

[0128] En un aspecto incluso más preferido, la célula huésped fúngica filamentosa es una célula de *Acremonium, Aspergillus, Aureobasidium, Bjerkandera, Ceriporiopsis, Chrysosporium, Coprinus, Coriolus, Criptococcus, Filibasidium, Fusarium, Humicola, Magnaporthe, Mucor, Myceliophthora, Neocallimastix, Neurospora, Paecilomyces, Penicillium, Phanerochaete, Phlebia, Piromyces, Pleurotus, Schizophyllum, Talaromyces, Thermoascus, Thielavia, Tolypocladium, Trametes, o Trichoderma*.

[0129] En un aspecto más preferido, la célula huésped filamentosa fúngica es 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 filamentosa fúngica 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 sarcochroum, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, o Fusarium venenatum*. En otro aspecto más preferido, la célula huésped filamentosa fúngica 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 brasilianum*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eringii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.

[0130] Las células micóticas se pueden transformar por un proceso que implica formación de protoplasto, transformación de los protoplastos, y regeneración de la pared celular de una forma conocida per se. Procedimientos adecuados para la transformación de células huésped de *Aspergillus* y *Trichoderma* están descritas en EP 238 023 y Yelton et al., 1984, Proceedings of the National Academy of Sciences USA 81: 1470-1474. Los métodos adecuados para la transformación de especies de *Fusarium* están descritas por Malardier et al., 1989, gen 78: 147-156, y WO 96/00787. La levadura puede ser transformada usando los procedimientos descritos por Becker y Guarente, en Abelson, J.N. y Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., Nueva York ; Ito et al., 1983, Journal of Bacteriology 153: 163 ; and Hinnen et al., 1978, Proceedings of the National Academy of Sciences USA 75: 1920 .

Métodos de producción

[0131] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención, que comprende: (a) el cultivo de una célula, que en su forma salvaje es capaz de producir el polipéptido, bajo condiciones propicias para la producción del polipéptido; y (b) la recuperación del polipéptido. En un aspecto preferido, la célula es del género *Myceliophthora*. En un aspecto más preferido, la célula es *Myceliophthora thermophila* CBS 117.65. La presente invención también se refiere a métodos para producir un polipéptido de la presente invención, que comprende: (a) e cultivo de una célula huésped bajo condiciones propicias para la producción del polipéptido; y (b) la recuperación del polipéptido.

[0132] La descripción también se refiere a métodos para producir un polipéptido de la presente invención, que comprende: (a) el cultivo de una célula huésped bajo condiciones propicias para la producción del polipéptido, donde la célula huésped comprende una secuencia mutante de nucleótidos con al menos una mutación en la secuencia codificante de polipéptido maduro de la SEC ID n°: 3, donde la secuencia mutante de nucleótidos codifica un polipéptido que comprende o consiste en el polipéptido maduro de la SEC ID n°: 4, y (b) la recuperación el polipéptido.

[0133] En un aspecto preferido, el polipéptido maduro de la SEC ID n°: 4 es aminoácidos 17 a 389.

[0134] En los métodos de producción de la presente invención, las células se cultivan en un medio nutritivo adecuado para la producción del polipéptido usando métodos bien conocido en la técnica. P.ej. la célula se puede cultivar mediante cultivo en matraz de agitación, y fermentación a gran o a pequeña escala (incluyendo continuo, lote, lote alimentado, o fermentaciones de estado sólido) en laboratorio o fermentadores industriales realizados en un medio adecuado y bajo condiciones que permitan al polipéptido ser expresado y/o aislado. El cultivo se desarrolla en un medio nutritivo adecuado que comprende fuentes de nitrógeno y carbono y sales inorgánicas, usando procedimientos conocidos en la técnica. Los medios adecuados están disponibles en proveedores comerciales o se pueden preparar según composiciones publicadas (p. ej. en catálogos de la American Type Culture Collection). Si el polipéptido se segrega en el medio nutritivo, el polipéptido se puede recuperar directamente del medio. Si el polipéptido no está segregado en el medio, esto se puede recuperar de los lisatos de célula.

[0135] Los polipéptidos pueden ser detectados 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. P.ej. se puede utilizar un ensayo enzimático para determinar la actividad del polipéptido como se describe en este caso.

[0136] El polipéptido resultante puede ser recuperado usando métodos conocidos en la técnica p.ej. el polipéptido puede ser recuperado del medio nutritivo por procedimientos convencionales incluyendo, pero no limitado a, centrifugado, filtración, extracción, secado por pulverización, evaporación, o precipitación.

[0137] Los polipéptidos de la presente invención se pueden purificar mediante una variedad de procedimientos conocidos en la técnica incluyendo, pero no limitado a, cromatografía (p. ej. intercambio iónico, afinidad, cromatografía hidrofóbica, y exclusión de tamaño), procedimientos electroforéticos (p. ej. isoelectroenfoque preparatorio), solubilidad diferencial (p. ej. precipitación de sulfato amónico), SDS-PAGE, o extracción (ver, p.ej. Protein Purification, J.-C. Janson and Lars Ryden, editors, VCH Publishers, Nueva York, 1989) para obtener polipéptidos sustancialmente puros.

Plantas

[0138] La presente invención también se refiere a plantas, p.ej. una planta transgénica, parte de una planta, o una célula vegetal, que comprende un polinucleótido aislado que codifica un polipéptido con actividad endoglucanasa de la presente invención para expresar y producir el polipéptido en cantidades recuperables. El polipéptido se puede recuperar de la planta o parte de planta. Alternativamente, la planta o parte de la planta que contiene el polipéptido recombinante

se puede utilizar como tal para mejorar la calidad de unos alimentos o pienso, p.ej. mejorando el valor nutritivo, palatabilidad, y propiedades reológicas, o para destruir un factor antinutritivo.

5 [0139] La planta transgénica puede ser dicotiledónea (un dicot) o monocotiledónea (un monocot). Ejemplos de monocotiledónea planta son hierbas, tal como poa de los prados (poa pratense, Poa), hierba forrajera tal como Festuca, Lolium, césped templado, tal como Agrostis, y cereales, p.ej. trigo, avena, centeno, cebada, arroz, sorgo, y maíz.

[0140] Ejemplos de plantas dicotiledóneas son tabaco, leguminosas, tales como altramuza, patata, remolacha azucarera, guisante, judía y judía de soja, y plantas crucíferas (de la familia *Brassicaceae*), tal como coliflor, semilla de colza, y el organismo modelo cercanamente relacionado *Arabidopsis thaliana*.

10 [0141] Ejemplos de partes de planta son tallo, callo, hojas, raíz, frutas, semillas, y tubérculos al igual que los tejidos individuales que comprenden estas partes, p.ej. epidermis, mesófilo, parénquima, tejidos vasculares, meristemas. Los compartimentos específicos de célula vegetal, tales como cloroplastos, apoplastos, mitocondria, vacuolas, peroxisomas y citoplasma también se consideran partes de una planta. Además, cualquier célula vegetal, cualquiera que sea el origen del tejido, se considera parte de una planta. Asimismo, partes de planta tales como tejidos específicos y células aisladas para facilitar el uso de la invención son también consideradas partes de planta, p.ej. embriones, endospermas, aleurona
15 y revestimientos de semillas.

[0142] También se incluyen dentro del campo de la presente invención la progenie de tales plantas, partes de planta, y células vegetales.

20 [0143] La planta transgénica o célula vegetal que exprese un polipéptido de la presente invención se puede construir conforme a métodos conocidos en la técnica. Resumiendo, la planta o célula vegetal es construido incorporando uno o más constructos de expresión que codifican un polipéptido de la presente invención en el genoma de la planta huésped o genoma de cloroplasto y propagando la planta modificada resultante o célula vegetal en una planta o célula vegetal transgénica.

25 [0144] Convenientemente el constructo de expresión es un constructo de ácidos nucleicos que comprende un polinucleótido que codifica un polipéptido de la presente invención enlazado operativamente con secuencias reguladoras apropiadas requeridas para la expresión de la secuencia de nucleótidos en la planta o parte de planta de elección. Además, el constructo de expresión puede comprender un marcador seleccionable útil para identificar células huésped en las cuales el constructo de expresión ha sido integrado y secuencias de ADN necesarias para la introducción del constructo en la planta en cuestión (éste depende del método de introducción de ADN que sea usada).

30 [0145] La elección de secuencias reguladoras, tales como secuencias promotoras y terminadoras y opcionalmente secuencias de señal o de tránsito se determinan, p.ej. basándose en cuando, dónde, y como se quiere expresar el polipéptido. P.ej. la expresión del gen que codifica un polipéptido de la presente invención puede ser constitutiva o inducible, o puede ser desarrollable, específico de fase o tejido, y el producto genético puede ser dirigido a un tejido específico o parte de planta tales como semillas u hojas. Secuencias reguladoras son, p.ej. descritos por Tague et al., 1988, *Plant Physiology* 86: 506.

35 [0146] Para una expresión constitutiva se pueden usar el 35S-CaMV, la ubiquitina 1 de maíz, y el promotor de actina 1 de arroz (Franck et al., 1980, *Cell* 21: 285-294, Christensen et al., 1992, *Plant Mo. Biol.* 18: 675-689; Zhang et al., 1991, *Plant Cell* 3: 1155-1165). Promotores específicos de un órgano pueden ser, p.ej. un promotor de tejidos sumideros de almacenamiento tales como semillas, tubérculos de patata, y frutas (Edwards & Coruzzi, 1990, *Ana. Rev. Genet.* 24: 275-303), o de tejidos sumidero metabólicos tales como meristemas (Ito et al., 1994, *Plant Mol. Biol.* 24: 863-878), un promotor específico de semilla tales como la glutelina, prolamina, globulina, o promotor de albúmina de arroz (Wu et al., 1998, *Plant and Cell Physiology* 39: 885-889), un promotor de *Vicia faba* de la legúmina B4 y el gen de proteína de semilla desconocido de *Vicia faba* (Conrad et al., 1998, *Journal of Plant Physiology* 152: 708-711), un promotor de una proteína de cuerpo de aceite de semilla (Chen et al., 1998, *Plant and Cell Physiology* 39: 935-941), el promotor de napA de proteína de almacenamiento de *Brassica napus*, o cualquier otro promotor específico de semilla conocido en la técnica, p.ej. como se describe en WO 91/14772. Además, el promotor puede ser un promotor específico de hoja tal como el promotor de eritrocitos de arroz o tomate (Kyojuka et al., 1993, *Plant Physiology* 102: 991-1000, el promotor de gen de metiltransferasa de adenina de virus de *chlorella* (Mitra y Higgins, 1994, *Plant Molecular Biology* 26: 85-93), o el promotor de gen *aldP* de arroz (Kagaya et al., 1995, *Molecular and General Genetics* 248: 668-674), o un promotor inducible por lesiones tal como el promotor pin2 de la patata (Xu et al., 1993, *Plant Molecular Biology* 22: 573-588). Asimismo, el promotor puede ser inducible por tratamientos abióticos tales como temperatura, sequía, o alteraciones en salinidad o inducido por sustancias exógenamente aplicadas que activan el promotor, p.ej. etanol, estrógenos, hormonas de planta tales como etileno, ácido abscísico, y ácido giberélico, y metales pesados.

55 [0147] También se puede usar un elemento intensificador del promotor para conseguir expresión más alta de un polipéptido de la presente invención en la planta. P.ej. el elemento intensificador del promotor puede ser un intrón que es colocado entre el promotor y la secuencia de nucleótidos que codifica un polipéptido de la presente invención. P.ej. Xu et al., 1993, supra, describe el uso del primer intrón del gen de actina 1 de arroz para mejorar la expresión.

[0148] El gen de marcador seleccionable y cualquier otra parte del constructo de expresión se puede elegir de aquellas disponibles en la técnica.

[0149] El constructo de ácido nucleico se incorpora en el genoma de la planta según técnicas convencionales conocidas en la técnica, incluyendo la transformación mediada por *Agrobacterium*, transformación mediada por virus, microinyección, bombardeo de partículas, transformación biolística, y electroporación (Gasser et al., 1990, Science 244: 1293; Potrykus, 1990, Bio/Technology 8: 535; Shimamoto et al., 1989, Nature 338: 274).

[0150] Actualmente, el método de elección para generar dicotiledóneas transgénicas es la transferencia de genes mediada por *Agrobacterium tumefaciens* (para conseguir una revisión, ver Hooykas y Schilperoort, 1992, Plant Molecular Biology 19: 15-38) y puede también ser usadas para transformar monocotiledóneas, aunque otros métodos de transformación son frecuentemente usados para estas plantas. Actualmente, el método de elección para generar monocotiledóneas transgénicas es bombardeo de partículas (oro microscópico o partículas de tungsteno revestidos con el ADN transformante) de callos embrionarios o embriones en desarrollo (Christou, 1992, Plant Journal 2: 275-281; Shimamoto, 1994, Current Opinion Biotechnology 5: 158-162; Vasil et al., 1992, Bio/Technology 10: 667-674). Un método alternativo para la transformación de monocotiledóneas se basa en transformación por protoplasto como se describe por Omirulleh et al., 1993, Plant Molecular Biology 21: 415-428.

[0151] Tras la transformación, los transformantes con el constructo de expresión incorporado son seleccionados y regenerados a plantas enteras según métodos bien conocido en la técnica. Frecuentemente el procedimiento de transformación se diseña para la eliminación selectiva de genes de selección bien durante la regeneración o en las siguientes generaciones usando, p.ej. cotransformación con dos constructos de t-ADN separados o un lugar específico de escisión del gen de selección mediante una recombinasa específica.

[0152] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención que comprende: (a) el cultivo de una planta transgénica o de una célula vegetal que comprende un polinucleótido que codifica un polipéptido con actividad endoglucanasa de la presente invención bajo condiciones propicias para la producción del polipéptido; y (b) recuperación el polipéptido.

Composiciones

[0153] La presente invención también se refiere a composiciones comprendientes un polipéptido de la presente invención. Preferiblemente, las composiciones se enriquecen en tal polipéptido. El término "enriquecido" indica que la actividad endoglucanasa de la composición ha sido aumentada, p.ej. con un factor de enriquecimiento de al menos 1.1.

[0154] La composición puede comprender un polipéptido de la presente invención como el componente mayor enzimático, p.ej. una composición monocomponente. Alternativamente, la composición puede comprender actividades enzimáticas múltiples, tal como una aminopeptidasa, amilasa, carbohidrasa, carboxipeptidasa, catalasa, celulasa, quitinasa, cutinasa, ciclodextrina glicosiltransferasa, desoxirribonucleasa, esterasa, alfa-galactosidasa, galactosidasa de beta, glucoamilasa, alfa-glucosidasa, beta-glucosidasa, haloperoxidasa, invertasa, lacasa, lipasa, manosidasa, oxidasa, enzima pectinolítica, peptidoglutaminasa, peroxidasa, fitasa, polifenoloxidasa, enzima proteolítica, ribonucleasa, transglutaminasa, o xilanasa. La/s enzima/s adicional/es puede/n ser producidos, p.ej. por un microorganismo de la especie *Aspergillus*, preferiblemente *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, o *Aspergillus oryzae*; de a especie *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 sarcochromum*, *Fusarium sulphureum*, *Fusarium toruloseum*, *Fusarium trichothecioides*, o *Fusarium venenatum*; de la especie *Humicola*, preferiblemente *Humicola insolens* o *Humicola lanuginosa*; o de la especie *Trichoderma*, preferiblemente *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.

[0155] Las composiciones de polipéptido se pueden preparar conforme a métodos conocidos en la técnica y puede tener forma líquida o de composición seca. P.ej. la composición de polipéptido puede estar en forma de granulado o un microgranulado. El polipéptido a ser incluido en la composición puede ser estabilizado conforme a métodos conocidos en la técnica.

[0156] Se dan ejemplos más abajo de los 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 cuales se pueden usar la composición pueden ser determinadas basándose en métodos conocidos en la técnica.

Usos

[0157] La presente invención también se refiere a métodos para degradar o convertir un material celulósico, que comprende: el tratado del material celulósico con una composición que comprende una cantidad eficaz de un polipéptido con actividad endoglucanasa de la presente invención. En un aspecto preferido, el método además comprende la recuperación del material celulósico degradado o convertido.

[0158] Los polipéptidos y células huésped de la presente invención se pueden utilizar en la producción de monosacáridos, disacáridos, y polisacáridos como materia prima química o de fermentación a partir de biomasa celulósica para de la producción de etanol, plástico, otros productos o intermedios. La composición que comprende el polipéptido con actividad endoglucanasa puede ser en forma de un caldo crudo de fermentación con o sin las células extraídas o en forma de una preparación enzimática semi-purificada o purificada. La composición puede también comprender otras proteínas y enzimas útiles en el tratamiento de biomasa, p.ej. celobiohidrolasa, beta-glucosidasa, enzimas hemicelulolíticas, potenciadores (WO 2005/074647, WO 2005/074656), etc. Alternativamente, la composición puede comprender una célula huésped de la presente invención como una fuente del polipéptido con actividad endoglucanasa en un proceso de fermentación con la biomasa. En particular, los polipéptidos y células huésped de la presente invención se pueden utilizar para aumentar el valor de los residuos de tratamiento (grano secos de destilería, grano gastado de cervecero, bagazo de caña de azúcar, etc.) mediante degradación completa o parcial de celulosa o hemicelulosa. La célula huésped también puede contener genes heterólogos o nativos que codifican otras proteínas y enzimas, mencionados arriba, útiles en el tratamiento de la biomasa.

[0159] En los métodos de la presente invención, cualquier material celulósico, tal como la biomasa, puede ser usado. Se entiende en la presente que el término "material celulósico" engloba la lignocelulosa. Biomasa puede incluir, pero no se limita a, recursos de madera, desperdicios sólidos municipales, papel usado, brotes, y residuos de cosecha (ver, p.ej. Wiseloge et al., 1995, in Handbook on Bioethanol (Charles E. Wyman, editor), pp. 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, in Advances in Biochemical Engineering/Biotechnology, T. Scheper, managing editor, Volume 65, pp.23-40, Springer-Verlag, Nueva York).

[0160] 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, producido después de que la célula ha detenido su crecimiento, también contiene polisacáridos y se refuerza mediante un enlace de lignina polimérica reticulada de manera covalente a la hemicelulosa. La celulosa es un homopolímero de anhidrocelobiosa y así un beta-(1-4)-D-glucano lineal, mientras que las hemicelulosas incluyen una variedad de compuestos, tal como xilanos, xiloglucanos, arabinoxilanas, y mananos en estructuras ramificado complejas con un espectro de sustituyentes. Aunque generalmente polimorfa, la celulosa se encuentra en el tejido vegetal principalmente como una matriz cristalina insoluble de cadenas paralelas de glucano. Las hemicelulosas normalmente tienen enlace de hidrógeno a celulosa, al igual que a otras hemicelulosas, que ayudan a estabilizar la matriz de la pared celular.

[0161] Se usan tres tipos de encimas para descomponer biomasa celular:

(1) Las "endo-1,4-beta-glucanasas" o 1,4-beta-D-glucano-4-glucanohidrolasas (EC 3,2,1,4), que actúan de forma aleatoria en 1,4-beta-glucan sustratos insolubles y solubles.

(2) Las "exo-1,4-beta-D-glucanasas" incluyendo las 1,4-beta-D-glucan glucohidrolasas (EC 3,2,1,74), que liberan D-glucosa de 1,4-beta-D-glucanos e hidrolizan D-celobiosa lentamente, y celobiohidrolasas (1,4-beta-D-glucano celobiohidrolasas, EC 3,2,1,91), que liberan D-celobiosa de 1,4-beta-glucanos.

(3) Las "beta-D-glucosidasas" o glucohidrolasas de beta-D-glucósidos (EC 3.2.1.21), que actúan para liberar unidades de D-glucosa de celobiosa y celodextrinas solubles, así como una variedad de glicósidos.

[0162] Los polipéptidos con actividad endoglucanasa de la presente invención son preferiblemente usadas en conjunción con otras proteínas celulolíticas, p.ej. exo-1,4-beta-D-glucanasas y beta-D-glucosidasas, para degradar el componente de celulosa del sustrato de biomasa, (ver, p.ej. Brigham et al., 1995, in Handbook on Bioethanol (Charles E. Wyman, editor), pp.119-141, Taylor & Francis, Washington D.C. ; Lee, 1997, Journal of Biotechnology 56: 1-24).

[0163] Las exo-1,4-beta-D-glucanasas y beta-D-glucosidasas se pueden producir mediante cualquier método conocido en la técnica (ver, p.ej. Bennett, J.W. and LaSure, L. (eds.), More Gene Manipulations in Fungi, Academic Press, CA, 1991).

[0164] Las cantidades óptimas de un polipéptido con actividad endoglucanasa y otras proteínas celulolíticas depende de diferentes factores incluyendo, pero no limitado a, la mezcla de proteínas celulolíticas de componente, el sustrato celulósico, la concentración de sustrato celulósico, e/los pretratamiento/s del sustrato celulósico, temperatura, tiempo, pH, e inclusión de un organismo de fermentación (p. ej. levadura para sacarificación y fermentación simultánea). El término "proteínas celulolíticas" se define en la presente como aquellas proteínas o mezclas de proteínas mostradas como capaces de hidrolizar o convertir o de degradar celulosa bajo las condiciones probadas.

[0165] La cantidad de polipéptido con actividad endoglucanasa por g de material celulósico puede ser aproximadamente de 0,5 a 50 mg, preferiblemente aproximadamente de 0,5 a 40 mg, más preferiblemente aproximadamente de 0,5 a 25 mg, más preferiblemente aproximadamente de 0,75 a 20 mg, más preferiblemente aproximadamente fr 0,75 a 15 mg, incluso más preferiblemente aproximadamente de 0,5 a 10 mg, y de la forma más preferible aproximadamente de 2,5 a

sobre 10 mg por g de material celulósico.

[0166] La cantidad de proteínas celulolíticas por g de material celulósico puede ser aproximadamente de 0,5 a 50 mg, preferiblemente aproximadamente de 0,5 a 40 mg, más preferiblemente aproximadamente de 0,5 a 25 mg, más preferiblemente aproximadamente de 0,75 a 20 mg, más preferiblemente aproximadamente de 0,75 a 15 mg, incluso más preferiblemente aproximadamente de 0,5 a 10 mg, y de la forma más preferible aproximadamente de 2,5 a 10 mg por g de material celulósico.

[0167] En los métodos de la presente invención, la composición puede ser suplementada por una o más actividades enzimáticas adicionales para mejorar la degradación del material celulósico. Enzimas preferidas adicionales son hemicelulasas, estererasas (p. ej. lipasas, fosfolipasas, y/o cutinasas), proteasas, lacasas, peroxidadas, o mezclas derivadas.

[0168] En los métodos de la presente invención, la/s enzima/s adicional/es pueden ser añadido/s antes de o durante la fermentación, incluso durante o después de la propagación de el/los microorganismo/s de fermentación.

[0169] Las enzimas puede ser derivadas u obtenidas de cualquier origen adecuado, incluyendo origen bacteriano, fúngico, levadura o mamífero. El término "obtenido" significa en la presente que la enzima puede haber sido aislada a partir de un organismo que produce naturalmente la enzima como una enzima nativa. El término "obtenida" también significa en la presente que la enzima puede haber sido producida recombinantemente en un organismo huésped, donde la enzima producida por recombinación es o extranjera o nativa al organismo huésped o tiene una secuencia de aminoácidos modificada, p.ej. teniendo uno o más aminoácidos que son eliminados, insertados y/o sustituidos, es decir, una enzima producida por recombinación que es un mutante y/o un fragmento de una secuencia de aminoácidos nativa o una enzima producida por procesos de redistribución de ácido nucleico conocidos en la técnica. Se engloba en el significado de una enzima nativa las variantes naturales y en el significado de enzimas extranjeras las variantes obtenidas mediante recombinación, tales como por mutagénesis dirigida o redistribución.

[0170] Las enzimas puede también ser purificadas. El término "purificada" tal y como se usa en este caso cubre enzimas libres de otros componentes del organismo del cual son derivadas. El término "purificada" también cubre las enzimas libres de otros componentes del organismo nativo del que son obtenidas. Las enzimas puede ser purificadas, sólo con cantidades menores de otra proteínas que estén presente. La expresión "otras proteínas" se refiere en particular a otras enzimas. El término "purificado" tal y como se usa en este caso también se refiere a la eliminación de otra componentes, particularmente de otras proteínas y más particularmente de otras enzimas presentes en la célula de origen de la enzima de la invención. La enzima puede ser "sustancialmente pura," es decir, libre de otros componentes del organismo en el que está producido, que es, p.ej. un organismo huésped para enzimas producidas por recombinación. En un aspecto preferido, las enzimas son al menos 75% (w/w), preferiblemente al menos 80%, más preferiblemente al menos 85%, más preferiblemente al menos 90%, más preferiblemente al menos 95%, más preferiblemente al menos 96%, más preferiblemente al menos 97%, incluso más preferiblemente al menos 98%, o de la forma más preferible al menos 99% puro. En otro aspecto preferido, la enzima es 100% puro.

[0171] Las enzimas usadas en la presente invención puede estar en cualquier forma adecuada para su uso en los procesos descritos en la presente, tal como, p.ej. un caldo de fermentación crudo con o sin células, un polvo seco o granulado, un granulado no polvoriento, un líquido, un líquido estabilizado, o una enzima protegida. Granulados puede ser producido, por ejemplo, como descritos en las patentes estadounidenses Nos. 4,106,991 y 4,661,452 y opcionalmente pueden ser revestidos por un proceso conocido en la técnica. Las preparaciones enzimáticas líquidas pueden, p.ej. ser estabilizadas añadiendo estabilizadores como un azúcar, un alcohol de azúcar u otro poliol, y/o ácido láctico u otro ácido orgánico según el proceso establecido. Las enzimas protegidas se pueden preparar según el proceso descrito en EP 238,216.

[0172] Los métodos de la presente invención se pueden utilizar para procesar de un material celulósico a muchos productos orgánicos útiles, productos químicos y combustibles. Además de etanol, algunos de los artículos de consumo y productos químicos de especialidad que se pueden producir a partir de celulosa incluyen xilosa, acetona, acetato, glicina, lisina, ácidos orgánicos (p. ej. ácido láctico), 1,3-propanediol, butanodiol, glicerol, etilenglicol, furfural, poli-hidroxi-alkanoatas, cis, ácido cis-mucónico, y pienso para animales Lynd, L. R., Wyman, C. E., and Gerngross, T. U., 1999, Biocommodity Engineering, Biotechnol. Prog., 15: 777-793 ; Philippidis, G. P., 1996, Cellulose bioconversion technology, in Handbook on Bioethanol: Production and Utilization, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212; and Ryu, D. D. Y., and Mandels, M., 1980, Cellulases: biosynthesis and applications, Enz. Microb. Technol., 2: 91-102). Los beneficios de coproducción potenciales se extienden más allá de la síntesis de productos orgánicos múltiples de carbohidrato fermentable. Residuos ricos de lignina restantes tras el tratamiento biológico se pueden convertir a productos químicos derivados de lignina, o usado para producción de energía.

[0173] Los métodos convencionales usados para procesar el material celulósico conforme a los métodos de la presente invención son bien entendidas a los expertos en la técnica. Los métodos de la presente invención pueden ser usados implementado cualquier aparato de tratamiento de biomasa convencional configurado para operar conforme a la

invención.

- [0174] Tal aparato puede incluir un reactor discontinuo de agitado, un reactor continuo de agitado con ultrafiltración, un reactor en flujo pistón de columna (Gusakov, A. V., and 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), an attrition reactor (Ryu, S. K., and Lee, J. M., 1983, Bioconversion of waste cellulose by using an attrition bioreactor, *Biotechnol. Bioeng.* 25: 53-65), or a reactor with intensive stirring induced by an electromagnetic field (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).
- [0175] Los métodos convencionales incluyen, pero no se limitan a, sacarificación, fermentación, hidrólisis y fermentación separada (SHF), sacarificación y fermentación simultánea (SSF), sacarificación y cofermentación simultánea (SSCF), hidrólisis y fermentación híbrida (HHF), y conversión microbiana directa (DMC).
- [0176] El SHF usa pasos separados en el proceso para primero hidrolizar enzimáticamente la celulosa a glucosa y luego fermentar la glucosa a etanol. En el SSF, la hidrólisis enzimática de la celulosa y la fermentación de glucosa a etanol se combina en un solo paso (Philippidis, G. P., 1996, Cellulose Bioconversion Technology, in *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212). El SSCF incluye la cofermentación de azúcares múltiples (Sheehan, J., and 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). El HHF incluye dos pasos separados realizados en el mismo reactor pero a temperaturas diferentes, es decir, sacarificación enzimática de alta temperatura seguida de SSF a una temperatura inferior de que la puede tolerar la cepa de fermentación. El DMC combina los tres procesos (producción de celulosa, hidrólisis de celulosa, y fermentación) en un paso (Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I.S., 2002, *Microbial cellulose utilization: Fundamentals and biotechnology*, *Microbiol. Mol. Biol. Reviews* 66: 506-577).
- [0177] "Proceso de fermentación" o "fermentación" se refiere a cualquier proceso de fermentación o cualquier proceso que comprende un paso de fermentación. Un proceso de fermentación incluye, sin limitarse a, procesos de fermentación usados para producir alcoholes incluyendo productos de fermentación (p. ej. arabinitol, butanol, etanol, glicerol, metanol, 1,3-propanediol, sorbitol, y xilitol); ácidos orgánicos (p. ej. ácido acético, ácido acetónico, ácido adípico, ácido ascórbico, ácido cítrico, ácido 2,5-diketo-D-glucónico, ácido fórmico, ácido fumárico; ácido glucárico, ácido glucónico, ácido glucurónico, ácido glutárico, ácido 3-hidroxypropionico, ácido itacónico, ácido láctico, ácido málico, ácido malónico, ácido oxálico, ácido propiónico, ácido succínico, y ácido xilónico); cetonas (p. ej. acetona); aminoácidos (p. ej. ácido aspártico, ácido glutámico, glicina, lisina, serina, y treonina); gases (p. ej. metano, hidrógeno (H₂), dióxido de carbono (CO₂), y monóxido de carbono (CO)). Los procesos de fermentación también incluyen los procesos de fermentación usados en la industria de alcohol consumible (p. ej. cerveza y vino), industria lechera (p. ej. productos lácteos fermentados), industria del cuero, e industria del tabaco.
- [0178] La presente invención además se refiere a métodos de producción de una sustancia, que comprende: (a) el sacarificado de un material celulósico con una composición que comprenda una cantidad eficaz de un polipéptido con actividad endoglucanasa; (b) el fermentado del material sacarificado celulósico del paso (a) con uno o más microorganismos fermentativos; y (c) recuperación de la sustancia de la fermentación. La composición que comprende el polipéptido con actividad endoglucanasa puede estar en forma de un caldo crudo de fermentación con o sin las células eliminadas o en forma de una preparación enzimática semi-purificada o purificada o la composición puede comprender una célula huésped de la presente invención como una fuente del polipéptido con actividad endoglucanasa en un proceso de fermentación con la biomasa.
- [0179] La sustancia puede ser cualquier sustancia derivada de la fermentación. En una forma de realización preferida, la sustancia es un alcohol. Se entiende que el término "alcohol" engloba una sustancia que contiene una o más fracciones de hidróxilo. En una forma de realización más preferida, el alcohol es arabinitol. En otra forma de realización más preferida, el alcohol es butanol. En otra forma de realización más preferida, el alcohol es etanol. En otra forma de realización más preferida, el alcohol es glicerol. En otra forma de realización más preferida, el alcohol es metanol. En otra forma de realización más preferida, el alcohol es 1,3-propanediol. En otra forma de realización más preferida, el alcohol es sorbitol. En otra forma de realización más preferida, el alcohol es xilitol. Ver, p.ej. Gong, C. S., Cao, N. J., Du, J., and Tsao, G. T., 1999, Ethanol production from renewable resources, in *Advances in Biochemical Engineering/Biotechnology*, Scheper, T., ed., Springer-Verlag Berlin Heidelberg, Germany, 65: 207-241 ; Silveira, M. M., and Jonas, R., 2002, The biotechnological production of sorbitol, *Appl. Microbiol. Biotechnol.* 59: 400-408 ; Nigam, P., and Singh, D., 1995, Processes for fermentative production of xylitol - a sugar substitute, *Process Biochemistry* 30 (2): 117-124 ; Ezeji, T. C., Qureshi, N. and Blaschek, H. P., 2003, Production of acetone, butanol and ethanol by *Clostridium beijerinckii* BA101 and in situ recovery by gas stripping, *World Journal of Microbiology and Biotechnology* 19 (6): 595-603.
- [0180] En otra forma de realización preferida, la sustancia es un ácido orgánico. En otra forma de realización más preferida, el ácido orgánico es ácido acético. En otra forma de realización más preferida, el ácido orgánico es ácido acetónico. En otra forma de realización más preferida, el ácido orgánico es ácido adípico. En otra forma de realización

- más preferida, el ácido orgánico es ácido ascórbico. En otra forma de realización más preferida, el ácido orgánico es ácido cítrico. En otra forma de realización más preferida, el ácido orgánico es ácido 2,5-diketeto-D-glucónico. En otra forma de realización más preferida, el ácido orgánico es ácido fórmico. En otra forma de realización más preferida, el ácido orgánico es ácido fumárico. En otra forma de realización más preferida, el ácido orgánico es ácido glucárico. En otra forma de realización más preferida, el ácido orgánico es ácido glucónico. En otra forma de realización más preferida, el ácido orgánico es ácido glucurónico. En otra forma de realización más preferida, el ácido orgánico es ácido glutárico. En otra forma de realización preferida, el ácido orgánico es ácido 3-hidroxipropiónico. En otra forma de realización más preferida, el ácido orgánico es ácido itacónico. En otra forma de realización más preferida, el ácido orgánico es ácido láctico. En otra forma de realización más preferida, el ácido orgánico es ácido málico. En otra forma de realización más preferida, el ácido orgánico es ácido malónico. En otra forma de realización más preferida, el ácido orgánico es ácido oxálico. En otra forma de realización más preferida, el ácido orgánico es ácido propiónico. En otra forma de realización más preferida, el ácido orgánico es ácido succínico. En otra forma de realización más preferida, el ácido orgánico es ácido xilónico. Ver, p.ej. Chen, R., and Lee, Y. Y., 1997, Membrane-mediated extractive fermentation for lactic acid production from cellulosic biomass, Appl. Biochem. Biotechnol. 63-65: 435- 448 .
- [0181] En otra forma de realización preferida, la sustancia es una cetona. Será entendido que el término "cetona" engloba una sustancia que contiene una o más fracciones de cetona. En otra forma de realización más preferida, la cetona es acetona. Ver, p.ej. Qureshi y Blaschek, 2003, supra.
- [0182] En otra forma de realización preferida, la sustancia es un aminoácido. En otra forma de realización más preferida, el ácido orgánico es ácido aspártico. En otra forma de realización más preferida, el aminoácido es ácido glutámico. En otra forma de realización más preferida, el aminoácido es glicina. En otra forma de realización más preferida, el aminoácido es lisina. En otra forma de realización más preferida, el aminoácido es serina. En otra forma de realización más preferida, el aminoácido es treonina. Ver, p.ej. Richard, A., and 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 .
- [0183] En otra forma de realización preferida, la sustancia es un gas. En otra forma de realización más preferida, el gas es metano. En otra forma de realización más preferida, el gas es H₂. En otra forma de realización más preferida, el gas es CO₂. En otra forma de realización más preferida, el gas es CO. Ver, p.ej. Kataoka, N., A. Miya, and K. Kiriya, 1997, Studies on hydrogen production by continuous culture system of hydrogen-producing anaerobic bacteria, Water Science and Technology 36 (6-7): 41-47 ; and Gunaseelan V.N. in Biomass and Bioenergy, Vol. 13 (1-2), pp. 83-114, 1997 , Anaerobic digestion of biomass for methane production: A review.
- [0184] La producción de una sustancia de material celulósico típicamente requiere cuatro pasos mayores. Estos cuatro pasos son pretratamiento, hidrólisis enzimática, fermentación, y recuperación. Abajo se ejemplifica un proceso para producir etanol, pero se entiende que procesos similares puede utilizarse para producir otras sustancias, p.ej. las sustancias anteriormente descritas.
- [0185] Pretratamiento. En el pretratamiento o paso de prehidrólisis, el material celulósico se calienta para descomponer la lignina y la estructura de carbohidrato, solubilizar la mayor parte de la hemicelulosa, y hacer la fracción de celulosa accesible a enzimas celulolíticas. El calentamiento se realiza o bien directamente con vapor o bien en compuesto acuoso donde un catalizador también puede ser añadido al material para acelerar las reacciones. Catalizadores incluyen ácidos fuertes, tales como ácido sulfúrico y SO₂, o álcali, tal como hidróxido sódico. El propósito de la fase previa de pretratamiento es la de facilitar la penetración de las enzimas y microorganismos. La biomasa celulósica puede también estar sujeto a una pretratamiento de explosión de vapor hidrotermal (ver solicitud de patente estadounidense N°. 20020164730).
- [0186] Sacarificación. En el paso de la hidrólisis enzimática, también conocido como sacarificación, las enzimas como se describen en este caso se agregan al material pretratado para convertir la fracción de celulosa a glucosa y/o a otros azúcares. La sacarificación generalmente se realiza en reactores de tanque de agitado o en fermentadores bajo pH, temperatura, y condiciones de mezcla controlados. Un paso de sacarificación puede durar hasta 200 horas. La sacarificación se puede realizar fuera a temperaturas de aproximadamente de 30°C a 65°C, en particular alrededor de 50°C, y a un pH en el intervalo de entre aproximadamente 4 y 5, especialmente alrededor de pH 4.5. Para producir glucosa que puede ser metabolizada por levadura, la hidrólisis típicamente es realizada en presencia de una beta-glucosidasa.
- [0187] Fermentación. En el paso de fermentación, los azúcares, liberados del material celulósico como resultado del pretratamiento y pasos de hidrólisis enzimática, se fermentan a etanol mediante un organismo de fermentación, tal como levadura. La fermentación puede también efectuarse simultáneamente con la hidrólisis enzimática en el mismo recipiente, nuevamente bajo pH, temperatura, y condiciones de mezcla controlados. Cuando la sacarificación y la fermentación se realizan simultáneamente en el mismo recipiente, el proceso es generalmente denominado sacarificación y fermentación simultánea o SSF.

[0188] Cualquier sustrato adecuado celulósico o materia prima se puede utilizar en un proceso de fermentación de la presente invención. El sustrato generalmente se selecciona basándose en el producto de fermentación deseado, es decir, la sustancia a ser obtenido de la fermentación, y el proceso empleado, como es bien conocido en la técnica. Ejemplos de sustratos adecuados para uso en los métodos de presente invención incluyen materiales con celulosa, tales como madera o residuos de planta o azúcares moleculares bajos DP1-3 obtenidos de material celulósico procesado que se puede metabolizar mediante el microorganismo fermentativo, y que se puede suministrar por adición directa al medio de fermentación.

[0189] El término "medio de fermentación" se entiende para referirse a un medio antes de que el/los microorganismo/s de fermentativo sea/sean añadidos, tal como, un medio que resulta de un proceso de sacarificación, al igual que un medio usado en un proceso de sacarificación y fermentación simultánea (SSF).

[0190] "Microorganismo fermentativo" se refiere a cualquier microorganismo adecuado para uso en un proceso deseado de fermentación. Microorganismos fermentativos adecuados según la invención son capaces de fermentar, es decir, de convertir, azúcares, tales como glucosa, xilosa, arabinosa, manosa, galactosa, u oligosacáridos directa o indirectamente en el producto de fermentación deseado. Ejemplos de microorganismos fermentativos incluyen organismos fúngicos, tal como la levadura. Las levaduras preferidas incluyen cepas de *Saccharomyces spp.*, y en particular, *Saccharomyces cerevisiae*. Las levaduras disponibles comercialmente incluyen, p. ej. Red Star®/™/Lesaffre Ethanol Red (disponible de Red Star/Lesaffre, EEUU) FALI (disponible de Fleischmann's Yeast, a division of Burns Philip Food Inc., EEUU), SUPERSTART (disponible de Alltech), GERT STRAND (disponible de Gert Strand AB, Suecia) and FERMIOL (disponible de DSM Specialties).

[0191] En una forma de realización preferida, la levadura es un *Saccharomyces spp.* En una forma de realización más preferida, la levadura es *Saccharomyces cerevisiae*. En otra forma de realización más preferida, la levadura es *Saccharomyces distaticus*. En otra forma de realización más preferida, la levadura es *Saccharomyces uvarum*. En otra forma de realización preferida, la levadura es un *Kluiveromyces*. En otra forma de realización más preferida, la levadura es *Kluiveromyces marxianus*. En otra forma de realización más preferida, la levadura es *Kluyveromyces fragilis*. En otra forma de realización preferida, la levadura es un *Candida*. En otra forma de realización más preferida, la levadura es *Candida brassicae*. En otra forma de realización preferida, la levadura es un *Clavispora*. En otra forma de realización más preferida, la levadura es *Clavispora lusitaniae*. En otra forma de realización más preferida, la levadura es *Clavispora opuntiae*. En otra forma de realización preferida, la levadura es un *Pachysolen*. En otra forma de realización más preferida, la levadura es *Pachysolen tannofilus*. En otra forma de realización preferida, la levadura es un *Bretannomyces*. En otra forma de realización más preferida, la levadura es *Bretannomyces clausenii* (Philippidis, G. P., 1996, Cellulose bioconversion technology, in Handbook on Bioethanol: Production and Utilization, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212).

[0192] Las bacterias que pueden fermentar eficazmente glucosa a etanol incluyen, p.ej. *Zimomonas mobilis* y *Clostridium thermocellum* (Philippidis, 1996, supra).

[0193] Es bien conocido en la técnica, que los organismos anteriormente descritos pueden también ser usados para producir otras sustancias como se describe en este caso.

[0194] La clonación de genes heterólogos en *Saccharomyces cerevisiae* Chen, Z., Ho, N. W. Y., 1993, Cloning and improving the expression of *Pichia stipitis* xylose reductase gene in *Saccharomyces cerevisiae*, Appl. Biochem. Biotechnol. 39-40: 135-147 ; Ho, N. W. Y., Chen, Z, Brainard, A. P., 1998, Genetically engineered *Saccharomyces* yeast capable of effectively cofermenting glucose and xylose, Appl. Environ. Microbiol. 64: 1852-1859), o en bacterias tales como *Escherichia coli* (Beall, D. S., Ohta, K., Ingram, L. O., 1991, Parametric studies of ethanol production from xylose and other sugars by recombinant *Escherichia coli*, Biotech. Bioeng. 38: 296-303), *Klebsiella oxytoca* (Ingram, L. O., Gomes, P. F., Lai, X., Moniruzzaman, M., Wood, B. E., Yomano, L. P., York, S. W., 1998, Metabolic engineering of bacteria for ethanol production, Biotechnol. Bioeng. 58: 204-214), y *Zimomonas mobilis* (Zhang, M., Eddy, C., Deanda, K., Finkelstein, M., and Picataggio, S., 1995, Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*, Science 267: 240-243 ; Deanda, K., Zhang, M., Eddy, C., and Picataggio, S., 1996, Development of an arabinose- fermenting *Zymomonas mobilis* strain by metabolic pathway engineering, Appl. Environ. Microbiol. 62: 4465-4470) ha llevado a la construcción de organismos capaces de convertir hexosas y pentosas a etanol (cofermentación).

[0195] Típicamente la levadura u otro microorganismo se añade a la celulosa degradada o hidrolizado y la fermentación se lleva a cabo durante aproximadamente entre 24 a 96 horas, como aproximadamente entre 35 a 60 horas. Típicamente la temperatura está aproximadamente entre 26°C y 40°C, en particular aproximadamente a 32°C, y aproximadamente entre pH 3 a pH 6, en particular alrededor de pH de 4-5.

[0196] En una forma de realización preferida, la levadura u otro microorganismo se aplica a la celulosa degradada o hidrolizado y la fermentación se lleva a cabo durante aproximadamente entre 24 y 96 horas, típicamente entre 35 y 60

- horas. En unas formas de realización preferidas, la temperatura está generalmente aproximadamente entre 26 y 40°C, en particular aproximadamente a 32°C, y el pH está generalmente aproximadamente entre pH 3 y pH 6, preferiblemente alrededor de pH 4-5. La levadura u otro microorganismo se aplica preferiblemente en cantidades de aproximadamente de 10^5 y 10^{12} , preferiblemente de aproximadamente entre 10^7 y 10^{10} , especialmente aproximadamente una cuenta viable de 5×10^7 a por ml de caldo de fermentación. Durante una fase de producción de etanol la cuenta de células de levadura debería preferiblemente estar en el intervalo de aproximadamente de 10^7 a 10^{10} , especialmente alrededor de aproximadamente 2×10^8 . Se pueden encontrar más directrices de guía para usar levadura para fermentación puede ser encontrada en, p.ej. "The Alcohol Textbook" (Editors K. Jacques, T.P. Lyons and D.R. Kelsall, Nottingham University Press, United Kingdom 1999), que se incluye en la presente por referencia.
- 5 [0197] El proceso más usado en la técnica es la sacarificación y fermentación simultánea (SSF) proceso dónde no hay fase de retención para la sacarificación, lo que significa que la levadura y la enzima se agregan juntas.
- [0198] Para la producción de etanol, tras la fermentación el triturado se destila para extraer el etanol. El etanol obtenido según el proceso de la invención puede ser utilizado como, p.ej. combustible de etanol; etanol bebible, es decir, licores bebibles neutrales, o etanol industrial.
- 10 [0199] Se puede usar un estimulador de fermentación en combinación con cualquiera de los procesos enzimáticos descritos en la presente para mejorar aún más el proceso de fermentación, y en particular, el rendimiento del microorganismo fermentativo, tal como, realce de índice y rendimiento de etanol. Un "estimulador fermentativo" se refiere a estimuladores para crecimiento de los microorganismos fermentativos, en particular, levadura. Los estimuladores de fermentación preferidos para 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. Ver, p.ej. Alfenore et al., Improving ethanol production and viability of *Saccharomyces cerevisiae* by a vitamin feeding strategy during fed-batch process, Springer- Verlag (2002), que se incluye en la presente por referencia. Ejemplos de minerales incluye minerales y sales minerales que pueden suministrar nutrientes que comprenden P, K, Mg, S, Ca, Fe, Zn, Mn, y Cu.
- 15 [0200] Recuperación. El alcohol es separado del material celulósico fermentado y purificado mediante métodos convencionales de destilación. Se puede obtener etanol con una pureza de hasta 96 vol.% etanol, que se puede usar como, p.ej. combustible de etanol, etanol bebible, es decir, licores bebibles neutrales, o etanol industrial.
- [0201] Para otras sustancias, se puede usar cualquier método conocido en la técnica incluyendo pero no limitado a, cromatografía (p. ej. intercambio iónico, afinidad, cromatografía hidrofóbica, y exclusión de tamaño), procedimientos electroforéticos (p. ej. isoelectroenfoque preparativo), solubilidad diferencial (p. ej. precipitación de sulfato amónico), SDS-PAGE, destilación, o extracción.
- 20 [0202] En los métodos de la presente invención, el polipéptido con actividad endoglucanasa y otra/s proteína/s celulolíticas se pueden suplementar con una o más actividades enzimáticas adicionales para mejorar la degradación del material celulósico. Enzimas preferidas adicionales son hemicelulasas, esterases (p. ej. lipasas, fosfolipasas, y/o cutinasas), proteasas, lacasas, peroxidases, o mezclas derivadas.
- 25 [0203] En los métodos de la presente invención, la/s enzima/s adicional/es puede/n ser añadida/s antes de o durante la fermentación, incluso durante o después de la propagación de el/los microorganismo/s fermentativo/s.
- [0204] La presente invención es descrita adicionalmente por los siguientes ejemplos.

Ejemplos

40 Materiales

[0205] Los productos químicos usados como tampones y sustratos fueron productos comerciales de al menos calidad reactiva.

Cepas

- 45 [0206] Como fuentes de los genes de endoglucanasa fueron usadas *Myceliophthora thermophila* CBS 117.65, *Basidiomycete* CBS 494.95, *basidiomycete* CBS 495.95, y *Penicillium brasilianum* cepa IBT 20888 (IBT Culture Collection of Fungi, Technical University of Denmark, Copenhagen, Denmark). Para la selección de librerías de expresión de *Myceliophthora thermophila* CBS 117.65 para actividad endoglucanasa se usaron cepas de *Saccharomyces cerevisiae* W3124 (MATa; ura 3-52 leu 2-3, 112 his 3-D200 pep 4-1137 prcl::HIS3 prb1::LEU2 cir⁺). Para la expresión del genes Cel5A se usó *Aspergillus oryzae* HowB104 (alfa-amilasa-negativo)

50 Medios y soluciones

[0207] El medio LB fue compuesto de 10 g de triptona, 5 g de extracto de levadura, y 5 g de cloruro sódico por litro.

[0208] El medio de ampicilina LB fue compuesto de 10 g de triptona, 5 g de extracto de levadura, 5 g de cloruro sódico

por litro, y 50 µg de ampicilina por ml (añadido tras esterilizar por autoclave).

[0209] Las placas de ampicilina LB fueron compuestas de medio de ampicilina LB y 15 g de bacto agar por litro.

[0210] El medio YPD fue compuesto por 1% extracto de levadura, 2% peptona, y 2% glucosa añadida tras esterilizar por autoclave.

5 [0211] El medio YPM fue compuesto por 1% extracto de levadura, 2% peptona, y 2% maltodextrina añadida tras esterilizar por autoclave.

[0212] El medio SC-URA con galactosa fue compuesto de 100 ml de 10X sales Basales, 28 ml de 20% ácidos de casamino sin vitaminas, 10 ml de 1% triptófano, 3.6 ml de 5% treonina (añadida tras esterilizar por autoclave) por litro, y 100 ml de 20% galactosa (añadida tras esterilizar por autoclave).

10 [0213] El medio SC-URA con glucosa fue compuesto de 100 ml de 10X solución de sales Basales, 28 ml de 20% ácidos de casamino sin vitaminas, 10 ml de 1% triptófano, 3.6 ml de 5% treonina (añadida tras esterilizado por autoclave) por litro, y 100 ml de 20% glucosa (añadida tras esterilizado por autoclave).

[0214] Una solución 10X de sales Basales fue compuesta de 75 g de base nitrogenada de levadura, 113 g de ácido succínico, y 68 g de NaOH por litro.

15 [0215] El SC-agar fue compuesto de medio SC-URA (con glucosa o galactosa como indicado) y 20 g de agar por litro.

[0216] Placas de SC-agar con AZCL HE celulosa 0.1% con galactosa fueron compuestos de SC-URA medio con galactosa, 20 g de agar por litro, y AZCL HE celulosa 0.1% (Megazyme, Wicklow, Irlanda).

[0217] El medio bA fue compuesta de 10 g de materia seca de licor de maíz fermentado, 10 g de NH₄NO₃, 10 g de KH₂PO₄; 0,75 g de MgSO₄·7H₂O; 0,1 ml de plurónico, y 0,5 g de CaCO₃. El pH fue ajustado a 6.5 antes del autoclave.

20 [0218] Placas de sales COVE fueron compuestas de 342,3 g de sacarosa, 25 g de agar Noble, 20 ml de solución de sales COVE, 10 mM acetamida, y 20 mM CsCl por litro. La solución fue ajustada a pH 7.0 antes del autoclave.

[0219] La solución de sales COVE fue compuestas de 26 g de KCl, 26 g de MgSO₄·7H₂O, 76 g de KH₂PO₄, y 50 ml de trazas de metálicas COVE por litro.

25 [0220] La solución de trazas metálicas COVE fue compuesto de 0,04 g de NaB₄O₇·10H₂O, 0,4 g de CuSO₄·5H₂O; 1,2 g de FeSO₄·7H₂O; 0,7 g de MnSO₄·H₂O; 0,8 g de Na₂MoO₄·2H₂O, y 10 g de ZnSO₄·7H₂O por litro.

[0221] El TE fue compuesto por 10 mM de Tris, pH 7.4 y 0,1 mM de EDTA.

Ejemplo 1: La construcción la expresión de librerías de cDNA de *Myceliophthora thermophila* CBS 117.65 en *Saccharomyces cerevisiae*.

30 [0222] *Myceliophthora thermophila* CBS 117.65 fue cultivado en 200 ml de medio de BA a 30°C durante cinco días a 200 r.p.m. Los micelios del cultivo de la matriz agitación fueron cosechados mediante el filtrado del contenido a través de un embudo forrado con Miracloth™ CalBiochem, San Diego, CA, EEUU). Los micelios fueron luego emparedados entre dos piezas de Miracloth™ y secados con toallas de papel absorbente. La masa micelial fue entonces transferida a tubos centrifugadores de plástico Falcon 1059 y congelados en nitrógeno líquido. Los micelios congelados fueron almacenados en un congelador a -80°C hasta su uso.

35 [0223] La extracción de ARN total fue realizada con tiocianato de guanidina seguido de un ultracentrifugado a través de una almohadilla de 5,7 M de CsCl, y se efectuó un aislamiento de poli(A)+RNA por cromatografía de afinidad de oligo(dT)-celulosa, usando los procedimientos descritos en WO 94/14953.

40 [0224] El ADNc bicatenario fue sintetizado a partir de 5 µg de poli(A)+ ARN por el método ribonucleasa H (Gubler and Hoffman, 1983, Gene 25: 263-269, Sambrook et al., 1989, Molecular cloning: A laboratory manual, Cold Spring Harbor lab., Cold Spring Harbor, NY, EEUU). El poli(A)⁺ARN (5 µg en 5 µl de DEPC (dietilpirocarbonato 0.1%)-de agua tratada) fue calentado a 70°C durante 8 minutos en un tubo de Eppendorf pre-siliconizado sin ribonucleasa, apagado en hielo, y combinado en un volumen final de 50 µl con tampón de transcriptasa inversa compuesto por 50 mM de Tris-HCl, pH 8.3, 75 mM de KCl, 3 mM de MgCl₂, 10 mM de ditiotreitól ((DTT) (Bethesda Research Laboratories, Bethesda, MD, EEUU), 1 mM de dATP, dGTP y dTTP, y 0,5 mM de 5-metil-dCTP (Pharmacia, Uppsala, Suecia), 40 unidades de inhibidor ribonucleasa de placenta humana RNasin, Promega, Madison, WI, EEUU), 1.45 µg de oligo(dT)₁₈ –Not I (Pharmacia) como cebador, y 1000 unidades de transcriptasa inversa RNase H Superscript II (Bethesda Research Laboratories). La primera cadena ADNc fue sintetizada incubando la mezcla reactiva a 45°C durante 1 hora. Después de la síntesis, la mezcla híbrida de mRNA:cDNA fue gel filtrada a través de un MicroSpin con una de columna de rotación s-400 HR (Pharmacia) según las instrucciones del fabricante.

50 [0225] Después filtración por gel, los híbridos fueron diluidos en 250 µl del tampón de la segunda cadena(20 mM Tris-HCl, pH 7.4, 90 mM KCl, 4,6 mM de MgCl₂, 10 mM de (NH₄)₂SO₄, 0,16 mM de NAD) que contienen 200 µM de cada

dNTP, 60 unidades de polimerasa I de ADN de *E. coli* (Pharmacia), 5,25 unidades de RNase H (Promega), y 15 unidades de ligasa de ADN de *E. coli* (Boehringer Mannheim, Mannheim, Alemania). El síntesis de la segunda cadena de ADN fue realizada incubando el tubo de reacción a 16°C durante 2 horas y 15 minutos adicionales 25°C. La reacción fue detenida mediante adición de EDTA a una concentración final de 20 mM seguida de extracciones de fenol y de cloroformo.

[0226] El ADNc bicatenario fue precipitado a -20°C durante 12 horas mediante adición de 2 volúmenes de etanol 96% y 0,2 volumen de 10 M de acetato amónico, recuperado mediante centrifugado a 13,000 x g, lavado en etanol 70%, secado, y resuspendido en 30 µl de tampón de nucleasa de judía Mung (30 mM acetato sódico pH 4,6, 300 mM NaCl, 1 mM ZnSO₄, 0,35 mM DTT, 2% glicerol) que contiene 25 unidades de nucleasa de judía Mung (Pharmacia). La horquilla de ADN monocatenaria fue recortado incubando la reacción a 30°C durante 30 minutos, seguido de la adición de 70 µl de 10 mM de Tris-HCl-1 mM de EDTA pH 7.5, extracción de fenol, y la precipitación con 2 volúmenes de etanol 96% y 0,1 volumen de 3 M acetato sódico pH 5.2 en hielo durante 30 minutos.

[0227] Los ADNc bicatenarios fueron recuperados mediante centrifugado a 13,000 x g y los extremos redondeados en 30 µl de Tampón de polimerasa de aDN T4 (20 mM Tris-acetato, pH 7,9, 10 mM acetato magnésico, 50 mM acetato de potasio, 1 mM DTT) que contiene 0,5 mM de cada dNTP y 5 unidades de polimerasa de ADN T4 (New England Biolabs, Ipswich, MA, EEUU) incubando la mezcla reactiva a 16°C durante 1 hora. La reacción fue detenida mediante adición de EDTA a una concentración final de 20 mM, seguido de extracciones de fenol y cloroformo, y precipitación durante 12 horas a -20°C añadiendo 2 volúmenes de etanol 96% y 0,1 volumen de 3 M acetato sódico pH 5.2.

[0228] Después de la reacción de relleno los ADNc fueron recuperados mediante centrifugado a 13,000 x g, lavado en etanol 70%, y secado. El ADNc granulado fue resuspendido en 25 µl de tampón de unión (30 mM Tris-HCl, pH 7.8, 10 mM de MgCl₂, 10 mM de DTT, 0,5 mM de ATP) conteniendo 2,5 µg de adaptadores Bst XI no-palindrómicos (Invitrogen, Carlsbad, CA, USA), mostrado abajo, y 30 unidades de ligasa T4 (Promega), y luego incubado a 16°C durante 12 horas. La reacción fue detenida calentando a 65°C durante 20 minutos y luego enfriada en hielo durante 5 minutos.

5'-CTTTCCAGCACA-3' (SEQ ID NO: 1)

3'-GAAAGGTC-5' (SEQ ID NO: 2)

[0229] El ADNc adaptado fue digerido con Not I, seguido de una incubación durante 2.5 horas a 37°C. La reacción fue detenida calentando a 65°C durante 10 minutos. Los ADNc fueron separados por tamaño mediante electroforesis de gel sobre un gel de agarosa de temperatura de fundido bajo Sea Plaque 0.8 GTG (Cambrex Corporation, East Rutherford, NJ, EEUU) en 44 mM de Tris Base, 44 mM de ácido bórico, tampón (TBE) 0,5 mM de EDTA para separar adaptadores no ligado y ADNc pequeños. El ADNc fue seleccionado por tamaño con un cortado a 0,7 kb y salvado del gel usando β-agarasa (New England Biolabs, Ipswich, MA, USA) según las instrucciones del fabricante y precipitado durante 12 horas a -20°C añadiendo dos volúmenes de etanol 96% y 0,1 volumen de 3 M de acetato sódico pH 5.2.

[0230] El ADNc dirigido, seleccionado por tamaño fue recuperado mediante centrifugado a 13.000 x g, lavado en etanol 70%, secado, y luego resuspendido en 30 µl de 10 mM de Tris-HCl-1 mM de EDTA pH 7.5. Los ADNc fueron desalados por filtración de gel a través de una columna de rotación MicroSpin S-300 HR según las instrucciones del fabricante. Tres ligaduras de prueba fueron llevadas a cabo en 10 µl de tampón de unión (30 mM Tris-HCl, pH 7.8, 10 mM de MgCl₂, 10 mM DTT, 0,5 mM de ATP) conteniendo 5 µl de ADNc de doble cadena (tubos de reacción #1 y #2), 15 unidades de ligasa T4 (Promega), y 30 ng (tubo #1), 40 ng (tubo #2), y 40 ng (tubo #3, el vector antecedente) vector piES2.0 de Bst XI-Not I escindido (Invitrogen, Carlsbad, CA, USA). Las reacciones de ligamiento fueron realizadas mediante incubación a 16°C durante 12 horas, luego calentamiento a 70°C durante 20 minutos, y finalmente añadiendo 10 µl de agua a cada tubo. Un µl de cada mezcla de ligadura fue sometida a electroporación en 40 µl de células electrocompetentes de *E. coli* DH10B (Bethesda Research Laboratories) como describe Sambrook et al., 1989, supra.

[0231] La genoteca de ADNc de *Myceliophthora thermophila* CBS 117.65 fue establecido en *E. coli* DH10B consistente en depósitos. Cada depósito fue hecha extendiendo *E. coli* transformada sobre placas de ampicilina LB, produciendo 15,000-30,000 colonias/placa tras la incubación a 37°C durante 24 horas. Veinte ml de medio de ampicilina LB se añadió a la placa y las células fueron suspendidas en ésta. La suspensión celular fue agitada a 100 r.p.m. en un 50 ml tubo durante 1 hora a 37°C.

[0232] La genoteca resultante de ADNc de *Myceliophthora thermophila* CBS 117.65 genoteca de ADNc consistió de aproximadamente 10⁶ clones individuales, con un vector antecedente de 1%. El ADN plásmido de algunos de los depósitos de librería fue aislado usando un Plasmid Midi Kit(QIAGEN Inc., Valencia, CA, USA), según las instrucciones del fabricante, y almacenado a -20°C.

Ejemplo 2: Rastreado de genotecas de expresión de *Myceliophthora thermophila* CBS 117.65 para detectar actividad endoglucanasa

[0233] Un ml de alícuotas de plásmido de ADN purificado (100 ng/ml) de algunos de los depósitos de librería (ejemplo 1) fueron transformados en *Saccharomyces cerevisiae* W3124 por electroporación (Becker and Guarante, 1991, Methods

Enzymol. 194: 182-187 y los transformantes fueron colocados sobre placas de agar-SC conteniendo 2% glucosa e incubado a 30°C. En total, 50- 100 placas conteniendo 250-400 colonias de levadura fueron obtenidas de cada depósito.

[0234] Después 3-5 días de incubación, las placas de agar SC fueron colocados mediante réplica de placas sobre un conjunto de placas de SC URA-agar con AZCL HE celulosa 0,1% con galactosa. Las placas fueron incubadas durante 2-4 días a 30°C y colonias positivas de endoglucanasa fueron identificadas como colonias rodeadas por un halo azul.

Ejemplo 3: Caracterización del gen *Cel5A* de *Myceliophthora thermophila* CBS 117.65.

[0235] Colonias de levadura de expresión de endoglucanasa fueron inoculadas en 20 ml de medio YPD en tubos de ensayo de vidrio 50 ml. Los tubos fueron agitados a 200 r.p.m. durante 2 días a 30°C. Las células fueron cosechadas mediante centrifugado durante 10 minutos a 3000 r.p.m. en un centrifugador Heraeus Megafuge 1.0R con un rotor 75002252 (Hanau, Alemania).

[0236] El ADN fue aislado según WO 94/14953 y disuelto en 50 µl de agua desionizada. El ADN fue transformado en células de *E. coli* DH10B mediante procedimientos estándar según Sambrook et al., 1989, supra. Un *E. coli* transformante posteriormente mostró contener el gen *Cel5A* de *Myceliophthora thermophila* CBS 117.65 fue designado pCIC161 (Figura 1) y usado como el material para depósito de material biológico. La cepa de *E. coli* pCIC161 fue depositado como *E. coli* NRRL B-30902 el 23 de febrero de 2006.

[0237] El ADN plásmido fue aislado del *E. coli* transformante de usando procedimientos estándar según Sambrook et al., 1989, supra. La secuencia total de ADNc de gen del *Cel5A* de *Myceliophthora thermophila* CBS 117.65 fue marcado con un Taq DyeDeoxy Terminator Cycle Sequencing Kit (Perkin Elmer, Wellesley, MA, EEUU) y cebadores oligonucleótidos sintéticos que usan un Secuenciador de ADN Applied Biosystems ABI PRISM™ 377 (ABI, Foster City, CA, EEUU) según las instrucciones del fabricante.

[0238] La secuencia de nucleótidos (SEC ID n°: 3) y la secuencia deducida de aminoácidos (SEC ID n°: 4) del gen *Cel5A* de *Myceliophthora thermophila* se muestran en la figura 2. La secuencia codificante es 1170 pb que incluyendo el codón de terminación. La proteína codificada predicha contiene 389 aminoácidos. El %G+C de la región de codificación del gen es 63.6% y la región de codificación del polipéptido maduro es 63.6%. Usando el programa SignalP, versión 3 (Nielsen et al., 1997, Protein Engineering 10: 1-6), un péptido señal de 16 residuos fue predicho. La proteína madura predicha contiene 373 aminoácidos con una masa molecular de 40,9 kDa.

[0239] El análisis de la secuencia de aminoácidos deducida del gen *cel5a* con el programa Interproscan (Zdobnov y Apweiler, 2001 Bioinformatics 17: 847-848) mostró que la proteína *Cel5A* contenía la secuencia núcleo típica de una hidrolasa de glicosil de familia 5, con una extensión de aproximadamente desde el residuo de aminoácido 77 hasta el residuo 350 del polipéptido maduro predicho. La proteína *Cel5A* también contenía la firma de una secuencia de dominio de un enlace de celulosa fúngica (CBMI) de tipo I. Esta secuencia de firma conocida como modelo Prosite PS00562 (Sigrist et al., 2002, Brief Bioinform. 3: 265-274) estuvo presente desde el residuo 8 de aminoácido hasta el residuo 35 del polipéptido maduro predicho.

[0240] Una alineación global comparativa de parejas de secuencias de aminoácidos fue determinada usando el algoritmo Needleman-Wunsch (Needleman and Wunsch, 1970, J. Mol. Biol. 48: 443-453) tal y como implementado en el programa Needle de EMBOSS con una penalización de hueco abierto de 10, y una penalización de extensión de hueco de 0.5, y el matriz EBLOSUM62. La alineación mostró que la secuencia deducida de aminoácidos del gen de *Myceliophthora thermophila* que codifica el polipéptido maduro *Cel5A* compartía 75% y 72% de identidad (excluyendo huecos) con las secuencias deducidas de aminoácidos de dos proteínas hidrolasas de glicosil de familia 5 de *Neurospora crassa* y *Humicola insolens*, respectivamente (números de registro Q7SDR1 y Q12624, respectivamente).

Ejemplo 4: La expresión del gen *cel5a* de *Myceliophthora thermophila* CBS 117.65 en *Aspergillus oryzae*

[0241] El gen *cel5a* de *Myceliophthora thermophila* CBS 117.65 fue cortado del vector pYES2.0 usando Hind III y Xba I, y ligado en el vector de expresión pHD414 de *Aspergillus* (EP 238 023, WO 93/11249) usando métodos estándar (Sambrook et al., 1989, supra). El vector de expresión fD414 de *Aspergillus* es un derivado de p775 (EP 238 023). El plásmido resultante fue designado pA2C161 (Figura 3).

[0242] Los protoplastos de *Aspergillus oryzae* HowB104 fueron preparados como se describe en WO 95/02043. Cien microlitros de suspensión de protoplasto fueron mezclados con 5-25 µg del vector de expresión pA2C161 de *Aspergillus* en 10 µl de STC compuesto por 1,2 M de sorbitol, 10 mM de Tris-HCl, pH 7.5, 10 mM de CaCl₂ y además mezclados con 5-25 µl de p3SR2, un gen *amdS* de *Aspergillus nidulans* que lleva plásmido (Christensen et al., 1988, Bio/Technology 6: 1419-1422). Se dejó la mezcla a temperatura ambiente durante 25 minutos. Doscientos microlitros de PEG 4000 60% (BDH, Poole, Inglaterra) (glicol de polietileno, peso molecular 4.000), 10 mM de CaCl₂, y 10 mM Tris-HCl pH 7,5 fue añadido y suavemente mezclado y finalmente 0,85 ml de la misma solución fue añadida y suavemente mezclada. Se dejó la mezcla a temperatura ambiente durante 25 minutos, fue centrifugado a 2.500 x g durante 15 minutos, y el granulado fue resuspendido en 2 ml de 1,2 M de sorbitol. Este proceso de sedimentación fue repetida, y los protoplastos fueron extendidos sobre placas de COVE. Tras la incubación durante 4-7 días a 37°C las esporas fueron

escogidas y extendidas para aislar colonias individuales. Este procedimiento fue repetido y las esporas de una única colonia después del segundo reaislamiento fueron almacenadas.

[0243] Cada uno de los transformantes fue inoculado en 10 ml de medio YPM. Después de 2-5 días de incubación a 30°C, 200 r.p.m., fue removido el sobrenadante. Se identificó actividad endoglucanasa aplicando 20 µl de caldo de cultivo a agujeros de 4 mm de diámetro fijados en una placa de SC-agar con AZCL HE celulosa 0.1% e incubado durante toda la noche a 30°C. Se identificó actividad endoglucanasa por un halo azul alrededor de una colonia. Diferentes caldos transformantes tenían actividad endoglucanasa significativamente mayor que un caldo de un control de antecedentes de *Aspergillus oryzae* no-transformado, que demostró una eficaz expresión de la endoglucanasa Cel5A de *Myceliophthora thermophila* CBS 117.65 en *Aspergillus oryzae*.

10 **Ejemplo 5: Construcción de una genoteca de expresión de ADNc de *basidiomycete* CBS 495.95 en *Saccharomyces cerevisiae***

[0244] Una genoteca de ADNc de *basidiomycete* CBS 495.95, que consiste en aproximadamente 10⁶ clones individuales fue construido en *E. coli* como se describe en el ejemplo 1, con un vector de antecedentes de 1 %.

15 **Ejemplo 6: Selección de genotecas de expresión de ADNc de *basidiomycete* CBS 495.95 para actividad endoglucanasa**

[0245] La selección de la genoteca de ADNc (ejemplo 5) fue realizada como se describe en el ejemplo 2.

Ejemplo 7: Caracterización de un gen codificante Cel5A de *basidiomycete* CBS 495.95

[0246] La clonación del gen Cel5A de *basidiomycete* CBS 495.95 se efectuó como se describe en el ejemplo 3. Posteriormente un transformante *E. coli* que mostró contener un gen Cel5A fue designado pCIC453 (Figura 4) y usado como el material para depósito de material biológico. La cepa de *E. coli* pCIC453 fue depositado como *E. coli* NRRL B-30903 el 23 de febrero del 2006.

[0247] El gen Cel5A del *basidiomycete* CBS 495.95 fue cortado de pCIC453 usando Hind III y Xba I y ligado en el vector de expresión *Aspergillus* pH414. El plásmido resultante fue designado pA2C453 (Figura 5).

[0248] La secuencia de nucleótidos (SEC ID n°: 5) y la secuencia deducida de aminoácidos (SEC ID n°: 6) del gen Cel5A CBS 495.95 se muestran en la figura 6. La secuencia codificante es 1194 par de bases incluyendo el codón de terminación. El %G+C de la región de codificación del gen es 59,8% y la región de codificación del polipéptido maduro es 60,1 %. La proteína codificada predicha contiene 397 aminoácidos. Usando el programa SignalP, versión 3 (Nielsen et al., 1997, supra), un péptido señal de 15 residuos fue predicho. La proteína madura predicha contiene 382 aminoácidos con una masa molecular de 40,1 kDa.

[0249] El análisis de la secuencia deducida de aminoácidos del gen Cel5A con el programa Interproscan (Zdobnov y Apweiler, 2001, supra) mostró que la proteína Cel5A contenía la secuencia núcleo típica de una hidrolasa de glicosil de familia 5, con extensión aproximadamente desde residuos 81 hasta 359 del polipéptido maduro predicho. La proteína Cel5A también contenía la firma de secuencia de un enlace de dominio de celulosa fúngica (CBMI) de tipo I. Esta firma de secuencia es conocida como modelo Prosite PS00562 (Sigrist et al., 2002, supra) estaba presente desde residuo de aminoácido 8 hasta residuo 36 del polipéptido maduro predicho.

[0250] Una alineación global comparativa de pareja de secuencias de aminoácidos fue determinada usando el algoritmo Needleman-Wunsch (Needleman y Wunsch, 1970, supra) como implementado en el programa Needle con una penalización de hueco abierto 10, una penalización de extensión de hueco de 0.5, y el matriz EBLOSUM62. La alineación mostró que la secuencia deducida de aminoácidos del gen CBS 495.95 que codifica el polipéptido maduro Cel5A compartía 82% y 79% identidad (excluyendo huecos) con las secuencias de aminoácidos deducidas de dos proteínas de hidrolasa de glicosil de familia 5 de *Irpelex lacteus* y *Trametes hirsuta*, respectivamente (números de registro Q5W7K4 y Q75UV6, respectivamente).

Ejemplo 8: Expresión del gen Cel5A de *basidiomycete* CBS 495.95 en *Aspergillus oryzae*

[0251] La expresión del gen Cel5A del *basidiomycete* CBS 495.95, y el análisis de actividad endoglucanasa fue llevado a cabo como se describe en el ejemplo 4.

Ejemplo 9: Construcción de una genoteca de expresión de ADNc de *basidiomycete* CBS 494.95 en *Saccharomyces cerevisiae*

[0252] Una genoteca de ADNc de *basidiomycete* CBS 494.95, que consiste en aproximadamente 10⁶ clones individuales fue construido en *E. coli* como se describe en el ejemplo 1, con un vector de antecedentes de 1%.

50 **Ejemplo 10: Selección de bibliotecas de expresión de ADNc de *basidiomycete* CBS 494.95 para actividad endoglucanasa**

[0253] La selección de la genoteca de ADNc (ejemplo 9) fue realizada como se describe en el ejemplo 2.

Ejemplo 11: Caracterización de un gen codificante CEL5B de *basidiomycete* CBS 494.95

[0254] La clonación del gen CEL5B de *basidiomycete* CBS 494.95 se efectuó como se describe en el ejemplo 3. Posteriormente un *E. coli* transformante que mostró contener el gen CEL5B CBS 494.95 fue designado pCIC486 (Figura 7) y usado como el material para depósito de material biológico. La cepa de *E. coli* pCIC486 fue depositado como *E. coli* NRRL B-30904 el 23 de febrero del 2006.

[0255] El gen CEL5B del *basidiomycete* CBS 494.95 fue cortado del pYES2.0 vector usando Kpn I y Xho I y ligado en el vector de expresión *Aspergillus* pHD423 Lassen et al., 2001, Appl Environ Microbiol 67: 4701-4707), un fD414 derivativo con un sitio Kpn I en el poliligador. El plásmido resultante fue designado pA2C486 (Figura 8). La secuencia de nucleótidos (SEC ID n°: 7) y secuencia deducida de aminoácidos (SEC ID n°: 8) gen CEL5B del CBS 494.95 se muestran en la figura 9. La secuencia codificante es 1290 pb incluyendo el codón de terminación. El %G+C de la región de codificación del gen es 56.0% G+C y la región de codificación de polipéptido maduro es 56,1%. La proteína codificada predicha contiene 429 aminoácidos. Usando el programa SignalP, versión 3 (Nielsen et al., 1997, supra), un péptido señal de 21 residuos fue predicho. La proteína madura predicha contiene 408 aminoácidos con una masa molecular de 43,1 kDa.

[0256] El análisis de la secuencia deducida de aminoácidos del gen CEL5B con el programa Interproscan (Zdobnov y Apweiler, 2001, supra) mostró que la proteína CEL5B contenía una secuencia de núcleo típica de una hidrolasa de glicosil de familia 5, con una extensión aproximadamente desde residuos de aminoácido 106 hasta residuos 385 del polipéptido maduro predicho. La proteína Cel5A también contenía la firma de secuencia de un dominio de enlace de celulosa fúngica (CBMI) tipo I. Esta firma de secuencia es conocida como modelo Prosite PS00562 (Sigrist et al., 2002, supra) estuvo presente desde el residuo de aminoácido 7 hasta el residuo 34 del polipéptido maduro predicho.

[0257] Una alineación global comparativa de parejas de secuencias de aminoácidos fue determinada usando el algoritmo Needleman-Wunsch (Needleman y Wunsch, 1970, supra) como está implementado en el programa Needle con una penalización de hueco abierto de 10, una penalización de extensión de espacio de 0,5 y el matriz EBLOSUM62. La alineación mostró que la secuencia deducida de aminoácidos del gen CBS 495.95 que codifica el polipéptido maduro Cel5A compartía 69% y 67% identidad (excluyendo huecos) con las secuencias deducidas de aminoácidos de dos proteínas de hidrolasa de glicosil de familia 5 de *Irpex lacteus* y *Trametes hirsuta*, respectivamente (números de registro Q5W7K4 y Q75UV6, respectivamente).

Ejemplo 12: Expresión de CEL5B de *basidiomycete* CBS 494.95 en *Aspergillus oryzae*

[0258] La expresión del gen CEL5B de *basidiomycete* CBS 494.95, y el análisis de actividad endoglucanasa fue realizado como se describe en el ejemplo 4.

Ejemplo 13: Purificación de endoglucanasas recombinantes de *Myceliophthora thermophila* CBS 117.65, *basidiomycete* CBS 494.95, y *basidiomycete* CBS 495.95

[0259] Las endoglucanasas de *Myceliophthora thermophila* CBS 117.65, *basidiomycete* CBS 494.95, y *basidiomycete* CBS 495.95, producidos recombinantemente en *Aspergillus oryzae* como se describe en ejemplos 4, 8, y 12, fueron purificados a homogeneidad usando un protocolo esencialmente tal y como se describe por Otzen et al., 1999, Protein Sci. 8: 1878-87.

[0260] La concentración proteínica en las preparaciones enzimáticas fue determinada usando el Bicinchoninic acid (BCA) Microplate Assay según las instrucciones del fabricante para un BCA Protein Assay Reagent Kit (Pierce Chemical Co., Rockford, IL, EEUU).

[0261] Las diluciones enzimáticas fueron preparadas frescas antes de cada experimento a partir de soluciones enzima almacenados, que fueron almacenados a -20°C.

Ejemplo 14: Aislamiento de ADN genómico de *Penicillium brasilianum* IBT 20888

[0262] Esporas de *Penicillium brasilianum* cepa IBT 20888 fueron propagadas sobre arroz según Carlsen, 1994, Ph.D. thesis, Department of Biotechnology, The Technical University of Denmark. Las esporas fueron recuperadas con 20 ml de Tween 20 al 0,1% e inoculados a una concentración de 1×10^6 esporas por ml en 100 ml de medio Mandels y Weben (Mandels y Weben, 1969, Adv. Chem. Ser. 95: 394-414) conteniendo 1% glucosa suplementada con 0,25 g de extracto de levadura y 0.75 g de bactopectona por litro en un 500 ml matraz de agitación con deflectores. Los micelios fúngicos fueron cosechados tras 24 horas de crecimiento aeróbico a 30°C, a 150 r.p.m.

[0263] Los micelios fueron recogidos por filtración a través de un filtro Nalgene DS0281-5000 (Nalge Nunc International Corporation, Rochester, NY, EEUU) hasta alcanzar sequedad y congelado en nitrógeno líquido. Los micelios congelados fueron molidos a polvo en un mortero frío de hielo seco y distribuidos a un tubo de ensayo con tapón de rosca. El polvo fue suspendido en un volumen total de 40 ml de 50 mM de tampón CAPS (3-(ciclohexilamino)-1-ácido propanesulfónico)-

NaOH pH 11 conteniendo 0,5% de sulfato de dodecilo de litio y 0,5 mM de EDTA. La suspensión fue colocada a 60°C durante 2 horas y periódicamente resuspendida por inversión. A la suspensión se le añadió un volumen igual de fenol:cloroformo (1:1 V/V) neutralizado con 0,1 M de Tris base, y el tubo fue mezclado en una rueda rotante a 37°C durante 2 horas. Después del centrifugado a 2500 r.p.m. durante 10 minutos en un rotor Sorvall H1000B, la fase acuosa (fase superior) fue reextraído de nuevo con fenol:cloroformo (1:1 V/V) y centrifugada a 15.000 x g durante 5 minutos. La fase acuosa de la segunda extracción fue portada a 2,5 M acetato amónico (materia prima 10 M) y colocada a -20°C hasta su congelación. Tras la descongelación, el extracto fue centrifugado a 15.000 x g durante 20 minutos en un rotor frío. La perla (principalmente ARNr) fue descartada y los ácidos nucleicos en el sobrenadante fueron precipitados mediante adición de 0.7 volúmenes de isopropanol. Después del centrifugado a 15.000 x g durante 15 minutos, el granulado fue enjuagado tres veces con 5 ml de etanol 70% (sin resuspensión), secado al aire casi al completo, y disuelto en 1,0 ml de 0,1X TE. La perla disuelta fue transferido a dos tubos de microfuga de 1,5 ml. Los ácidos nucleicos fueron precipitados mediante adición de acetato amónico (0,125 ml) a 2,0 M y etanol a 63% (1,07 ml) y centrifugado a velocidad máxima durante 10 minutos en una microcentrifugadora Sorvall MC 12V (Kendro Laboratory Products, Asheville, NC, EEUU). El granulado fue enjuagado dos veces con etanol 70%, secado al aire completamente, y disuelto en 500 µl de 0.1X TE.

Ejemplo 15: Preparación de una genoteca de ADN genómico de *Penicillium brasilianum* IBT 20888

[0264] Las genotecas genómicas fueron construidas usando un equipo de subclonado TOPO Shotgun Subcloning Kit (Invitrogen, Carlsbad, CA, EEUU). Brevemente, el ADN celular total fue cortado por nebulización bajo nitrógeno psi 10 durante 15 segundos y separado por tamaño sobre geles de agarosa 1% usando un tampón de 40 mM Tris base, 20 mM sodio acetato, 1 mM de (TAE) y disodio de EDTA. Los fragmentos de ADN migratorios en el intervalo de tamaño 3-6 kb fueron cortados y eluidos usando un equipo de extracción MiniElute™ Gel Extraction Kit (QIAGEN Inc, Valencia, CA, EEUU). Los fragmentos eluidos fueron divididos por tamaño nuevamente usando un gel de agarosa 1% como mencionado arriba y los fragmentos de ADN migratorios en el intervalo de tamaño 3-6 kb fueron cortados y eluido usando un MiniElute™ Gel Extraction Kit.

[0265] Los fragmentos de ADN eluidos fueron reparados en sus extremos romos y defosforilado y fosfatasa alcalina de gamba (Roche Applied Science, Mannheim, Alemania). Los fragmentos de ADN de extremo romo fueron clonados en el pCR4Blunt-TOPO vector (Invitrogen, Carlsbad, CA, USA) según las instrucciones del fabricante, transformados en células *E. coli* TOP10 electrocompetentes por electroporación, y colocados en placas sobre placas de ampicilina LB. La electroporación resultó en 15.300 clones.

Ejemplo 16: Purificación de endoglucanasa CEL5C nativo de *Penicillium brasilianum* IBT 20888

[0266] La endoglucanasa fue purificada y evaluada como se describe en Jørgensen et al., 2003 (Enzyme Microb. Technol. 32: 851-861). El sustrato fue azo-carboximetilo celulosa (Megazyme International Ireland Ltd., Bray, Irlanda) y el período de incubación fue de 15 minutos. La enzima purificada fue almacenada congelada a -20°C.

Ejemplo 17: Secuenciación N-terminal de endoglucanasa CEL5C de *Penicillium brasilianum* IBT 20888

[0267] La muestra purificada endoglucanasa CEL5C de *Penicillium brasilianum* (ejemplo 16) fue descongelada. Un alícuota de 100 µl de la muestra se añadió a 100 µl de Tampón de muestra de SDS-PAGE (4 ml de 0,5 M de TRIS-HCl, pH 6.8, 20 ml de 10% SDS, 20 ml de glicerol (87%), 56 ml de agua ultrapura Milli-Q®, y 15 granos de bromofenol azul) en un tubo de Eppendorf y calentado a 95 °C durante 4 minutos. Tras el calentamiento de cuatro alícuotas de 20 µl de la muestra diluida fueron aplicadas separadamente a un gel prefabricado 4-20% de poliacrilamida SDS (Invitrogen, Carlsbad, CA, EEUU). Además de las cuatro vías con la muestra, una mezcla de estándar de proteína Mark 12 (Invitrogen, Carlsbad, CA, EEUU) fue aplicada al gel.

[0268] El gel fue pasado por un aparato de gel Xcell SureLock™ (Invitrogen, Carlsbad, CA, EEUU) durante 90 minutos con ajustes de potencia inicial de 40 mA a 135 V máximo. Tras la electroforesis el gel fue incubado durante 5 minutos en una solución secante que consiste en 10 mM CAP pH 11 conteniendo metanol 6%. Una membrana ProBlott (Applied Biosystems, Foster City, CA, EEUU) fue humedecido durante 1 minuto en metanol puro antes de ser colocado en la solución secante durante 5 minutos para saturar la membrana con 10 mM CAP pH 11 conteniendo metanol 6%.

[0269] La electrotransferencia se efectuó en un aparato de secado Semi Dry Blotter II (KemEnTec, Copenhagen, Dinamarca) de la siguiente manera. Seis piezas de papel mojado Whatman n°. 1 humedecidos en la solución secante fueron colocadas en el electrodo positivo del aparato secante seguida de la membrana ProBlott, el gel de poliacrilamida, y seis piezas de papel mojado Whatman n°. 1 humedecidos en solución secante. El aparato secante fue ensamblado poniendo el electrodo negativo en contacto con la pila superior de papel Whatman n°. 1. Un peso de 11,3 kg fue colocado encima del aparato secante. La electrotransferencia fue realizada a una corriente de 175 mA durante 180 minutos.

[0270] Tras la electrotransferencia la membrana ProBlott fue manchada durante 1 minuto en Azul Brillante Coomassie R-250 al 0,1% (p/v) disuelto en metanol al 60%, ácido acético al 1%, H₂O al 39%. La decoloración de la membrana

ProBlott fue realizada en metanol 40% acuoso durante 5 minutos previamente la membrana fue enjuagada en agua desionizada. Finalmente la membrana ProBlott fue secado al aire.

5 [0271] Para la secuenciación N-terminal del aminoácido dos piezas de la membrana ProBlott que consisten en un banda 65 kDa fueron recortados y colocados en el cartucho secante de un secuenciador de proteínas Procise de Applied Biosystems (Applied Biosystems, Foster City, CA, EEUU). La secuenciación N-terminal se efectuó usando el método de correr filas para muestras de membrana PVDF (Pulsed liquid PVDF) según las instrucciones del fabricante.

[0272] La secuencia N-terminal de aminoácidos fue deducida de los cromatogramas resultantes comparando los tiempos de retención de los picos en los cromatogramas a los tiempos de retención de los aminoácidos PTH en el cromatograma estándar.

10 [0273] La secuencia N-terminal de aminoácidos de la endoglucanasa CEL5C de *Penicillium brasilianum* purificado fue determinada directamente usando un sistema secuenciador Procise 494 HT Sequencing System (Applied Biosystems, Foster City, CA, USA). La secuencia N-terminal se determinó como Ala-Ser-Ser-Phe-Val-Trp-Phe-Gly-Thr-Ser-Glu-Ser-Gly-Ala-Glu-Phe-Gly-Asn- Gln-Asn (aminoácidos 25 a 44 en SEC ID n°: 10).

Ejemplo 18: Amplificación PCR del gen de endoglucanasa cel5c de *Penicillium brasilianum* IBT 20888

15 [0274] Basado en la secuencia de aminoácidos N-terminal de la endoglucanasa purificada de *Penicillium brasilianum*, un cebador directo fue diseñado usando la estrategia CODEHOP (Rose et al., 1998, Nucleic Acids Res. 26: 1628-35). De la información de base de datos de otras endoglucanasas, dos cebadores inversos fueron diseñados como se muestra abajo usando la estrategia CODEHOP.

Cebador directo, Fwd:

20 5'-TTCGGTACCTCTGAGTCTGGNGCNGARTT-3' (SEQ ID NO: 11)

Cebador inverso, Rev1:

5'-TGATCCATATCGTGGTACTCGTTRTTNGTRTCRAA-3' (SEQ ID NO: 12)

Cebador inverso, Rev2:

5'-CCGTTGTAGCGACCGTARTTRTGNGGRTC-3' (SEQ ID NO: 13)

25 donde R=A a G, Y=C a T, K=G a T y N=A, C, G a T

[0275] Las reacciones de amplificación (30 µl) fueron preparadas usando aproximadamente 1 µg de ADN genómico de *Penicillium brasilianum* como molde. Además, cada reacción contenía los siguientes componentes: 30 pmol del cebador directo, 30 pmol del cebador inverso, 200 µM cada uno de dATP, dCTP, dGTP, y dTTP, un tampón 1X AmpliTaq de polimerasa (Applied Biosystems, Foster City, CA, EEUU), y 0.5 unidad de AmpliTaq polimerasa (5,0 U/µl, Applied Biosystems, Foster City, CA, EEUU). Las reacciones fueron incubadas en un Robocycler (Stratagene, La Jolla, CA, EEUU) programado durante 1 ciclo a 96°C durante 3 minutos y a 72°C durante 3 minutos; 34 ciclos cada una a 95°C durante 0,5 minutos, 56°C durante 0,5 minutos, y 72°C durante 1,5 minutos; 1 ciclo a 72°C durante 7 minutos; y un ciclo de remojo a 6°C. Se añadió polimerasa Taq a 72°C en el primer ciclo.

35 [0276] Los productos de reacción PCR fueron separados en un gel de agarosa 2% (Amresco, Solon, OH, EEUU) usando un tampón TAE. Una banda de aproximadamente 650 pb (Fwd y Rev1 cebadores) y bandas de aproximadamente 320 y 380 pb (cebadores Fwd y Rev2) fueron cortadas del gel y purificadas usando un equipo de extracción de gel MiniElute™ Gel Extraction Kit (QIAGEN Inc., Valencia, CA, USA) según las instrucciones del fabricante. Los productos PCR purificados fueron posteriormente analizados mediante secuenciación del ADN. Se encontró que el producto de 320 pb codificaba una parte de un polipéptido de hidrolasa de glicosil de familia 5 y fue designado CEL5C.

40 Ejemplo 19: Selección de banco genómico de *Penicillium brasilianum* IBT 20888

[0277] Las transferencias de colonias de la librería descritos en ejemplo 15 fueron realizados (Maniatis et al., 1982, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Cold Spring Harbor, Nueva York) y el ADN fue reticulada sobre membranas Hibond N+ (Amersham, Arlington Heights, IL, EEUU) durante 2 horas a 80°C. Las membranas de las transferencias de colonias fueron prehumedecidos usando 0.2X SSC (30 mM NaCl, 3 mM citrato sódico), SDS 0,2%. Los filtros prehumedecidos fueron colocados en un vaso de precipitación con 7.5 ml de 6X SSPE (0,9 M NaCl, 0,06 M de NaH₂PO₄, y 6 mM de EDTA), SDS 7%) por filtro a 68°C en un baño maría con agitación durante 30 minutos.

50 [0278] Aproximadamente 40 ng del producto PCR descrito en el ejemplo 18 fue marcado por cebador aleatorio usando un equipo Stratagene Prime-It II Kit (Stratagene, La Jolla, CA) según las instrucciones del fabricante. El fragmento de gen radiomarcado fue separado del nucleótido no-incorporado usando un equipo de purificación PCR MinElute (QIAGEN Inc., Valencia, CA, EEUU).

[0279] La sonda radiactiva fue desnaturalizada añadiendo 5,0 M de NaOH a una concentración final de 0,5 M, y añadido a la solución de hibridación a una actividad de aproximadamente $0,5 \times 10^6$ cpm por ml de solución de hibridación. La mezcla fue incubada durante 10 horas a 68°C en un baño maría con agitación. Tras la incubación, las membranas fueron lavadas tres veces en 0.2X SSC, 0,2% de SDS a 68°C. Las membranas fueron luego secadas con papel secante durante 15 minutos, envueltos en SaranWrap™, y expuesto a película radiográfica durante toda la noche a -80°C con filtros intensificadores (Kodak, Rochester, NY, USA).

[0280] Colonias produciendo señales de hibridación con la sonda fueron inoculadas en 1 ml de medio ampicilina LB y cultivado durante toda la noche a 37°C. Diluciones de cada solución fueron hechas y 100 µl fueron colocadas en placas sobre placas de ampicilina LB. La dilución para cada positivo que produjo aproximadamente 40 colonias por placa fue elegida para transferencias secundarias. Las transferencias, fueron preparadas, hibridadas, y evaluadas como descritos arriba. Dos colonias de cada placa positiva fueron inoculadas en 3 ml de medio ampicilina LB y cultivado durante toda la noche a 37°C.

[0281] ADN miniprep fue obtenido a partir de cada colonia usando un Bio Robot 9600 (QIAGEN Inc, Valencia, CA, USA) según el protocolo del fabricante. El tamaño de cada inserto fue determinado por restricción Eco RI y electroforesis en gel de agarosa. Dos clones contenían cada uno aproximadamente 5,5 kb de inserto. La secuenciación reveló que los clones eran idénticos, y se refiere a ellos de ahora en adelante como pKKAH1 (véase ejemplo 20).

Ejemplo 20: Caracterización de la secuencia genómica *cel5c* que codifica la endoglucanasa CEL5C de *Penicillium brasilianum* IBT 20888

[0282] La secuenciación del ADN del gen de endoglucanasa de *Penicillium brasilianum* de pKKAH1 fue realizada con un secuenciador de ADN automatizado de Applied Biosystems modelo 3700 (Applied Biosystems, Foster City, CA, USA) usando la técnica de paseo de cebador con química de colorante determinador de (Giesecke et al., 1992, J. Virol. Métodos 38: 47-60).

[0283] La secuencia de nucleótidos (SEC ID nº: 9) y la secuencia deducida de aminoácidos (SEC ID nº: 10) del gen *cel5c* de *Penicillium brasilianum* se muestran en las Figuras 10A y 10B. La secuencia codificante genómica de 1471 pb (incluyendo el codón de terminación) codifica un polipéptido de 421 aminoácidos, interrumpido por 4 intrones de 51 pb (89-139 pb), 47 pb (352-398 pb), 55 pb (464-518 pb), y 52 pb (617-668 pb). El %G+C contenido de la región de codificación del gen es 51.2% y la región de codificación de polipéptido maduro es 50,8%. Usando el programa de software SignalP (Nielsen et al., 1997, Protein Engineering 10: 1-6), un péptido señal de 16 residuos fue predicho. Basado en la secuencia N-terminal de la endoglucanasa, residuos desde 17 a 24 aparecen para constituir una pro-región que es proteolíticamente dividido durante la maduración. La proteína madura predicha contiene 397 aminoácidos y tiene una masa predicha de 42,6 kDa.

[0284] El análisis de la secuencia deducida de aminoácidos del gen *cel5c* con el programa Interproscan (Zdobnov y Apweiler, 2001, supra) mostró que la proteína CEL5C contenía una secuencia de núcleo típica de una hidrolasa de glicosil de familia 5, con extensión de aproximadamente desde residuos 32 hasta 307 del polipéptido predicho en toda su longitud. La proteína CEL5C también contenía la firma de secuencia de un dominio de enlace de celulosa fúngica tipo I (CBMI). Esta firma de secuencia conocida como modelo Prosite PS00562 (Sigrist et al., 2002, supra) estuvo presente desde el residuo de aminoácido 393 hasta el residuo 420 del polipéptido predicho.

[0285] Una alineación global comparativa de secuencias de aminoácidos pareja en bases de datos públicos fue determinada usando el algoritmo Needleman-Wunsch (Needleman y Wunsch, 1970, supra) como implementado en el programa Needle con una penalización de hueco abierto de 10, una penalización de extensión de hueco de 0.5, y el matriz EBLOSUM62. La alineación mostró que la secuencia deducida de aminoácidos del gen *cel5c* de *Penicillium brasilianum* que codifica el polipéptido maduro CEL5C compartían 74.9% y 74.5% identidad (excluyendo huecos) con las secuencias deducidas de aminoácidos de dos proteínas predichas de hidrolasa de glicosil de familia 5 de *Neosartoria fischeri* y *Aspergillus fumigatus*, respectivamente (números de registro A1DAP7 y Q4WM09, respectivamente).

Ejemplo 21: Construcción de una expresión de plásmido de *Aspergillus oryzae* para el gen de endoglucanasa *cel5c* de *Penicillium brasilianum* IBT 20888

[0286] El plásmido de expresión de *Aspergillus* pJaL721 (WO 03/008575) consiste en un cassette de expresión basado en la amilasa neutra II de *Aspergillus niger* promotor fundido a la secuencia líder triosa fosfata isomerasa no traducido de *Aspergillus nidulans* (NA2-tpi) y el terminador amilglicosidasa de *Aspergillus niger*. También presente en el plásmido está el marcador selectivo *amdS* de *Aspergillus nidulans* que permite el crecimiento sobre acetamida como una sola fuente de nitrógeno y el marcador URA3 de *Saccharomyces cerevisiae* que permite el crecimiento del cepa DB6507 de *Escherichia coli* pirF defectuosa (ATCC 35673). La transformación en *E. coli* DB6507 fue realizada usando el gen URA3 de *Saccharomyces cerevisiae* como marcador selectivo como se describe abajo.

[0287] *E. coli* DB6507 fue hecho competente por el método de Mandel e Higa, 1970, J. Mol. Biol. 45: 154. Los transformantes fueron seleccionados en medio sólido M9 (J. Sambrook, E.F. Fritsch, and T. Maniatis, 1989, Molecular

Cloning, A Laboratory Manual, 2d edition, Cold Spring Harbor, Nueva York) suplementado con 1 g de ácidos de casamino, 500 µg de tiamina, y 10 mg de canamicina por litro.

[0288] El gen de endoglucanasa fue clonada en pJaL721 como se describe abajo. El gen *cel5c* de endoglucanasa de *Penicillium brasilianum* fue amplificado por PCR que usa los siguiente dos cebadores oligonucleótidos:

5 PCR hacia adelante:

5'-AATTGGATCCACCATGAAATACCCTCTACTCCTGGCAAC-3' (SEQ ID NO: 14)

PCR inverso:

5'- TTAACGCGAGTTACAGACACTGCGAATAATACGCATTC-3'(SEQ ID NO: 15)

10 [0289] Para facilitar la clonación un sitio de restricción enzima fue insertada en el extremo 5' de cada cebador donde el cebador directo contenía un sitio Bam HI y el cebador inverso contenía un sitio Xho I.

15 [0290] El ADN genómico preparado como en el ejemplo 14 fue usado como molde en la reacción PCR. La reacción fue realizada en un volumen de 50 µl conteniendo 1,0 unidad de Phusion (Finnzymes Oy, Espoo, Finlandia), un tampón 1X Phusion HF (Finnzymes Oy, Espoo, Finlandia), 500 ng de genómico, 250 µM de cada dNTP, y 50 pmol de cada uno de los dos cebadores anteriormente descritos. La amplificación se efectuó en un motor PTC-220 DNA Engine Dyad Peltier Thermal Cycler (MJ Research, Inc., Waltham, MA, EEUU) programado durante 1 ciclo a 95°C durante 5 minutos; 24 ciclos cada una a 94°C durante 0.5 minutos, 58°C durante 0.5 minutos, y 68°C durante 4,0 minutos; y 1 ciclo a 68°C durante 15 minutos. La técnica de PCR de inicio caliente (Chou et al., 1992, Nucleic Acids Res. 20: 1717) fue usado y la polimerasa Phusion fue añadida 1 minuto después del primer ciclo.

20 [0291] La reacción PCR produjo un único fragmento de ADN de aproximadamente 1500 pb de longitud. El fragmento fue digerido con Bam HI y Xho I y aislado mediante electroforesis en gel de agarosa, purificado, y clonado en pJaL721 digerido con Bam HI y Xho I, dando como resultado un plásmido designado pKBK03 (Figura 11). La secuencia del gen de endoglucanasa en pKBK03 fue verificada por secuenciación con un analizador de ADN 3730xl de Applied Biosystems

25 [0292] Para crear un plásmido para el depósito de material biológico, el pKBK03 fue digerido con Bam HI y Xho I y purificado. Se repararon los extremos romos usando enzima Klenow (Roche Applied Science, Mannheim, Alemania) durante 30 minutos a 25°C. El plásmido pUC13 fue digerido con Sma I y defosforilada con proteasa Intestinal de Ternero (Roche Applied Science, Mannheim, Alemania; CIP) durante 1 hora a 37°C y el CIP fue inactivada calentando la muestra a 80°C durante 15 minutos.

30 [0293] El fragmento con los extremos romos reparados y fragmento desfosforilado pUC13 fueron ligados durante toda la noche a 16°C usando ADN ligasa T4 (Roche Applied Science, Mannheim, Alemania). Un muestra de 0,25 µg del ligada producto fue transformado en *Escherichia coli* DH5α (Invitrogen, Carlsbad, CA, USA). Tras la incubación durante toda la noche a 37°C en placas de ampicilina LB, los transformantes fueron transferidos a 2 ml de medio LB e incubado a 37°C. Un plásmido designado pPBCEl5C (Figura 12) fue purificado usando plásmido Jetquick Plasmid Miniprep (Genomed, Löhne, Alemania). La secuencia del gen de endoglucanasa fue verificada por secuenciación con un analizador de ADN 3730xl de Applied Biosystems. Las células de *E. coli* TOP10 células (Invitrogen, Carlsbad, CA) que contenían el plásmido pPBCEl5C (designación de cepa PBCEl5C) fueron depositadas con la Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, as NRRL B-30900N, con una fecha de depósito del 23 de febrero del 2006.

Ejemplo 22: Expresión de la endoglucanasa CEL5C de *Penicillium brasilianum* IBT 20888 en *Aspergillus oryzae*

40 [0294] *Aspergillus oryzae* BECh2 (WO 00/30322) fue transformado con 5 µg de pKBK03 como describe por Christensen et al., 1988, Biotechnology 6: 1419-1422.

45 [0295] Los transformantes fueron cultivados en tubos de 50 ml durante 4 días a 30°C en 10 ml de medio YPM. Los caldos enteros fueron centrifugados a 12.100 x g y los sobrenadantes quitados. Los sobrenadantes fueron analizados por SDS-PAGE usando un Gel prefabricado Criterion XT, Gel de Bis-Tris 10% en un tampón XT ms (Laboratorios BioRad, Hercules, CA, EEUU) según las instrucciones del fabricante. Un volumen de 10 µl de sobrenadante fue mezclado con 9 µl del tampón de muestra (0,125 M de Tris-HCl pH 6.8, 20% de glicerol y 4,6% de SDS), y 1 µl de 1 M ditiotreitol, y calentado a 96°C durante 5 minutos. En 16 de 20 sobrenadantes, una banda de aproximadamente 65 kDa era visible en el intervalo de los estándares de 35 kDa a 150 kDa por SDS-PAGE. Los sobrenadantes dando que resultaban una banda visible tras SDS-PAGE también contenían actividad de endoglucanasa, evaluado como se describe en el ejemplo 3. Cuanto más alta es la intensidad de la banda, más alta media la actividad endoglucanasa en el mismo sobrenadante.

50 [0296] Un transformante fue designado *Aspergillus oryzae* KBK03.

Ejemplo 23: Producción y purificación de endoglucanasa CEL5C recombinante de *Penicillium brasilianum* IBT

20888

[0297] *Aspergillus oryzae* transformante KBK03 fue cultivado en veinte matraces de agitación de 500 ml con 200 ml de YPM medio.

5 [0298] La biomasa fue eliminada de 4.0 litros de caldo de fermentación mediante centrifugado y filtración. El análisis de sDS-PAGE fue realizado como se describe en el ejemplo 9. La solución de endoglucanasa fue cargada sobre una columna XK 50 (Amersham Biosciences, Uppsala, Sweden) conteniendo 110 g de Avicel Ph 101 (Merck KGaA, Darmstadt, Alemania) pre-equilibrada con 25 mM Tris pH 7.5 antes de la carga y la enzima ligada fue eluida con 25 mM Tris, trietanolamina 1% a pH 11.6. La elución de la endoglucanasa fue monitoreada a 280 nm. La proteína eluida que
10 contenía fracciones fue agrupada inmediatamente y el pH ajustado a 7.5. Las fracciones con la endoglucanasa fueron agrupadas.

[0299] El contenido de proteína fue determinada a partir la absorbencia a 280 nm y el coeficiente de extinción calculado a partir de la estructura primaria de la endoglucanasa.

15 [0300] La purificación fue seguida de SDS-PAGE. Las muestras fueron hervidas durante 2 minutos con un volumen igual de 2X tampón de muestra y 1/5 volumen de PMSF 1% y cargado sobre un gel de Tris-glicina 4-20% (Invitrogen, Carlsbad, CA, EEUU). El gel fue manchado con reactivo azul GelCode Blue Stain Reagent (Pierce, Rockford, IL, EEUU) y desmanchado con agua. SDS-PAGE reveló una banda de aproximadamente 65 kDa.

Ejemplo 24: Caracterización de endoglucanasa CEL5C recombinante purificado de *Penicillium brasilianum* IBT 20888

20 [0301] La endoglucanasa CEL5C recombinante purificado de *Penicillium brasilianum* descrita en ejemplo 23 fue caracterizada con respecto a pH óptimo, temperatura óptima, y estabilidad de temperatura. La actividad de endoglucanasa fue medida como se describe en el ejemplo 16 a temperaturas de 20°C a 80°C y a valores de pH de 3.0 a 10.0. La endoglucanasa purificada fue diluida en agua ultrapura Milli-Q® (Millipore, Billerica, MA, EEUU) para asegurar que la actividad se mantuviera en la curva estándar. Para el pH óptimo, el sustrato fue disuelto en un tampón Britton-Robinson (50 mM ácido bórico, 50 mM ácido acético, 50 mM ácido fosfórico) ajustado al deseado pH. La estabilidad de
25 temperatura fue determinada durante 20 horas a 50°C en el margen de pH de 4.0 a 6.0. Todos los ensayos experimentales fueron realizados por duplicado.

[0302] Los determinación de los resultados del pH óptimo se muestra en la figura 13. El óptimo pH estuvo cerca de 4.0 a 50°C con muy poca actividad a pH 3.0 y aproximadamente 80% de actividad de valor máximo a pH 5.0.

30 [0303] Los resultados de la determinación óptima de temperatura se muestra en la figura 14. La temperatura óptima a pH 4.8 fue aproximadamente de 70°C con más de un 75% de actividad máxima de 60°C a 80°C.

[0304] Los resultados de la determinación de estabilidad de temperatura se muestra en la figura 15. Cuando preincubada en ausencia de sustrato durante 20 horas a 25°C y 50°C en el margen de pH de 4.0 a 6.0, la endoglucanasa retuvo más del 80% de su actividad inicial.

Ejemplo 25: Preparación de sustratos

35 [0305] El rastrojo de maíz pretratado (PCS) fue preparado por el Laboratorio Nacional de Energía Renovable del Departamento de Energía de los EEUU (NREL) usando ácido sulfúrico diluido. Las siguientes condiciones fueron usadas para el pretratamiento: 1,4 peso % ácido sulfúrico a 165°C y 107 psi durante 8 minutos. El análisis composicional fue realizado en el NREL. Se determinó celulosa y hemicelulosa mediante una hidrólisis de dos fases de ácido sulfúrico con análisis posterior de azúcares mediante cromatografía de líquidos de alto rendimiento (hPLC) el Procedimiento Analítico Estándar #002 de la NREL. Se determinó lignina gravimétricamente después de la hidrolizar las fracciones de celulosa y de hemicelulosa con ácido sulfúrico (Procedimiento Analítico Estándar #003 de la NREL). Se determinaron sólidos insolubles en agua en el rastrojo de maíz preparado (PCS) siendo 56,5% de celulosa, 4,6% de hemicelulosa, y 28,4% de lignina.

45 [0306] El PCS fue lavado con un volumen grande de agua desionizada en un embudo Kimax con un filtro de vidrio de porosidad gruesa (Fisher Scientific, Pittsburg, PA, EEUU). Los PCS lavados en agua fueron molidos en un molinillo de café y adicionalmente lavado con agua desionizada en un filtro de Millipore de 22 µm con una membrana Express 6P (Millipore, Bedford, MA, USA). El peso en seco del PCS molido fue 32,4%.

50 [0307] Un 10 mg/ml de suspensión de materia prima de celulosa hinchado de ácido fosfórico (PASC) en agua desionizada fue preparada usando el siguiente procedimiento. Ciento cincuenta ml de O-fosfórico ácido 85% helado se añadió a 5 g de Avicel PH101 (FMC Corp., Philadelphia, PA, EEUU) humedecido con agua. La suspensión fue lentamente agitada en un baño de hielo durante una hora, y 100 ml de acetona de frío de hielo se añadió a la suspensión mediante agitación constante. El compuesto acuoso fue transferido a un embudo Kimax con un filtro de vidrio de porosidad gruesa, lavado tres veces con 100 ml de acetona de helada, y drenados tanto como fuera posible después

cada lavado. Finalmente, el compuesto acuoso fue lavado dos veces con 500 ml de agua, y nuevamente drenado tanto como fuera posible después de cada lavado. El PASC fue mezclado con agua a un volumen total de 500 ml. Se añadió azida sódica a una concentración final de 0,02% para prevenir el crecimiento microbiano. El compuesto acuoso fue homogenizado usando una mezcladora y almacenado a 4°C hasta un mes.

5 [0308] Se obtuvo carboximetilcelulosa (CMC, sal sódica, tipo 7L2) con un grado medio de sustitución (DS) de 0.7 obtenida de Aqualon Division of Hercules Inc., Wilmington, DE, EEUU. Una solución de 6.25 mg/ml CMC en 50 mM de acetato sódico pH 5.0 fue preparada añadiendo lentamente CMC al tampón enérgicamente agitado. El compuesto acuoso fue calentado a aproximadamente 60°C bajo agitación continua hasta que el CMC fue completamente disuelto.

10 [0309] Se obtuvo celulosa bacteriana (BC) a partir de Nata de Coco, una celulosa apta para uso alimentario comercial (Fujicco Co., Kobe, Japón), como se describe en Boisset et al., 1999, Biochemical Journal, 340: 829-835. Una suspensión de 1 mg/ml de celulosa bacteriana en agua desionizada con azida sódica (w/v) 0.01 % fue almacenada a 4°C.

[0310] Avicel f101 se obtuvo de FMC Corporation, Philadelphia, PA, EEUU.

15 [0311] Xilano de madera de abedul was obtained from Sigma, St. Louis, MO, EEUU. Xiloglucano de semilla de tamarindo (amiloido, cantidad 00401), arabinoxilano de trigo (viscosidad media, 27 cSt, cantidad 90601), 1,4-beta-D-manano (borohidruro reducido, Man:Gal=97:3, grado de polimerización DP - 15, cantidad 90302), y galactomanano de algarrobo (viscosidad baja, borohidruro reducido, cantidad 90301) fueron obtenidos de Megazyme, Bray, Irlanda.

Ejemplo 26: Ensayo de hidrácido de ácido p-hidroxibenzoico para la determinación de azúcares reductores.

20 [0312] Los azúcares reductores (RS) fueron determinadas por un ensayo de hidrácido de ácido p-hidroxibenzoico (PHBAH) (palanca, 1972, anal. Biochem. 47: 273-279), que fue adaptado al formato de una placa microtituladora de 96 pocillos.

25 [0313] Un alícuota de 90-µl de la muestra diluida fue colocada en cada pocillo de un placa microtituladora de 96 pocillos de fondo cónico (Costar, clear polycarbonate, Corning Inc., Acton, MA, EEUU). El ensayo fue iniciado añadiendo 60 µl de 1.25% PHBAH en hidróxido sódico 2% a cada pocillo. La placa descubierta fue calentada en un bloque de calentamiento hecho a medida durante 10 minutos a 95°C. Tras el calentamiento, la placa microtituladora fue enfriada a temperatura ambiente, y 35 µl de agua desionizada se añadió a cada pocillo. Un alícuota de 100 µl fue quitada de cada pocillo y transferida a una placa microtituladora de fondo plano (Costar, medium binding polystyrene, Corning Inc., Acton, MA, EEUU). La absorbencia a 410 nm (A410) fue medida usando un lector de microplacas SpectraMAX (Molecular Devices, Sunnyvale, CA, EEUU). El valor A410 fue traducido en equivalentes de glucosa usando una curva estándar.

30 [0314] La curva estándar se obtuvo con seis estándares de glucosa (0,005; 0,010; 0,025; 0,050; 0,075 y 0,100 mg/ml), que fueron tratados de forma similar a las muestras. Los estándares de glucosa fueron preparados diluyendo 10 mg/ml solución de glucosa de materia prima con agua desionizada.

[0315] Para todos sustratos salvo xilano y arabinoxilano, el grado de conversión (%) fue calculado usando la siguiente ecuación:

35

$$\text{Conversión}_{(\%)} = \frac{\text{RS}_{(\text{mg/ml})} \times 100 \times 162}{(\text{Concentración inicial de sustrato}_{(\text{mg/ml})} \times 180)} = \frac{\text{RS}_{(\text{mg/ml})} \times 100}{(\text{Concentración inicial de sustrato}_{(\text{mg/ml})} \times 1,111)}$$

[0316] Para xilano y arabinoxilano, porcentaje de sustrato hidrolizado a RS fue calculado usando la siguiente ecuación:

40

$$\text{Conversión}_{(\%)} = \frac{\text{RS}_{(\text{mg/ml})} \times 100 \times 132}{(\text{Concentración inicial de sustrato}_{(\text{mg/ml})} \times 150)} = \frac{\text{RS}_{(\text{mg/ml})} \times 100}{(\text{Concentración inicial de sustrato}_{(\text{mg/ml})} \times 1,136)}$$

45 [0317] En estas ecuaciones, RS es la concentración de azúcares reductores en una solución medida en equivalentes de glucosa (mg/ml), y los factores 1,111 e 1,136 reflejan el aumento de peso en la de conversión de los polisacáridos correspondientes a azúcares hexosa (PM 180) o pentosa (PM 150).

Ejemplo 27: Actividad relativa de endoglucanasas en carboximetilcelulosa a 50°C

[0318] La tabla 1 muestra la actividad relativa de las endoglucanasas purificadas de *Myceliophthora thermophila* CBS 117.65, *basidiomycete* CBS 494.95, y *basidiomycete* CBS 495.95 hacia el soluble de sal sódica de carboximetilcelulosa (CMC). La actividad relativa se muestra como porcentaje de la actividad de la endoglucanasa de *basidiomycete* CBS 495.95. La actividad fue determinada midiendo la concentración de azúcares reductores (RS) producidas de CMC (5 mg/ml) después de 30 minutos de hidrólisis en 50 mM de acetato sódico pH 5.0 a 50°C. La hidrólisis se efectuó sin agitación en presencia de 0,5 mg/ml albúmina de suero bovino (BSA, Sigma, St. Louis, MO, USA). Los azúcares reductores fueron determinados usando el ensayo hidrácido de ácido p-hidroxibenzoico (PHBAH) descrito en el ejemplo 26.

Tabla 1. Actividad Relativa de endoglucanasas en carboximetilcelulosa (5 mg/ml) a pH 5.0 y 50°C

Endoglucanasa	Actividad en CMC, %
Cel5A de <i>Myceliophthora thermophila</i>	28
Cel5B de <i>basidiomycete</i> CBS 494.95	44
Cel5A de <i>basidiomycete</i> CBS 495.95	100

Ejemplo 28: Termoestabilidad de las endoglucanasas a 40°C-80°C

[0319] La termoestabilidad de las endoglucanasas purificadas de *Myceliophthora thermophila* CBS 117.65, *basidiomycete* CBS 494.95, y *basidiomycete* CBS 495.95 fue determinado incubando soluciones enzimáticas a cinco temperaturas (40°C, 50°C, 60°C, 70°C, y 80°C), y midiendo la actividad residual de enzimas en carboximetilcelulosa (CMC).

[0320] Las enzimas fueron diluidas en 50 mM de acetato sódico pH 5.0, que contenían 3,0 mg/ml de BSA, e incubados durante 3 horas en microtubos ImmunoWare de 1.1-ml arreglados en una placa microtituladora de formato 8x12 (Pierce, Rockford, IL, EEUU). Se añadió BSA para prevenir la posible adsorción de enzima sobre las paredes plásticas de los microtubos. La concentración de proteína en las mezclas de incubación fue elegida de modo que cada enzima daría menos de una conversión de CMC de 1% en un ensayo posterior para actividad CMCase.

[0321] Después de que un 3 incubación de hora, partes alícuotas de 15 µl fueron sustraídas usando una pipeta de 8 canales, y añadidos a 75 µl de solución CMC (6 mg/ml en 50 mM de acetato sódico pH 5.0) en una placa microtituladora de 96 pocillos de fondo cónico (Costar, clear polycarbonate, Corning Inc., Acton, MA, USA). La actividad residual CMCase después fue medido como se describe en el ejemplo 27, y expresado como un porcentaje inicial de la actividad CMCase (tabla 2).

[0322] A 40°C y 50°C, las tres endoglucanasas eran estables y retenían un 98-100% inicial de la actividad CMCase tras 3 horas de incubación. A 60°C y 70°C, el *Myceliophthora thermophila* Cel5A mostró mejor estabilidad que las dos otras endoglucanasas, y reteniendo el 100% y 49.3% de la actividad inicial de CMCase tras una incubación de 3 horas, respectivamente. Ninguna de las endoglucanasas eran estables a 80°C.

Tabla 2. Actividad residual de CMCase de endoglucanasas tras una incubación de tres horas a pH 5.0 y a 40-80°C

Endoglucanasa	Actividad residual CMCase, % de actividad inicial				
	40°C	50°C	60°C	70°C	80°C
Cel5A de <i>Myceliophthora thermophila</i>	100	100	100	49,3	3,8
Cel5B de <i>basidiomycete</i> CBS 494.95	100	98	28,0	5,0	4,1
Cel5A de <i>basidiomycete</i> CBS 495.95	100	99	6,4	2,4	0,8

Ejemplo 29: Actividad relativa de endoglucanasas sobre celulosa hinchado de ácido fosfórico a 40-70°C

[0323] La actividad de las endoglucanasas purificadas de *basidiomycete* CBS 494.95 y *basidiomycete* CBS 495.95 sobre celulosa hinchado de ácido fosfórico (PASC) fue determinada midiendo la concentración de los azúcares reductores (RS) liberados durante la hidrólisis inicial de PASC (2 mg/ml) en 50 mM de acetato sódico pH 5.0. La hidrólisis se efectuó sin agitación en presencia de 0,5 mg/ml albúmina de suero bovino (BSA, Sigma, St. Louis, MO, EEUU). Las enzimas fueron

diluidas de modo que la concentración de RS aumentara linealmente durante el los primeros 30 a 90 minutos de hidrólisis, y el grado de conversión PASC no excediera el 2% durante este tiempo. Los azúcares reductores fueron determinados usando el ensayo de hidrácido de ácido p-hidroxibenzoico (PHBAH) como se describe en el ejemplo 26.

- 5 [0324] La actividad relativa como una función de temperatura de las endoglucanasas de *basidiomycete* CBS 494.95 y *basidiomycete* CBS 495.95 se muestra en la figura 16. La actividad se muestra como un porcentaje de la actividad de la endoglucanasa de *basidiomycete* CBS 495.95 a 60°C. Para ambas endoglucanasas, la actividad en PASC alcanzó el valor máximo a T opt. = 60°C.

Ejemplo 30: Hidrólisis a largo plazo de celulosa hinchada de ácido fosfórico a 40-70°C

- 10 [0325] La hidrólisis de celulosa hinchado de ácido fosfórico (PASC, 2 mg/ml) por las endoglucanasas purificadas de *basidiomycete* CBS 494.95 y *basidiomycete* CBS 495.95 se efectuó durante 45 horas en 50 mM acetato sódico pH 5.0 que contenía 0,5 mg/ml BSA a cuatro temperaturas, 40°C, 50°C, 60°C, y 70°C. Las endoglucanasas fueron usadas en tres cargas de proteína, 0,056, 0,167, o 0,5 mg por g de PASC. Las reacciones con el volumen inicial de 1 ml fueron realizadas sin agitación en microtubos ImmunoWare de 1.1-ml arreglados en una placa de microtitulado en formato 8x12 (Pierce, Rockford, IL, EEUU).

- 15 [0326] Alícuotas de cien microlitros fueron sustraídas de las reacciones en diferentes puntos temporales (1, 1.5, 3, 6, 21, 27, y 45 horas) usando una pipeta de 8 canales, y añadidos a 25 µl de NaOH 2% en placa de filtración MultiScreen HV de 96 pocillos (Millipore, Bedford, MA, USA). Las muestras recogidas fueron filtradas por vacío en una placa microtitulada de fondo plano para eliminar el residuo PASC. El filtrado fue analizado para azúcares reductores mediante el ensayo de hidrácido de ácido p-hidroxibenzoica (PHBAH) como se describe en el ejemplo 26.

- 20 [0327] La Figura 17 muestra la conversión relativa de PASC después de una incubación de 45 horas con las endoglucanasas de *basidiomycete* CBS 494.95 y *basidiomycete* CBS 495.95 (0.5 mg proteína por g de PASC) como una función de temperatura. La conversión relativa se muestra como un porcentaje de la conversión obtenida después una incubación de 45 horas con *basidiomycete* CBS 495.95 a 50°C. Los perfiles de temperatura obtuvieron en otras dos cargas de proteína, 0,056 y 0,167 mg proteína por g de PASC, tenían formas similares. Para ambas endoglucanasas, la temperatura óptima para la hidrólisis a largo plazo de PASC fue 50°C.

Ejemplo 31: Caracterización de endoglucanasas en varios sustratos de polisacáridos a 50°C

- 30 [0328] Las endoglucanasas purificadas de *Myceliophthora thermophila* CBS 117.65, *basidiomycete* CBS 494.95, y *basidiomycete* CBS 495.95 fueron evaluados en la hidrólisis de varios polisacáridos a pH 5.0 (50 mM de tampón de acetato sódico) y 50°C. Los resultados fueron comparados con aquellos para endoglucanasa recombinante de *Trichoderma reesei* Cel7B (EGI). La endoglucanasa recombinante de *Trichoderma reesei* Cel7B (EGI) producida por *Aspergillus oryzae* se puede preparar según Takashima et al., 1998, Journal of Bacteriology 65: 163-171.

- 35 [0329] Los polisacáridos incluyendo rastrojo de maíz pretratado (PCS), celulosa hinchado de ácido fosfórico (PASC), carboximetilcelulosa (CMC), celulosa bacteriana (BC), Avicel, xilano, xiloglucano, arabinoxilano, manano y galactomanano. Todos los sustratos fueron usados a 5 mg/ml, con la excepción de la celulosa bacteriana, que fue usado a 0,9 mg/ml.

- [0330] Reacciones con un volumen inicial de 1 ml fueron llevados a cabo durante 24 horas agitación intermitente en placas Eppendorf 96 DeepWell (1.2 ml, VWR Scientific, West Chester, PA, EEUU) cubierto con una alfombrilla de sellado Eppendorf 96 DeepWell (VWR Scientific, West Chester, PA, EEUU). A menos que se especifique, las enzimas fueron cargadas a 5 mg de proteína por g de sólidos.

- 40 [0331] Después de 24 horas, alícuotas 20 µl fueron sustraídas de las reacciones de hidrólisis usando una pipeta de 8 canales, y añadidos a 180 µl de 102 mM Na₂CO₃-58 mM NaHCO₃) en una placa de filtración MultiScreen HV de 96 pocillos (Millipore, Bedford, MA, EEUU) para terminar la hidrólisis. Las muestras fueron filtradas por vacío en una placa microtitulada de fondo plano. Tras de la dilución apropiada, el filtrado fue analizado para azúcares de reducción que usando el ensayo de la hidracida de ácido p-hidroxibenzoica (PHBAH) como se describe en el ejemplo 26.

- 45 [0332] La tabla 3 muestra la conversión relativa de varios polisacáridos por las endoglucanasas después una incubación de 24 horas. La conversión relativa fue calculada como un porcentaje de conversión obtenida después una hidrólisis de 24 horas de 1,4-β-D-manano por endoglucanasa Cel5A de *basidiomycete* CBS 495.95. Las endoglucanasas de hidrolasa glucósida (GH) de familia 5 tenían una actividad relativamente alta en manano y galactomanano, pero actividad baja en xilano, xiloglucano y arabinoxilano. En cambio, Cel7B de *Trichoderma reesei* tuvo actividad relativamente alta en xilano, xiloglucano y arabinoxilano, pero actividad baja en manano y galactomanano. Las endoglucanasas GH5 mostraron mejor hidrólisis de PASC (celulosa amorfa insoluble no sustituida) que de CMC (derivado de celulosa soluble sustituida). Las endoglucanasas GH5 tenían actividad baja en sustratos insolubles con una grado alto de cristalinidad: celulosa bacteriana, Avicel, y PCS.

Tabla 3. Conversión Relativa de varios sustratos polisacáridos (5 mg/ml) por endoglucanasas (5 mg proteína por g sólidos); pH 5.0, 50°C, 24 horas

Substrato	Cel5A de <i>Myceliophthora thermophila</i>	CEL5B de <i>basidiomycete</i> CBS 494.95	Cel5A de <i>basidiomycete</i> CBS 495.95	Cel7B de <i>Trichoderma reesei</i>
Rastrojo de maíz pretratado (PCS)	2	6	6	10
Celulosa hinchada de ácido fosfórica (PASC)**	10	40	70	38
Carboximetilcelulosa (CMC)**	12	12	12	14
Celulosa bacteriana (BC)*	1	5	5	5
Avicel (celulosa microcristalina)	1	2	3	5
Xilano de madera de abedul	0	3	2	51
Xiloglucano de tamarindo	0	1	1	87
Arabinoxilano de trigo	6	7	100	2
1,4-β-D-Manano	70	73	100	2
Galactomanano de algarrobo	52	54	62	3
*La concentración inicial de celulosa bacteriana fue de 0.9 mg/ml				
**Todas las endoglucanasas fueron usadas en 0,25 mg proteína por g sólidos para hidrólisis de PASC y CMC				

Ejemplo 32: Hidrólisis de beta-glucano soluble de cebada por endoglucanasas a 60°C

- 5 [0333] La actividad de las endoglucanasas de *Myceliophthora thermophila* CBS 117.65, *basidiomycete* CBS 494.95, y *basidiomycete* CBS 495.95 en beta-glucano soluble de cebada (viscosidad media, 230 kDa, Megazyme International Ireland Ltd., Bray, Irlanda) fue determinada a pH 5.5 (50 mM acetato sódico con 0,02% de azida sódica) y 60°C. Los resultados fueron comparados con aquellos para endoglucanasa de *Trichoderma reesei* Cel7B (EG1). Se puede preparar endoglucanasa recombinante de *Trichoderma reesei* Cel7B (EG1) como se describe en el ejemplo 31.
- 10 [0334] La concentración inicial de beta-glucano en las reacciones de hidrólisis fue 1.0% (w/v). Un ml de reacciones fueron realizados sin agitación en placas Eppendorf 96 DeepWell (1,2 ml, VWR Scientific, West Chester, PA, EEUU). Las enzimas fueron usadas en tres cargas de proteína, 0,05, 0,1, y 0,2 mg por g de glucano. En un control de tampón, las endoglucanasas fueron sustituidas con 50 mM acetato sódico pH 5.5 conteniendo azida sódica 0,02%.
- 15 [0335] Las partes alícuotas fueron quitadas de las reacciones de hidrólisis a las 2 horas y a las 24 horas, diluidos con agua desionizada, y analizados para azúcares reductores usando el ensayo de hidracida de ácido p-hidroxibenzoica (PHBAH) como se describe en el ejemplo 26. La conversión relativa de beta-glucano como función de carga de proteína a dos tiempos de incubación, 2 horas y 24 horas, se muestra en las Figuras 18 y 19, respectivamente. La conversión relativa se muestra como un porcentaje de conversión obtenida después una hidrólisis de 24 horas de beta-glucano de endoglucanasa Cel5A de *Myceliophthora thermophila* CBS 117.65 (0,2 mg de proteína por g de glucano).
- 20 [0336] Las endoglucanasas de *basidiomycete* CBS 494.95 y *basidiomycete* CBS 495.95 mostraron un rendimiento similar en beta-glucano hidrolizante, y no produjeron aumento adicional alguno en la concentración de azúcares reductores después una hidrólisis de 2 horas. En cambio, las endoglucanasas de *Myceliophthora thermophila* y *Trichoderma reesei* continuaron produciendo nuevos grupos terminales reductores tras el período de incubación de 2 horas. La endoglucanasa de *Myceliophthora thermophila* mostró mejor rendimiento en beta-glucano hidrolizante que la
- 25 endoglucanasa CEL5B de *basidiomycete* CBS 494.95 y endoglucanasa Cel5A de *basidiomycete* CBS 495.95.

Depósito de material biológico

[0337] Los siguientes materiales biológicos han sido depositados según las condiciones del tratado de Budapest con la Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, y dado los siguientes números de registro:

Depósito	Número de registro	Fecha de depósito
Cepa de <i>E. coli</i> pCIC161	NRRL B-30902	23 de Febrero del 2006

5

Las cepas han sido depositadas bajo condiciones que aseguran que el acceso a los cultivos estará disponible durante la tramitación de esta solicitud de patente a una determinada por el Comisario de Patentes y Marcas Registradas a ser autorizado a ellas bajo 37 C.F.R. §1.14 y 35 U.S.C. §122. Los depósitos representan cultivos sustancialmente puros de las cepas depositadas. Los depósitos están disponibles según requieren leyes de patente extranjeras en países donde los homólogos del sujeto de la aplicación, o sus descendientes son solicitados. No obstante, debe ser entendido que la disponibilidad de un depósito no constituye una licencia para practicar e de sujeto de la invención en derogación del derecho resultante de las patentes concedidas por acción gubernamental.

10

LISTADO DE SECUENCIAS

[0338]

15

<110> Novozymes, Inc.
 Novozymes A/S
 Lassen, Søren F
 Harris, Paul
 Vlasenko, Elena

20

<120> Polipéptidos con actividad endoglucanasa y polinucleótidos codificantes de los mismos

<130> 10890.204-WO

25

<150> 60784088 <151> 2006-03-20

<160> 15

<170> Patent In versión 3.4

30

<210> 1
 <211> 12
 <212> DNA
 <213> *Myceliophthora thermophila*

35

<400> 1
 cttccagca ca 12

40

<210> 2
 <211> 8

<212> DNA

<213> *Myceliophthora thermophila*

<400> 2

5 gaaaggtc 8

<210> 3

<211> 1188

<212> DNA

10 <213> *Myceliophthora thermophila*

<400> 3

ES 2 368 608 T3

cgacttgaaa cgccccaaat gaagtcctcc atcctcgcca gcgtcttcgc cacggggccc 60
 gtygctcaaa gtggtcctgt gcagcaatgt ggtggcatcg gatggcaagg atcgaccgac 120
 tgytgctcgg gctaccactg cgtctaccag aacgattggt acagccagtg cgtgcctggc 180
 ggggctoga caacgctgca gacatcgacc acgtccaggc ccaccgccac cagcaccgcc 240
 cctccgtcgt ccaccacctc gcctagcaag ggcaagctga agtggctcgg cagcaacgag 300
 tcgggctcgg agttcgggga gggcaattac cccggcctct ggggcaagca ctccatcttc 360
 ccgtcgactt cggcgattca gacgctcacc aatgatggat acaacatctt ccggatcgac 420
 ttctcgatgg agcgtctggt gcccaccag ttgacgtcgt ccttcgacca gggttacctc 480
 cycaacctga ccgaggtggt caacttcgtg acgaacgcgg gcaagtacgc cgtcctggac 540

ccgcacaact acggccggta ctacggcaac atcatcacgg acacgaacgc gttccggacc 600
 ttctggacca aectggccaa gcagttcgcc tccaaactcgc tcgtcatctt cgacaccaac 660
 aacgagtaca acacgatgga ccagaccctg gtgctcaacc tcaaccaggc cggcatcgac 720
 ggcacccggg ccgcccgggc gacctcgag tacatcttcg tcgagggcaa cgcgtggagc 780
 ggggctgga gctggaacac gaccaacacc aacatggccg cctgacgga cccgcagaac 840
 aagatcgtgt acgagatgca ccagtacctc gactcggaca gctcgggac ccacgccgag 900
 tgcgtcagca gcaccatcgg cgcaccagcc gtcgtcggag ccaccctagt gctccgcgcc 960
 aacggcaagc tcggcgtcct cggcgagttc gccggggcg ccaacgcctt ctgccagcag 1020
 gccgtcaccg gctcctcga ccacctccag gacaacagcg acgtctggct gggcgccctc 1080
 tggggggcgg ccggtccctg gtggggcgac tacatgtact cgttcgagcc tccttcgggc 1140
 accggctatg tcaactacaa ctcgatcttg aagaagtact tgccgtaa 1188

<210> 4

<211> 389

5

<212> PRT

<213> *Myceliophthora thermophila*

<400> 4

ES 2 368 608 T3

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

Gln Ser Gly Pro Trp Gln Gln Cys Gly Gly Ile Gly Trp Gln Gly Ser
 20 25 30

Thr Asp Cys Val Ser Gly Tyr His Cys Val Tyr Gln Asn Asp Trp Tyr
 35 40 45

Ser Gln Cys Val Pro Gly Ala Ala Ser Thr Thr Leu Gln Thr Ser Thr
 50 55 60

Thr Ser Arg Pro Thr Ala Thr Ser Thr Ala Pro Pro Ser Ser Thr Thr
 65 70 75 80

Ser Pro Ser Lys Gly Lys Leu Lys Trp Leu Gly Ser Asn Glu Ser Gly
 85 90 95

Ala Glu Phe Gly Glu Gly Asn Tyr Pro Gly Leu Trp Gly Lys His Phe
 100 105 110

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

Asn Ile Phe Arg Ile Asp Phe Ser Met Glu Arg Leu Val Pro Asn Gln
 130 135 140

Leu Thr Ser Ser Phe Asp Gln Gly Tyr Leu Arg Asn Leu Thr Glu Val
 145 150 155 160

Val Asn Phe Val Thr Asn Ala Gly Lys Tyr Ala Val Leu Asp Pro His
 165 170 175

Asn Tyr Gly Arg Tyr Tyr Gly Asn Ile Ile Thr Asp Thr Asn Ala Phe
 180 185 190

Arg Thr Phe Trp Thr Asn Leu Ala Lys Gln Phe Ala Ser Asn Ser Leu
 195 200 205

Val Ile Phe Asp Thr Asn Asn Glu Tyr Asn Thr Met Asp Gln Thr Leu
 210 215 220

Val Leu Asn Leu Asn Gln Ala Ala Ile Asp Gly Ile Arg Ala Ala Gly
 225 230 235 240

Ala Thr Ser Gln Tyr Ile Phe Val Glu Gly Asn Ala Trp Ser Gly Ala
 245 250 255

Trp Ser Trp Asn Thr Thr Asn Thr Asn Met Ala Ala Leu Thr Asp Pro
 260 265 270

Gln Asn Lys Ile Val Tyr Glu Met His Gln Tyr Leu Asp Ser Asp Ser
 275 280 285

Ser Gly Thr His Ala Glu Cys Val Ser Ser Thr Ile Gly Ala Gln Arg
 290 295 300

Val Val Gly Ala Thr Gln Trp Leu Arg Ala Asn Gly Lys Leu Gly Val
 305 310 315 320

Leu Gly Glu Phe Ala Gly Gly Ala Asn Ala Val Cys Gln Gln Ala Val
 325 330 335

Thr Gly Leu Leu Asp His Leu Gln Asp Asn Ser Asp Val Trp Leu Gly
 340 345 350

Ala Leu Trp Trp Ala Ala Gly Pro Trp Trp Gly Asp Tyr Met Tyr Ser
 355 360 365

ES 2 368 608 T3

Phe Glu Pro Pro Ser Gly Thr Gly Tyr Val Asn Tyr Asn Ser Ile Leu
 370 375 380

Lys Lys Tyr Leu Pro
 385

<210> 5

<211> 1232

<212> ADN

5 <213> BASIDIOMYCETE CBS 495. 95

<400> 5

```

ggatccactt agtaacggcc gccagtggtc tggaaagcat gaagtctctc ttcctgtcac    60
ttgtagcgac cgtcgcgctc agctcgccag tattctctgt cgcagtctgg gggcaatgcg    120
gcggcattgg cttcagcggg agcaccgtct gtgatgcagg cgcggctgt gtgaagctca    180
acgactatta ctctcaatgc caaccggcg ctccactgc tacatccgcg ggcceaagta    240
gcaacgcacc gtcggcact tcgacggcct cggcccctc ctccagcctt tgccttgcca    300
gccgcacgcc gttccagttc ttcgggtgca acgaatccgg cggggagttc ggcaacctga    360
acatccccgg tgttctgggc accgactaca cctggccgtc gccatccagc attgacttct    420
tcattgggcaa gggaaatgaat accttccgta ttccttctct catggagcgt cttgtcccc    480
ctgccactgg catcacagga cctctcgacc agacgtactt gggcggcctg cagacgattg    540
tcaactacat caccggcaaa ggcggctttg ctctcattga ccgcacaaac tttatgatct    600
acaatggcca gacgatctcc agtaccagcg acttccagaa gttctggcag aacctcgcag    660
gagtgtttaa atcgaacagt caggtcatct tcgatgttat gaacgagcct cagcatattc    720
ccgccagac cgtgttccaa ctgaaccaag ccgctgtcaa tggcatccgt gcgagcggtg    780
cgacgtcgca gctcattctg gtcgagggca caagctggac tggagcctgg acctggacga    840
cctctggcaa cagcgatgca ttcggtgcca ttaaggatcc caacaacaac gtcgcgatcc    900
agatgcatca gtacctgat agcgatggtc ctggcacttc gcagacctgc gttctccca    960
ccatcgggtc cgagcgggtg caggctgcga ctcaatggtt gaagcagaac aacctcaagy   1020
gcttccctggg cgagatcggc gccggctcta actccgcttg catcagcgtc gtgcagggtg   1080
cgtctgttcc gatgcagcaa tctgggtgtg ggctcggcgc tctctggtgg gctgcccccc   1140
cgtggtgggg cgactactac cagtccatcg agccgcctc tggcccggcg gtgtccgcga   1200
tcttcccgca ggccctgctg ccgttcgcgt aa                                  1232
    
```

<210> 6

<211> 397

<212> PRT

5 <213> BASIDIOMYCETE CBS 495.95

<400> 6

ES 2 368 608 T3

Met Lys Ser Leu Phe Leu Ser Leu Val Ala Thr Val Ala Leu Ser Ser
1 5 10 15

Pro Val Phe Ser Val Ala Val Trp Gly Gln Cys Gly Gly Ile Gly Phe
20 25 30

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

Asp Tyr Tyr Ser Gln Cys Gln Pro Gly Ala Pro Thr Ala Thr Ser Ala
50 55 60

Ala Pro Ser Ser Asn Ala Pro Ser Gly Thr Ser Thr Ala Ser Ala Pro
65 70 75 80

Ser Ser Ser Leu Cys Ser Gly Ser Arg Thr Pro Phe Gln Phe Phe Gly
85 90 95

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

Leu Gly Thr Asp Tyr Thr Trp Pro Ser Pro Ser Ser Ile Asp Phe Phe
115 120 125

Met Gly Lys Gly Met Asn Thr Phe Arg Ile Pro Phe Leu Met Glu Arg
130 135 140

Leu Val Pro Pro Ala Thr Gly Ile Thr Gly Pro Leu Asp Gln Thr Tyr
145 150 155 160

Leu Gly Gly Leu Gln Thr Ile Val Asn Tyr Ile Thr Gly Lys Gly Gly
165 170 175

Phe Ala Leu Ile Asp Pro His Asn Phe Met Ile Tyr Asn Gly Glu Thr
180 185 190

Ile Ser Ser Thr Ser Asp Phe Gln Lys Phe Trp Gln Asn Leu Ala Gly
195 200 205

Val Phe Lys Ser Asn Ser His Val Ile Phe Asp Val Met Asn Glu Pro
210 215 220

His Asp Ile Pro Ala Gln Thr Val Phe Gln Leu Asn Gln Ala Ala Val

ES 2 368 608 T3

```

ggaaagcgtc agtatgggga aatttgcgct tgggcaact gtcggcgcaa tcttgagcgc      50
ttctgcggcc aatgcggcct ctatctacca gcaatgtgga ggcattggat ggtctgggtc      120
cactgtttgc gacgcgggtc tcgcttgcgf tatcctcaat gcgtactact ttcagtgcct      180
gacgcccgcc gcgggccaga caacgacggg ctcgggcgca ccggcgta caacaacctc      240
tcaetcaacg gtcactacgg ggagctcaca ctaacaacc gggacgacgg cgacgaaaac      300
aactaccact ccgtcgacca ccacgacct acccgccatc tctgtgtctg gtcggtctg      360

ctctggctcc aggacgaagt tcaagttctt cgggtggaat gaaagcggcg ccgaattcgg      420
gaacactgct tggccagggc agctcgggaa agactataca tggccttcgc ctacgagcgt      480
ggactacttc atgggggctg gattcaatac attccgtatc accttcttga tggagcgtat      540
gagccctcgg gctaccggac tcaactggccc attcaaccag acgtacctgt cgggectcac      600
caccattgtc gactacatca cgaacaagg aggatacgtc cttattgacc cccacaactt      660
catgcgttac aacaacggca taatcagcag cacatctgac ttcgcgactt ggtggagcaa      720
tttggcactc gtattcaaat ccacgaaga cggcatcttc gacatccaga acgagccgta      780
cggaatcgat gcgcagaccg tatacgaact gaatcaagct gccatcaatt cgatccgcgc      840
cgtcggcgtc acgtcacagt tgattctggt tgaaggaacg tcatacactg gagcttggac      900
gtgggtctcg tccggaaacg gagctgcttt cgcggccggt acggatcctt acaacaacac      960
ggcaattgaa atgcaccaat acctcgacag cgcgggtctt gggacaaacg aagactgtgt      1020
ctctccacc attgggtcgc aacgtctcca agctgccact gcgtggctgc aacaacagg      1080
actcaagggg ttctcggag agacgggtgc tgggtcgaat tcccagtgca tcgacgccgt      1140
gttcgatgaa ctttctata tgcaacagca aggcggctcc tggatcggty cactctggtg      1200
ggctgcgggt cctctgtygy gcacgtacat ttactcgatt gaacctcga gcggtgccgc      1260
tatccagaa gtccttctc agggctcgc tccattctc tag      1303

```

<210> 8

<211> 429

5

<212> PRT

<213> BASIDIOMYCETE CBS 494.95

<400> 8

ES 2 368 608 T3

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

Ser Ala Ala Asn Ala Ala Ser Ile Tyr Gln Gln Cys Gly Gly Ile Gly
20 25 30

Trp Ser Gly Ser Thr Val Cys Asp Ala Gly Leu Ala Cys Val Ile Leu
35 40 45

Asn Ala Tyr Tyr Phe Gln Cys Leu Thr Pro Ala Ala Gly Gln Thr Thr
50 55 60

Thr Gly Ser Gly Ala Pro Ala Ser Thr Ser Thr Ser His Ser Thr Val
65 70 75 80

Thr Thr Gly Ser Ser His Ser Thr Thr Gly Thr Thr Ala Thr Lys Thr

Ser Ser Thr Ile Gly Ser Gln Arg Leu Gln Ala Ala Thr Ala Trp Leu
 340 345 350

Gln Gln Thr Gly Leu Lys Gly Phe Leu Gly Glu Thr Gly Ala Gly Ser
 355 360 365

Asn Ser Gln Cys Ile Asp Ala Val Phe Asp Glu Leu Cys Tyr Met Gln
 370 375 380

Gln Gln Gly Gly Ser Trp Ile Gly Ala Leu Trp Trp Ala Ala Gly Pro
 385 390 395 400

Trp Trp Gly Thr Tyr Ile Tyr Ser Ile Glu Pro Pro Ser Gly Ala Ala
 405 410 415

Ile Pro Glu Val Leu Pro Gln Gly Leu Ala Pro Phe Leu
 420 425

<210> 9

<211> 1498

<212> ADN

5 <213> pENICILLIUM BRASILLANUM

<400> 9

ES 2 368 608 T3

atgaaatacc ctctactcct ggcaaccagc gctgcctcgg ctcttgctgg gcctcaaggc 50
 ttctccaagc gagcctcttc ttttgttgt atgtgtagag cccataagaa taagagctgg 120
 aactgacctg aatttttagg gtttgytacc agcgagtctg gggcggagtt tggcaatcag 180
 aacattccgg gtgttttggg gacggactac atttggcccg acacttcagc cattcaaacg 240
 ctacgcaacg cgggtatgaa catctttcgg gtggcatttt tgatggagcg attggtccc 300
 accacattga cctcaactcc ggattcgacg tatctgcaag acctgaagag tgtacgttat 360
 cctcgtcaac aatgcaacct gtgctaaacg atctccagac agttgactac atcaacgtcga 420
 ccggcgcata tgctattgtg gacccccata actttggacg atagtaagtt gaacctttca 480
 tctaaccaat attgaaaate tcttcacggg tgtattagct atggcaatat aatcaactcc 540
 acaagtgact ttgctgcatt ttggaccacc gtcgcaaagc agtttgcatc gaatgacaag 600
 gtcactcttg acacgagtga gcatcactac ttcactaact ttgttcccc actaactgac 660
 atctgcagac aacgaattca acacagagga ccaaacgctt gtactgaacc tcaatcaage 720
 ggccattaat gccatccgag ctgccggagc cacctctcag tatatctttg tggaaaggaa 780
 ttcgtggagt ggtgcctyga cgtggacgic gytcaatacc aatctcgtca gcttgacgga 840
 ccccaacaac aagatagctt acgagatgca ccagtatctt gactcggatg gatctggcac 900

 atccgacaca tgtgtcagct cgaccatcgg ccaggagcgt gtacaatcgg cgaccgagtg 960
 gctgaaaagc aatggcaaac ttggattctt gggcgaattt gcgggcgggtg ctaactcggg 1020
 ctgtcagagc gccgtgactg gaatgctyga ctatatgca gcaaatagtg acgtctggct 1080
 cggcgcattc ttgtgggcag caggaccatg gtggggaacc tatatatatt cgattgagcc 1140
 gccatcgggg actgcttatt cttactatct caacatcttg tctgcctact tccccccag 1200
 ctcgggaagt tccacaacca caactacctc cactaccacc cgctctacat cgacaagcac 1260
 tacagtatcc accacaaaat caacaagcac cactacaagc gccacgaaat cgacaagcac 1320
 cacgacaagc accaccagca ccgggtctac tgcactyca acagcatctc actyggcaca 1380
 gtgtggcggc attggctyga cgggggcgac gacgtgtgcc agcccgtata cctgccaggt 1440
 ccagaatcgg tattattcgc agtgtctgta atgcagaagt atcagaaa 1488

<210> 10

<211> 421

<212> PRT

5

<213> PENICILLIUM BRASILIANUM

<400> 10

Met Lys Tyr Pro Leu Leu Leu Ala Thr Ser Ala Ala Leu Ala Leu Ala
 1 5 10 15

Gly Pro Gln Gly Phe Ser Lys Arg Ala Ser Ser Phe Val Trp Phe Gly
 20 25 30

Thr Ser Glu Ser Gly Ala Glu Phe Gly Asn Gln Asn Ile Pro Gly Val
 35 40 45

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

Arg Asn Ala Gly Met Asn Ile Phe Arg Val Ala Phe Leu Met Glu Arg
 65 70 75 80

Leu Val Pro Thr Thr Leu Thr Ser Thr Pro Asp Ser Thr Tyr Leu Gln
 85 90 95

Asp Leu Lys Ser Val Leu Asp Tyr Ile Thr Ser Thr Gly Ala Tyr Ala
 100 105 110

Ile Val Asp Pro His Asn Phe Gly Arg Tyr Tyr Gly Asn Ile Ile Asn
 115 120 125

ES 2 368 608 T3

Ser Thr Ser Asp Phe Ala Ala Phe Trp Thr Thr Val Ala Lys Gln Phe
 130 135 140

Ala Ser Asn Asp Lys Val Ile Phe Asp Thr Asn Asn Glu Phe Asn Thr
 145 150 155 160

Glu Asp Gln Thr Leu Val Leu Asn Leu Asn Gln Ala Ala Ile Asn Ala
 165 170 175

Ile Arg Ala Ala Gly Ala Thr Ser Gln Tyr Ile Phe Val Glu Gly Asn
 180 185 190

Ser Trp Ser Gly Ala Trp Thr Trp Thr Ser Val Asn Thr Asn Leu Val
 195 200 205

Ser Leu Thr Asp Pro Asn Asn Lys Ile Val Tyr Glu Met His Gln Tyr
 210 215 220

Leu Asp Ser Asp Gly Ser Gly Thr Ser Asp Thr Cys Val Ser Ser Thr
 225 230 235 240

Ile Gly Gln Glu Arg Val Gln Ser Ala Thr Glu Trp Leu Lys Ser Asn
 245 250 255

Gly Lys Leu Gly Phe Leu Gly Glu Phe Ala Gly Gly Ala Asn Ser Val
 260 265 270

Cys Gln Ser Ala Val Thr Gly Met Leu Asp Tyr Met Gln Ala Asn Ser
 275 280 285

Asp Val Trp Leu Gly Ala Ser Trp Trp Ala Ala Gly Pro Trp Trp Gly
 290 295 300

Thr Tyr Ile Tyr Ser Ile Glu Pro Pro Ser Gly Thr Ala Tyr Ser Tyr
 305 310 315 320

Tyr Leu Asn Ile Leu Ser Ala Tyr Phe Pro Ser Ser Ser Gly Ser Ser
 325 330 335

Thr Thr Thr Thr Thr Ser Thr Thr Thr Arg Ser Thr Ser Thr Ser Thr
 340 345 350

Thr Val Ser Thr Thr Lys Ser Thr Ser Thr Thr Thr Ser Ala Thr Lys
 355 360 365

Ser Thr Ser Thr Thr Thr Ser Thr Thr Ser Thr Gly Ser Thr Ala Thr

ES 2 368 608 T3

370

375

380

Ala Thr Ala Ser His Trp Ala Gln Cys Gly Gly Ile Gly Trp Thr Gly
385 390 395 400

Ala Thr Thr Cys Ala Ser Pro Tyr Thr Cys Gln Val Gln Asn Ala Tyr
405 410 415

Tyr Ser Gln Cys Leu
420

<210> 11

<211> 29

<212> DNA

5 <213> PENICILLIUM BRASILIANUM

<220>

<221> misc_feature

<222> (21)..(21)

10 <223> N= A,C,G,OR T

<220>

<221> misc_feature

<222> (24)..(24)

15 <223> N= A,C,G,OR T

<220>

<221> misc_feature

<222> (27)..(27)

20 <223> R= A OR G

<100> 11

ttcggtagct ctgagtctgg ngcaartt□ □ □ 29

25 <210> 12

<211> 35

<212> ADN

<213> PENICILLIUM BRASILIANUM

30 <220>

<221> misc_feature
 <222> (24)..(24)
 <223> R=A OR G

5

<220>
 <221> misc_feature
 <222> (27)..(27)
 <223> N=A,C,G, OR T

10

<220>
 <221> misc_feature
 <222> (30)..(30)
 <223> R=A OR G

15

<220>
 <221> misc_feature
 <222> (33)..(33)
 <223> R=A OR G

20

<400> 12
 tgatccatat cgtgtactc gtrrtngtr tcr aa □ □ □ 35

<210> 13
 <211> 29

25

<212> DNA
 <213> PENICILLIUM BRASILIANUM

<220>
 <221> misc_feature

30

<222> (18)..(18)
 <223> R=A OR G

<220>
 <221> misc_feature

35

<222> (21)..(21)
 <223> R=A OR G

<220>
 <221> mise_feature

ES 2 368 608 T3

<222> (24)..(24)

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

<220>

5 <221> misc_feature

<222> (27)..(27)

<223> R=A OR G

<400> 13

10 ccgtttagc gaccgtarrtt rtgnggrtc □ □ □ 29

<210> 14

<211> 39

<212> ADN

15 <213> PENICILLIUM BRASILIANUM

<400> 14

aattggarcc accatgaaat acccttact cctggcaac □ □ □ 39

20 <210> 15

<211> 38

<212> ADN

<213> PENICILLIUM BRASILIANUM

25 <400> 15

ttaactcgag ttacagacac tgccaataat acgcattc □ □ □ 38

REIVINDICACIONES

1. Polipéptido aislado con actividad endoglucanasa, seleccionado del grupo que consiste en:
 - 5 (a) un polipéptido que comprende una secuencia de aminoácidos que tiene al menos un grado de identidad del 85% con el polipéptido maduro de la SEC ID n°: 4;
 - (b) un polipéptido codificado por un polinucleótido que hibridiza bajo al menos condiciones de astringencia altas con (i) la secuencia codificante de polipéptido maduro de la SEC ID n°: 3, (ii) la secuencia de ADNc contenida en la secuencia codificante de polipéptido maduro de la SEC ID n°: 3, o (iii) una cadena complementaria de longitud completa de (i) o (ii), donde una astringencia alta se define como hibridación en 10 5xSSPE y formamida al 50% a 42°C y lavado de 15 min tres veces usando 2x SSC, SDS al 0,2% al menos a 65°C; y
 - (c) un polipéptido codificado por un polinucleótido con al menos un grado de identidad del 85% con la secuencia codificante de polipéptido maduro de la SEC ID NO:3;

15 y donde la actividad endoglucanasa del polipéptido aislado es al menos un 70% de la actividad endoglucanasa del polipéptido descrito como SEC ID n°. 4 determinado usando hidrólisis de carboximetilcelulosa y donde 1 unidad se define como 1.0 micro mol de azúcares reductores producidos por minuto a 50°C, pH 4.8.
2. Polipéptido según la reivindicación 1, que es codificado por un polinucleótido con preferiblemente al menos una identidad del 90%, y de la forma más preferible al menos una identidad del 95% con la secuencia codificante de polipéptido maduro de la SEC ID n°: 3.
- 20 3. Polipéptido según cualquiera de las reivindicaciones 1 ó 2, que es codificado por el polinucleótido contenido en el plásmido pCIC161 contenido en *E. coli* NRRL B-30902.
4. Polipéptido según cualquiera de las reivindicaciones de 1 a 3, donde el polipéptido maduro es los aminoácidos 17 a 389 de la SEC ID n°: 4.
- 25 5. Polipéptido según cualquiera de las reivindicaciones de 1 a 4, teniendo además actividad enzimática hacia uno o más sustratos seleccionados del grupo que consiste en xilano, xiloglucano, arabinoxilano, 1,4-beta-D-manano, y galactomanano.
6. Polinucleótido aislado que comprende una secuencia de nucleótidos que codifica el polipéptido de cualquiera de las reivindicaciones de 1 a 5.
- 30 7. Constructo de ácidos nucleicos que comprende el polinucleótido según la reivindicación 6 operativamente enlazado a una o más secuencias de control que dirigen la producción del polipéptido en un huésped de expresión.
8. Vector de expresión recombinante que comprende el constructo de ácidos nucleicos según la reivindicación 7.
9. Célula huésped recombinante que comprende el constructo de ácidos nucleicos según la reivindicación 7.
- 35 10. Método para la producción del polipéptido según cualquiera de las reivindicaciones de 1 a 5 que comprende: (a) el cultivo de una célula huésped que comprende un constructo de ácidos nucleicos que comprende una secuencia de nucleótidos que codifica el polipéptido bajo condiciones propicias para la producción del polipéptido; y (b) la recuperación del polipéptido.
- 40 11. Método para la producción del polipéptido según cualquiera de las reivindicaciones de 1 a 5, que comprende: (a) el cultivo de una planta transgénica o de una célula vegetal que comprende un polinucleótido que codifica un polipéptido con actividad endoglucanasa según las reivindicaciones 1-5, bajo condiciones propicias para la producción del polipéptido; y (b) la recuperación el polipéptido.
12. Planta transgénica, parte de planta o célula vegetal, que comprende un constructo de ácidos nucleicos según la reivindicación 7.
- 45 13. Método de degradación o conversión de un material celulósico, que comprende: el tratamiento del material celulósico con una composición que comprende una cantidad eficaz del polipéptido aislado con actividad endoglucanasa de cualquiera según cualquiera de las reivindicaciones de 1 a 5.
14. Método según la reivindicación 13, donde la composición además comprende una cantidad eficaz de una endo-1,4-beta-glucanasa, exo-1,4-beta-D-glucanasa, y/o beta-D-glucosidasa.
15. Método según la reivindicación 13 ó 14, donde la composición además comprende una cantidad eficaz de una o más enzimas seleccionadas del grupo que consiste en una hemicelulasa, esterasa, proteasa, lacasa y peroxidasa.

16. Método según cualquiera de las reivindicaciones 13-15, donde el método es un proceso de pretratamiento; un paso en un proceso de sacarificación y fermentación simultáneas (SSF); o un paso de un proceso de hidrólisis y fermentación híbridas (HHF).
- 5 17. Método para producir una sustancia, que comprende: (a) la sacarificación de un material celulósico con una composición que comprende una cantidad eficaz del polipéptido aislado con actividad endoglucanasa según cualquiera de las reivindicaciones de 1-5, (b) la fermentación del material celulósico sacarificado del paso (a) con uno o más microorganismos fermentativos; y (c) la recuperación de la sustancia a partir de la fermentación.
18. Método según la reivindicación 17, donde la composición además comprende una cantidad eficaz de endo-1,4-beta-glucanasa, exo-1,4-beta-D-glucanasa, y/o beta-D-glucosidasa.
- 10 19. Método según la reivindicación 17 ó 18, donde la composición además comprende una cantidad eficaz de una o más enzimas seleccionadas del grupo que consiste en una hemicelulasa, esterasa, proteasa, lacasa y peroxidasa.
20. Método según cualquiera de las reivindicaciones 17-19, donde los pasos (a) y (b) se realizan simultáneamente en una sacarificación y fermentación simultáneas.
- 15 21. Método según cualquiera de las reivindicaciones 17-20, donde la sustancia es un alcohol, ácido orgánico, cetona, aminoácido o gas.

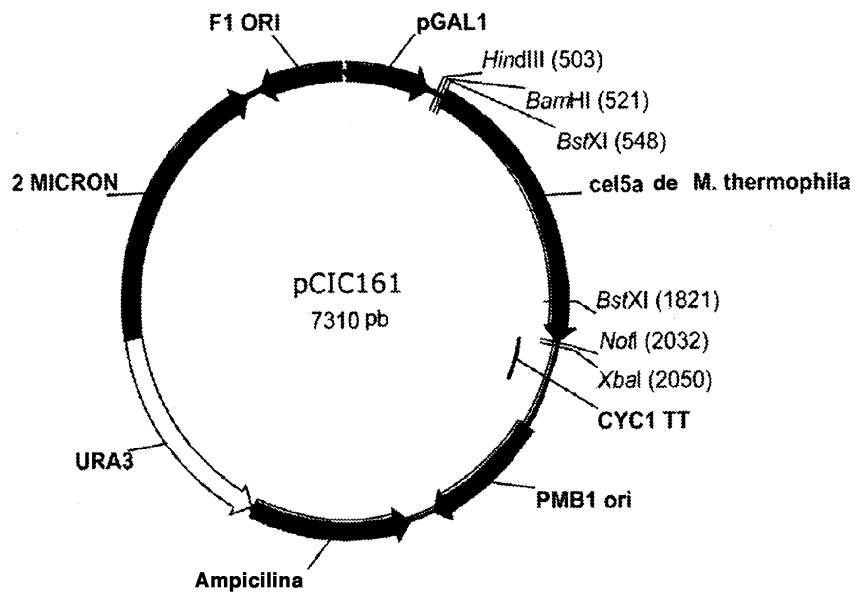


Fig. 1

CGACTTGAAACGCCCCAAATGAAGTCCATCCTCGCAGCGGTCCTCCACCGGGCCGCTGGCTCAAAGTGGTCGGTGGCAGCAATGTGGTGGC
M K S S I L A S V F A T G A V A Q S G P W Q O C G G
ATCGATGGCAAGGATCGACCGACTGTGTGGCTACCATCGTGTACCAGCATGGTGTACAGCCAGTGGTCCCTGGCGCGGCTGCAC
I G W Q G S T D C V S G Y H C V Y Q N D W Y S Q C V P G A A S T
AGGCTGAGACATCGACCGTCCAGCCCGCCACCGCCCTCGTCCACACCTCGCTAGCAGGGCAAGCTGAAGTGGCTC
T L Q T S T T S R P T A T S T A P P S T T S P S K G K L K W L
GGCAGCAACGAGTCGGGCGGAGTTCGGGAGGGCAATACCCCGGCTCTGGGCAAGCACTTCACTTCCCGTCCACTCCGGGATTCAGACG
G S N E S G A E F G E G N Y P G L W G K H F I F P S I S A I Q T
CTCATCAATGATGATACAACTTCCGGATCGACTTCTCGATGGAGGCTGTGGTCCCAACCACTTGACCTGTCTTCCGACCCAGGGTTACCTC
L I N D G Y N I F R I D F S M E R L V P N Q L T S S F D Q G Y L
CGCAACTGACCGAGGTGTCAACTTGTGACGAACCGGGCACTACCGCCGCTCTGGACCCGACACTAGCCCGCTACTACGGCAACATCATC
R N L T E V V N F V T N A G K Y A L D P H N Y G R Y G N I I
ACGGACAGGACCGGTTCCGGACCTTCTGGACCAACCTGGCCCAAGCTCCCACTCGTCCGTCATCTTCGACACCAACCAACGAGTACCAAC
T D T N A F R T F W T N L A K Q F A S N S L V I F D T N E Y N
ACGATGGACCAACCTTGTCTCAACCTCAACCGCCCGCCATCCGGCCCGGGCCGCGGACTCGGCACTACATCTTCCGTCGAGGGC
T M D Q T L V L N L N Q A A I D G I R A A G A T S Q Y I F V E G
AAGCGTGGAGCGGGCCCTGGAGCTGGAACACGACCAACCAACATGGCCCGCTGACGGACCCCGCAGAACAAAGATCGTGTACGAGATGCACCCAG
N A W S G A W S W N T T N M A A L T D P Q N K I V Y E M H Q
TACCTCGACTGGACAGCTCGGGCACCCACCGGAGTGCCTCAGCAGCACCATCGGGCCAGCGCTGTGGAGCCACCCAGTGGCTCCGCGCC
Y L D S D S S G T H A E C V S S T I G A Q R V V G A T Q W L R A
AAGCGCAAGCTCGGCGTCTCGGCGAGTTCGCGGGCGCCCAACCGCTGTGCAGAGCGCGTCCCGCCCTCCGACCACTCCAGGACAAAC
N G K L G V L G E F A G G A N A V C Q Q A V T G L L D H L Q D N
AGCGACGTCTGGTGGTCCCTCTGGTGGGCGCGGCTCCCTGGTGGGCGACTACATGTACTGTGGAGCCCTCTTCGGGACCCGGCTATGTC
S D V W L G A L W W A A G P W W G D Y M Y S F E P P S G T G Y V
AACTACACTCGATCTGAGAAGTACTTCCCTAA
N Y N S I L K K Y L P

Fig. 2

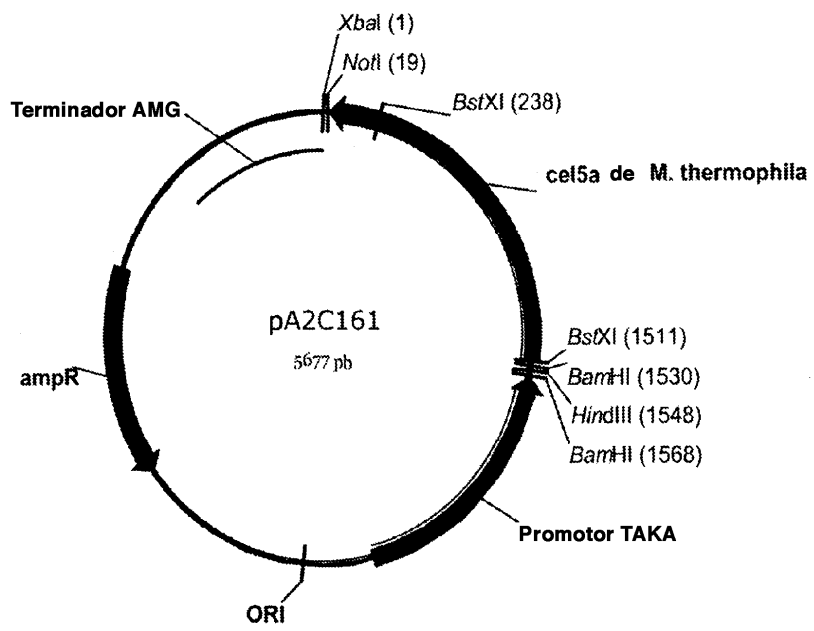


Fig. 3

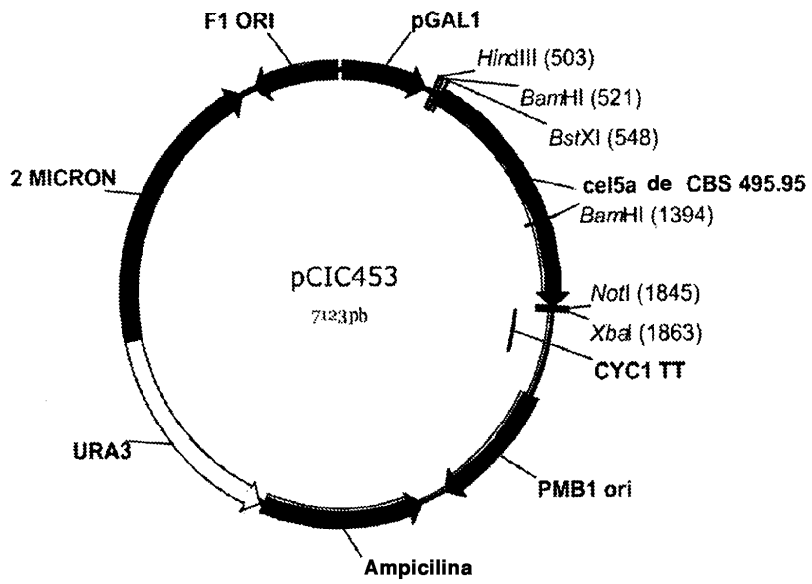


Fig. 4

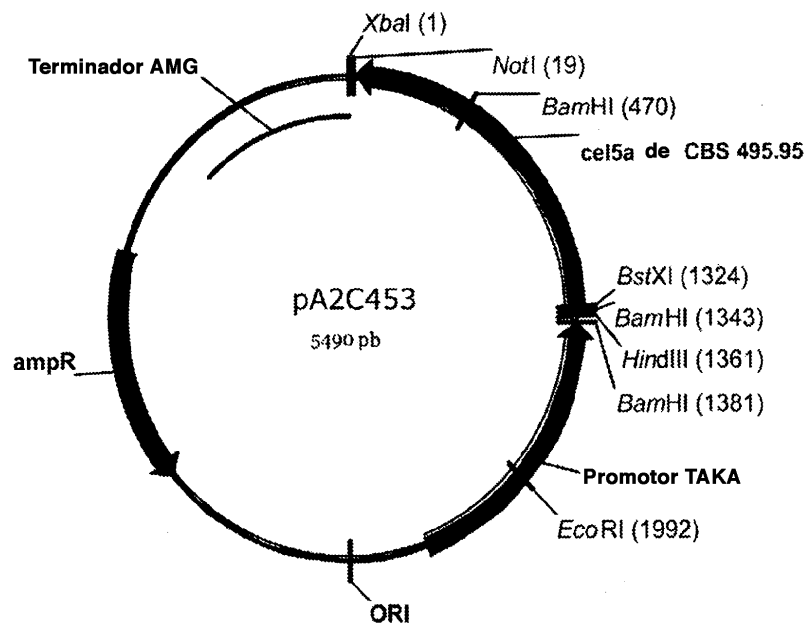


Fig. 5

GGATCCACTTAGTAACGGCCCGCAGCTGTGCTGGAAAGCATGAAGTCTCTCTTCCCTGTCTCACTTTAGCGACCGGTCCGGCTCAGCTCCG
CAGTATTCCTGTCCGACGCTTGGGGCAAIUGCGCGGATGGCTTAGCGGAGCACCCTCTGTGATGCAGCGCGCGGCTGTGTGA
P V F S V A V W G Q C G G I G F S G S T V C D A G A G C C V
AGCTCAACGACTATTACTCTCAAIUGCAACCGCGGCTCCCACTGCTACATCCGCGGCGAAGTAGCAAGCCACCGTCCGGCACATT
K L N D Y S Q C Q P G A P T A T S A A P S S N A P S G T
CGACGGCTCCGGCCCTCCACGCTTIGCTTGGGACCGCACCGCTTCCAGTCTCTCGGTGTCAACGAAACCGGGGGAGT
S T A S A P S S L C S G S R T P F Q F F G V N E S G A E
TCGGCAACCTGAACATCCCGGTGTCTGGGACCGACTACACTGGCGTGGCCATCCAGCATTCGACTTCTTCAITGGGCAAGGGAA
F G N L N I P G V L G T D Y T W P S P S I D F F M G K G
TGAATACCTCCGTAATCCGTTCCATGGAGCGTCTGTCCCGCTGCCACTGGCATCACAGGACCTTCGACCCAGACCGTACTTGG
M N T F R I P F L M E R L V P P A T G I T G P L D Q T Y L
GGGCTGCAGACGATGTCAACTACATCACCGGCAAGGGGCTTGTCTCTCATTTGACCCGCAACTTTATGATCTACAATGGCC
G G L Q T I V N Y I T G K G F A L I D P H N F M I Y N G
AGACGATCCAGTACCGACFTCCAGAGTTCGGCAGAACCTCCAGGAGTGTAAATCGAACACTCACGTCATCTTCGATG
Q T I S S T S D F Q K F W Q N L A G V F K S N S H V I F D
TTATGAACGAGCCTCAGGATATTCGCGCCAGACCGTGTCCCAACTGAACCAAGCGCTGTCAATGGCATCCGCTGGAGCGGTGGGA
V M N E P H D I P A Q T V F Q L N Q A A V N G I R A S G A
CGTCGAGCTCATCTGTGAGGGCACAAGCTGGACTGGAGCTGGACCTGGCAACAGCGATGCATTCGGTGGCA
T S Q L I L V E G T S W T G A W T T S G N S D A F G A
TTAAGGATCCCAACAACAGCTCCGATCCAGATGCATCAGTACATGGATGGCTCTGGCACTTCGACATTCGACAGACTCGCTGTCTC
I K D P N N V A I Q M H Q Y L D S D G S G T S Q T C V S
CCACCATCGGTCCGAGCGGTTGAGGCTGGACTCAATGGTTGAGCAGACAACTCAAGGGTTCCTGGGGCGAGATCCGGCCCG
P T I G A E R L Q A A T Q W L K Q N N L K G F L G E I G A
GCTCIAACTCCGCTTGCATCAGCGCTGTGCAGGGTGGTGTTCGATGCAGCAATCTGGTGTGGCTGGCGCTCTCTGGTGGG
G S N S A C I S A V Q G A L C S M Q S G V W L G A L W W
CTGGGGCCCGTGTGGGGCGACTACTACCACTCCATCGAGGCGCCCTCTGGCCCGGGGTGTCCGGCGATCCTCCCGCAGGCCCTGC
A A G P W W G D Y Y Q S I E P P S G P A V S A I L P Q A L
TGCCGTTCCGGTAA
L P F A

Fig. 6

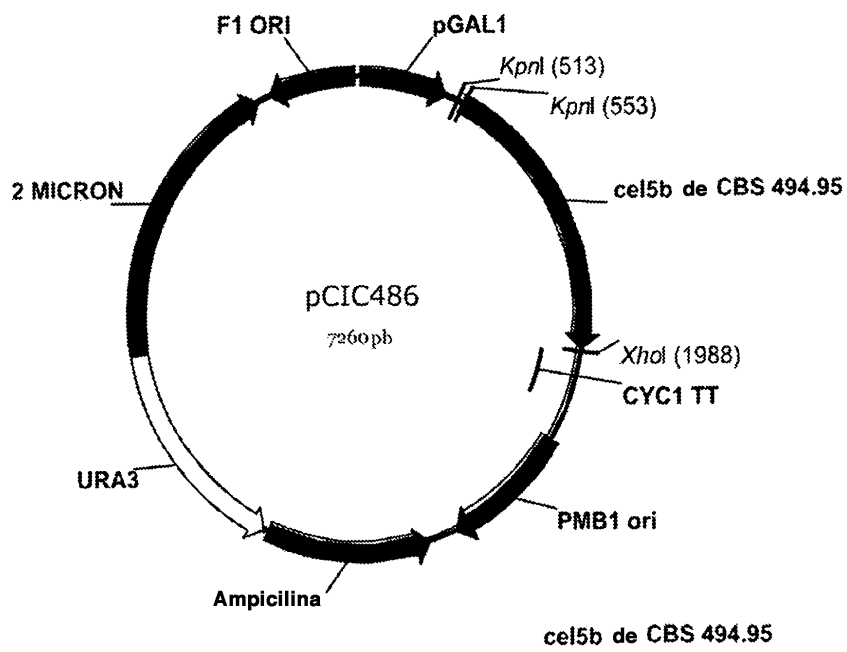


Fig. 7

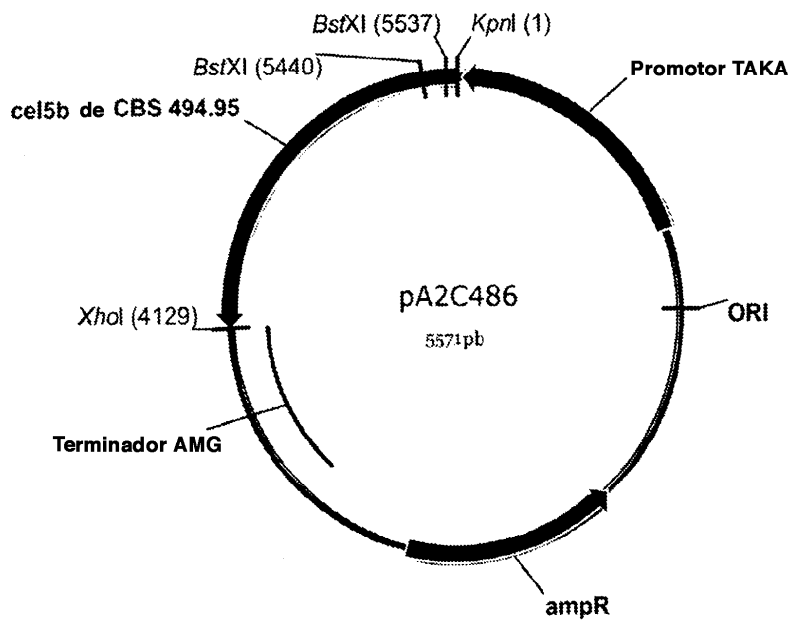


Fig. 8

GGAAAGCCTCAGTATGGTGAATTTGGCTTGGCAACTCTGGCGCAATCTTGAGCGCTTTCGGCCCAATCGCGCTTCTATCTA
M V K F A L V A T V G A I L S A S A A N A A S I Y
CCAGCAATGTGGAGGCATTTGGTCCACTGTTTCGACCGCGGTCTGGCTTATCTCAATGGTACTACTTTCA
Q Q C G I G W S G S T V C D A G L A C V I L N A Y Y F Q
GTCTTGGCCCGCGGACAGCAACGACGGCTCGGCGCACCGGCTCAACCTCTCAGTCAACGGTCTACTACGGG
C L T P A A G Q T T G S G A P A S T S H S T V T G
GAGCTCAGCTCAACAACCGGACGACCGGCAAACTACCTCCGTGACCAACGACCTACCCGCCATCTCTGTGTC
S S H S T T G T A T K T T P S T T L P A I S V S
TGGTCCGCTCTGGCTCCAGGACGAAGTTCAAGTTCTGGTGAATGAAGCGGCGCGAATTCGGGAACACTGCTTGGCC
G R V C S G S R T K F K F G V N E S G A E F G N T A W P
AGGCAGCTGGGAAGACTATACATGGCCTTCGCTAGAGCTGGACTCTTCATGGGGCTGGATTCATATACATTCGGTATCAC
G Q L G K D Y T W P S P S V D Y F M G A G F N T F R I T
CTTCTTGTGGAGCGTATGAGCCCTCCGGCTACCGGACTACTGGCCCAATCAACGAGCTACTGTGCGGCCCTCACCCATTGT
F L M E R M S P A T G L T G P F N Q T Y L S G L T I V
CGACTACATCACGAAAGGAGGATACGCTCTTATTGACCCCAACTTCAATCGTTTACAACAACCGCATATTCAGCAGCACATC
D Y I T N K G Y A L I D P H N F M R Y N N G I I S S T S
TGACTTCGGACTTGGTGGCAATTTGGCCACTGTATTCAATCCAGAAAGCCCAITCTTCGACATCCAGAACGACCGGTACGG
D F A T W S N L A T V F K S T K N A I F D I Q N E P Y G
AATCGATGGCAGACCGTATACGAACGTGAATCAAGCTGCCATCAATTCGATCCGCGCGGTGCTACGTACAGTTGATTCGGT
I D A Q T V Y E L N Q A A I N S I R A A G A T S Q L I L V
TGAAGAACGTACACTGGAGCTTGGACGTGGTCTCGTCCGGAACCGGAGCTGCTTTCGCGCGCGTTACGGATCCTTACAACAA
E G T S Y T G A W T W V S S G N G A A F A A V T D P Y N N
CAGGCCAATGAAATGCACCAATACCTCGACAGCGGTTCTGGGACAAACGAAAGACTGTCTCTCCACCCATGGGTCCGCAACG
T A I E M H Q Y L D S D G S G T N E D C V S T I G S Q R
TCTCCAGCTGCCACTGGCTGCACAAACAGGACTCAAGGATTCCTCGGAGACCGGGTGTGGTCCGAATCCCAATGCCAGTGCAT
L Q A A T A W L Q Q T G L K G F L G E T G A G S N S Q C I
CGAGCCGGTTCGATGAACCTTGTATATGCAACAGCGGCTCCTGGATCGGTGCACTCTGGTGGCTGGGGTCCCTGGTG
D A V F D E L C Y M Q Q Q G G S W I G A L W A A G P W W
GGCACGTACATTTACTCGATTGAACCTCCGAGCGGTGCCGTATCCCAAGTCTCTCCAGGGTCTGGCTCCCATTCCTCTAG
G T Y I Y S I E P S G A A I P E V L P Q G L A P F E L

Fig. 9

ATGAATACCCCTACTCTGGCAACAGCGCTGGCCCTCTGCTGGCCCTCAAGCTTCTCCAAAGCGAGCCTTCTTTTGTGTATG
 M K Y P L L A T S A A L A L A G P O G F S K R A S S F V
 TGTAGACCCATAAGAAATAGAGCTGGAATCTGAACTTTTGGTGTGGTACCGAGCTCTGGGGCGGAGTTTGGCAATCAGAACATTT
 W F G T S E S G A E F G N Q N I
 CCGGTGTTTGGGACGGACTACATTTGGCCCGACATTCAGCCATTCAAACCGCTACGCAAGCGGTATGAACATCTTTCCGGTGGCAITTT
 P G V L G T D Y I W P D T S A I Q T L R N A G M N I F R V A F
 TTGATGGAGCGATTTCCACCACATTTGACCTCAACTCCGGATTCGAGTATCTGCAAGACCTGCAAGAGTGTACCTTATCTCCGTCACAA
 L M E R L V P T T L T S T P D S T Y L Q D L K S V
 TGCACCCGGTCTAAACGATCTCCACAGACTGACTACATCACGTCGACCGCCGATATGCTATTTGGACCCCCCTAACTTTGGACGATAGT
 L D Y I T S T G A Y A I V D P H N F G R Y
 AAGTTGAACCTTTCATACCAATATTGAAAATCTCTCAGCGGTGTATTAGCTATGGCAATATAATCAACTCCACAAGTACATTTGTGCTGCA
 Y G N I I N S T S D F A A
 TTTTGGACCCCGTCCGCAACGAGTTTGGCATCGAATGACAAGTTCATCTTTGACACCGAGTGCATCTCAATCAAGCGGCCAATTAATGCCATCCGAGCTGC
 F W T T V A K Q F A S N D K V I F D T
 CTAACGTGACATCTGCACACAGCAANTCAACACAGAGGACCAACCGCTTGTACTGAACCTCAATCAAGCGGCCAATTAATGCCATCCGAGCTGC
 N N E F N T E D Q T L V L N L N Q A A I N A I R A A
 CGGAGCCACCTCTCAGTATATCTTTGTGGAGGAAATTCGTGGAGTGTGCTGGACCTGAGCTGAGTCAATCCCAATCTCGTACGCTTGAC
 G A T S Q Y I F V E G N S W S G A W T W T S V N T N L V S L T
 GGACCCCAACAAAGATAGTCTACGAGATGCACCGATATCTTGGACTCGGATGGATCTGGCACATCCGACACATGTGTACGCTCGACCATCGG
 D P N N K I V Y E M H Q Y L D S D G S G T S D T C V S S T I G
 CCAGGCGGTSTACAAATCGGACCGAGTGGCTGAAAAGCAAGCAAACTTGGATTTTGGGGGAATTTGGGGGGTGTAACTCGGCTCG
 Q E R V Q S A T E W L K S N G K L G F L G E F A G A N S V C
 TCAGAGCCCTGACTGGAATGCTGGAATATGCAAGCCAAATAGTACGCTGGCTCGGCGCATCTGTGGCAGCAGGACCATGGTGGGG
 Q S A V T G M L D Y M Q A N S D V W L G A S W A A G P W W G
 AACCTATATATTCGATTCAGCCCGCATCGGGACTGCTTATTTACTATCTCAACATCTTGTCTCCCTACTTCCCTCCAGCTCGGGAG
 T Y I Y S I E P S G T A Y S Y Y L N I L S A Y F P S S S G S

Fig. 10A

TTCCACAACCAACTACCTCCACTACCCACCCGGCTCTACATCGACAAGCAGCTACAGTATCCAGCCACAATAACAAGCACCACCTACACAGGGC
S T T T T S T T R S T S T S T T V S T T K S T S T T S A
CACGAAATCGACAAGCACCGACAAGCACCCAGCAGCGGGTCTACTGTGCAACAGCATCTCACTGGGGCACAGTGTGGCGGCATTGG
T K S T S T T S T T S T G S T A T A T A S H W A Q C G G I G
CTGGACGGGGCGACGACGTGTGGCAGCCGTAACCTGCCAGTCCAGATCGGTATTTTCGACGTGTCGTATATGCAGAAAGTATCAGAAA
W T G A T T C A S P Y T C Q V Q N A Y Y S Q C L

Fig. 10B

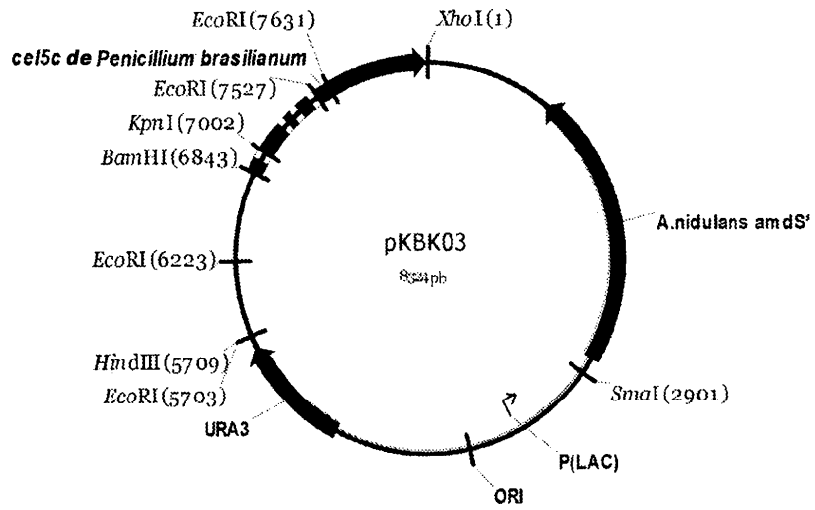


Fig. 11

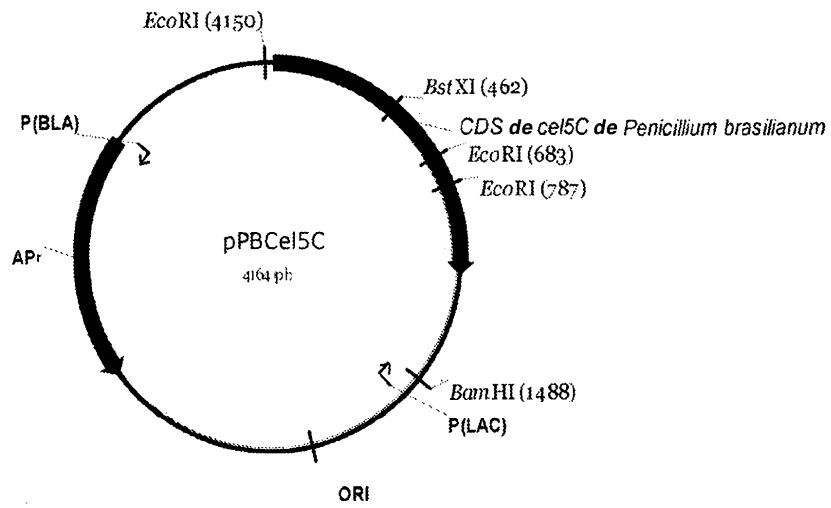


Fig. 12

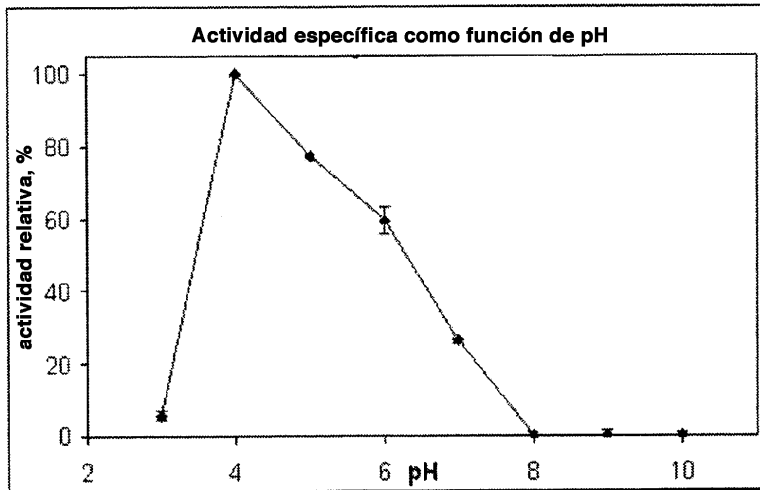


Fig.13

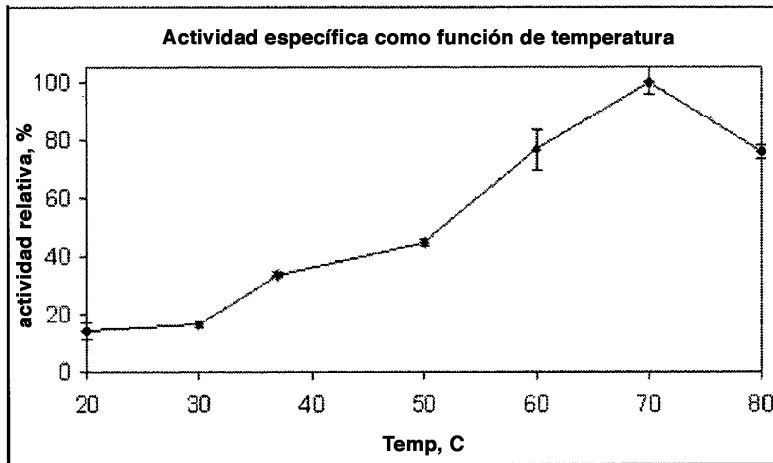


Fig.14

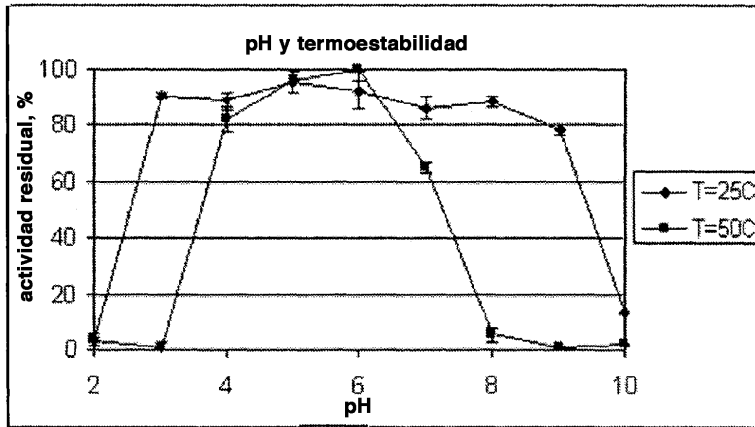


Fig.15

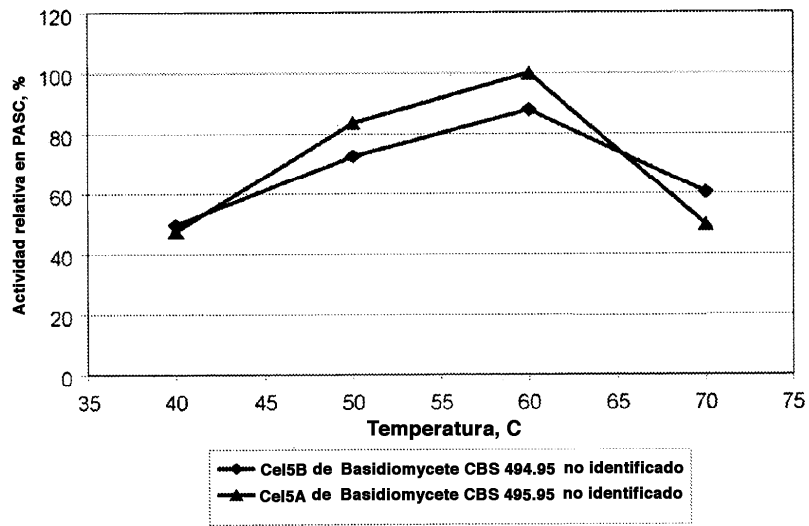


Fig.16

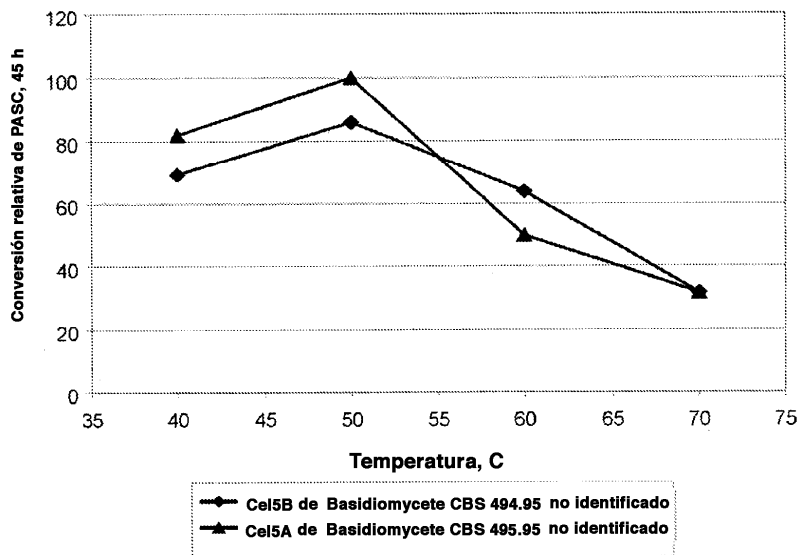


Fig.17

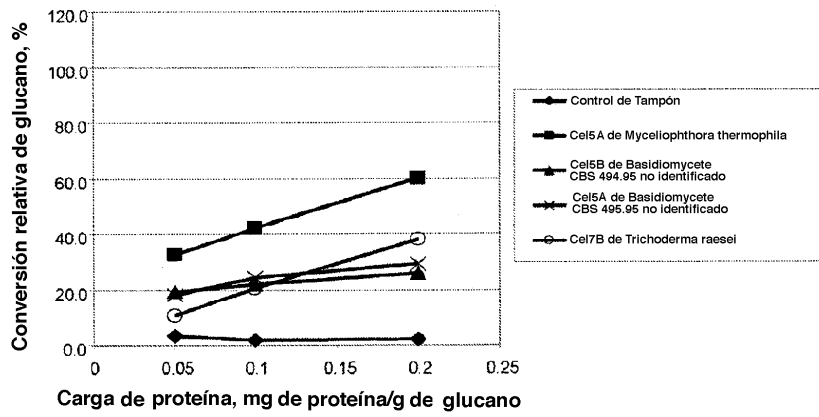


Fig.18

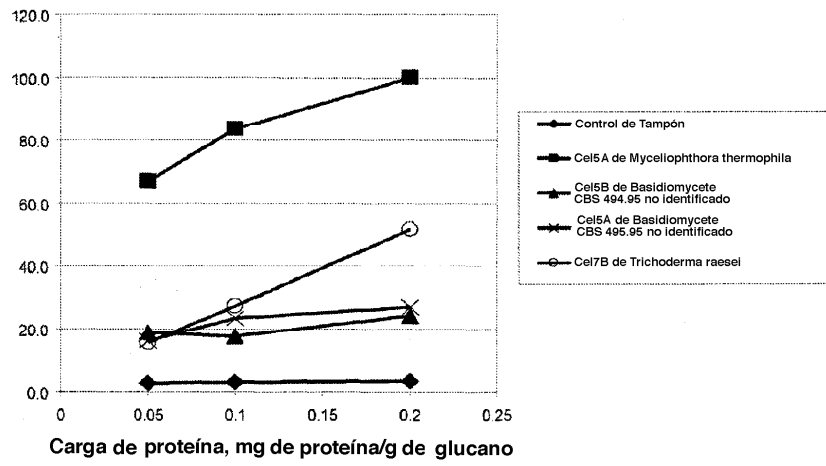


Fig.19