

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 430 825**

51 Int. Cl.:

C12N 9/42 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **26.03.2004 E 04785732 (1)**

97 Fecha y número de publicación de la concesión europea: **10.07.2013 EP 1627050**

54 Título: **CBH1.1 de Humicola grisea variante**

30 Prioridad:

01.04.2003 US 459734 P

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

21.11.2013

73 Titular/es:

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

72 Inventor/es:

**LARENAS, EDMUND A.;
GOEDEGEBUUR, FRITS;
GUALFETTI, PETER y
MITCHINSON, COLIN**

74 Agente/Representante:

RIZZO, Sergio

ES 2 430 825 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).

CBH1.1 DE HUMICOLA GRISEA VARIANTE

Descripción

CAMPO DE LA INVENCION

5 [0001] La presente invención se refiere a polipéptidos que presentan actividad de celobiohidrolasa I (también denominados CBH I o CBH1) y polinucleótidos que tienen una secuencia de nucleótidos que codifica para los polipéptidos. La invención también hace referencia a construcciones de ácido nucleico, vectores y células huésped que contienen construcciones de ácido nucleico, así como métodos para producir y utilizar los polipéptidos.

ANTECEDENTES DE LA INVENCION

10 [0002] La celulosa es una materia prima industrial importante y una fuente de energía renovable. La estructura física y la morfología de celulosa nativa son complejas y ha sido difícil determinar los detalles precisos de su estructura de forma experimental. No obstante, la composición química de la celulosa es simple, compuesta de residuos de D-glucosa enlazados por uniones de beta-1,4-glicosídicas para formar cadenas lineales.

15 [0003] Con el fin de ser eficaz, la digestión de celulosa requiere que muchos tipos de enzimas actúen de modo cooperativo. Al menos tres categorías de enzimas son necesarias para convertir celulosa en glucosa: endo(1,4)-beta-D-gluconasas (EC 31.4) que cortan las cadenas de celulosa al azar; celobiohidrolasas (EC 31.91) que dividen unidades de celobiosil de los extremos de la cadena de celulosa y beta-glucosidasas (EC 31.21) que convierten la celobiosa y las celodextrinas solubles en
20 glucosa. Entre estas tres categorías de enzimas implicadas en la biodegradación de celulosa, las celobiohidrolasas son las enzimas clave para la degradación de la celulosa cristalina nativa.

[0004] Las exo-celobiohidrolasas (Celobiohidrolasa 1 o CBH 1) se refieren a las celobiohidrolasas que degradan celulosa hidrolizando la celobiosa del extremo no reductor de las cadenas de polímeros de celulosa.

25 [0005] Es un objetivo de la presente invención lograr polipéptidos mejorados que tienen actividad de celobiohidrolasa I y polinucleótidos que codifican polipéptidos. Los polipéptidos mejorados pueden haber mejorado la actividad específica y/o la estabilidad, en concreto la termoestabilidad.

[0006] Takashima S *et al.*, JOURNAL OF BIOTECHNOLOGY, vol. 50, nº 2, 1 de octubre de 1996, págs.137-147, presenta la clonación, secuenciación y expresión de genes de celulasa de una cepa de
30 *Humicola grisea var. thermoidea*.

[0007] Oliveira Azevedo De Metal, NUCLEIC ACIDS RESEARCH, vol. 18, nº 3, 1 de enero de 1990, pág. 668, divulga la secuencia de un gen de cbh-1 (núm. de acceso del Laboratorio Europeo de Biología Molecular (EBML) X17258) de una cepa de *Humicola grisea var. thermoidea*.

[0008] Aunque las composiciones de celulasa se han descrito previamente, todavía se necesitan
35 nuevas y mejoradas composiciones de celulasas para su uso en detergentes domésticos, composiciones para el lavado a la piedra o detergentes de lavandería, etc. Son de gran interés las celulasas que muestran un mejor rendimiento.

REFERENCIAS

[0009]

- Altschul, S. F., *et al. J. Mol. Biol.* 215:403-410, 1990.
- Altschul, S. F., *et al., Nucleic Acids Res.* 25:3389-3402, 1997.
- 5 Aro, N., *et al., J. Biol. Chem.*, 10.1074/ M003624200, 13 de abril, 2001.
- Aubert, *et al.*, Ed., pág. 11 y ss., Academic Press, 1988.
- Ausubel G. M., *et al. CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, John Wiley & Sons, Nueva York, N.Y., 1993.
- Baker *et al., Appl. Biochem. and Biotechnol.* 45/46:245 - 256, 1994.
- 10 Bhikhabhai, R. *et al., J. Appl. Biochem.* 6:336, 1984.
- Boel *et al. EMBO J* 3:1581-1585 1984.
- Brumbauer, A. *et al., Bioseparation* 7:287-295, 1999.
- Deutscher, M.P., *Methods Enzymol.* 182:779-80, 1990.
- Ellouz, S. *et al., J. Chromatography* 396:307, 1987.
- 15 Filho, *et al. Can. J. Microbiol.* 42:1-5, 1996.
- Fliess, A., *et al., Eur. J. Appl. Microbiol. Biotechnol.* 17:314, 1983.
- Goedegebuur *et al., Curr. Genet.* 41 :89 - 98, 2002.
- Goyal, A. *et al. Bioresource Technol.* 36:37, 1991.
- Hazell, B. W. *et al., Lett. Appl. Microbiol.* 30:282-286, 2000.
- 20 Herr *et al., Appl. Microbiol. Biotechnol.* 5:29-36, 1978.
- Hu *et al., Mol Cell Biol.* 11:5792-9, 1991.
- Jeeves *et al., Biotechnol. Genet. Eng. Rev.* 9:327 -369, 1991.
- Kawaguchi, T *et al., Gene* 173(2):287-8, 1996.
- Kelley *et al. EMBO J.* 4 :475-479, 1985.
- 25 Knowles, J. *et al., TIBTECH* 5, 255-261, 1987.
- Krishna, S. *et al., Bioresource Tech.* 77:193-196, 2001.
- Kuhls K. *et al., Proc. Natl. Acad. Sci. EE. UU.* 93(15): 7755 - 7760, 1996.
- Kumar, A., *et al., Textile Chemist and Colorist* 29:37-42, 1997.
- Medve, J. *et al., J. Chromatography A* 808:153, 1998.
- 30 Mohagheghi, A. *et al., Int. J. Syst. Bacteriol.* 36:435 - 443, 1986.
- Nieves *et al., Appl. Biochem. and Biotechnol.* 51/52 211 - 223, 1995.
- Nunberg *et al. Mol. Cell Biol.* 4:2306-2315 1984.
- Ohmiya *et al., Biotechnol. Gen. Engineer. Rev.* 14:365-414, 1997.
- Okada, M. *et al., Appl. Environ. Microbiol.*, 64:555-563, 1988.
- 35 Ooi *et al., Nucleic Acids Res.* 18:5884, 1990.
- Penttila *et al., Gene* 45:253-263, 1986.

- Penttila *et al.*, *Gene* 61: 155 -164, 1987.
- Penttila *et al.*, *Gene* 63: 103-112, 1988.
- Pere, J., *et al.*, *In Proc. Tappi Pulping Conf.*, Nashville, TN, 27-31, pp. 693-696, 1996.
- Saarilahti *et al.*, *Gene* 90:9-14, 1990.
- 5 Sakamoto *et al.*, *Curr. Genet.* 27:435-439, 1995.
- Saloheimo M, *et al.*, *Gene* 63:11-22, 1988.
- Saloheimo, A. *et al.*, *Molecular Microbiology*, 13:219-228, 1994.
- Saloheimo, M. *et al.*, *Eur. J. Biochem.*, 249:584-591, 1997.
- Sambrook *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL* (Segunda Edición), Cold
 10 Spring Harbor Press, Plainview, N.Y., 1989.
- Schulein, *Methods Enzymol.*, 160, 25, páginas 234 y ss, 1988.
- Scopes, *Methods Enzymol.* 90 Pt E:479-90, 1982.
- Shoemaker *et al.*, *Biochem. Biophys. Acta.* 523 :133-146 1978.
- Shoemaker, S. *et al.*, *Bio/Technology*, 1:691-696, 1983.
- 15 Srisodsuk, M. *et al. J. Biol. Chem.* 268(28): 20756 - 20761, 1993.
- Strathern *et al.*, eds. (1981) *The Molecular Biology of the Yeast Saccharomyces*.
- Suumakki, A. *et al.*, *Cellulose* 7:189-209, 2000.
- Teeri, T. *et al.*, *Gene*, 51:43-52, 1987.
- Tilbeurgh, H. *et al.*, *FEBS Lett.* 16:215, 1984.
- 20 Tomaz, C. y Queiroz, J., *J. Chromatography A* 865:123-128, 1999.
- Tomme, P. *et al.*, *Eur. J. Biochem.* 170:575-581, 1988.
- Van Tilbeurgh, H. *et al.*, *FEBS Lett.* 204:223-227, 1986.
- Ward, M. *et al.*, *Appl. Microbiol. Biotechnol.* 39:738 - 743, 1993.
- Wood, *Biochem. Soc. Trans.*, 13, pp. 407-410, 1985.
- 25 Wood *et al.*, *METHODS IN ENZYMOLOGY*, 160, 25, p. 87 y ss., Academic Press, Nueva York, 1988.

BREVE RESUMEN

30 **[0010]** En un primer aspecto, la presente divulgación se refiere a un polipéptido que muestra actividad de celobiohidrolasa I, seleccionada del grupo que consta de:

- a) variante de CBH1 de *H. grisea* derivada de CBS 225.36.
- b) variante de CBH1.1 de *H. grisea* que presenta la secuencia mostrada en la Figura 3.
- c) variante de CBH1.1 de *H. grisea* que presenta la secuencia mostrada en la Figura 4.
- d) variante de CBH1 de *Hypocrea jecorina* como se describe aquí; y
- 35 e) CBH1 de *Scytalidium thermophilum* derivada de CBS 671.88

[0011] En el segundo aspecto, la presente invención hace referencia a un polinucleótido que codifica una variante de CBH1.1 de *H. grisea*, una variante de CBH1 de *H. jecorina* o una CBH1 de *S.*

thermophilium. En un modo de realización, el polinucleótido que codifica una variante de CBH1.1 de *H. grisea* se deriva de CBS 225.63. En otro modo de realización, el polinucleótido codifica una variante de CBH1.1 de *H. grisea* mostrada en la Figura 3. En otro modo de realización, el polinucleótido codifica una variante de CBH1.1 de *H. grisea* presentada en la Figura 4.

- 5 **[0012]** En un modo de realización general, un polinucleótido que codifica una variante de CBH1.1 de *H. grisea*, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilium* tiene al menos 90%, preferentemente 95%, 98% o más de identidad de secuencia con las secuencias que codifican CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una variante de CBH1.1 de *H. grisea* presentadas en el presente documento utilizando un programa de alineamiento de secuencias.
- 10 **[0013]** En un tercer aspecto, la presente invención se refiere a una construcción de ácido nucleico que comprende la secuencia de nucleótidos, que codifica para la variante de CBH1.1 inventiva, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilium*, unida de modo funcional a una o más secuencias de control que dirigen la producción de la variante de CBH1.1, variante de CBH1 de *H. jecorina* o CBH1 de *S. thermophilium* en un huésped adecuado.
- 15 **[0014]** En una cuarta realización, la presente invención hace referencia a un vector de expresión recombinante que comprende la construcción de ácido nucleico de la invención.
- [0015]** En un quinto aspecto, la presente divulgación se refiere a una célula huésped recombinante que comprende la construcción de ácido nucleico de la invención.
- [0016]** En un sexto aspecto, la presente invención hace referencia a un método para la producción de
- 20 CBH1, el método comprende:
- a) la transformación de una célula huésped con un ácido nucleico que codifica una variante de CBH1.1 de *H. grisea*, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilium* descrita en este documento.
 - b) el cultivo de la célula huésped en condiciones para producir el polipéptido; y
 - 25 c) la recuperación del polipéptido.

BREVE DESCRIPCIÓN DE LOS DIBUJOS

- [0017]** La Figura 1 es la secuencia de ADN genómico para CBH1.1 de *H. grisea* (SEQ ID N°:1). El intrón putativo aparece subrayado y en negrita.
- [0018]** La Figura 2 es la secuencia de ADNc para CBH1 de *H. grisea* (SEQ ID N°:2). Se ha eliminado
- 30 la secuencia de intrón putativo (nucleótidos 413-472 en la Figura 1).
- [0019]** La Figura 3 es la secuencia señal y la secuencia de aminoácidos madura para CBH1.1 de *H. grisea* (SEQ ID N°:3). La secuencia señal aparece subrayada y en negrita.
- [0020]** La Figura 4 es la secuencia de aminoácidos madura para CBH1.1 de *H. grisea* (SEQ ID N°:4).
- [0021]** La Figura 5 muestra un alineamiento de dos secuencias maduras públicas y una CBH1.1 de
- 35 *Humicola grisea* variante. Las dos secuencias públicas son X17258 (SEQ ID N°:5) y D63515 (SEQ ID N°:6).
- [0022]** La Figura 6 representa el plásmido pRAX1. Este vector se basa en el plásmido pGAPT2 a excepción de que se insertó un fragmento HindIII de 5259bp de la secuencia AMA1 del fragmento de ADN genómico de *Aspergillus nidulans* (*Molecular Microbiology* 1996 19:565-574). Las bases 1 a

1134 contienen el promotor génico de glucoamilasa de *Aspergillus niger*. Las bases 3098 a 3356 y 4950 a 4971 contienen el terminador de glucoamilasa de *Aspergillus niger*. El gen *pyrG* de *Aspergillus nidulans* se insertó de 3357 a 4949 como marcador para la transformación fúngica. Existe un sitio de clonación múltiple (MCS) en el que se pueden insertar genes.

5 **[0023]** La Figura 7 es la estructura principal del vector pRAXdes2. Este vector se basa en el vector de plásmido pRAX1. Se ha insertado un casete Gateway en el vector pRAX1 (indicado por la flecha en el interior del plásmido circular). Este casete contiene las secuencias de recombinación attR1 y attR2 y los marcadores de selección *catH* y *ccdB*. El vector se ha hecho según el manual proporcionado en Gateway™ Cloning Technology: versión 1 páginas 34-38 y sólo se puede replicar en *E. coli* DB3.1 de
 10 Invitrogen; en otros huéspedes de *E. coli*, el gen de *ccdB* es letal. En primer lugar, se realiza un fragmento de PCR con cebadores que contienen las secuencias de recombinación attB1/2. Este fragmento se recombina con pDONR201 (comercialmente disponible de Invitrogen); este vector contiene las secuencias de recombinación attP1/2 con *catH* y *ccdB* entre los sitios de recombinación. Las enzimas de BP clonasa de Invitrogen se utilizan para recombinar el fragmento de PCR en este
 15 vector llamado ENTRY, los clones con el fragmento de PCR pueden seleccionarse a 50 µg/ml de kanamicina porque los clones que expresan *ccdB* no sobreviven. Ahora las secuencias att están alteradas y se denominan attL1 y attL2. El segundo paso es recombinar este clon con el vector pRAXdes2 (que contiene attR1 y attR2 *catH* y *ccdB* entre los sitios de recombinación). Las enzimas de LR clonasa de Invitrogen se utilizan para recombinar el inserto del vector ENTRY en el vector de
 20 destino. Sólo los vectores pRAXCBH1 se seleccionan usando 100 µg/ml de ampicilina porque *ccdB* es letal y el vector ENTRY es sensible a la ampicilina. Con este método, se prepara ahora el vector de expresión y se puede utilizar para transformar *A. niger*. La expresión de CBH1.1 de *H. grisea* está controlada por el terminador y el promotor de glucoamilasa de *Aspergillus*. El gen *pyrG* del marcador de transformación y la secuencia AMA1 son de *Aspergillus nidulans*.

25 **[0024]** La Figura 8 ofrece un ejemplo del vector pRAXdes2cbh1 que se utilizó para la expresión de los ácidos nucleicos que codifican las variantes de CBH1.1 en *Aspergillus*. Un ácido nucleico que codifica una variante de enzima de CBH1.1 se clonó en el vector mediante recombinación homóloga de las secuencias att.

30 **[0025]** La Figura 9 ilustra la actividad de varias celulasas CBH1 en un ensayo de conversión de celulosa en un día, a 38°C y 15 mg de enzima total/g de celulosa usando el forraje de maíz pretratado (PCS) como un sustrato. Este ensayo combina a aproximadamente una ratio de la masa de 50:50 de la muestra de CBH1.1 para ser probada con sobrenadante de cultivos de una cepa de *T. reesei* que se ha eliminado para CBH1. A 38°C, CBH1 de *H. grisea* des irrelevante en este ensayo.

35 **[0026]** Leyenda: delCBH1 es una cepa a la que se le ha eliminado su CBH1 endógena; es deficiente en CBH1. rCBH1 es una enzima purificada a partir de una cepa de *A. niger* a la que se le ha insertado un gen de CBH1 de *T. reesei*. *A. niger* es una CBHB purificada de una cepa de *A. niger* que sobreexpresa su CBHB de celulasa endógena. Hschweinitzii/An es una CBH1 purificada a partir de una cepa de *A. niger* que expresa un gen de CBH1 heterólogo insertado a partir de *Hypocrea schweinitzii*. Tpseudokoni/An es una CBH1 purificada de una cepa de *A. niger* que expresa un gen de
 40 CBH1 heterólogo insertado a partir de *Trichoderma pseudokoningii*. Hgrisea/An-1 es una CBH1.1 purificada a partir de un primer clon de *A. niger* que expresa un gen de CBH1.1 heterólogo insertado

a partir de Hgrisea. Hgrisea/AN-2 es una CBH1.1 purificada a partir de un segundo clon de *A. niger* que expresa un gen de CBH1.1 heterólogo insertado a partir de *H. grisea*. Hgrisea/An-1 y Hgrisea/An-2 son dos clones de la misma transformación de *A. niger* con el gen de CBH1.1 variante de *H. grisea*.

[0027] La figura 10 muestra la actividad de celulasas de CBHI en un ensayo de conversión de celulosa en un día, a 65°C, y 15 mg de enzima total/g de celulosa usando forraje de maíz pretratado como sustrato. Este ensayo combina a aproximadamente una ratio de masa de 50:50, la muestra de CBHI a analizar con sobrenadante del cultivo de una cepa de *T. reesei* que se ha eliminado para CBHI. A diferencia del resultado a 38°C, ambas muestras de CBH1.1 de *H. grisea* son claramente mejores que otras muestras. Leyenda: la misma que para la Figura 9.

[0028] La Figura 11 ilustra la actividad de celulasas de CBH1/*T. reesei* y CBH1.1 de *H. grisea* en un ensayo de conversión de celulosa en un día, a 38°C y 15 mg de enzima total/g de celulosa usando forraje de maíz pretratado (PCS) como un sustrato. Este ensayo combina a aproximadamente un ratio de masa de 50:50, la muestra de CBH1.1 a analizar con sobrenadante del cultivo de una cepa de *T. reesei* que ha sido eliminada para CBH1. A 38°C, la CBH1.1 de *H. grisea* es irrelevante en este ensayo. CBH1 nativa (nCBH1) se purificó de una celulasa completa de *T. reesei* empleando métodos conocidos en la técnica (véase Métodos para purificar CBH, a continuación).

[0029] La Figura 12 muestra los resultados de la actividad para las enzimas rCBH1 de *T. reesei* y *H. grisea* sólo en un ensayo de conversión de celulosa en un día, a 65°C y 15 mg de enzima total/g de celulosa utilizando forraje de maíz pretratado como sustrato. Este ensayo combina a aproximadamente una ratio de masa de 50:50, la muestra de CBHI a analizar con sobrenadante del cultivo de una cepa de *T. reesei* que ha sido eliminada para CBHI. A diferencia del resultado a 38°C, la muestra de CBH1.1 de *H. grisea* es mejor que la de *T. reesei*.

[0030] La Figura 13A-C muestra la actividad de celulasa CBHI en un ensayo de conversión de celulosa en celulosa hinchada con ácido fosfórico (PASC) en diferentes períodos. En la tabla A, la temperatura es 38°C con mediciones tomadas durante un periodo de 120 minutos. En la tabla B, las mediciones se tomaron a 65°C de temperatura. En la tabla C, las mediciones se tomaron a la temperatura de 70°C. Se puede ver como la CBH1.1 de *H. grisea* variante libera más celobiosa que la CBH1 de *T. reesei* en cualquier momento.

[0031] La Figura 14A-14C muestra las secuencias de ADN genómico (SEQ ID N°:7), ADNc (SEQ ID N°:8) y aminoácidos (SEQ ID N°:9) de la CBH1 de *Scytalidium thermophilum*. La secuencia de aminoácidos incluye la secuencia señal.

[0032] La Figura 15 muestra el alineamiento de las formas maduras (es decir, sin una secuencia señal) de CBH1.1 de *H. grisea*, CBH1 de *H. jecorina* (SEQ ID N°:10) y CBH1 de *Scytalidium thermophilum* (SEQ ID N°:11). También se muestra la secuencia consenso (SEQ ID N°:12). Están subrayados y en negrita seis residuos en el dominio catalítico de CBH1 de *H. jecorina* que indican que los sitios pueden ser importantes para la estabilidad mejorada.

[0033] La Figura 16A y B muestran los perfiles de termoestabilidad para CBH1.1 de *H. grisea* y CBH1 de *S. thermophilum*.

DESCRIPCIÓN DETALLADA

[0034] La invención se describirá con detalle a modo de referencia utilizando sólo las siguientes definiciones y ejemplos.

[0035] A no ser que se defina lo contrario, todos los términos técnicos y científicos utilizados en este documento tienen el mismo significado que el que tendrían para cualquier experto del campo de la presente invención. Singleton, *et al.*, DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 2ª ed., John Wiley y Sons, Nueva York (1994), y Hale y Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, NY (1991) ofrecen al experto un diccionario general de muchos de los términos que se usan en el presente documento. Aunque se puede utilizar cualquier método y material similar o equivalente a los descritos en este documento en la práctica o ensayo de la presente invención, se describen los materiales y métodos preferidos. Los intervalos numéricos incluyen los números que definen el intervalo. A no ser que se indique lo contrario, los ácidos nucleicos se escriben de izquierda a derecha en orientación de 5' a 3'; las secuencias de aminoácidos se escriben de izquierda a derecha en la orientación de amino a carboxilo, respectivamente. En concreto, los profesionales se dirigen a Sambrook *et al.*, 1989, y Ausubel FM *et al.*, 1993, para las definiciones y términos de la técnica. Debe entenderse que la divulgación no se limita a la metodología, protocolos y reactivos concretos descritos, ya que pueden variar.

[0036] Los títulos descritos en el presente documento no son limitaciones de los diversos aspectos o modos de realización que se pueden tener por referencia a la memoria en general. En consecuencia, los términos definidos a continuación se definen con mayor detalle en referencia a la memoria en general.

I. DEFINICIONES

[0037] “Celulosa”, “enzimas celulolíticos” o “enzimas de celulasa” significa exocelobiohidrolasas o exoglucanasas fúngicas o bacterianas, y/o endoglucanasas, y/o β -glucosidasas.

[0038] El término “celulasa” se refiere a una categoría de enzimas capaces de hidrolizar polímeros de celulosa a oligómeros celooligosacáridos más cortos, celobiosa y/o glucosa. Se han obtenido numerosos ejemplos de celulasas, como exoglucanasas, exocelobiohidrolasas, endoglucanasas y glucosidasas a partir de organismo celulolíticos, incluyendo en particular hongos, plantas y bacterias. Las enzimas con estos microbios son mezclas de proteínas con tres tipos de acciones útiles en la conversión de celulosa en glucosa: endoglucanasas (EG), celobiohidrolasas (CBH) y beta-glucosidasa. Estos tres tipos diferentes de enzimas de celulasas actúan de forma sinérgica para convertir la celulosa y sus derivados en glucosa.

[0039] Muchos microbios crean enzimas que hidrolizan celulosa, también el hongo *Trichoderma*, las bacterias del compost *Thermomonospora*, *Bacillus* y *Cellulomonas*; *Streptomyces*; y el hongo *Humicola*, *Aspergillus* y *Fusarium*.

[0040] “Actividad de celobiohidrolasa 1” se refiere a la actividad de una celulosa 1,4-beta-celobiosidasa (también llamada Exo-glucanasa, Exo-celobiohidrolasa, CBH1 o 1,4-beta-celobiohidrolasa), como se define en el número de enzima EC3.2.1.91 que cataliza la hidrólisis de uniones de 1,4-beta-D-glucosídicas en celulosa y celotetraosa, que liberan celobiosa de los extremos

no reductores de las cadenas de celulosa. Para la presente invención, la actividad de CBH1 puede determinarse según el procedimiento descrito en el Ejemplo 4.

[0041] Por el término “célula huésped” se entiende una célula que contiene un vector y soporta la replicación y/o la transcripción o la transcripción y la traducción (expresión) de la construcción de expresión. Las células huésped utilizadas en la presente invención pueden ser células procariotas, como *E.coli*, o células eucariotas, como células de levadura, planta, insecto, anfibio o mamífero. En general, las células huésped son hongos filamentosos.

[0042] El término “recombinante” cuando se utiliza con referencia a, p.ej., una célula, ácido nucleico, proteína o vector, indica que la célula, ácido nucleico, proteína o vector, han sido modificados por la introducción de un ácido nucleico o proteína heterólogos o la alteración de una proteína o ácido nucleico nativo, o que la célula se deriva de una célula modificada de ese modo. En consecuencia, por ejemplo, las células recombinantes expresan genes que no se encuentran en la forma original (no recombinante) de la célula o expresan genes nativos que de otra manera se expresan anormalmente, se expresan poco o no se expresan en absoluto.

[0043] El término “secuencia señal secretora” indica una secuencia de ADN que codifica un polipéptido (un “péptido secretor” o “péptido señal secretor”) que, como componente de un polipéptido más grande, dirige al polipéptido más grande a través de la vía secretora de una célula en la que se sintetiza. El péptido más largo se divide en general para eliminar el péptido secretor durante el transporte a través de la vía secretora para producir el péptido señal secretor y un péptido más pequeño normalmente designado como el polipéptido maduro.

[0044] Como se utiliza en el presente documento, las frases “preparación de celulasa completa” y “composición de celulasa completa” se utilizan de modo intercambiable y se refieren a composiciones de origen natural y no natural. Una composición de “origen natural” es la producida por una fuente de origen natural y que comprende uno o más componentes de tipo celobiohidrolasa, uno o más de tipo endoglucanasa y uno o más de tipo β -glucosidasa donde se encuentra cada uno de estos componentes en la ratio producida por la fuente. Una composición de origen natural es la producida por un organismo no modificado con respecto a las enzimas celulolíticas de modo que la ratio de las enzimas del componente no se altera por la producida por el organismo nativo.

[0045] Una composición de “origen no natural” abarca aquellas composiciones producidas por: (1) la combinación de enzimas celulolíticas del componente en una ratio de origen natural o de origen no natural, es decir, ratio alterada; o (2) la modificación de un organismo para sobreexpresar o subexpresar una o más enzimas celulolíticas; o (3) la modificación de un organismo de modo que al menos se elimine una enzima celulolítica o (4) la modificación de un organismo para expresar una enzima celulolítica del componente heterólogo.

[0046] Como se utiliza en este documento, el término “promotor” hace referencia a una secuencia de ácido nucleico que funciona para dirigir la transcripción de un gen en dirección 3'. En general, el promotor será adecuado para la célula huésped en la que se expresa el gen diana. El promotor, junto con otras secuencias de ácido nucleico reguladoras de la transcripción y de la traducción (también llamadas “secuencias de control”) es necesario para expresar un gen determinado. En general, entre las secuencias reguladoras de la transcripción y de la traducción se incluyen, sin carácter limitativo, secuencias promotoras, sitios de unión a ribosoma, secuencias de inicio y terminación de la

transcripción, secuencias de inicio y terminación de la traducción y secuencias potenciadoras o activadoras. El promotor puede ser el promotor normalmente asociado con el gen corriente abajo o puede ser heterólogo, es decir, de otro gen u otro microorganismo siempre y cuando funcione para dirigir el gen. En un enfoque, el promotor es un promotor inducible. En otro enfoque, el promotor es el promotor de *cbh1* de *T. reesei* el cual está depositado en el GenBank con el número de acceso D86235. En otro aspecto, el promotor es un promotor de *cbh II* o xilanasa de *T. reesei*.

[0047] Los ejemplos incluyen el promotor de los genes de glucoamilasa de *A. niger* o *A. awamori* (Nunberg, J. H. *et al.* (1984) *Mol. Cell. Biol.* 4, 2306-2315; Boel, E- *et al.* (1984) *EMBO J.* 3, 1583-1585), el gen de proteasa carboxilo de *Mucor miehei*, el gen de celobiohidrolasa I de *Trichoderma reesei* (Shoemaker, S. P. *et al.* (1984) Solicitud de patente europea núm. EPO0137280A1), el gen *trpC* de *A. nidulans* (Yelton, M. *et al.* (1984) *Proc. Natl. Acad. Sci. EE.UU.* 81, 1470-1474; Mullaney, E. J. *et al.* (1985) *Mol. Gen. Genet.* 199, 37-45), el gen *alcA* de *A. nidulans* (Lockington, R. A. *et al.* (1986) *Gene* 33, 137-149), el gen *tpiA* de *A. nidulans* (McKnight, G: L. *et al.* (1986) *Cell* 46, 143-147), el gen *amdS* de *A. nidulans* (Hynes, M. J. *et al.* (1983) *Mol. Cell Biol.* 3, 1430-1439), el gen *xln1* de *T. reesei*, el gen *cbh2* de *T. reesei*, el gen *egl1* de *T. reesei*, el gen *egl2* de *T. reesei*, el gen *egl3* de *T. reesei* y promotores de eucariotas en cantidades altas como el promotor temprano SV40 (Barclay, S. L. y E. Meller (1983) *Molecular and Cellular Biology* 3, 2117-2130).

[0048] Un ácido nucleico está “unido de modo funcional” cuando está situado en una relación funcional con otra secuencia de ácidos nucleicos. Por ejemplo, el ADN que codifica una secuencia señal está unido de modo funcional al ADN para un polipéptido si éste se expresa como una preproteína que participa en la secreción del polipéptido; un promotor o potenciador está unido de modo funcional a una secuencia de codificación si afecta a la transcripción de la secuencia; o un sitio de unión a ribosoma está unido de modo funcional a una secuencia de codificación si se coloca con el fin de facilitar la traducción. En general, “unido de modo funcional” significa que las secuencias de ADN que se unen son contiguas y, en el caso de una secuencia señal, contiguas y en el marco de lectura. Sin embargo, los potenciadores no tienen que ser contiguos. La unión se realiza mediante el enlace en los sitios de restricción convenientes. Si no existen dichos sitios, se utilizan adaptadores de oligonucleótidos sintéticos o enlaces de acuerdo con la práctica convencional.

[0049] Como se utiliza en este documento, el término “gen” se refiere al segmento de ADN que participa en la producción de una cadena de polipéptidos, que puede o no incluir las regiones anterior y posterior a la región de codificación, p.ej., secuencias 5' no traducidas (5' UTR) o secuencias “líder” y secuencias 3' UTR o “tráiler”, así como secuencias intercalada (intrones) entre los segmentos de codificación individuales (exones).

[0050] El gen puede codificar péptidos o proteínas terapéuticamente significantes, como factores de crecimiento, citocinas, ligandos, receptores e inhibidores, así como vacunas y anticuerpos. El gen puede codificar péptidos o proteínas industriales comercialmente importantes, como las enzimas, p.ej., celulasas como una variante de CBH1.1 de *H. grisea*, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilum*. El gen de interés puede ser un gen de origen natural, un gen mutado o un gen sintético.

[0051] Los “hongos filamentosos” de la presente invención son microorganismos eucariotas e incluyen todas las formas filamentosas de la subdivisión Eumicotina (véase Alexopoulos, C. J. (1962),

Introductory Mycology, Nueva York: Wiley). Estos hongos se caracterizan por un micelio vegetativo que tiene una pared celular compuesta de quitina, celulosa y otros polisacáridos complejos. Los hongos filamentosos de la presente invención son distintos morfológica, fisiológica y genéticamente de las levaduras. El crecimiento vegetativo mediante hongos filamentosos es por la elongación hifal y catabolismo de carbono es obligadamente aeróbico. Por el contrario, el crecimiento vegetativo mediante levaduras como *S. cerevisiae* es por gemación de un talo unicelular y el catabolismo de carbono puede ser fermentativo. *S. cerevisiae* tiene una fase diploide muy estable y prominente, mientras que los diploides existen sólo brevemente antes de la meiosis en los hongos filamentosos, p.ej., *Aspergilli* y *Neurospora*. *S. cerevisiae* tiene 17 cromosomas, a diferencia de los 8 y 7 para *A. nidulans* y *N. crassa* respectivamente. Entre los ejemplos recientes de diferencias entre *S. cerevisiae* y los hongos filamentosos se incluyen: la incapacidad de *S. cerevisiae* para procesar intrones de *Aspergillus* y *Trichoderma* y la incapacidad para reconocer muchos reguladores de la transcripción de hongos filamentosos (Innis, M. A. *et al.* (1985) *Science*, 228, 21-26).

[0052] El término “heterólogas” cuando se utiliza en referencia a las porciones de un ácido nucleico indica que el ácido nucleico comprende dos o más subsecuencias que no se encuentran normalmente en la misma relación entre sí en la naturaleza. Por ejemplo, normalmente se produce el ácido nucleico de modo recombinante y tiene dos o más secuencias, p.ej., de genes no relacionados colocados para hacer un nuevo ácido nucleico funcional, p.ej., un promotor de una fuente y una región de codificación de otra fuente. De modo similar, una proteína heteróloga hará normalmente referencia a dos o más subsecuencias que no se encuentran normalmente en la misma relación entre sí en la naturaleza (p.ej., una proteína de fusión).

[0053] Los términos “aislado” o “purificado” como se utiliza en el presente documento se refieren a un ácido nucleico o aminoácido que se elimina de al menos un componente con el que se asocia naturalmente.

[0054] En este contexto, el término “polipéptido sustancialmente puro” significa una preparación de polipéptido que contiene como máximo el 10% en peso de otro material de polipéptido con el que se asocia de modo natural (se prefieren porcentajes más bajos de otro material de polipéptido, p.ej., como máximo un 8% en peso, como máximo un 6% en peso, como máximo un 5% en peso, como máximo un 4% como máximo un 3% en peso, como máximo un 2% en peso, como máximo un 1% en peso y como máximo un 1/2% en peso). En consecuencia, se prefiere que el polipéptido sustancialmente puro sea al menos un 92% puro, es decir, que el polipéptido constituya al menos un 92% en peso del material del polipéptido total presente en la preparación, y se prefieren porcentajes más altos como al menos un 94% puro, al menos un 95% puro, al menos un 96% puro, al menos un 97% puro, al menos un 98% puro, al menos un 99% puro y al menos un 99,5% puro. Los polipéptidos presentados en este documento tienen preferentemente una forma sustancialmente pura. En concreto, se prefiere que los polipéptidos de esta invención tengan “forma esencialmente pura”, es decir, que la preparación de polipéptido esté esencialmente libre de otro material de polipéptidos con el que se asocia de modo natural. Esto puede llevarse a cabo, por ejemplo, mediante la preparación de polipéptidos con métodos recombinantes bien conocidos. En este documento, el término “polipéptido sustancialmente puro” es sinónimo de los términos “polipéptido aislado” y “polipéptido en forma aislada”.

[0055] En general, las moléculas de ácido nucleico que codifican la CBH1 de *H. grisea* variante se hibridarán, en condiciones de alta astringencia, con la secuencia proporcionada en este documento como la SEQ ID N°:1 (la CBH1.1 de *H. grisea* variante). No obstante, en algunos casos, se emplea una secuencia de nucleótidos que codifica CBH1 que posee un uso de codones sustancialmente diferente, mientras que la proteína codificada por la secuencia de nucleótido que codifica CBH1 tiene la misma o sustancialmente la misma secuencia de aminoácidos que la proteína nativa. Por ejemplo, la secuencia de codificación puede modificarse para facilitar la expresión más rápida de CBH1 en un sistema de expresión eucariota o procariota concreto, según la frecuencia con la que se utiliza un codón concreto por el huésped. Por ejemplo, Te'o, *et al.*, (2000) describe la optimización de los genes para la expresión en hongos filamentosos.

[0056] Una secuencia de ácido nucleico se considera que es “selectivamente hibridable” a una secuencia de ácido nucleico de referencia, si las dos secuencias se hibridan específicamente entre sí en condiciones de hibridación de moderada a alta astringencia y de lavado. Las condiciones de hibridación se basan en la temperatura de fusión (T_m) de la sonda o del complejo de unión a ácidos nucleicos. Por ejemplo, la “máxima astringencia” normalmente se produce a aproximadamente $T_m - 5^\circ\text{C}$ (5°C por debajo de la T_m de la sonda); la “alta astringencia” a aproximadamente $5-10^\circ\text{C}$ por debajo de la T_m ; “astringencia intermedia” aproximadamente a $10-20^\circ\text{C}$ por debajo de la T_m de la sonda; y “baja astringencia” aproximadamente a $20-25^\circ\text{C}$ por debajo de la T_m . Funcionalmente, pueden usarse las condiciones de máxima astringencia para identificar secuencias que tienen identidad estricta o casi estricta con la sonda de hibridación; mientras que las condiciones de alta astringencia se utilizan para identificar secuencias que tienen aproximadamente un 80% o más de identidad con la sonda.

[0057] En la técnica, son bien conocidas las condiciones de hibridación de moderada a alta astringencia (véanse, por ejemplo, Sambrook, *et al.*, 1989, capítulos 9 y 11, y Ausubel, F.M., *et al.*, 1993). Un ejemplo de las condiciones de alta astringencia incluye la hibridación a aproximadamente 42°C en 50% de formamida, 5X de SSC, 5X de solución Denhardt, 0,5% de SDS y 100 mg/ml de ADN portador desnaturalizado seguido de un lavado en dos veces en 2X de SSC y 0,5% de SDS a temperatura ambiente y dos veces más en 0,1X de SSC y 0,5% de SDS a 42°C .

[0058] El término “% de homología” se utiliza en el presente documento de modo intercambiable con el término “% de identidad” aquí y se refiere al nivel de identidad de secuencia de aminoácidos o ácidos nucleicos entre la secuencia de ácidos nucleicos que codifica una variante de CBH1.1 de *H. grisea*, una variante de CBH1 de *H. jecorina* o una secuencia de aminoácidos de CBH1 de *S. thermophilum*, cuando se alinean usando un programa de alineamiento de secuencias.

[0059] Por ejemplo, como se utiliza en este documento, 80% de homología es sinónimo de una identidad de secuencia de 80% determinada por un algoritmo definido y, en consecuencia, un homólogo de una secuencia determinada presenta más de un 80% de identidad de secuencia a lo largo de una secuencia dada. Los niveles de ejemplo de identidad de secuencia incluyen, sin carácter limitativo, el 80, 85, 90, 95, 98% o más de identidad de secuencia con una secuencia determinada, p. ej., la secuencia de codificación para una variante de CBH1.1 de *H. grisea*, una variante de CBH1 de *H. jecorina* o CBH1 de *S. thermophilum*, como se describe en este documento.

[0060] Entre los programas de ordenador que pueden utilizarse para determinar la identidad entre dos secuencias se incluyen, por ejemplo, sin carácter limitativo, el paquete de programas BLAST, p.ej., BLASTN, BLASTX y TBLASTX, BLASTP y TBLASTN, disponibles al público en Internet en <http://www.ncbi.nlm.nih.gov/BLAST/>. Véanse también, Altschul, *et al.*, 1990 y Altschul, *et al.*, 1997.

5 **[0061]** Las búsquedas de secuencias se llevan a cabo normalmente empleando el programa BLASTN cuando se evalúa una secuencia de ácido nucleico en relación con secuencias de ácidos nucleico en las secuencias de ADN del GenBank y otras bases de datos públicas. Se prefiere el programa BLASTX para la búsqueda de secuencias de ácidos nucleicos que se han traducido en todos los marcos de lectura frente a secuencias de aminoácidos en las secuencias de proteínas del GenBank y
10 otras bases de datos públicas. Tanto BLASTN como BLASTX se ejecutan utilizando parámetros por defecto de una penalización por hueco abierto de 11,0 y una penalización por hueco extendido de 1,0 y utilizan la matriz BLOSUM-62. (Véase, p.ej., Altschul, *et al.*, 1997).

[0062] Un alineamiento de secuencias seleccionadas con el fin de determinar el “% de identidad” entre dos o más secuencias se realiza empleando, por ejemplo, el programa CLUSTAL-W en
15 MacVector versión 6.5, ejecutado con los parámetros por defecto, incluyendo la penalización por hueco abierto de 10,0, una penalización por hueco extendido de 0,1 y una matriz de similitud BLOSUM 30.

[0063] Como se usa en el presente documento, los términos “transformada”, “transformada de modo estable” o “transgénica”, con referencia a una célula, significan que la célula tiene una secuencia de
20 ácidos nucleicos no nativa (heteróloga) integrada en su genoma o como un plásmido episomal que se mantiene a través de múltiples generaciones.

[0064] Según su uso en este documento, el término “expresión” hace referencia al proceso mediante el cual se produce un polipéptido sobre la base de la secuencia de ácido nucleico de un gen. El proceso incluye tanto la transcripción como la traducción.

25 **[0065]** El término “introducido” en el contexto de la inserción de una secuencia de ácido nucleico en una célula significa “transfección” o “transformación” o “transducción” e incluye la referencia a la incorporación de una secuencia de ácido nucleico en una célula eucariota o procariota, donde la secuencia de ácido nucleico puede incorporarse en el genoma de la célula (por ejemplo, cromosoma, plásmido, plástido o ADN mitocondrial), convertido en un replicón autónomo, o expresado de forma
30 transitoria (por ejemplo, ARNm transfectado).

[0066] En consecuencia, el término “expresión de CBH1.1” hace referencia a la transcripción y traducción de gen de celulasa CBH1.1 variante, cuyos productos incluyen precursores de ARN, ARNm, polipéptido, polipéptidos procesados de manera postraducciona. A modo de ejemplo, entre los ensayos para la expresión de CBH1.1 se incluyen: transferencia Western blot para proteínas de
35 CBH1.1, análisis de transferencia Northern blot y ensayos de reacción en cadena de polimerasa con transcriptasa inversa (RT-PCR) para ARNm de CBH1, y ensayos de actividad de endoglucanasa como se describe en Shoemaker S.P. y Brown R.D. Jr. (*Biochim. Biophys. Acta*, 1978, 523:133-146) Y Schulein (1988). De modo similar, como se utiliza en este documento, la “expresión CBH1” hace referencia a la transcripción y traducción de una variante del gen de celulasa CBH1 de *S. thermophilum* o una variante de CBH1 de *H. jecorina*, cuyos productos incluyen precursores de ARN,
40 ARNm, polipéptido, polipéptido procesados de manera postraducciona.

[0067] Como se utiliza en el presente documento, el término “surfactante” se refiere a cualquier compuesto generalmente reconocido en la materia por tener cualidades tensoactivas. De este modo, por ejemplo, los surfactantes comprenden surfactantes aniónicos, catiónicos y no iónicos, como los que se encuentran habitualmente en detergentes. Los surfactantes aniónicos incluyen alquilbenzenosulfonatos lineales o ramificados; sulfatos de éter de alquilo o alquenilo que tienen grupos alquilo lineales o ramificados o grupos alquenilos; sulfatos de alquilo o alquenilo; olefinsulfonatos y alcanosulfonatos. Los surfactantes anfóliticos incluyen sulfonatos de sal de amonio cuaternario y surfactantes anfóliticos de tipo betaína. Dichos surfactantes anfóliticos tienen tanto grupos de carga positiva como negativa en la misma molécula. Los surfactantes no iónicos pueden comprender éteres de polioxialquilenos, así como alcanolamidas de ácidos grasos superiores o aductos de óxido de alquilenos, monoésteres de glicerina de ácidos grasos y similares.

[0068] Como se usa aquí, el término “tejido que contiene celulosa” se refiere a cualquier tejido cosido o no cosido, hilos o fibras de algodón o no de algodón que contienen celulosa o mezclas de celulosa que contienen o no algodón, incluidos celulósicos naturales y celulósicos artificiales (como yute, lino, ramio, rayón y liocel).

[0069] Como se utiliza en este documento, el término “tejido que contiene algodón” se refiere a tejidos cosidos o no cosidos, hilos o fibras de algodón puro o mezclas de algodón, entre los que se incluye algodón en tejidos blanqueados de algodón, tricotados de algodón, vaqueros de algodón, hilos de algodón, algodón en rama y similares.

[0070] Tal y como se usa aquí, el término “composición de lavado a la piedra” se refiere a una formulación para uso en tejidos que contienen celulosa para el lavado a la piedra. Las composiciones de lavado a la piedra se utilizan para modificar los tejidos que contienen celulosa antes de su venta, es decir, durante el proceso de fabricación. Por el contrario, las composiciones detergentes están destinadas a la limpieza de prendas sucias y no se utilizan durante el proceso de fabricación.

[0071] Como se utiliza en el presente documento, el término “composición detergente” se refiere a una mezcla que está destinada para su uso en un medio de lavado de tejidos sucios que contienen celulosa. En el contexto de esta invención, dichas composiciones pueden incluir, además de celulasas y surfactantes, enzimas hidrolíticas adicionales, agentes de relleno, agentes de blanqueo, activadores de blanqueo, agentes de azulado y tintes fluorescentes, inhibidores de la aglomeración, agentes de enmascaramiento, activadores de celulasas, antioxidantes y solubilizantes.

[0072] Tal y como se utiliza en el presente documento, el término “reducción o eliminación de la expresión del gen *cbh1.1*” significa que el gen *cbh1.1* ha sido eliminado del genoma y, por tanto, no puede ser expresado por el microorganismo huésped recombinante, o que el gen *cbh1.1* ha sido modificado de manera que el microorganismo huésped recombinante no produce una enzima CBH1.1 funcional.

[0073] Los términos “*gen cbh1.1* variante” o “CBH1.1 variante” significan, respectivamente, que la secuencia de ácido nucleico del gen *cbh1.1* de *H. grisea* se ha alterado mediante la eliminación, adición y/o manipulación de la secuencia de codificación o que la secuencia de aminoácidos de la proteína expresada se ha modificado de acuerdo con la invención descrita en este documento.

[0074] Los términos “*gen cbh1* variante” o “CBH1 variante” significan, respectivamente, que la secuencia de ácido nucleico del gen *cbh1* de *H. jecorina* se ha alterado mediante la eliminación,

adición y/o manipulación de la secuencia de codificación o que la secuencia de aminoácidos de la proteína expresada se ha modificado de acuerdo con la invención descrita en este documento.

[0075] Tal como se usa aquí, los términos “activo” y “biológicamente activos” se refieren a una actividad biológica asociada a una proteína determinada y se utilizan de modo intercambiable. Por ejemplo, la actividad enzimática asociada con una proteasa es proteólisis y, por tanto, una proteasa activa tiene actividad proteolítica. Se deduce que la actividad biológica de una proteína determinada se refiere a cualquier actividad biológica que suele atribuirse a esa proteína por parte de los expertos en la materia.

[0076] Cuando se utiliza en soluciones enzimáticas, el componente de CBH1 de *S. thermophilum*, variante de CBH1 de *H. jecorina* o CBH1.1 variante u homóloga se añade generalmente en una cantidad suficiente para lograr el porcentaje más alto de liberación de azúcares solubles a partir de la biomasa. La cantidad de componente de CBH1 de *S. thermophilum*, variante de CBH1 de *H. jecorina* o CBH1.1 variante u homóloga añadida depende del tipo de biomasa que debe ser sacarificada que puede determinarse fácilmente por parte del experto en la materia. No obstante, cuando se utiliza, el porcentaje en peso de componente de CBH1 de *S. thermophilum*, variante de CBH1 de *H. jecorina* o CBH1.1 variante u homóloga con respecto a cualquier componente de tipo EG presente en la composición de celulasa es preferentemente de aproximadamente 1, preferentemente aproximadamente 5, preferentemente aproximadamente 10, preferentemente aproximadamente 15 o preferentemente aproximadamente 20 por ciento en peso a preferentemente 25 aproximadamente, preferentemente 30 aproximadamente, preferiblemente 35 aproximadamente, preferentemente 40 aproximadamente, preferentemente aproximadamente 45 o preferiblemente aproximadamente 50 por ciento en peso. Además, los intervalos preferidos pueden ser de aproximadamente 0,5 a aproximadamente 15 por ciento en peso, de aproximadamente 0,5 a aproximadamente 20 por ciento en peso, de aproximadamente 1 a aproximadamente 10 por ciento en peso, de aproximadamente 1 a aproximadamente 15 por ciento en peso, de aproximadamente 1 a aproximadamente 20 por ciento en peso, de aproximadamente 1 a aproximadamente 25 por ciento en peso, de aproximadamente 5 a aproximadamente 20 por ciento en peso, de aproximadamente 5 a aproximadamente 25 por ciento en peso, de aproximadamente 5 a aproximadamente 30 por ciento en peso, de aproximadamente 5 a aproximadamente 35 por ciento en peso, de aproximadamente 5 a aproximadamente 40 por ciento en peso, de aproximadamente 5 a aproximadamente 45 por ciento en peso, de aproximadamente 10 a aproximadamente 20 por ciento en peso, de aproximadamente 10 a aproximadamente 25 por ciento en peso, de aproximadamente 10 a aproximadamente 30 por ciento en peso, de aproximadamente 10 a aproximadamente 35 por ciento en peso, de aproximadamente 10 a aproximadamente 40 por ciento en peso, de aproximadamente 10 a aproximadamente 45 por ciento en peso, de aproximadamente 15 a aproximadamente 20 por ciento en peso, de aproximadamente 15 a aproximadamente 25 por ciento en peso, de aproximadamente 15 a aproximadamente 30 por ciento en peso, de aproximadamente 15 a aproximadamente 35 por ciento en peso, de aproximadamente 15 a aproximadamente 45 por ciento en peso, de aproximadamente 15 a aproximadamente 50 por ciento en peso.

II. ORGANISMOS HUÉSPED

[0077] Los hongos filamentosos incluyen todas las formas filamentosas de la subdivisión Eumycota y Oomycota. Los hongos filamentosos se caracterizan por un micelio vegetativo que tiene una pared celular compuesta de quitina, glucano, quiosano, manano y otros polisacáridos complejos, con crecimiento vegetativo por la elongación hifal y catabolismo de carbono que es obligadamente aeróbico.

[0078] En la presente invención, la célula precursora de los hongos filamentosos puede ser una célula de una especie de, sin carácter limitativo, *Trichoderma*, por ejemplo, *Trichoderma longibrachiatum* (*reesei*), *Trichoderma viride*, *Trichoderma koningii*, *Trichoderma harzianum*; *Penicillium sp*; *Humicola sp.*, incluida *Humicola insolens*; *Chrysosporium sp.*; incluida *C. lucknowense*; *Gliocladium sp.*; *Aspergillus sp.*; *Fusarium sp*, *Neurospora sp.*, *Hypocrea sp.* y *Emericella sp.* Como se utiliza en el presente documento, el término “*Trichoderma*” o “*Trichoderma sp*” se refiere a cualquier cepa de hongos que ha sido previamente clasificada como *Trichoderma* o está actualmente clasificada como *Trichoderma*.

[0079] En un modo de realización preferido, la célula precursora de los hongos filamentosos puede ser una célula de *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus aculeatus* o *Aspergillus nidulans*.

[0080] En una realización preferida, la célula precursora de los hongos filamentosos es una célula de *Trichoderma reesei*.

III. CELULASAS

[0081] Las celulasas se conocen en la técnica como enzimas que hidrolizan celulosa (enlaces beta-1,4-glucano o beta D-glucosídicos), lo que resulta en la formación de glucosa, celobiosa, celooligosacáridos y similares. Como se ha señalado anteriormente, las celulasas se han dividido tradicionalmente en tres clases principales: endoglucanasas (EC 3.2.1.4) (“EG”), exoglucanasas o celobiohidrolasas (EC 3.2.1.91) (“CBH”) y beta-glucosidasas (CE 3.2.1.21) (“BG”). (Knowles *et al.*, 1987; Schulein, 1988).

[0082] Algunos hongos producen sistemas de celulasas completos que incluyen exocelobiohidrolasas o celulasas de tipo CBH, endoglucanasas o celulasas de tipo EG y beta-glucosidasas o celulasas de tipo BG (Schulein, 1988). No obstante, algunas veces estos sistemas carecen de celulasas de tipo CBH y las celulasas bacterianas también incluyen normalmente pocas o ninguna celulasa de tipo CBH. Además, se ha demostrado que los componentes EG y los componentes CBH interactúan sinérgicamente para degradar la celulosa de modo más eficaz. Véase, por ejemplo, Wood, 1985. Los diferentes componentes, es decir, las diversas endoglucanasas y exocelobiohidrolasas en un sistema de celulasas multicomponentes o completo, generalmente tienen propiedades diferentes, como el punto isoeléctrico, peso molecular, grado de glucosilación, especificidad de sustrato y patrones de acción enzimáticos.

[0083] También se ha demostrado que las composiciones de celulasa degradan tejidos que contienen algodón, lo que resulta en una pérdida de resistencia reducida en el tejido (Patente de EE. UU. N°: 4.822.516), lo que contribuye a la resistencia al uso de composiciones de celulasas en aplicaciones comerciales de detergente. Se ha sugerido que las composiciones de celulasas formadas por

componentes de endoglucanasa presentan una pérdida de resistencia reducida para los tejidos que contienen algodón en comparación con las composiciones que comprenden un sistema de completo de celulasas.

[0084] También se ha demostrado que las composiciones de celulasas son útiles en la degradación de biomasa de celulasas en etanol (en la que la celulasa degrada celulosa a glucosa y levadura u otros microbios fermentan además la glucosa en etanol), en el tratamiento de pulpa mecánica (Pere *et al.*, 1996), para su uso como aditivo de piensos (WO 91/04673) y en la molienda húmeda del grano.

[0085] La mayoría de CBHs y EGs tienen una estructura multidominio que consiste en un dominio de núcleo separado de un dominio de unión a celulosa (CBD) mediante un péptido de enlace (Suumakki *et al.*, 2000). El dominio de núcleo contiene el sitio activo mientras que el CBD interacciona con la celulosa mediante la unión de la enzima al mismo (van Tilbeurgh *et al.*, 1986; Tomme *et al.*, 1988). Los CBDs son particularmente importantes en la hidrólisis de la celulosa cristalina. Se ha demostrado que la capacidad de las celobiohidrolasas para degradar la celulosa cristalina disminuye claramente cuando el CBD está ausente (Linder y Teeri, 1997). Sin embargo, la función exacta y el mecanismo de acción de los CBD siguen siendo un tema de especulación. Se ha sugerido que el CBD mejora la actividad enzimática simplemente gracias al aumento de la concentración efectiva de la enzima en la superficie de celulosa (Stahlberg *et al.*, 1991) y/o el aflojamiento de cadenas de celulosa simples de la superficie de celulosa (Tormo *et al.*, 1996). La mayoría de estudios sobre los efectos de los dominios de celulasa en diferentes sustratos se han realizado con proteínas de núcleo de celobiohidrolasas, ya que sus proteínas de núcleo pueden producirse fácilmente mediante proteólisis limitadas con papaína (Tomme *et al.*, 1988). Numerosas celulasas han sido descritas en publicaciones científicas, ejemplos de los cuales incluyen: de *Trichoderma reesei*: Shoemaker, S. *et al.*, *Bio/Technology*, 1:691-696, 1983, que presenta la CBHI; Teeri, T. *et al.*, *Gene*, 51:43-52, 1987, que describe la CBHII. Las celulasas de especies diferentes de *Trichoderma* también se han descrito, por ejemplo, Ooi *et al.*, 1990, que describe la secuencia de ADNc que codifica la endoglucanasa F1-CMC producida por *Aspergillus aculeatus*; Kawaguchi T *et al.*, 1996, que describe la clonación y la secuenciación de ADNc que codifica la beta-glucosidasa de *Aspergillus aculeatus*; Sakamoto *et al.*, 1995, que describe la secuencia de ADNc que codifica endoglucanasa CMCasa 1 de *Aspergillus kawachii* IFO 4308; Saarilahti *et al.*, 1990 que presenta una endoglucanasa de *Erwinia carotovora*; Spilliaert R, *et al.*, 1994, que describe la clonación y la secuenciación de *bgIA*, que codifica una beta-glucanasa de *Rhodothermus marinus*; y Halldorsdottir S *et al.*, 1998, que describe la clonación, la secuenciación y la sobreexpresión de un gen de *Rhodothermus marinus* que codifica una celulasa termoestable de la familia 12 de glicosil hidrolasa. No obstante, sigue existiendo la necesidad de identificación y caracterización de nuevas celulasas, con propiedades mejoradas, como un rendimiento mejorado en condiciones de tensión térmica o en presencia de surfactantes, una actividad específica aumentada, una alteración del patrón de división de sustrato y/o un nivel elevado de expresión *in vitro*.

[0086] El desarrollo de nuevas y mejoradas composiciones de celulasa que constan de cantidades variables de celulasas de tipo CBH, de tipo EG y de tipo BG es de interés para su uso: (1) en composiciones detergentes que muestran una mayor capacidad de limpieza, que funcionan como un agente suavizante y/o de mejor del tacto de los tejidos de algodón (por ejemplo, "lavado a la piedra" o

“biopulido”); (2) en composiciones para degradar la pulpa de madera u otras biomásas en azúcares (por ejemplo, para la producción de bioetanol) y/o (3) en composiciones de piensos.

IV. BIOLOGÍA MOLECULAR

[0087] En un modo de realización, esta invención trata la expresión de genes de celulasa CBH1.1 de *H. grisea* variante controlada por un promotor operativo en un hongo filamentoso. Por tanto, esta divulgación confía en las técnicas rutinarias en el campo de la genética recombinante. Entre los textos básicos que relatan los métodos de uso generales se incluyen: Sambrook *et al.*, *Molecular Cloning, A Laboratory Manual* (2ª ed., 1989); Kriegler, *Gene Transfer and Expression: A Laboratory Manual* (1990); y Ausubel *et al.*, eds., *Current Protocols in Molecular Biology* (1994)).

10 A. Métodos para identificar los genes de CBH1.1 variante

[0088] Dos secuencias de ácido nucleico de CBH1.1 de *H. grisea* disponibles públicamente se muestran en la Figura 5. La invención, en un aspecto, comprende una molécula de ácido nucleico que codifica una CBH1.1 de *H. grisea* variante descrita en la presente invención. El ácido nucleico puede ser una molécula de ADN.

15 [0089] Las técnicas que pueden utilizarse para aislar secuencias de ADN que codifican CBH1.1 variante son muy conocidas en la técnica e incluyen, sin carácter limitativo, cribado de biblioteca genómica y/o ADNc con un cribado de la expresión y sondas de ADN homólogo con ensayos de la actividad o anticuerpos frente a CBH1. Cualquiera de estos métodos puede encontrarse en Sambrook *et al.*, o en *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*,, F. Ausubel, *et al.*, ed. Greene Publishing and Wiley-Interscience, Nueva York (1987) (“Ausubel”).

20 B. Vectores de expresión/construcciones de ácidos nucleicos

[0090] Se pueden incorporar fragmentos de polinucleótidos sintéticos o naturales que codifican una celulasa CBH1.1 de *H. grisea* variante (“secuencias de ácido nucleico que codifican celulasa de CBH1.1 de *H. grisea* variante”) en construcciones o vectores de ácido nucleico heterólogos, capaces de la introducción y la replicación en una célula de hongo filamentoso o una célula de levadura. Los vectores y métodos descritos aquí son útiles para usar en células huésped para la expresión de una celulasa CBH1.1 variante. Se puede usar cualquier vector siempre y cuando sea replicable y viable en las células en las que se introduce. Los expertos en la materia conocen un gran número de vectores y promotores adecuados, y estos están disponibles en el mercado. Los vectores de clonación y expresión también se describen en Sambrook *et al.*, 1989, Ausubel FM *et al.*, 1989 y Strathern *et al.*, 1981. Los vectores de expresión apropiados para los hongos se describen en van den Hondel, C.A.M.J.J. *et al.* (1991) en: Bennett, J.W. y Lasure, L.L. (eds.) *More Gene Manipulations in Fungi*. Academic Press, págs. 396-428. La secuencia de ADN apropiada puede insertarse en un plásmido o vector (en el presente documento denominados “vectores”) mediante una variedad de procedimientos. En general, la secuencia de ADN se inserta en un sitio o sitios de endonucleasas de restricción apropiados por procedimientos estándares. Estos procedimientos y los de subclonación relacionados se consideran parte de los conocimientos de los expertos en la materia.

[0091] Los hongos filamentosos recombinantes que comprenden la secuencia de codificación para una CBH1.1 variante, una CBH1 de *H. jecorina* variante o una celulasa CBH1 de *S. thermophilium* pueden producirse por la introducción de una construcción de ácido nucleico heteróloga que comprende la secuencia que codifica CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una
5 celulasa CBH1 de *S. thermophilium* en las células de una cepa seleccionada de los hongos filamentosos.

[0092] Una vez que se obtiene la forma deseada de una secuencia de ácido nucleico de celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante, se puede modificar en una variedad de maneras. Cuando la secuencia implica regiones flanqueantes no
10 codificantes, las regiones flanqueantes pueden someterse a resección, mutagénesis, etc. De este modo, pueden realizarse transiciones, transversiones, eliminaciones e inserciones en la secuencia de origen natural.

[0093] Puede insertarse una secuencia que codifica una secuencia que codifica celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante seleccionada en un
15 vector adecuado según técnicas recombinantes conocidas y utilizarse para transformar los hongos filamentosos capaces de expresar celulasa. Debido a la degeneración inherente del código genético, pueden utilizarse otras secuencias de ácidos nucleicos que codifican sustancialmente la misma secuencia de aminoácidos o una secuencia funcionalmente equivalente para clonar y expresar una
20 CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. thermophilium*. Por tanto, se entiende que dichas sustituciones en la región de codificación están dentro de las variantes de secuencia descritas en la presente invención.

[0094] La presente invención también incluye construcciones de ácido nucleico recombinante que comprenden una o más secuencias de ácido nucleico que codifican una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o CBH1.1 variante tal como se describe
25 anteriormente. Las construcciones están compuestas de un vector, como un plásmido o vector viral, en el que se ha insertado una secuencia, en una orientación directa o inversa

[0095] Las construcciones de ácido nucleico heterólogas pueden incluir la secuencia que codifica para una CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. Thermophilium*: (i) en aislamiento, (ii) en combinación con secuencias de codificación adicionales; como secuencias que codifican péptidos señal o proteínas de fusión, en las que la secuencia
30 deseada que codifica celulasa es la secuencia de codificación dominante; (iii) en combinación con secuencias no codificantes, como intrones y elementos de control, como elementos de promotor y de terminación o regiones 5' o 3' no traducidas, eficaces para la expresión de la secuencia de codificación en un huésped apropiado; y/o (iv) en un vector o medio de huésped en el que la
35 secuencia que codifica celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina*, o la CBH1.1 variante es un gen heterólogo.

[0096] En un aspecto, se describe una construcción de ácido nucleico heteróloga que se utiliza para transferir una secuencia de ácido nucleico que codifica una CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. thermophilium* en una célula *in vitro*, pero se prefieren las
40 líneas de levaduras y hongos filamentosos establecidas. Para la producción de alto rendimiento de CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. thermophilium* a

largo plazo, se prefiere la expresión estable. De lo que se deduce que se puede utilizar cualquier método eficaz para generar transformantes estables.

[0097] Los vectores apropiados están normalmente equipados con una secuencia de ácido nucleico que codifica un marcador seleccionable, sitios de inserción y elementos de control adecuados, como secuencias de promotores y de terminación. El vector puede comprender secuencias reguladoras, incluyendo, por ejemplo, secuencias no codificantes, como intrones y elementos de control, es decir, elementos promotores y de terminación o regiones 5' y/o 3' no traducidas, eficaces para la expresión de la secuencia de codificación en las células huésped (y/o en un vector o medio de células huésped en el que una secuencia de codificación de antígeno de proteína soluble modificada no se expresa normalmente), unida de modo funcional a la secuencia de codificación. Los expertos en la técnica conocen un gran número de vectores y promotores adecuados, muchos de los cuales están disponibles en el mercado y/o se describen en Sambrook, *et al.*, (*supra*).

[0098] Entre los promotores de ejemplo se incluyen promotores constitutivos y promotores inducibles, cuyos ejemplos incluyen un promotor de CMV, un promotor temprano de SV40, un promotor de RSV, un promotor de EF-1 α , un promotor que contiene un elemento de respuesta a tetraciclina (TRE) en el sistema tet-on o tet-off, como se ha descrito (ClonTech y BASF), el promotor de beta-actina y el promotor de metalotionina que puede estimularse mediante la adición de determinadas sales de metales. Una secuencia promotora es una secuencia de ADN que es reconocida por el hongo filamentoso concreto para los propósitos de expresión. Está unida de modo funcional a una secuencia de ADN que codifica un polipéptido CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o la CBH1.1 variante. Dicha unión comprende la colocación del promotor en relación con el codón de iniciación de la secuencia de ADN que codifica el polipéptido CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante en los vectores de expresión descritos. La secuencia promotora contiene secuencias de control de la transcripción y de la traducción que median la expresión del polipéptido CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante. Entre los ejemplos se incluyen los promotores de los genes que codifican glucoamilasa, alfa-amilasa o alfa-glucosidasa de *Aspergillus niger*, *A. awamori* o *A. oryzae*; los genes *gpdA* o *trpC* de *A. nidulans*; los genes *cbh1* o *trp1* de *Neurospora crassa*; los genes que codifican la aspártico proteinasa de *A. niger* o *Rhizomucor miehei*; los genes *cbh1*, *cbh2*, *egl1*, *egl2* de *T. reesei*, u otros genes que codifican celulasas.

[0099] La elección del marcador seleccionable adecuado dependerá de la célula huésped, y los marcadores adecuados para los diferentes huéspedes son bien conocidos en la técnica. Los genes de marcadores seleccionables típicos incluyen *argB* de *A. nidulans* o *H.jecorina*, *amdS* de *A. nidulans*, *pyr4* de *Neurospora crassa* o *H.jecorina*, *pyrG* de *Aspergillus niger* o *A. nidulans*. Entre los marcadores seleccionables de ejemplo se incluyen, sin carácter limitativo, *trpc*, *trp1*, *oliC31*, *niaD* o *leu2*, que se incluyen en construcciones de ácido nucleico heterólogas utilizadas para transformar una cepa mutante como *trp-*, *pyr-*, *leu-* y similares.

[0100] Dichos marcadores seleccionables confieren a los transformantes la capacidad de utilizar un metabolito que normalmente no es metabolizado por los hongos filamentosos. Por ejemplo, el gen *amdS* de *H. jecorina* que codifica la enzima acetamidasa que permite que las células transformantes se desarrollen en acetamida como fuente de nitrógeno. El marcador seleccionable (p.ej., *pyrG*) puede

restablecer la capacidad de una cepa mutante auxotrófica para crecer en un medio mínimo selectivo o el marcador seleccionable (p.ej. *ollc31*) puede conferir a los transformantes la capacidad de crecer en presencia de un fármaco o antibiótico inhibidor.

[0101] La secuencia que codifica el marcador seleccionable se clona en cualquier plásmido adecuado utilizando métodos empleados generalmente en la técnica. Entre los plásmidos de ejemplo se incluyen: pUC18, pBR322, pRAX y pUC100. El plásmido pRAX contiene secuencias de AMA1 de *A.nidulans*, lo que posibilita el duplicado en *A. niger*.

[0102] La práctica de la presente invención empleará, a no ser que se indique lo contrario, técnicas convencionales de biología molecular, microbiología, ADN recombinante e inmunología, que son parte de la técnica. Estas técnicas se explican con detalle en la bibliografía. Véanse, por ejemplo, Sambrook *et al.*, 1989; Freshney, 1987; Ausubel, *et al.*, 1993; y Coligan *et al.*, 1991.

C. Métodos para transformar una célula huésped

[0103] La célula precursora de los hongos filamentosos puede ser una célula de una especie de, sin carácter limitativo, *Trichoderma*, por ejemplo, *Trichoderma longibrachiatum (reesei)*, *Trichoderma viride*, *Trichoderma koningii*, *Trichoderma harzianum*; *Penicillium sp*; *Humicola sp.*, incluida *Humicola insolens*; *Chrysosporium sp.*; incluida *C. lucknowense*; *Gliocladium sp.*; *Aspergillus sp.*; *Fusarium sp*, *Neurospora sp.*, *Hypocrea sp.* y *Emericella sp.* Como se utiliza en el presente documento, el término “*Trichoderma*” o “*Trichoderma sp*” se refiere a cualquier cepa de hongos que ha sido previamente clasificada como *Trichoderma* o está actualmente clasificada como *Trichoderma*.

[0104] Entre los ejemplos de líneas celulares parentales que pueden ser tratadas y/o modificadas para expresión de CBH1 de *S. thermophilum* CBH1.1 de *H. grisea* variante o una variante de CBH1 de *H. jecorina* se incluyen, sin carácter limitativo, células de hongos filamentosos. Ejemplos de tipos de células primarias apropiadas para su uso en la práctica de la invención incluyen, sin carácter limitativo, *Aspergillus* y *Trichoderma*.

[0105] En un modo de realización, la célula precursora de los hongos filamentosos puede ser una célula de *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus aculeatus* o *Aspergillus nidulans*.

[0106] En un segundo modo de realización, la célula parental de hongos filamentosos es una célula de *Hypocrea jecorina*. Esta célula se denominaba anteriormente *T. reesei*.

[0107] Después de las secuencias de ADN que codifican las variantes de CBH1.1, variante de CBH1 de *H. jecorina* o CBH1 de *S. thermophilum* se han clonado en construcciones de ADN, el ADN se utiliza para transformar microorganismos. El microorganismo a transformar para los propósitos de expresión una CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilum* pueden comprender de modo favorable una cepa derivada de *Trichoderma sp*. Por tanto, un modo preferido para preparar celulasas de una CBH1.1 de *S. thermophilum* variante, una variante de CBH1 de *H. jecorina* o una CBH1 variante comprende transformar una célula huésped de *Trichoderma sp*. con una construcción de ADN que comprenden al menos un fragmento de ADN que codifica una porción o toda la CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilum*. La construcción de ADN generalmente se adjuntará funcionalmente, es decir, unida de modo funcional, a un promotor. La célula huésped transformada se cultiva entonces en

condiciones para expresar la CBH1.1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1 de *H. grisea* variante. En consecuencia, la CBH1.1 de *H. grisea* variante, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilium* pueden aislarse. Sería deseable tener la CBH1.1 de *H. grisea* variante, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilium* en una forma sustancialmente pura. De modo similar, sería deseable tener la CBH1.1 de *H. grisea* variante, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilium* en una forma esencialmente pura.

[0108] No obstante, puede de hecho ser que el mejor vehículo de expresión para un ADN determinado que codifica una CBH1.1 variante puede diferir de *H. jecorina* (es decir, *T. reesei*). Por tanto, puede ser que sea más ventajoso expresar una proteína en un huésped de transformación que lleva similitud filogenética para el organismo de origen para la CBH1.1 variante. En un modo de realización alternativo, puede utilizarse *Aspergillus niger* como vehículo de expresión. Para obtener una descripción de técnicas de transformación con *A. niger*, véase WO 98/31821.

[0109] En consecuencia, la presente descripción de un sistema de expresión de *Trichoderma spp.* se proporciona sólo a modo de ejemplo y como una opción para expresar CBH1.1 variante de la invención. Un experto en la técnica, sin embargo, puede estar predispuesto a expresar el ADN que codifica la CBH1.1 variante en una célula huésped diferente si es apropiado y debe entenderse que la fuente de la CBH1.1 variante debe considerarse para determinar el huésped de expresión óptimo. Además, el trabajador cualificado en la materia será capaz de seleccionar el mejor sistema de expresión para un gen concreto con técnicas rutinarias utilizando las herramientas disponibles en la técnica.

D. Métodos para la expresión de una CBH1.1 variante, una variante de CBH1 de *H. jecorina* o de CBH1 de *S. thermophilium*

[0110] Los métodos confían en el uso de células para expresar una celulasa CBH1.1 variante, sin un método concreto de expresión requerido.

[0111] La invención proporciona células huésped que se han transducido, transformado o transfectado con un vector de expresión que comprende una secuencia de ácido nucleico que codifica celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante. Las condiciones de cultivo, como la temperatura, el pH y similares, son las utilizadas anteriormente para la célula huésped original antes de la transducción, transformación o transfección y serán claras para los expertos en la materia.

[0112] En una estrategia, una célula de hongos filamentosos o célula de levadura se transfecta con un vector de expresión que tiene un promotor o fragmento de promotor biológicamente activo o uno o más (p.ej., una serie de) potenciadores que funcionan en la línea celular huésped, unidos de modo funcional a un segmento de ADN que codifica una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante, de manera que la celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante se expresa en la línea celular.

[0113] En consecuencia, la presente invención presenta hongos filamentosos que comprenden células que han sido modificadas, seleccionadas y cultivadas de un modo eficaz para lograr producción o expresión de una celulasa CBH1 de *S. thermophilum*, una CBH1.1 variante, una variante de CBH1 de *H. jecorina* en relación con los hongos parentales no transformados correspondientes.

[0114] Entre los ejemplos de especies de hongos filamentosos originales que se pueden tratar y o modificar para mejorar la expresión de celulasa se incluyen, sin carácter limitativo, *Trichoderma*, *Penicillium sp.*, *Humicola sp.*, incluidas *Humicola insolens*; *Aspergillus sp.*, incluidas *Aspergillus niger* *Cryosporium sp.*, *Fusarium sp.*, *Hypocrea sp.* y *Emericella sp.*

[0115] Las células que expresan una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante se cultivan en condiciones que se emplean habitualmente para el cultivo de la línea fúngica precursora. En general, las células se cultivan en un medio estándar que contiene sales fisiológicas y nutrientes, como se describe en Pourquie, J. *et al.*, *Biochemistry and Genetics of Cellulose Degradation*, eds. Aubert, J.P. *et al.*, Academic press, págs. 71-86, 1988 e Ilmen, M. *et al.*, *Appl. Environ. Microbiol.* 63:1298-1306, 1997. Las condiciones de cultivo son también estándares, por ejemplo, los cultivos se incuban a 28°C en cultivos agitadores o fermentadores hasta lograr los niveles deseados de expresión de celulasa CBH1.1 variante.

[0116] Las condiciones de cultivo preferidas para un hongo filamentosos determinado se pueden encontrar en la bibliografía científica y/o en la fuente de hongos como la American Type Culture Collection (ATCC; <<http://www.atcc.org/>>). Después de establecer el crecimiento fúngico, se exponen las células a condiciones eficaces para producir o permitir la sobreexpresión de una celulasa CBH1.1 variante.

[0117] En los casos en los que una secuencia que codifica una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante está controlada por un promotor inducible, el agente inductor, p.ej., un azúcar, sal de metal o antibióticos, se añade al medio en una concentración eficaz para inducir la expresión de una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante.

[0118] En un modo de realización, la cepa comprende *Aspergillus niger*, que es una cepa útil para obtener proteína sobreexpresada. Por ejemplo, se sabe que *A. niger var awamori* dgr246 secreta cantidades elevadas de celulasas secretadas (Goedegebuur *et al*, *Curr. Genet* (2002) 41: 89-98). Son conocidas otras cepas de *Aspergillus niger var awamori* como GCDAP3, GCDAP4 y GAP3-4 Ward *et al* (Ward, M, Wolson, L.J. y Kodama, K.H., 1993, *Appl. Microbiol. Biotechnol.* 39:738-743).

[0119] En otro modo de realización, la cepa comprende *Trichoderma reesei*, que es una cepa útil para obtener proteína sobreexpresada. Por ejemplo, se sabe que RL-P37, descrita por Sheir-Neiss, *et al.*, *Appl. Microbiol. Biotechnol.* 20:46-53 (1984) secreta cantidades elevadas de enzimas celulasas. Los equivalentes funcionales de RL-P37 incluyen la cepa RUT-C30 (ATCC núm, 56765) y cepa QM9414 (ATCC núm. 26921) de *Trichoderma reesei*. Se contempla que estas cepas también serían útiles en la sobreexpresión de CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilum*,

[0120] Cuando se desea obtener la celulasa en ausencia de actividad celulolítica natural potencialmente perjudicial, es útil obtener una cepa de célula huésped en la que se ha eliminado uno

o más genes de celulasa antes de la introducción de una construcción de ADN o plásmido que contiene el fragmento de ADN que codifica CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. thermophilum*. Dichas cepas pueden prepararse por el método presentado en la Patente de EE. UU. N° 5.246.853 y WO 92106209. Al expresar una CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. thermophilum* en un microorganismo huésped que falta en uno o más genes de celulasa, se simplifican los procedimientos de identificación y posterior purificación. Cualquier gen de *Trichoderma sp.* que ha sido clonado puede eliminarse, por ejemplo, los genes *cbh1*, *cbh2*, *egl1* y *egl2* así como aquellos que codifican una proteína CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante (véanse, p.ej., la Patente de EE. UU. N° 5.475.101 y WO 94/28117, respectivamente).

[0121] La eliminación de genes se puede realizar insertando una forma del gen deseado para ser eliminado o alterado en un plásmido con métodos conocidos en la técnica. A continuación, el plásmido de deleción se corta en un sitio o sitios de enzimas de restricción apropiados, internos a la región de codificación génica deseada, y la secuencia que codifica el gen o parte de la misma se sustituye con un marcador seleccionable. Las secuencias de ADN flanqueante del locus del gen a eliminar o alterar, preferentemente entre aproximadamente 0,5 y 2,0 kb, se mantienen en ambos lados del gen del marcador seleccionable. Un plásmido de deleción apropiado generalmente tendrá sitios únicos de enzimas de restricción presentes en el mismo para permitir que se elimine el fragmento que contiene el gen eliminado, incluyendo las secuencias de ADN flanqueante, y el gen del marcador seleccionable para una pieza lineal simple.

[0122] Debe elegirse un marcador seleccionable de manera que permita la detección del microorganismo transformado. Será adecuado cualquier gen marcador seleccionable que se expresa en el microorganismo seleccionado. Por ejemplo, con *Aspergillus sp.*, el marcador seleccionable se elige de modo que la presencia del marcador seleccionable en los transformantes no afectará significativamente a las propiedades de los mismos. Dicho marcador seleccionable puede ser un gen que codifica un producto detectable. Por ejemplo, se puede utilizar una copia funcional de un gen *Aspergillus sp.* si carece de la cepa huésped da lugar a la cepa huésped en la cepa huésped que expresa un fenotipo auxotrófico.

[0123] En un modo de realización, se transforma una cepa derivada de *pyrG* de *Aspergillus sp.* con un gen *pyrG* funcional, que proporciona de este modo un marcador seleccionable para la transformación. Se puede obtener una cepa derivada de *pyrG* mediante la selección de cepas de *Aspergillus sp.* que son resistentes al ácido fluoroorótico (FOA). El gen *pyrG* codifica orotidina.5'-monofosfato descarboxilasa, una enzima necesaria para la biosíntesis de uridina. Las cepas con un gen *pyrG* intacto crecen en un medio sin uridina, pero son sensibles al ácido fluoroorótico. Es posible seleccionar cepas derivadas de *pyrG* sin enzima de orotidin-monofosfato descarboxilasa funcional y que necesitan uridina para crecer mediante la selección de resistencia a FOA. Con el uso de la técnica de selección de FOA, también es posible obtener cepas que necesitan uridina que carecen de orotato pirofosforibosil transferasa funcional. Es posible transformar estas células con una copia funcional del gen que codifica esta enzima (Berges & Barreau, *Curr. Genet.* 19:359-365 (1991), y van Hartingsveldt *et al* (1986) "Development of a homologous transformation system for *Aspergillus niger* based on the *pyrG* gene." *Mol. Gen. Genet.* 206:71-75). La selección de cepas derivadas se lleva a

cabo fácilmente empleando la técnica de resistencia a FOA señalada anteriormente y, de este modo, el gen *pyrG* se utiliza preferentemente como marcador seleccionable. En otro modo de realización, se transforma una cepa derivada de *pyr4* de *Trichoderma sp.* con un gen *pyr4* funcional, que proporciona así un marcador seleccionable para la transformación. Aunque a continuación se presenta el sistema de *Aspergillus*, se pueden utilizar otros procedimientos similares para *Trichoderma* y otros sistemas fúngicos como un experto en la materia entenderá.

[0124] Con el fin de transformar *pyrG* de *Aspergillus sp.* para que carezca de la capacidad de expresar uno o más genes de celulasas, a continuación un único fragmento de ADN que comprende un gen de celulasa eliminado o alterado se aísla del plásmido de delección y se utiliza para transformar un huésped *pyr* *Aspergillus* apropiado. Después, se identifican los transformantes y se seleccionan en base a su capacidad de expresar el producto génico de *pyrG* y de este modo cumplir con la auxotrofia de uridina de la cepa huésped. Entonces, se lleva a cabo el análisis de transferencia Southern blot en los transformantes resultantes para identificar y confirmar un caso de integración de doble cruzamiento que reemplaza parte o toda la región de codificación de la copia genómica del gen a eliminar con los marcadores seleccionables *pyr4*.

[0125] Aunque los vectores de plásmidos específicos descritos anteriormente se refieren a la preparación de transformantes de *pyr*, la presente invención no se limita a estos vectores. Se pueden eliminar o reemplazar varios genes en la cepa de *Aspergillus sp.* empleando las técnicas anteriores. Además, se puede utilizar cualquier marcador seleccionable disponible, como se indica anteriormente. De hecho, se puede eliminar cualquier gen de *Aspergillus sp.* que ha sido clonado, y por tanto identificado, del genoma utilizando la estrategia descrita anteriormente.

[0126] Como se describe anteriormente, las cepas huésped utilizadas son derivadas de *Aspergillus sp.* que carecen o tienen un gen o genes no funcionales en relación al marcador seleccionable elegido. Por ejemplo, si se elige *pyrG* como marcador seleccionable, después se utiliza una cepa derivada de *pyrG* como receptor en el procedimiento de transformación. De modo similar, se pueden utilizar los marcadores seleccionables que comprende genes de *Aspergillus sp.* equivalentes a los genes de *Aspergillus nidulans* como *amdS*, *argB*, *trpC* y *niaD*. Por lo tanto, la cepa receptora correspondiente debe ser una cepa derivada, como *argB*, *trpC*, *niaD*, respectivamente.

[0127] A continuación, se prepara el ADN que codifica la celulasa CBH1.1 variante para la inserción en un microorganismo apropiado. Según la presente invención, el ADN que codifica celulasa CBH1.1 variante comprende el ADN necesario para codificar para una proteína que muestra actividad celulolítica funcional. El fragmento de ADN que codifica la CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. thermophilium* puede ser funcionalmente unido a una secuencia promotora fúngica, por ejemplo, el promotor del gen *glaA*.

[0128] También se contempla que más de una copia de ADN que codifica CBH1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. thermophilium* se puede recombinar en la cepa para facilitar la sobreexpresión. El ADN que codifica la CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. thermophilium* se puede preparar mediante la construcción de un vector de expresión que porta el ADN que codifica la celulasa. El vector de expresión que porta el fragmento de ADN insertado que codifica la celulasa CBH1.1 variante puede ser cualquier vector que sea capaz de replicarse autónomamente en un organismo huésped determinado o de integrarse

en el ADN del huésped, normalmente un plásmido. En modos de realización preferidos, se contemplan dos tipos de vectores de expresión para obtener la expresión de genes. El primero contiene secuencias de ADN en las que el promotor, la región de codificación génica y la secuencia terminadora se originan todas del gen a expresar. El truncamiento de genes puede obtenerse cuando se desee mediante la eliminación de las secuencias de ADN no deseadas (p.ej., las que codifican para dominios no deseados) para dejar que el dominio se exprese controlado por sus propias secuencias reguladoras de la transcripción y de la traducción. El vector también contiene un marcador seleccionable que permite la selección para la integración en el huésped de múltiples copias de las nuevas secuencias génicas.

5
10 **[0129]** El segundo tipo de vector de expresión está preensamblado y contiene secuencias necesarias para la transcripción a nivel elevado y un marcador seleccionable. Se contempla que la región de codificación para un gen o parte de la misma se puede insertar en este vector de expresión de propósito general, de modo que esté bajo control transcripcional de las secuencias de promotor y terminador de los casetes de expresión. Por ejemplo, pRAX es uno de dichos vectores de expresión de propósito general. Los genes o partes de los mismos se pueden insertar en dirección 3' del promotor de *glaA* fuerte. Un ejemplo de un vector de expresión integrador es el vector pTrex. Los genes o parte de los mismos se pueden insertar en dirección 3' del promotor de *cbh1* fuerte.

15
20 **[0130]** En el vector, la secuencia de ADN que codifica la CBH1.1 variante, una variante de CBH1 de *H. jecorina*, una celulasa CBH1 de *S. thermophilum* debe estar unida de modo funcional a secuencias de traducción y de transcripción, es decir, una secuencia promotora y una secuencia señal adecuadas en el marco de lectura del gen estructural. El promotor puede ser cualquier secuencia de ADN que muestra actividad transcripcional en la célula huésped y puede derivar de genes que codifican proteínas homólogas o heterólogas a la célula huésped. Un péptido señal opcional mantiene la producción extracelular de la CBH1.1 variante, una variante de CBH1 de *H. jecorina* o una celulasa CBH1 de *S. thermophilum*. Se prefiere el ADN que codifica la secuencia señal que se asocia de modo natural con el gen a expresa, sin embargo se contempla en la invención la secuencia señal de cualquier origen apropiado.

25
30 **[0131]** Se conocen bien en la técnica los procedimientos empleados para fusionar las secuencias de ADN que codifican para una celulasa CBH1 de *S. thermophilum* una variante de CBH1 de *H. jecorina* o la CBH1.1 variante con el promotor en vectores adecuados.

35 **[0132]** Se pueden utilizar varios métodos para suministrar un vector de expresión, vector de ADN o construcción en células *in vitro*. Los profesionales conocen los métodos generales para introducir ácidos nucleicos en células para expresar secuencias de ácido nucleico heterólogas. Entre estos métodos se incluyen, sin carácter limitativo: electroporación; microinyección nuclear o microinyección directa en células simples; fusión de protoplastos bacterianos con células intactas; uso de policones, por ejemplo, polibreno o poliornitina; fusión de membrana con liposomas, transfección mediada por lipofectamina, o lipofección; bombardeo de alta velocidad con microproyectiles recubiertos de ADN; incubación con precipitación del ADN-fosfato de calcio; transfección mediada con DEAE-dextrano; infección con ácidos nucleicos virales modificados; transferencia de ADN mediada por Agrobacterios y similares. Además, las construcciones de ácido nucleico heterólogas que constan de una secuencia de ácido nucleico que codifica CBH1.1 variante pueden transcribir *in vitro*, e

introducir el ARN resultante en la célula huésped por métodos conocidos, por ejemplo, mediante inyección.

[0133] El método preferido para preparar *Aspergillus sp.* para la transformación conlleva la preparación de protoplastos a partir de micelio fúngico. Véase Campbell *et al.* "Improved transformation efficiency of *A. niger* using homologous *niaD* gene for nitrate reductase." *Curr. Genet.* 16:53-56; 1989. El micelio se puede obtener de esporas vegetales germinadas. El micelio se trata con una enzima que digiere la pared celular que produce protoplastos. A continuación, se protegen los protoplastos con la presencia de un estabilizador osmótico en el medio de suspensión. Estos estabilizadores incluyen sorbitol, manitol, potasio, cloruro, sulfato de magnesio y similares. Normalmente, la concentración de estos estabilizadores varía entre 0,8 M y 1,2 M. Se prefiere utilizar una solución de aproximadamente 1, 2 M de sorbitol en el medio de suspensión.

[0134] La captación del ADN en la cepa de *Aspergillus sp.* huésped, depende de la concentración de iones de calcio. Por lo general, se utiliza entre aproximadamente 10 mM de CaCl_2 y 50 mM de CaCl_2 en una solución de captación. Además de la necesidad de ion de calcio en la solución de captación, generalmente se incluyen otros elementos como el tampón TE (10 mM de Tris, pH 7,4; 1 mM de EDTA) o 10 mM de tampón MOPS, pH 6,0 (ácido morfolinopropanosulfónico) y polietilenglicol (PEG). Se cree que el polietilenglicol actúa para fusionar las membranas celulares permitiendo así la entrega del contenido del medio en el citoplasma de la cepa de *Aspergillus sp.* y el ADN plásmido se transfiere al núcleo. Frecuentemente, esta fusión deja múltiples copias del ADN plásmido integradas de modo delicado en el cromosoma huésped.

[0135] Normalmente, en la transformación se utiliza una suspensión que contiene las células o protoplastos de *Aspergillus sp.* que se ha sometido a un tratamiento de permeabilidad a una densidad de 10^5 a 10^6 /mL, preferentemente de 2×10^5 /mL. Se mezcla un volumen de 100 μL de estos protoplastos o células en una solución apropiada (p.ej., 1,2 M de sorbitol; 50 mM de CaCl_2) con el ADN deseado. Generalmente, se añade una concentración elevada de PEG a la solución de captación. Se puede añadir de 0,1 a 1 volumen de PEG 4000 al 25% a la suspensión de protoplastos. No obstante, es preferible añadir aproximadamente 0,25 volúmenes a la suspensión de protoplastos. Se pueden añadir también aditivos, como dimetil sulfóxido, heparina, espermidina, cloruro de potasio y similares a la solución de captación y ayudar en la transformación. Existen otros procedimientos similares para otras células huésped fúngicas. Véase, por ejemplo, la Patente de EE. UU. N° 6.268.328.

[0136] En general, la mezcla se incuba después a aproximadamente 0°C durante un periodo de entre 10 y 30 minutos. A continuación, se añade PEG adicional a la mezcla para aumentar más la captación de la secuencia de ADN o gen deseados. Generalmente, se añade PEG 4000 al 25% en volúmenes de 5 a 15 veces el volumen de la mezcla de transformación; sin embargo, pueden ser adecuados volúmenes menores o mayores. El PEG 4000 al 25% es preferentemente aproximadamente 10 veces el volumen de la mezcla de transformación. Después de añadir el PEG, la mezcla de la transformación se incuba entonces a temperatura ambiente o en hielo antes de añadir una solución de CaCl_2 y sorbitol. A continuación, la solución de protoplastos se añade también a las alícuotas fundidas de un medio de cultivo. Este medio de cultivo sólo permite el cultivo de transformantes. Cualquier medio de cultivo puede usarse en la presente invención que sea adecuado para el cultivo

de los transformantes deseados. No obstante, si se seleccionan los transformantes *Pyr⁺*, se prefiere utilizar un medio de cultivo que no contenga uridina. Las colonias posteriores se transfieren y purifican en un medio de cultivo sin contenido de uridina.

5 [0137] En esta fase, los transformantes estables se pueden distinguir de los transformantes inestables por su velocidad de cultivo más rápida y la formación de colonias circulares con un perfil liso, en lugar de irregular en un medio de cultivo sólido que carece de uridina. Además, en algunos casos, se puede realizar una prueba adicional de estabilidad mediante el cultivo de los transformantes en un medio de cultivo sólido no selectivo (es decir, que contiene uridina), la recogida de esporas de este medio de cultivo y la determinación del porcentaje de estas esporas que germinarán y crecerán
10 posteriormente en medios selectivos que carecen de uridina. Como alternativa, se pueden utilizar otros métodos conocidos en la materia para seleccionar transformantes.

[0138] En un modo de realización concreto del método anterior, una celulasa o celulasas CBH1 de *S. thermophilium*, una variante de *H. jecorina* o la CBH1.1 variante se recuperan en forma activa de la célula huésped después del cultivo en medio líquido como resultado del proceso de post traducción
15 apropiado de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o la CBH1.1 variante,.

E. Métodos para el análisis de expresión de proteínas y/o secuencias que codifican ácidos nucleicos de CBH1

[0139] Con el fin de evaluar la expresión de una celulasa CBH1 de *S. thermophilium*, una variante de
20 CBH1 de *H. jecorina* o una CBH1.1 variante por una línea celular que ha sido transformada con una construcción de ácido nucleico que codifica una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante, pueden realizarse ensayos a nivel de proteínas, a nivel del ARN o mediante el uso de bioensayos funcionales en particular de la actividad y/o producción de celobiohidrolasa.

25 [0140] En una aplicación de ejemplo de las secuencias de proteínas y ácido nucleico de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante descritas en la presente invención, se diseña una cepa genéticamente modificada de hongos filamentosos, p.ej., *Trichoderma reesei*, para producir una mayor cantidad de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante. Estos hongos filamentosos
30 genéticamente modificados serían útiles para producir un producto de celulasa con capacidad celulolítica aumentada. En un enfoque, esto se logra mediante la introducción de la secuencia de codificación para una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante en un huésped adecuado, por ejemplo, un hongo filamentosos como *Aspergillus niger*.

35 [0141] Por consiguiente, la divulgación incluye procedimientos para expresar una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante en un hongo filamentosos u otro huésped adecuado mediante la introducción de un vector de expresión que contiene la secuencia de ADN que codifica una celulasa CBH1 de *S. thermophilium*, una variante de

CBH1 de *H. jecorina* o una CBH1.1 variante en células del hongo filamentoso u otro huésped adecuado.

[0142] En otro aspecto, la invención también incluye métodos para modificar la expresión de una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante en un hongo filamentoso u otro huésped adecuado. Esta modificación incluye una reducción o eliminación de la expresión de la CBH endógena.

[0143] En general, entre los ensayos empleados para analizar la expresión de una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante se incluyen: transferencia Northern blot, dot blot (análisis de ADN o ARN), RT-PCR (reacción en cadena de la polimerasa con transcriptasa inversa) o hibridación *in situ*, utilizando una sonda marcada de manera apropiada (basada en la secuencia que codifica el ácido nucleico) y transferencia Southern blot y autorradiografía convencionales.

[0144] Además, la producción y/o la expresión de una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante se pueden medir en una muestra directamente, por ejemplo, mediante ensayos de la actividad, expresión y/o producción de la celobiohidrolasa. Dichos ensayos se describen, por ejemplo, en Becker *et al.*, *Biochem J.* (2001) 356:19-30 y Mitshuishi *et al.*, *FEBS* (1990) 275:135-138. La capacidad de CBH1 para hidrolizar sustratos aislados solubles e insolubles puede medirse con el uso de los ensayos descritos en Srisodsuk *et al.*, *J. Biotech.* (1997) 57:49-57 y Nidetzky y Claeyssens *Biotech. Bioeng.* (1994) 44:961-966. Los sustratos útiles para ensayar las actividades de celobiohidrolasa, endoglucanasa o β -glucosidasa incluyen celulosa cristalina, papel de filtro, celulosa hinchada con ácido fosfórico, celooligosacáridos, metilumbeliferil lactósido, metilumbeliferil celobiósido, ortonitrofenil lactósido, paranitrofenil lactósido, ortonitrofenil celobiósido, paranitrofenil celobiósido.

[0145] Además, la expresión de proteínas se puede evaluar con métodos inmunológicos, como la tinción inmunohistoquímica de células, secciones de tejido o inmunoensayo del medio de cultivo de tejidos, por ejemplo, por análisis Western blot o ELISA. Pueden utilizarse estos ensayos para evaluar cualitativa y cuantitativamente la expresión de una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante. Los detalles de dichos métodos son conocidos por aquellos expertos en la materia y muchos reactivos para la práctica de dichos métodos están disponibles en el mercado

[0146] Se puede usar una forma purificada de celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante para producir anticuerpos monoclonales o policlonales específicos para la proteína expresada para su uso en varios inmunoensayos. (Véase, p.ej., Hu *et al.*, 1991). Los ensayos de ejemplo incluyen ELISA, inmunoensayos competitivos, radioinmunoensayos, transferencia Western blot, ensayos de inmunofluorescencia indirecta y similares. En general, se pueden utilizar anticuerpos y/o kits disponibles comercialmente para el inmunoensayo cuantitativo del nivel de expresión de las proteínas de celobiohidrolasas.

F. Métodos para purificar una CBH1

[0147] En general, una proteína de celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante producida en cultivo celular es secretada en el medio y se puede purificar o aislar, por ejemplo, mediante la eliminación de los componentes no deseados del medio de cultivo celular. No obstante, en algunos casos, se puede producir una proteína de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante en una forma celular que necesita la recuperación de un lisado celular. En dichos casos, la proteína de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante se purifica a partir de las células en las que se produjo utilizando técnicas habitualmente empleadas por los expertos en la técnica. Entre los ejemplos se incluyen, sin carácter limitativo, cromatografía de afinidad (Tilbeurgh *et al.* 1984), métodos de cromatografía de intercambio iónico (Goyal *et al.*, 1991; Fliess *et al.*, 1983; Bhikhabhai *et al.*, 1984; Ellouz *et al.*, 1987), incluyendo materiales que usan el intercambio iónico con alto poder de resolución (Medve *et al.*, 1998), cromatografía de interacción hidrófoba (Tomaz and Queiroz, 1999) y separación en dos fases (Brumbauer, *et al.*, 1999).

[0148] Normalmente, la proteína de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante se fracciona para segregar proteínas que tienen las propiedades seleccionadas, tales como la afinidad de unión a agentes de unión particulares, p.ej., anticuerpos o receptores; o que tienen un intervalo de peso molecular seleccionado o un intervalo de puntos isoeléctricos.

[0149] Una vez lograda la expresión de una proteína de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante determinada, la proteína de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante producida de este modo es purificada a partir de las células o cultivo celular. Los procedimientos de ejemplo adecuados para esta purificación incluyen los siguientes: cromatografía de columna de afinidad con anticuerpos, cromatografía de intercambio iónico; precipitación con etanol; HPLC en fase inversa; cromatografía en sílice o en una resina de intercambio iónico, como DEAE; cromatoenfoco; SDS-PAGE; precipitación con sulfato de amonio y filtración en gel empleando, por ejemplo, Sephadex G-75. Se pueden utilizar diversos métodos de purificación de proteínas y estos métodos son conocidos en la técnica y descritos, p.ej., en Deutscher, 1990; Scopes 1982. La etapa o etapas de purificación seleccionadas dependerán, por ejemplo, de la naturaleza del proceso de producción utilizado y de la proteína concreta producida

V. Utilidad de *cbh1* y CBH1

[0150] Se puede observar que los ácidos nucleicos de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante, la proteína de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante y las composiciones que contienen actividad de la proteína una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante son útiles en una amplia variedad de aplicaciones, algunas de las cuales se describen a continuación.

[0151] Las composiciones nuevas y mejoradas de celulasas que comprenden cantidades variables de una celulasa CBH1 de *S. thermophilium*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante

son útiles en composiciones detergentes que presentan una capacidad de limpieza mejorada, funcionan como agente suavizante y/o mejoran el tacto de los tejidos de algodón (por ejemplo, “lavado a la piedra” o “biopulido”), en las composiciones para degradar la pulpa de madera en azúcares (por ejemplo, para la producción de bioetanol) y/o en composiciones de piensos. El aislamiento y la caracterización de las celulasas de cada tipo proporcionan la capacidad de controlar los aspectos de dichas composiciones.

[0152] Las celulasas CBH1 de *S. thermophilum*, variante de CBH1 de *H. jecorina* o CBH1.1 variante con termoestabilidad disminuida es útil, por ejemplo, en áreas donde se requiere que la actividad de la enzima sea neutralizada a temperaturas inferiores, de modo que otras enzimas que pueden estar presentes no se vean afectadas. Además, las enzimas pueden ser útiles en la conversión limitada de celulasas, por ejemplo, en el control del grado de cristalinidad o de la longitud de cadena celulósica. Después de alcanzar el grado deseado de conversión, la temperatura de sacarificación se puede elevar por encima de la temperatura de supervivencia de la CBH1.1 desestabilizada, una variante de CBH1 de *H. jecorina* o una CBH1 de *S. thermophilum*. Puesto que la actividad de CBH1 es esencial para la hidrólisis de la celulosa cristalina, la conversión de celulosa cristalina cesará a temperatura elevada.

[0153] En un enfoque preferido, una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante es útil para las composiciones detergentes o para tratamientos de tejidos para mejorar el tacto y el aspecto.

[0154] Puesto que la velocidad de hidrólisis de los productos celulósicos puede incrementarse mediante el uso de un transformante que tenga al menos una copia adicional del gen de celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante como un plásmido replicativo o insertado en el genoma, los productos que contienen celulosa o heteroglucanos pueden degradarse a un ritmo más rápido y en mayor medida. Los productos producidos a base de celulosa como el papel, algodón, pañales de celulosa y similares se pueden degradar de manera más eficaz en un vertedero. De este modo, el producto de fermentación que se puede obtener de los transformantes o los transformantes solos se pueden utilizar en composiciones para ayudar a degradar mediante licuefacción de una gran variedad de productos de celulosa que se añaden a los vertederos abarrotados.

[0155] La sacarificación y la fermentación separadas es un proceso por el que la celulosa presente en la biomasa, por ejemplo, forraje de maíz, se convierte en glucosa y, posteriormente, las cepas de levadura convierten la glucosa en etanol. La sacarificación y fermentación simultáneas es un proceso por el que la celulosa presente en la biomasa, por ejemplo, forraje de maíz, se convierte en glucosa y, al mismo tiempo y en el mismo reactor, las cepas de levadura convierten la glucosa en etanol. Por lo tanto, en otra estrategia preferida, la celulasa deseada de la invención es útil en la degradación de la biomasa en etanol. La producción de etanol a partir de fuentes fácilmente disponibles de celulosa proporciona una fuente de combustible renovable estable.

[0156] Las materias primas con base de celulosa comprenden residuos agrícolas, pastos y bosques y otra biomasa de bajo valor, como los residuos municipales (p.ej., papel reciclado, hierba cortada de los jardines, etc.) El etanol puede producirse a partir de la fermentación de cualquiera de estas materia primas de celulosa. No obstante, primero debe convertirse la celulosa en azúcares antes de

que pueda producirse la conversión en etanol. Por tanto, la celulasa CBH1 descrita en este documento es útil en la conversión de biomasa en azúcares.

[0157] Se puede utilizar una gran variedad de materias primas con una celulasa o celulasas CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante y la seleccionada para su uso puede depender de la región en la que se realiza la conversión. Por ejemplo, en el Medio Oeste de Estados Unidos los residuos agrícolas como paja del trigo, forraje de maíz y bagazo pueden predominar, mientras que en California puede predominar la paja de arroz. No obstante, debe entenderse que puede utilizarse cualquier biomasa disponible de celulosa en cualquier región.

[0158] Una composición de celulasa que contiene una cantidad mayor de celobiohidrolasa es útil en la producción de etanol. El etanol obtenido en este proceso también se puede utilizar como un potenciador del octanaje o directamente como combustible en lugar de gasolina, lo que es una ventaja porque el etanol como fuente de combustible es más ecológico que los productos derivados del petróleo. Se sabe que el uso de etanol mejorará la calidad del aire y posiblemente reducirá los niveles locales de ozono y la contaminación. Además, la utilización de etanol en lugar de gasolina puede ser una estrategia importante en la amortiguación del impacto de los cambios repentinos en los suministros de energía no renovable y petroquímica.

[0159] El etanol se puede producir a través de procesos de sacarificación y fermentación a partir de biomasa de celulosa, como árboles, plantas herbáceas, residuos sólidos municipales y residuos agrícolas y forestales. No obstante, la proporción de las enzimas de celulasa individuales en una mezcla de celulasa de origen natural producida por un microbio puede no ser la más eficiente para la conversión rápida de celulosa de la biomasa en glucosa. Se sabe que las endoglucanasas actúan para producir nuevos extremos en la cadena de celulosa que por si mismos son sustratos para la acción de celobiohidrolasas y, por lo tanto, mejoran la eficacia de la hidrólisis del sistema de celulasa completo. Por eso, la utilización de una actividad de celobiohidrolasas aumentada o mejorada aumentaría considerablemente la producción de etanol.

[0160] De este modo, la(s) celobiohidrolasa(s) es útil en la hidrólisis de la celulosa en sus componentes de azúcar. En un modo de realización, una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante se añade a la biomasa antes de la adición de un organismo de fermentación. En un segundo modo de realización, se añade una celulasa CBH1 de *S. thermophilum*, una variante de CBH1 de *H. jecorina* o una CBH1.1 variante a la biomasa al mismo tiempo que un organismo de fermentación. De modo opcional, puede haber otros componentes de celulasa presentes en ambos modos de realización. Puede lograrse la conversión de celulosa mejorada a temperaturas superiores usando los polipéptidos CBH1 descritos en el presente documento.

[0161] En otro modo de realización, la materia prima de celulosa se puede pretratar. El tratamiento previo puede ser mediante la adición de una solución de ácido diluido, ácido concentrado o álcali diluida a temperatura elevada. La solución de tratamiento previo se añade durante un tiempo suficiente para, por lo menos parcialmente, hidrolizar los componentes de hemicelulosa y después neutralizarlos.

[0162] Las composiciones detergentes de la invención pueden emplear, además de la composición de celulasa (con independencia del contenido de celobiohidrolasa, es decir, libre de celobiohidrolasa,

sustancialmente libre de celobiohidrolasa o celobiohidrolasa mejorada), un surfactante, incluyendo surfactantes aniónicos, no iónicos y anfólicos, una hidrolasa, agentes de relleno, agentes blanqueantes, agentes azulantes y tintes fluorescentes, inhibidores de la aglomeración, solubilizantes, surfactantes catiónicos y similares. Todos estos componentes son conocidos en el sector de los
 5 detergentes. La composición de celulasa descrita anteriormente puede añadirse a la composición detergente, ya sea en un diluyente líquido, en gránulos, en emulsiones, en geles, en pastas y similares. Estas formas son bien conocidas por parte de los expertos en la materia. Cuando se utiliza una composición detergente sólida, la composición de celulasa se formula preferentemente en gránulos. Preferentemente, los gránulos pueden formularse de manera que contengan un agente
 10 protector de celulasa. Para una discusión más completa, véase la Patente de EE. UU. Nº: 6.162.782, titulada "Detergent compositions containing cellulase compositions deficient in CBH1 type components".

[0163] Preferentemente, se emplean composiciones de celulasa a partir de aproximadamente un 0,00005 por ciento en peso a aproximadamente un 5 por ciento en relación con la composición
 15 detergente total. Más preferiblemente, se utilizan composiciones de celulasa de aproximadamente un 0,0002 por ciento en peso a aproximadamente un 2 por ciento en peso en relación al total de la composición detergente.

[0164] Además, la secuencia de ácidos nucleicos de celulasa es útil en la identificación y caracterización de secuencias de ácido nucleico relacionadas. Una serie de técnicas útiles para
 20 determinar (predecir o confirmar) la función de genes o productos génicos relacionados incluyen sin carácter limitativo: (a) análisis de ADN/ARN, como (1) la sobreexpresión, expresión ectópica y expresión en otras especies; (2) *knock out* génico (genética inversa, *knock-out* dirigido, silenciamiento génico inducido viral (VIGS, véase Baulcombe, 1999); (3) análisis del estado de metilación del gen, especialmente en regiones reguladoras flanqueantes e (4) hibridación *in situ*; (B) análisis de
 25 productos génicos como (1) expresión de proteína recombinante; (2) producción de antisueros; (3) inmunolocalización; (4) ensayos bioquímicos de la actividad catalítica u otra actividad; (5) estado de fosforilación; e (6) interacción con otras proteínas a través de análisis de dos híbridos de levaduras; (C) análisis de la vía, como la colocación de un gen o producto génico en una vía bioquímica particular o de señalización, basado en su fenotipo de sobreexpresión o mediante homología de la
 30 secuencia con genes relacionados; y (D) otros análisis que también se pueden llevar a cabo para determinar o confirmar la participación del gen aislado y su producto en una vía metabólica o de señalización concreta, y ayudar a determinar la función del gen.

EJEMPLOS

[0165] La presente invención describe con mayor detalle en los siguientes ejemplos que de ningún
 35 modo se pretende que limiten el alcance de la invención como se reivindica. Las figuras adjuntas deben considerarse como parte integrante de la especificación y descripción.

Ejemplo 1

Identificación de variantes de CBH1.1

[0166] Este ejemplo ilustra el aislamiento y caracterización del ácido nucleico que codifica una CBH1.1 de *H. grisea* variante.

Aislamiento de ADN genómico

[0167] El ADN genómico se puede aislar empleando cualquier método conocido en la técnica. En esta serie de experimentos, se obtuvo una muestra de *Humicola grisea var thermoidea* (CBS 225.63). No obstante, debe utilizarse el siguiente protocolo:

[0168] Se cultivan las células a 45°C en 20 ml de caldo de dextrosa de papa (PDB) durante 24 horas. Las células se diluyen en una ratio 1:20 en medio PBD fresco y se cultivan durante la noche. Se centrifugan dos mililitros de células y el sedimento se lava en 1 ml de KC (60g de KCl, 2g de ácido cítrico por litro, pH ajustado a 6, 2 con 1 M de KOH). El sedimento celular se resuspendió en 900 µl de KC. 100 µl (20 mg/ml) de Novozyme® se añaden, mezclan con cuidado y la protoplastización siguió microscópicamente a 37°C hasta que más del 95% de protoplastos se forman durante un máximo de 2 horas. Las células se centrifugan a 1500 rpm (460g) durante 10 minutos. Se añaden, mezclan e incuban 200 µl de TES/SDS (10mM de Tris, 50mM de EDTA, 150mM de NaCl, 1% de SDS) a temperatura ambiente durante 5 minutos. El ADN se aísla empleando un kit de aislamiento de mini-preparación de Qiagen (Qiagen). La columna se eluyó con 100 µl de agua mili-Q y el ADN recogido.

[0169] Este es el método empleado para el aislamiento de ADN genómico de *H. grisea var thermoidea* a partir de placas de PDA cultivadas 45°C. Podría ser deseable un método alternativo empleando el método Fast-Prep®. El sistema consiste en el instrumento Fast-Prep® así como los kits FastPrep® para el aislamiento de ácidos nucleicos. FastPrep® está disponible a través de Qbiogene, Inc. Qbiogene, Inc., 2251 Rutherford Road, Carlsbad, CA 92008.

Construcción de cebadores

[0170] La PCR se llevó a cabo en una máquina de PCR estándar como el termociclador Peltier PTC-200 (Peltier Thermal Cycler) de MJ Research Inc. en las siguientes condiciones:

- 1) 1 minuto a 96°C durante 1 ciclo
- 2) 30 segundos a 94°C

60 segundos a 55°C

2 minutos a 72°C

- 3) Repetir el paso 2 durante 30 ciclos
- 4) 7 minutos a 72°C durante 1 ciclo, y
- 5) temperatura inferior a 15°C para el almacenaje y análisis posterior.

[0171] Se construyeron los siguientes cebadores de ADN para su uso en la amplificación de genes de CBH1 homólogos a partir de ADN genómico aislado de varios microorganismos. Todos los símbolos utilizados en el presente documento para las secuencias de ADN y proteínas se corresponden con los códigos de la Comisión de Nomenclatura Bioquímica IUPAC IUB.

[0172] Se desarrollaron cebadores homólogos 3' (PVS204) y 5' (PVS203) en base a la secuencia de CBH1 de *Humicola grisea var thermoidea* (IFO9854). Esta cepa expresa la secuencia dada en la figura de alineamiento como D63515 (Figura 5). Ambos cebadores contenían secuencias de clonación Gateway de Invitrogen® en el 5' del cebador. El cebador PVS203 contenía la secuencia attB1 y el
5 cebador PVS204 contenía la secuencia attB2.

Secuencia de PVS203 sin la attB1:

5' ATGCGTACCGCCAAGTTCGC 3' (secuencia señal de CBH1)

Secuencia de PVS204 sin la attB2:

5' TTACAGGCACTGAGAGTACCAG 3' (módulo de unión a celulosa de CBH1)

10 [0173] Las condiciones de la PCR fueron las siguientes: 20µl de solución amortiguadora de la reacción 5X (solución amortiguadora de la reacción 5X que comprende 50mM de Tris HCl, pH 8,5; 87,5 mM de (NH₄)₂SO₄, 6,25 mM de MgCl₂; 2,5% de Teen20 (v/v) 7,5 % de DMSO (v/v)); 0,2 mM de cada uno de dATP, dTTP, dGTP, dCTP (concentración final), 1 µl de 100 ng/µL de ADN genómico, 1 µL de Tgo polimerasa (Roche Diagnostic GmbH, Cat#3186199) en 1 unidad por µL, 0,2µM de cada
15 cebador, PVS203 y PVS204, (concentración final), y agua a 100 µL.

Aislamiento de la secuencia del gen de CBH1.1 de *H. grisea* variante

[0174] La secuencia de longitud completa se obtuvo directamente utilizando los cebadores del extremo N terminal (PVS203) y del extremo C terminal (PVS204). Se tradujo la secuencia de ADN de longitud completa en tres marcos abiertos de lectura utilizando el programa Vector NTI. Se llevaron a
20 cabo dos comparaciones de ADN y secuencias de proteína con dos secuencias conocidas para *H. grisea* Cel7A (X17258 y D63515) con el fin de identificar cualquier secuencia de intrón putativa. La traducción de la secuencia de ADN genómico sin las secuencias de intrones reveló la secuencia de proteínas de la CBH1.1 de *H. grisea* variante. Se ha obtenido el gen de longitud completa y se proporciona en la Figura 1 (ADN genómico). El ADNc putativo se presenta en la Figura 2.

25 **Ejemplo 2**

Expresión de variantes de CBH1.1

[0175] El siguiente ejemplo detalla cómo se realizó la expresión del gen de CBH1.1 de *H. grisea*.

[0176] Los genes de longitud completa del Ejemplo 1 se transfirieron al vector de destino compatible Gateway de *A. niger*, desarrollado por Genencor. Este vector se construyó utilizando el pRAX1 como
30 estructura principal, que se muestra en la Figura 6, según el manual proporcionado por Gateway™ Cloning Technology: versión 1 pág. 34-38.

[0177] El vector de expresión desarrollado recientemente se muestra en la Figura 7; es un producto para transferir el nuevo gen al vector de destino pRAXdes2. Lo que dio lugar a los vectores de expresión finales denominados pRAXdesCBH1.1 (véase la Figura 8).

35 [0178] La construcción se transformó en *A. niger* var. *Awamori* según el método descrito por Cao *et al* (Cao Q-N, Stubbs M, Ngo KQP, Ward M, Cunningham A, Pai EF, Tu G-C y Hofmann T (2000)

“Penicillopepsin-JT2 a recombinant enzyme from *Penicillium janthinellum* and contribution of a hydrogen bond in subsite S3 to kcat” *Protein Science* 9:991-1001).

[0179] Los transformantes se extendieron en placas con medio mínimo (Ballance DJ, Buxton FP, y Tumer G (1983) “Transformation of *Aspergillus nidulans* by the orotidine-5'-phosphate decarboxylase gene of *Neurospora crassa*” *Biochem Biophys Res Commun* 112:284-289) y se desarrollaron durante 5 4 días a 30 °C. Se recogieron las esporas utilizando métodos bien conocidos en la técnica (Véase <www.fgsc.net/fgn48/Kaminskyj.htm>). Se recogieron conidios de *A. nidulans* en agua (frotando la superficie de un cultivo de conidios con una varilla de vidrio estéril curvada para desplazar esporas) y pueden almacenarse de semanas a meses a 4 °C sin una pérdida seria de viabilidad. Sin embargo, 10 las esporas recogidas recientemente germinan de un modo más reproducible. Para el almacenamiento a largo plazo, las esporas se pueden conservar en 50% glicerol a -20 °C, o en 15-20% glicerol a -80 °C. El glicerol se pipetea más fácilmente como una solución 80% en agua. Se utilizan 800 µl de una suspensión acuosa de conidios (producida para el almacenamiento a 4 °C) añadidos a 200 µl 80% de glicerol para una solución de reserva a -80 °C; se utilizan 400 µl de 15 suspensión añadidos a 600 µl de 80% glicerol para una solución de reserva a -20 °C. Se agita antes de congelarse. Para las recogidas mutantes, se pueden escindir piezas pequeñas de cultivos de conidios y se colocan en 20% de glicerol, se agitan y se congelan como soluciones de reserva a -80 °C. En este caso, se almacenaron en 50% glicerol a -80 °C.

[0180] Se desarrollaron transformantes de *A. niger var awamori* en medio mínimo que carecía de 20 uridina (Ballance *et al.* 1983). Se analizaron los transformantes para encontrar la actividad de celulasa mediante la inoculación de 1cm² de suspensión de esporas de la placa de agar cultivado esporulado en matraces de agitación de 100 ml durante 3 días a 37 °C, como se describe en Cao *et al.* (2000).

[0181] El ensayo de actividad de CBHI se basa en la hidrólisis de la 4-metilumbeliferil-β-lactosa no fluorescente en los productos lactosa y 7-hidroxi-4-metilcoumarina, éste último es responsable de la 25 señal fluorescente. Se pipetea 170 µl de tampón de 50 mM de NaAc pH4.5 en una placa de microtitulación (MPT) de 96 pocillos (Greiner, Fluotrac 200, art. nr. 655076) apropiada para la fluorescencia. Se añaden 10 µl de sobrenadante y 10 µl de MUL (1 mM de 4-metilumbeliferil-β-lactosa (MUL) en agua miliQ) y se pone la MTP en el Fluostar Galaxy (BMG Labtechnologies; D-77656 Offenburg). Se mide la cinética durante 16 min. (8 ciclos de 120 s cada uno) utilizando λ_{320 nm} 30 (excitación) y λ_{460 nm} (emisión) a 50 °C. A continuación, los sobrenadantes que muestran actividad de CBH1 se sometieron a cromatografía de interacción hidrofóbica, como se describe en el Ejemplo 3 a continuación.

[0182] Se dedujeron las secuencias de aminoácidos como se indica anteriormente en el Ejemplo 1. La secuencia de aminoácidos de la CBH1.1 variante se muestra en la Figura 3 con una secuencia 35 señal y en la Figura 4 sin una secuencia señal. La secuencia señal se muestra en la Figura 3 en fuente subrayada y en negrita.

Ejemplo 3

Termoestabilidad de las variantes de CBH1

[0183] El siguiente ejemplo detalla cómo la CBH1.1 de *H. grisea* variante difiere en termoestabilidad de la enzima de celulasa CBH1 de *T. reesei*.

[0184] Las variantes de celulasa CBH I se clonaron y expresaron como se indica anteriormente (véase Ejemplo 2). A continuación, Cel7A de tipo salvaje y variantes se purificaron de sobrenadantes
5 libres de células de estos cultivos por cromatografía en columna. Las proteínas se purificaron utilizando la cromatografía de interacción hidrofóbica (HIC). Las columnas se desarrollaron en un sistema de cromatografía de perfusión BioCad® Sprint utilizando resinas Poros® 20 HP2, ambos de Applied Biosystems.

[0185] Las columnas HIC se equilibraron con 5 volúmenes de columna de 0,020 de fosfato de sodio, 10 0,5 M de sulfato de amonio a pH 6,8. Se añadió sulfato de amonio a los sobrenadantes hasta una concentración final de aproximadamente 0,5 M y el pH se ajustó a 6,8. Después de la filtración, el sobrenadante se cargó en la columna. Después de cargarse, se lavó la columna con 10 volúmenes de columna de tampón de equilibrio y a continuación se eluyó con una gradiente de 10 volúmenes de columna desde 0,5 M de sulfato de amonio a 0,02 de fosfato de sodio pH
15 6,8. La Cel7A se eluyó aproximadamente a la mitad del gradiente. Se recogieron las fracciones y se agruparon en base de un análisis de gel SDS-PAGE reducido.

[0186] Los puntos de fusión se determinaron según los métodos de Luo, *et al.*, *Biochemistry* 34:10669 y Gloss, *et al.*, *Biochemistry* 36:5612.

[0187] Se recogieron los datos en un espectrofotómetro del dicroísmo circular de Aviv 215. Los
20 espectros de las variantes entre 210 y 260 nanómetros se tomaron a 25°C. Las condiciones de tampón fueron 50 mM de Bis Tris Propan/ 50 mM de acetato de amonio/ácido acético glacial a pH 5,5. La concentración de proteína se mantuvo entre 0,25 y 0,5 mgs/mL. Después de determinar la longitud de onda para monitorizar el despliegue, las muestras se desnaturalizaron térmicamente mediante un aumento gradual de la temperatura de 25°C a 75°C en las mismas condiciones de tampón. Se
25 recogieron datos durante 5 segundos cada 2 grados. Se monitorizó el despliegue parcialmente reversible a 230 nanómetros en una célula de longitud de paso de 0,1 centímetros.

[0188] La celulasa CBH1.1 de *H. grisea* variante tiene un perfil de termoestabilidad mejorada en comparación con la CBH1 de *T. reesei* de tipo salvaje. La CBH1.1 de *H. grisea* variante muestra una temperatura de fusión (T_m) de 72,5 °C; *T. reesei* tiene una T_m de 62,3°C.

30 **Ejemplo 4**

Actividad de variantes de CBH1

[0189] El siguiente ejemplo detalla cómo se evaluó la actividad de la variante de CBH1.1 de *H. grisea*.

[0190] La conversión de celulosa se evaluó con técnicas conocidas en la materia. Véase, por ejemplo, Baker *et al*, *Appl Biochem Biotechnol* 1998 Spring; 70-72(): 395-403.

35 [0191] Se utilizó un ensayo de conversión celulósica estándar en los experimentos. En este ensayo, se colocaron la enzima y el sustrato tamponado en recipientes y se incubaron a una temperatura durante un tiempo. La reacción se templó con suficiente glicina 100 mM, pH 11,0 para llevar el pH de la mezcla de reacción a al menos pH 10. Una vez que se templó la reacción, se filtró un alícuota de la mezcla de la reacción a través de una membrana de 0,2 micras para eliminar los sólidos. A

continuación, se probó la solución filtrada para determinar azúcares solubles por HPLC según los métodos descritos por Baker et al., *Appl. Biochem. Biotechnol.* 70-72:395 - 403 (1998).

[0192] Forraje de maíz pretratado (PCS) – Se pretrató el forraje de maíz con 2% m/m de H₂SO₂ como se describe en Schell, D. *et al.*, *J. Appl. Biochem. Biotechnol.* 105:69 - 86 (2003), a lo que le siguieron múltiples lavados con agua desionizada hasta lograr un pH de 4,5. Se añadió acetato de sodio para conseguir una concentración final de 50mM y se ajustó la dosis para pH 5,0. La concentración de celulosa en la mezcla de reacción era de aproximadamente el 7%.

[0193] Celulosa hinchada con ácido fosfórico (PASC) – La PASC se preparó a partir de Avicel según el método descrito en Walseth (1971) *Tappi* 35: 228 (1971) y Wood *Biochem J.* 121:353 (1971). Este material se diluyó con tampón y agua para lograr una mezcla de 0,63 % m/v de manera que la concentración final de acetato de sodio era 50mM, pH 5,0. Las enzimas se dosificaron a 1,6 mg de proteína total por gramo de celulosa.

[0194] El primer grupo de experimentos examinó el porcentaje de expresión a 38°C durante 1 día en 12,66% forraje de maíz pretratado (PCS) (véase Schell, D. *et al.*, *J. Appl. Biochem. Biotechnol.* (2003) 105:69 - 86) empleando 15,5 mg de enzima/ g de celulosa. Las muestras se agitaron a 700rpm. Se realizaron las comparaciones entre: 1) una celulasa de una cepa de *Trichoderma* de CBH1 eliminada (delCBH1); 2) una cepa de *A. niger* que tenía un gen de CBH1 de *T. reesei* insertado (rCBH1); 3) una cepa de *A. niger* que sobreexpresa su CBHB nativa (Aniger); 4) una cepa de *A. niger* que tenía un gen de CBH1 de *H. schweinitzii* insertado (Hschweinitzii/An); 5) una cepa de *A. niger* que tenía un gen de CBH1 de *T. pseudokoningii* insertado (Tpseudokoni/An); 6) un cepa de *A. niger* que tenía un gen de CBH1.1 variante de *H. grisea* insertado (Hgrisea/An-1); 7) una cepa de *A. niger* que tenía un gen de CBH1.1 variante de *H. grisea* insertado (Hgrisea/An-2). Hgrisea/An-1 y Hgrisea/An-2 son dos clones de la misma transformación de *A. niger* con el gen de CBH1.1 variante de *H. grisea*. Los resultados de la primera serie de experimentos se presentan en la Figura 9. Como se puede observar, la CBH1.1 variante de *H. grisea* no supera cualquiera de las otras CBH analizadas. La Figura 11 es una comparación de la CBH1.1 variante de *H. grisea* con los resultados de nCBH1 de *T. reesei* en las mismas condiciones experimentales.

[0195] La segunda serie de experimentos examinó el porcentaje de conversión en condiciones similares a las de la primera serie de experimentos, excepto la temperatura de incubación que era de 65°C, no 38°C. Los resultados de esta serie de experimentos se presentan en la Figura 10. Como se puede observar, la CBH1.1 variante de *H. grisea* supera las otras CBH analizadas. La Figura 12 es una comparación de la CBH1.1 variante de *H. grisea* con los resultados de rCBH1 de *T. reesei* en las mismas condiciones experimentales.

[0196] La tercera serie de experimentos examinó la velocidad de generación de celobiosa a partir de PASC en condiciones de ensayo similares a las de los experimentos anteriores; las temperaturas utilizadas fueron 38°C, 65°C y 70°C. Los resultados se presentan en la Figura 13. Como se puede observar, la CBH1.1 variante de *H. grisea* supera la CBH1 de *T. reesei* a todas las temperaturas analizadas.

Ejemplo 5

Asilamiento de CBH1 de *Scytalidium thermophilum*

[0197] Este ejemplo ilustra el aislamiento y caracterización del ácido nucleico que codifica una CBH1 de *S. thermophilum*.

Aislamiento de ADN genómico y construcciones de cebadores

5 [0198] Los métodos para la clonación, expresión y purificación de CBH1 de *Scytalidium thermophilum* fueron todos como los descritos para la CBH1.1 de *Humicola grisea* var. *Thermoidea* (véase el Ejemplo 1). El gen de CBH1 de *Scytalidium thermophilum* se amplificó a partir de ADN genómico de la entrada de la colección cepas pública CBS 671.88, utilizando los mismos cebadores, PVS203 y PVS204, empleados para amplificar la CBH1.1 de *Humicola grisea* var. *Thermoidea*.

10 Aislamiento de la secuencia de gen de CBH1 de *S. thermophilum*

[0199] La secuencia de longitud completa se obtuvo directamente utilizando los cebadores del extremo N terminal (PVS203) y C terminal (PVS204). La secuencia de ADN de longitud completa se tradujo en tres marcos de lectura abiertos usando el programa Vector NTI. La traducción de la secuencia de ADN genómico sin las secuencia de intrones mostró la secuencia de proteína de la
15 CBH1 de *S. thermophilum*. Se ha obtenido el gen de longitud completa, y se muestra en la Figura 14A (ADN genómico). El ADNc putativo se presenta en la Figura 14B. La secuencia de aminoácidos con y sin secuencia señal se presenta en las Figuras 14C y 15, respectivamente.

[0200] Se alinearon las secuencias de proteína de la CBH1 de *H. jecorina*, CBH1.1 DE *Humicola grisea* var. *Thermoidea* y CBH1 de *S. thermophilum*. Véase la Figura 15.

20 [0201] Con la identificación de los sitios de diferencia entre CBH1.1 de *Humicola grisea* var. *Thermoidea* y la CBH1 considerablemente más estable de *Scytalidium thermophilum* y, a continuación, con el uso del alineamiento de la Figura 15 para identificar el sitio correspondiente en CBH1 de *Hypocrea jecorina*, que es mucho menos estable que la CBH1.1 de *Humicola grisea* var. *Thermoidea* o CBH1 de *Scytalidium thermophilum*.) Se identifican como importantes para la
25 estabilidad los siguientes sitios en CBH1 de *Hypocrea jecorina* (numeración de residuos de proteína madura):

- Thr(T)55, preferentemente Thr55Glu (T55E) y Thr55Lys (T55K)
- Ser(S)58, preferentemente Ser58Thr (S58T)
- Gln(Q)101, preferentemente Gln101Tyr (Q101Y) y Gln101His (Q101H)
- 30 Asn(N)250, preferentemente Asn250Asp (N250D) y Asn250Glu (N250E)
- Pro(P)265, preferentemente Pro265Ala (P265A) y Pro265Ser (P265S)
- Leu(L)288, preferentemente Leu288Ile (L288I).

Ejemplo 6

Termoestabilidad de CBH1 de *Scytalidium thermophilum*

[0202] Este ejemplo describe la medición de la termoestabilidad de CBH1 de *S. thermophilum* y CBH1.1 de *H. grisea var. thermoidea*.

[0203] Se expresaron la CBH1 de *S. thermophilum* y la CBH1.1 de *H. grisea var. thermoidea* empleando métodos como en el Ejemplo 2, purificadas como en el Ejemplo 3 y analizadas mediante calorimetría de barrido diferencial (DSC) con el uso del calorímetro Microcal VP-DSC. Se analizaron las muestras de 30 a 95 grados C a 90 grados C/hora. Las proteínas purificadas se desalaron en 50 mM de acetato de amonio y 50 mM de propano bistris a pH 5.5. La concentración final de proteína fue entre 0,05 y 0,25 mg/mL para todas las muestras.

[0204] La transición térmica observada para CBH1 de *S. thermophilum* tenía un punto medio de CBH1.1 de 78,3 grados C. CBH1.1 de *H. grisea* tiene una transición con un punto medio de 76,0 grados C: Véase la Figura 16.

[0205] Se entiende que los ejemplos y modos de realización descritos en el presente documento tienen únicamente propósitos ilustrativos y que se sugerirán diversas modificaciones o cambios a la luz de las mismas a los expertos en la técnica tal como definen las reivindicaciones adjuntas.

15

Reivindicaciones

1. Un polipéptido que tiene actividad de celobiohidrolasa I, donde el polipéptido es una variante de CBH1.1 de *H. grisea* que presenta la secuencia mostrada en la Figura 4 (SEQ ID N°:4).
2. El polipéptido de la reivindicación 1, que también comprende una secuencia señal.
- 5 3. El polipéptido de la reivindicación 2, compuesto por una variante de CBH1.1 de *H. grisea* que presenta la secuencia de la Figura 3 (SEQ ID N°:3)
4. Un polinucleótido que codifica una celobiohidrolasa que consta de una variante de CBH1.1 de *H. grisea* que comprende la SEQ ID N°:4.
- 10 5. Un polinucleótido según la reivindicación 4, en el que la variante de CBH1.1 de *H. grisea* comprende la SEQ ID N°:3.
6. Una construcción de ácido nucleico que comprende una secuencia de nucleótidos según la reivindicación 4, unida de modo funcional a una o más secuencias de control.
7. Un vector de expresión recombinante que comprende la construcción de ácido nucleico de la reivindicación 6.
- 15 8. Una célula huésped recombinante que comprende la construcción de ácido nucleico de la reivindicación 6.
9. Un método para la producción de un polipéptido de CBH1, dicho método comprende:
 - a) la transformación de una célula huésped con un ácido nucleico que consta de un polinucleótido según la reivindicación 4;
 - 20 b) el cultivo de una célula huésped en condiciones para producir el polipéptido; y
 - c) la recuperación del polipéptido.
10. Una composición que comprende un polipéptido que presenta actividad de celobiohidrolasa I según la reivindicación 1.
- 25 11. Un método para convertir biomasa en azúcares que comprende la puesta en contacto de la biomasa con un polipéptido que presenta actividad de celobiohidrolasa I, en el que dicho polipéptido tiene actividad mejorada con respecto a *T. reesei*, ya sea en un ensayo de conversión de PCS a 65°C o en un ensayo de PASC a 38°C, y en el que el polipéptido comprende una variante de CBH1.1 de *H. grisea* que presenta la secuencia de la Figura 4 (SEQ ID N°:4).

30

CBH1.1 de *H. grisea*

Figura 1: Secuencia genómica total (1638 nucleótidos)

1	ATGCGTACCG	CCAAGTTCGC	CACCCTCGCC	GCCCTTGTGG	CCTCGGCCGC	50
51	CGCCAGCAG	GCGTGCAGT	TCACCACCGA	GAGGCACCC	TCCCTCTCTT	100
101	GGAAGAAGT	CACCGCCGG	GGCCAGTGCC	AGACCGTCCA	GGCTTCCATC	150
151	ACTCTCGACT	CCAAGTGGCG	CTGGACTCAC	CAGGTGTCTG	GCTCCACCAA	200
201	CTGCTACACG	GGCAACAAGT	GGGATACTAG	CATCTGCACT	GATGCCAAGT	250
251	CGTGGCTCA	GAAGTGTGC	GTCGATGGTG	CCGACTACAC	CAGCACCTAT	300
301	GGCATCACCA	CCAACGGTGA	TTCCCTGAGC	CTCAAGTTCG	TCACCAAGGG	350
351	CCAGCACTCG	ACCACGTCG	GCTCGCGTAC	CTACCTGATG	GACGGCGAGG	400
401	ACAAGTATCA	<u>GAGTACGTT</u>	<u>TATCTTCAGC</u>	<u>CTTCTCGCGC</u>	<u>CTTGAATCCT</u>	450
451	<u>GGCTAACGTT</u>	<u>TACACTTCAC</u>	<u>AGCCTTCGAG</u>	<u>CTCCTCGGCA</u>	<u>ACGAGTTCAC</u>	500
501	CTTCGATGTC	GATGTCTCCA	ACATCGGCTG	CGGTCTCAAC	GGCGCCCTGT	550
551	ACTTCGTCTC	CATGGACGCC	GATGGTGGTC	TCAGCCGCTA	TCCTGGCAAC	600
601	AAGGCTGGTG	CCAAGTACGG	TACCGGCTAC	TGCGATGCTC	AGTGCCCCCG	650
651	TGACATCAAG	TTCATCAACG	GCGAGGCCAA	CATTGAGGGC	TGGACCGGCT	700
701	CCACCAACGA	CCCCAACGCC	GGCGCGGGCC	GCTATGGTAC	CTGCTGCTCT	750
751	GAGATGGATA	TCTGGGAAGC	CAACAACATG	GCTACTGCCT	TCACTCCTCA	800
801	CCCTFGCACC	ATCATTTGCC	AGAGCCGCTG	CGAGGGCGAC	TCGTGCGGTG	850
851	GCACCTACAG	CAACGAGCGC	TACGCCGGCG	TCTGCGACCC	CGATGGCTGC	900
901	GACTTCAACT	CGTACCGCCA	GGGCAACAAG	ACCTTCTACG	GCAAGGGCAT	950
951	GACCGTCGAC	ACCACCAAGA	AGATCACTGT	CGTCACCCAG	TTCTCAAGG	1000
1001	ATGCCAACGG	CGATCTCGGC	GAGATCAAGC	GCTTCTACGT	CCAGGATGGC	1050
1051	AAGATCATCC	CCAAGTCCGA	GTCCACCATC	CCCGGCGTCG	AGGGCAATTC	1100
1101	CATCACCCAG	GACTGGTGCG	ACCGCCAGAA	GGTTGCCTTT	GGCGACATTG	1150
1151	ACGACTTCAA	CCGCAAGGGC	GGCATGAAGC	AGATGGGCAA	GGCCTCGCC	1200
1201	GGCCCCATGG	TCCTGGTCAT	GTCCATCTGG	GATGACCACG	CCTCCAACAT	1250
1251	GCTCFGGCTC	GACTCGACCT	TCCCTGTCGA	TGCCGCTGGC	AAGCCCGGCG	1300
1301	CCGAGCGCGG	TGCCTGCCCG	ACCACCTCGG	GTGTCCCTGC	TGAGGTTGAG	1350
1351	GCCGAGGCC	CCAACAGCAA	CGTCGTCTTC	TCCAACATCC	GCTTCGGCCC	1400
1401	CATCGGCTCG	ACCGTTGCTG	GTCTCCCCGG	CGCGGGCAAC	GGCGGCAACA	1450
1451	ACGGCGGCAA	CCCCCGGCC	CCCACCACCA	CCACCTCCTC	GGCTCCGGCC	1500
1501	ACCACCACCA	CCGCCAGCGC	TGGCCCCAAG	GCTGGCCGCT	GGCAGCAGTG	1550
1551	CGGCGGCATC	GGCTTCACTG	GCCCGACCCA	GTGCGAGGAG	CCCTACACTT	1600
1601	GCACCAAGCT	CAACGACTGG	TACTCTCAGT	GCCTGTAA		1638

**Figura 2: Secuencia de intrón putativo eliminada (GTACGTT....CAG = 413-472)
Proporciona la secuencia de ADNc (1578 nucleótidos)**

1	ATGCGTACCG	CCAAGTTCGC	CACCCTCGCC	GCCCTTGTGG	CCTCGGCCGC	50
51	CGCCAGCAG	GCGTGCAGTC	TCACCACCGA	GAGGCACCCT	TCCTCTCTTT	100
101	GGAAGAAGTG	CACCGCCGGC	GGCCAGTGCC	AGACCGTCCA	GGCTTCCATC	150
151	ACTCTCGACT	CCAACTGGCG	CTGGACTCAC	CAGGTGTCTG	GCTCCACCAA	200
201	CTGCTACACG	GGCAACAAGT	GGGATACTAG	CATCTGCACT	GATGCCAAGT	250
251	CGTGCCTCA	GAAGTGTGC	GTCGATGGTG	CCGACTACAC	CAGCACCTAT	300
301	GGCATACCA	CCAACGGTGA	TTCCCTGAGC	CTCAAGTTCG	TCACCAAGGG	350
351	CCAGCACTCG	ACCAACGTCG	GCTCGCGTAC	CTACCTGATG	GACGGCGAGG	400
401	ACAAGTATCA	GACCTTCGAG	CTCCTCGGCA	ACGAGTTCAC	C TTCGATGTC	450
451	GATGTCTCCA	ACATCGGCTG	CGGTCTCAAC	GGCGCCCTGT	ACTTCGTCTC	500
501	CATGGACGCC	GATGGTGGTC	TCAGCCGCTA	TCCTGGCAAC	AAGGCTGGTG	550
551	CCAAGTACGG	TACCGGCTAC	TGCGATGCTC	AGTGCCCCCG	TGACATCAAG	600
601	TTCATCAACG	GCGAGGCCAA	CATTGAGGGC	TGGACCGGCT	CCACCAACGA	650
651	CCCCAACGCC	GGCGCGGGCC	GCTATGGTAC	CTGCTGCTCT	GAGATGGATA	700
701	TCTGGGAAGC	CAACAACATG	GCTACTGCCT	TCACTCCTCA	CCCTTGCAAC	750
751	ATCATTGGCC	AGAGCCGCTG	CGAGGGCGAC	TCGTGCGGTG	GCACCTACAG	800
801	CAACGAGCGC	TACGCCGGCG	TCTGCGACCC	CGATGGCTGC	GACTTCAACT	850
851	CGTACCGCCA	GGCAACAAG	ACCTTCTACG	GCAAGGGCAT	GACCGTCGAC	900
901	ACCACCAAGA	AGATCACTGT	CGTCACCCAG	TTCCTCAAGG	ATGCCAACGG	950
951	CGATCTCGGC	GAGATCAAGC	GCTTCTACGT	CCAGGATGGC	AAGATCATCC	1000
1001	CCAACTCCGA	GTCCACCATC	CCCGGCCTCG	AGGGCAATTC	CATCACCAG	1050
1051	GACTGGTGGC	ACCGCCAGAA	GGTTGCCTTT	GGCGACATTG	ACGACTCAA	1100
1101	CCGCAAGGGC	GGCATGAAGC	AGATGGGCAA	GGCCCTCGCC	GGCCCCATGG	1150
1151	TCTTGGTCA	GTCCATCTGG	GATGACCACG	CCTCCAACAT	GCTCTGGCTC	1200
1201	GACTCGACCT	TCCCTGTGCA	TGCCGCTGGC	AAGCCCGGCG	CCGAGCGCGG	1250
1251	TGCCTGCCCG	ACCACCTCGG	GTGTCCCTGC	TGAGTTGAG	GCCGAGGCC	1300
1301	CCAACAGCAA	CGTCGTCTTC	TCCAACATCC	GCTTCGGCCC	CATCGGCTCG	1350
1351	ACCGTTGCTG	GTCTCCCCGG	CGCGGGCAAC	GGCGGCAACA	ACGGCGGCAA	1400
1401	CCCCCGCCCC	CCCACCACCA	CCACCTCCTC	GGCTCCGGCC	ACCACCACCA	1450
1451	CCGCCAGCGC	TGGCCCCAAG	GCTGGCCGCT	GGCAGCAGTG	CGGCGGCATC	1500
1501	GGCTTCACTG	GCCCGACCCA	GTGCGAGGAG	CCCTACACTT	GCACCAAGCT	1550
1551	CAACGACTGG	TACTCTCAGT	GCCTGTAA			1578

Figura 3: La traducción de la secuencia de ADNc proporciona la secuencia de proteínas (525 aminoácidos) del precursor de CBH1 de *H. grisea* var *thermoidea* (es decir, con secuencia señal):

1	<u>MRTAKFATLA</u> <u>ALVASAAAQ</u>	ACSLTTERHP	SLSWKKCTAG	GQCQTVQASI	50
51	TLDSNWRWTH	QVSGSTNCYT	GNKWDTSICT	DAKSCAQNCC	100
101	GITTNGDSL	LKFVTKGQHS	TNVGSRTYLM	DGEDKYQTFE	150
151	DVSNIGCGLN	GALYFVSM	DGGLSRYPGN	KAGAKYGTGY	200
201	FINGEANIEG	WTGSTNDPNA	GAGRYGTCCS	EMDIWEANNM	250
251	IIGQSRCEGD	SCGGTYSNER	YAGVCDPDGC	DFNSYRQGNK	300
301	TTKKITVVTQ	FLKDANGDLG	EIKRFYVQDG	KIIPNSESTI	350
351	DWCDRQKVAE	GDIDDFNRKG	GMKQMGKALA	GPMVLVMSIW	400
401	DSTFPVDAAG	KEGAERGACP	TTSGVPAEVE	AEAPNSNVVF	450
451	TVAGLPGAGN	GGNNGGNPPP	PTTTTSSAPA	TTTASAGPK	500
501	GFTGPTQCEE	PYTCTKLNDW	YSQCL		525

Figura 4: Secuencia de proteínas (507 aminoácidos) maduras (es decir, proteína expresada con la secuencia señal putativa eliminada)

1	QQACSLTTER	HPSLSWKKCT	AGGQCQTVQA	SITLDSNWRW	THQVSGSTNC	50
51	YTGKWDTSI	CTDAKSCAQN	CCVDGADYTS	TYGITTINGDS	LSLKFVTKGQ	100
101	HSTNVGSRTY	LMDGEDKYQT	FELLGNEFTF	DVDVSNIGCG	LNGALYFVSM	150
151	DADGGLSRYP	GNKAGAKYGT	GYCDAQCPRD	IKFINGEANI	EGWTGSTNDP	200
201	NAGAGRYGTC	CSEMDIWEAN	NMATAFTPHP	CTIIGQSRCE	GDSCGGTYSN	250
251	ERYAGVCDPD	GCDFNSYRQG	NKTFYGKGMT	VDTTKKITVV	TQFLKDANGD	300
301	LGEIKRFYVQ	DGKIIPNSES	TIPGVEGNSI	TQDWCDRQKV	AFGDIDDFNR	350
351	KGGMKQMGKA	LAGPMVLVMS	IWDDHASNML	WLDSTFPVDA	AGKPGAERGA	400
401	CPTTSGVPAE	VEAEAPNSNV	VFSNIRFGPI	GSTVAGLPGA	GNGGNNGGNP	450
451	PPPTTTTSSA	PATTTTASAG	PKAGRWQQCG	GIGFTGPTQC	EEPYTCTKLN	500
501	DWYSQCL					507

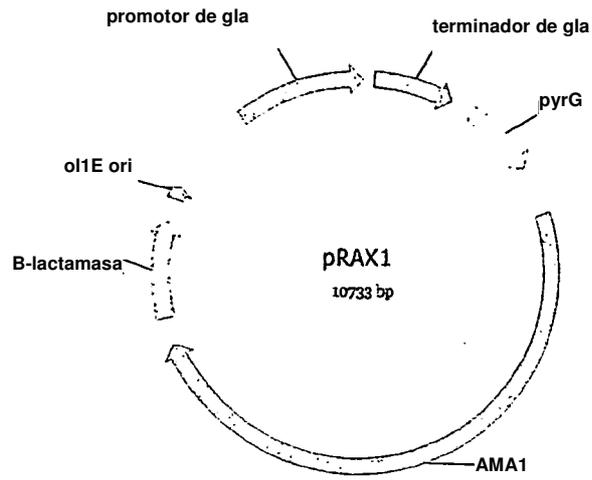


Figura 6: pRAX1

Figura 7: Vector de destino pRAXdes2 para la expresión en *A. niger*

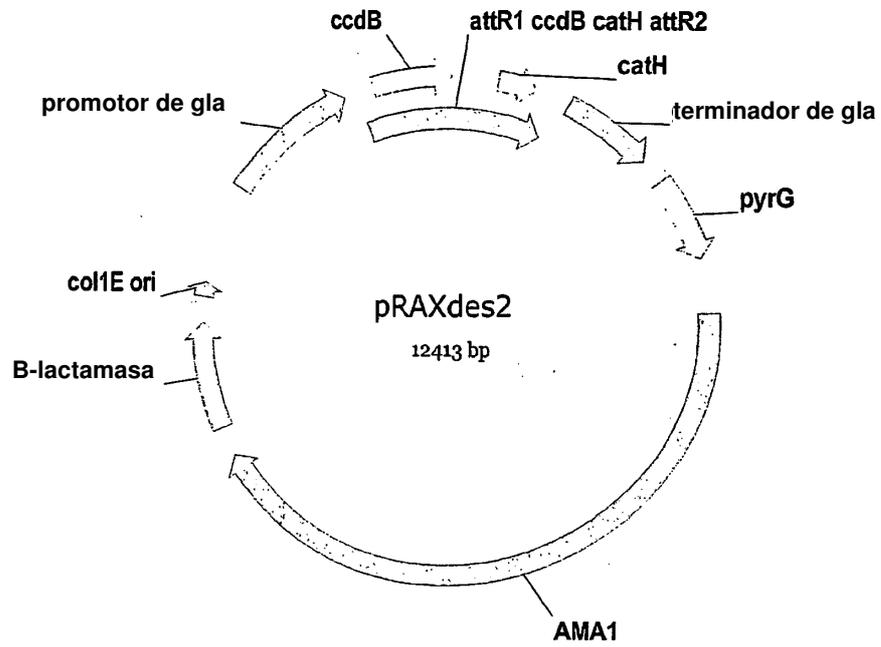
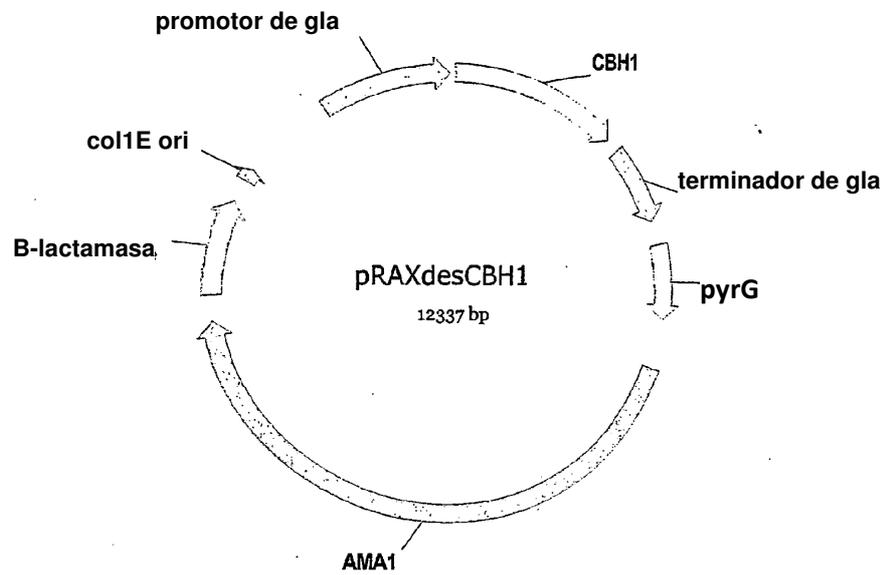
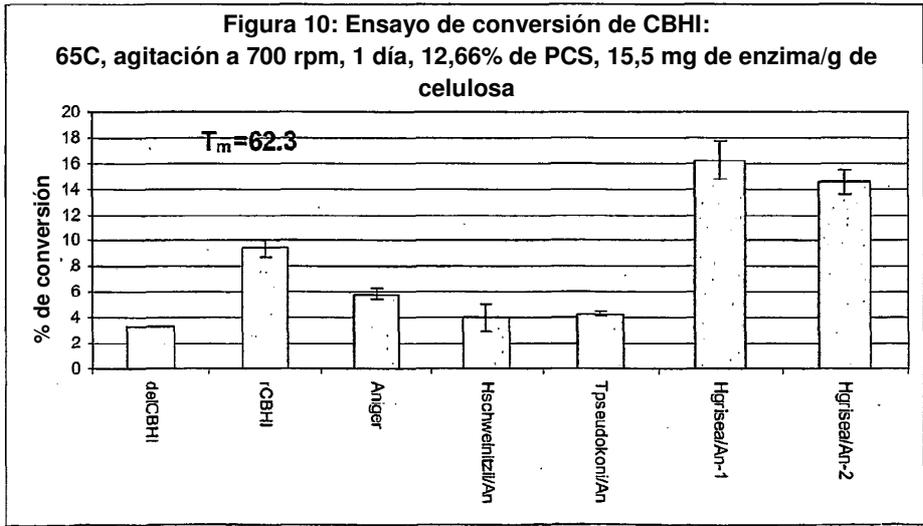
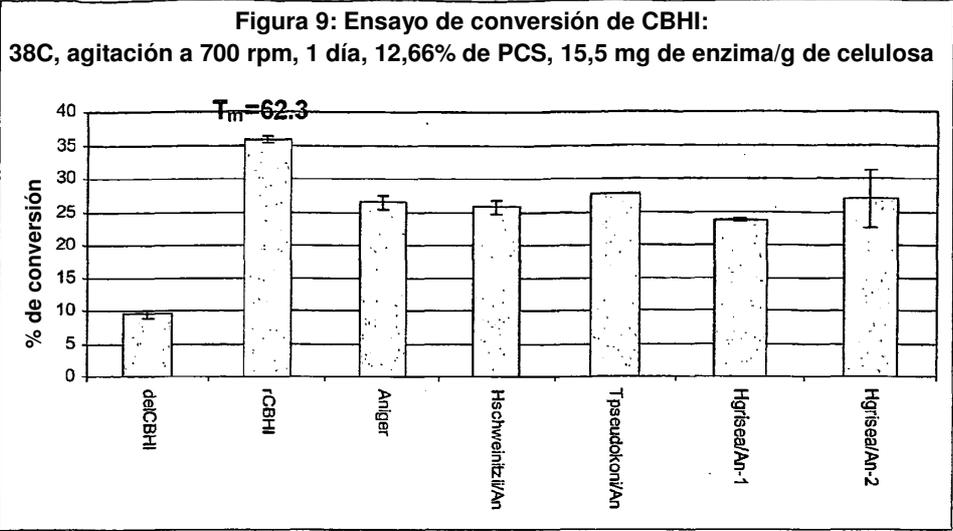


Figura 8: Vector pRAXdesCBH1 de expresión replicativa de genes CBH1 bajo control del promotor de glucoamilasa.





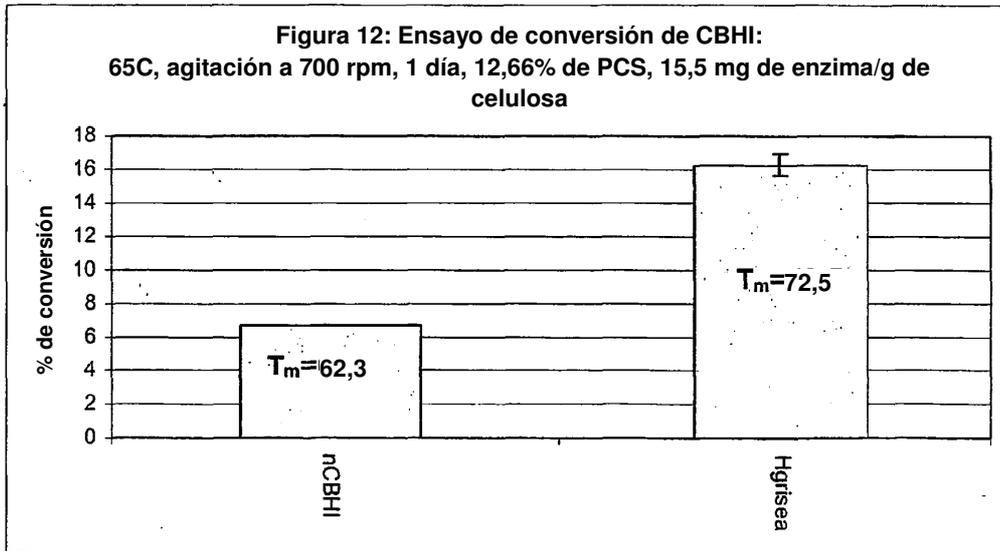
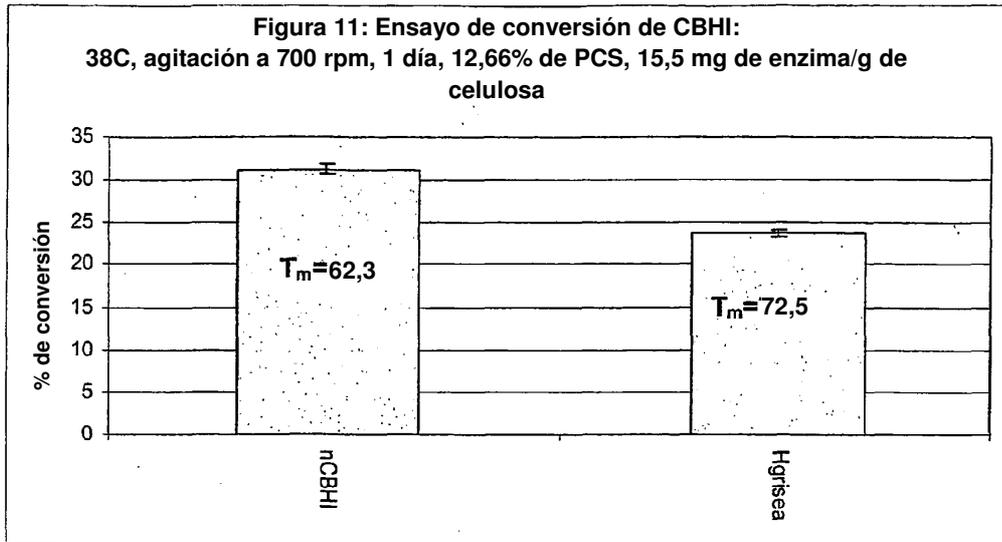


FIGURA 13A
diferencia de velocidad de 3 veces b/n *H. grisea* y rCBHI

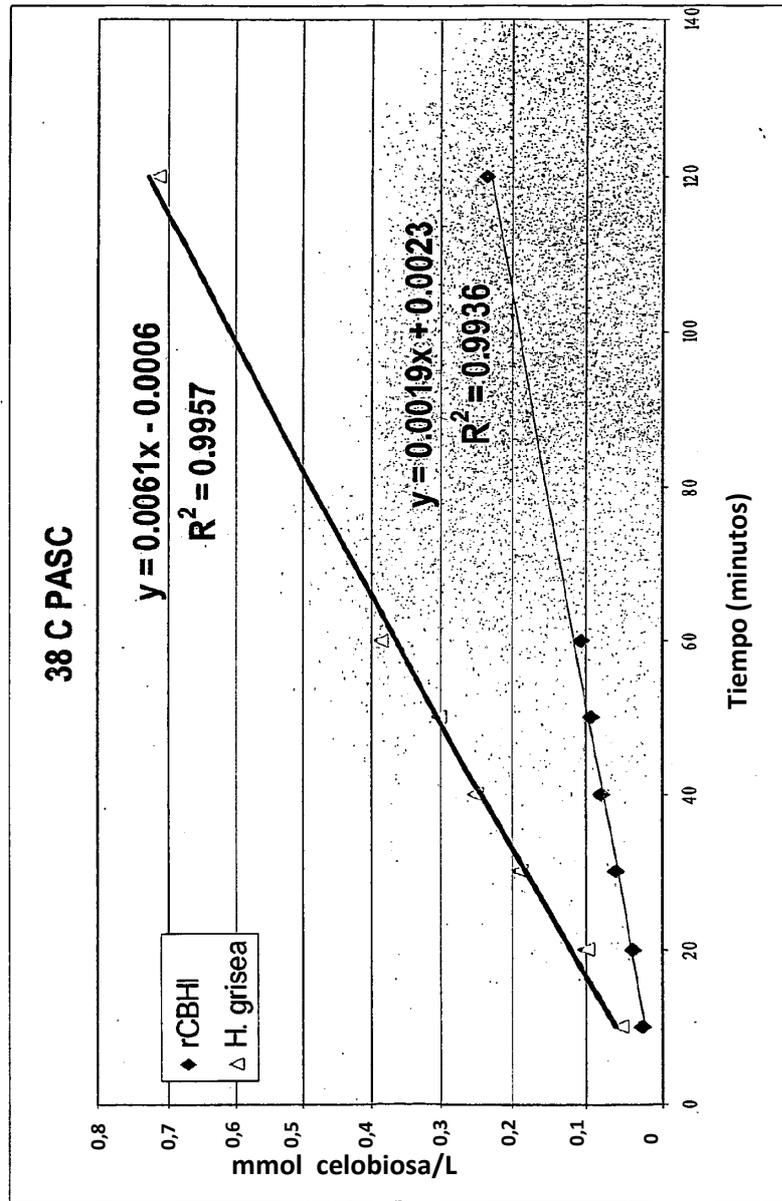


FIGURA 13B
diferencia de velocidad de 4,8 veces

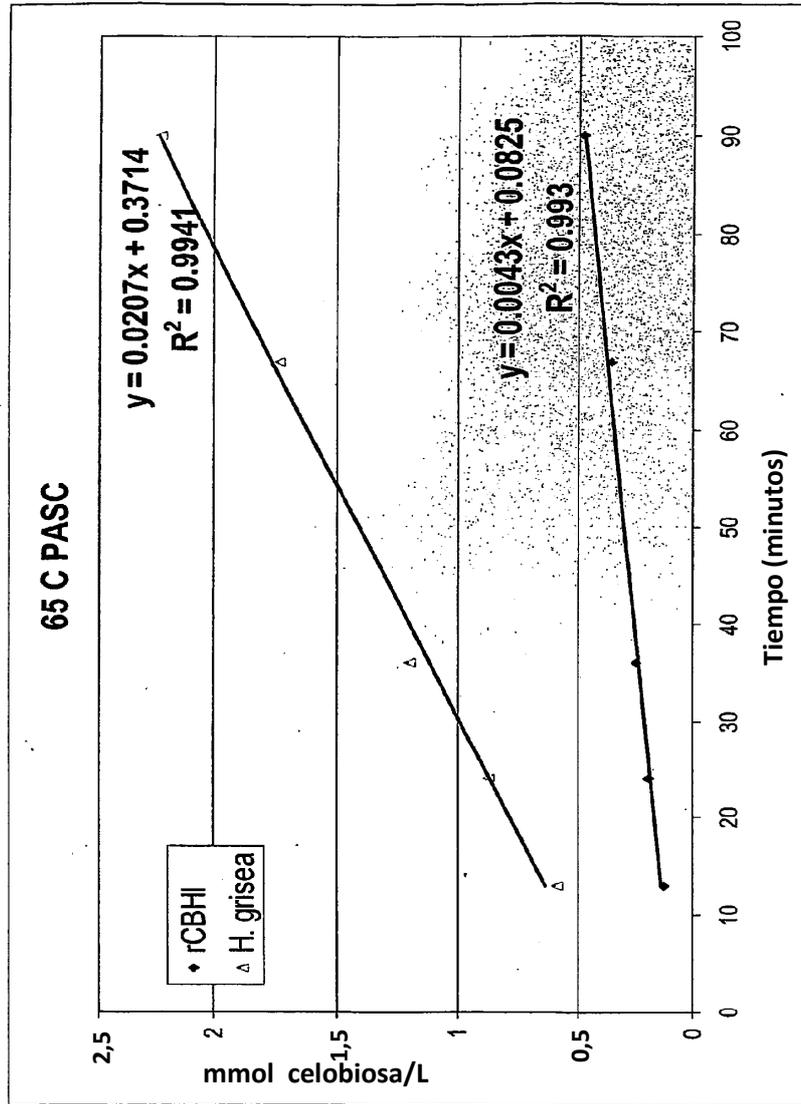
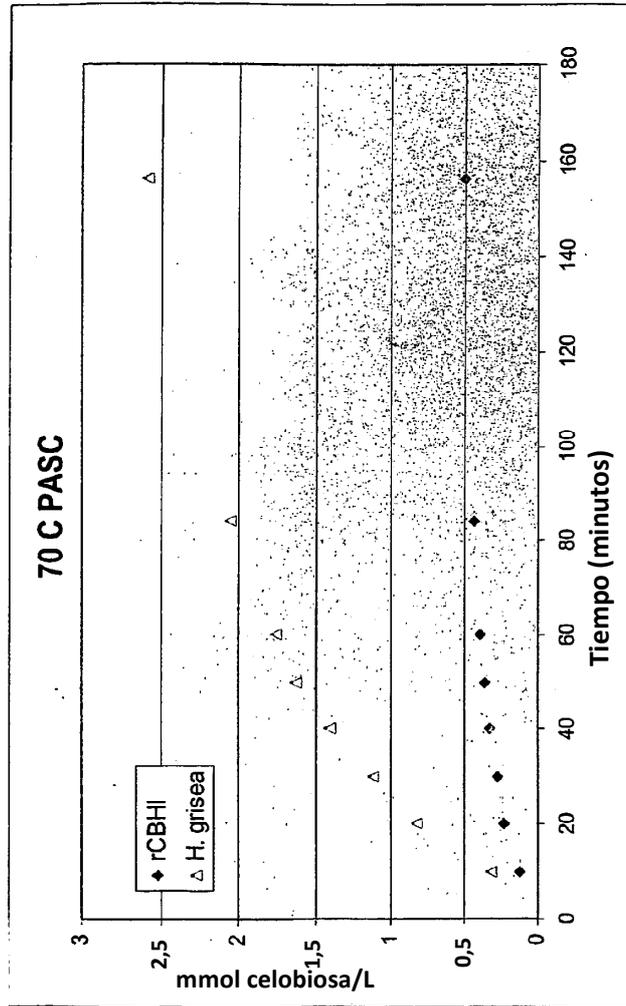


FIGURA 13C



```

1 ATGCGTACCG CCAAGTTCGC CACCCTCGCC GCCCTTGTGG CCTCGGCCGC
51 CGCCCAGCAG GCGTGCAGCC TCACCACCGA GAGGCACCCCT TCCCTCTCCT
101 GGAAGAAGTG CACCGCCGGC GGCCAGTGCC AGACCGTCCA GGCTTCCATC
151 ACTCTCGACT CCAACTGGCG CTGGACTCAC CAGGTGTCTG GCTCCACCAA
201 CTGCTACACG GGC AACGAGT GGGATTCTAG CATCTGCACT GATGCCAAGT
251 CGTGCCTCA GAACTGCTGC GTCGATGGTG CTGACTACAC CAGCACCTAT
301 GGCATCACCA CCAACGGTGA TTCCCTGAGC CTCAAGTTCG TCACCAAGGG
351 CCAGTACTCG ACCAACGTCG GCTCGCGTAC CTACCTGATG GACGGCGAGG
401 ACAAGTATCA GAGTAGGTTT TATCTTCAGC CTTCTCGCGC CTTGAATCCT
451 GGCTAACTTT TACACTTCAC AGCCTTCGAG CTCCTCGGCA ACGAGTTCAC
501 CTTTCGATGC GATGTCTCCA ACATCGGCTG CGGTCTCAAC GGGCCCTGT
551 ACTTCGTCTC CATGGACGCC GATGGTGGTC TCAGCCGCTA TCCTGGCAAC
601 AAGGCTGGTG CCAAGTACGG TACC CGCTAC TGCGATGCTC AGTGCCCCCG
651 TGACATCAAG TTCAATCAAG GCGAGGCCAA CATTGAGGGC TGGACCGGCT
701 CCACCAACGA CCCAACGCC GCGCGGGCC GCTATGGTAC CTGCTGCTCT
751 GAGATGGATA TCTGGGAGGC CAACAACATG GCTACTGCCT TCACTCCTCA
801 CCCTTGCACT ATCATTGGCC AGAGCCGCTG CGAGGGCGAC TCGTGCGGTG
851 GCACCTACAG CAACGACCGC TACGCCGGCG TCTGCGACCC CGATGGCTGC
901 GACTTCAACG CGTATCGCCA GGGCAACAAG ACCTTCTACG GCAAGGGCAT
951 GACCGTCGAC ACCACCAAGA AGCTCACCGT CGTCACCCAG TTCCTCAAGG
1001 ACGCCAACGG CGATCTCGGC GAGATCAAGC GCTTCTACGT CCAGGATGGG
1051 AAGATCATCC CCAACTCCGA GTCCACCATC CCCGGCGTCG AGGGCAACTC
1101 CATACCCAG GATTGGTGCG ACCGCCAGAA GGTTCCTTTT GGCGACATTG
1151 ACGACTTCAA CCGCAAGGGC GGCAATGAAGC AGATGGGCAA GGCCCTCGCC
1201 GGCCCATGG TCCTGGTCAT GTCCATCTGG GATGACCACG CCTCCAACAT
1251 GCTCTGGCTC GACTCGACCT TCCCTGTCTG TGCCGCTGGC AAGCCCGCGC
1301 CCGAGCGCGG TGCTGCCCCG ACCACCTCGG GTGTCCCTGC TGAGGTTGAG
1351 GCCGAGGCC CCAACAGCAA CGTCGTCTTC TCCAACATCC GCTTCGGCCC
1401 CATCGGCTCG ACCGTTGCCG GCCTTCCAG CGATGGCGGC AACACCGCGC
1451 GCAACACCAC CGTCCAGCCC CCGCCCAGCA CCACCACCAC CTCTGCCAGC
1501 AGCAGCACCA CCTCGGCTCC TGCCACCACC ACCACCGCCA GCGCTGGCCC
1551 CAAGGCTGGC CGCTGGCAGC AGTGGCGCGG CATCGGCTTC ACTGGCCCGA
1601 CCCAGTGC GAAGCCCTAC ACTTGCACCA AGCTCAACGA CTGGTACTCT
1651 CAGTGCCTGT AA

```

Figura 14A

ADN genómico de CBH1 de *Scytalidium thermophilum*

```

1  ATGCGTACCG CCAAGTTCGC CACCCTCGCC GCCCTTGTGG CCTCGGCCGC
   CGCCAGCAG GCGTGCAGCC TCACCACCGA GAGGCACCCT TCCCTCTCCT
101 GGAAGAAGTG CACCGCCGGC GGCCAGTGCC AGACCGTCCA GGCTTCCATC
   ACTCTCGACT CCAACTGGCG CTGGACTCAC CAGGTGTCTG GCTCCACCAA
201 CTGCTACACG GGCAACGAGT GGGATTCTAG CATCTGCACT GATGCCAAGT
   CGTGCCTCA GAACTGCTGC GTCGATGGTG CTGACTACAC CAGCACCTAT
301 GGCATCACCA CCAACGGTGA TTCCCTGAGC CTCAAGTTCG TCACCAAGGG
   CCAGTACTCG ACCAACGTCG GCTCGCGTAC CTACCTGATG GACGGCGAGG
401 ACAAGTATCA GACCTTCGAG CTCCTCGGCA ACGAGTTCAC CTTTCGATGC
   GATGTCTCCA ACATCGGCTG CGGTCTCAAC GGCGCCCTGT ACTTCGTCTC
501 CATGGACGCC GATGGTGGTC TCAGCCGCTA TCCTGGCAAC AAGGCTGGTG
   CCAAGTACGG TACCGGCTAC TCGCATGCTC AGTGCCCCCG TGACATCAAG
601 TTCATCAACG GCGAGGCCAA CATTGAGGGC TGGACCGGCT CCACCAACGA
   CCCAACGCC GCGCGGGGCC GCTATGGTAC CTGCTGCTCT GAGATGGATA
701 TCTGGGAGGC CAACAACATG GCTACTGCCT TCACTCCTCA CCCTTGCACT
   ATCATTGGCC AGAGCCGCTG CGAGGGCGAC TCGTGCGGTG GCACCTACAG
801 CAACGACCGC TACGCCGGCG TCTGCACCC CGATGGCTGC GACTTCAACG
   CGTATCGCCA GGGCAACAAG ACCTTCTACG GCAAGGGCAT GACCGTCGAC
901 ACCACCAAGA AGCTCACCGT CGTCACCCAG TTCCTCAAGG ACGCCAACGG
   CGATCTCGGC GAGATCAAGC GCTTCTACGT CCAGGATGGG AAGATCATCC
1001 CCAACTCCGA GTCCACCATC CCCGGCGTCG AGGGCAACTC CATCACCAG
   GATTGGTGCG ACCGCCAGAA GGTTCCTTT GCGGACATTG ACGACTTCAA
1101 CCGCAAGGGC GGCATGAAGC AGATGGGCAA GGCCCTCGCC GGCCCATGG
   TCCTGGTCAT GTCCATCTGG GATGACCACG CCTCCAACAT GCTCTGGCTC
1201 GACTCGACCT TCCCTGTGCG TGCCGCTGGC AAGCCCGGCG CCGAGCGCGG
   TGCCCTGCCC ACCACCTCGG GTGTCCCTGC TGAGGTTGAG GCCGAGGCC
1301 CCAACAGCAA CGTCGTCTTC TCCAACATCC GCTTCGGCCC CATCGGCTCG
   ACCGTTGCCG GCCTTCCCAG CGATGGCGGC AACACGGCG GCAACACCAC
1401 CGTCCAGCCC CCGCCAGCA CCACCACCAC CTCTGCCAGC AGCAGCACCA
   CCTCGGCTCC TGCCACCACC ACCACCGCCA GCGCTGGCCC CAAGGCTGGC
1501 CGTGGCAGC AGTGCGGCGG CATCGGCTTC ACTGGCCCGA CCCAGTGCGA
   GGAGCCCTAC ACTTGCACCA AGCTCAACGA CTGGTACTCT CAGTGCCTGT
1601 AA

```

Figura 14B
ADNc de CBH1 de *Scytalidium thermophilum*

ES 2 430 825 T3

1 MRTAKFATLAALVASAAAQQACSLTTERHPSLSWKKCTAGGQCQTVQASI 50
51 TLDSNWRWTHQVSGSTNCYTGNEWDSSICTDAKSCAQNCCVDGADYTSTY 100
101 GITTNGDSLKLVTKGQYSTNVGSRTYLMGDGEDKYQTFELLGNEFTFDV 150
151 DVSNIGCGLNGALYFVSMADGGLSRYPGNKAGAKYGTGYCDAQCPRDIK 200
201 FINGEANIEGWTGSTNDPNAGAGRYGTCCSEMDIWEANNMATAFTPHPCT 250
251 IIGQSRCEGDSCGGTYSNDRYAGVCDPDGCDFNAYRQGNKTFYKGGMTVD 300
301 TTKKLTVVVTQFLKDANGDLGEIKRFYVQDGKII PNSESTIPGVEGNSITQ 350
351 DWCDRQKVAFGDIDDFNRKGGMKQMGKALAGPMVLVMSIWDDHASNMLWL 400
401 DSTFPVDAAGKPGAERGACPTTSGVPAEVEAEAPNSNVVFSNIRFGPIGS 450
451 TVAGLPSDGGNNGGNTTVQPPPSTTTTSASSSTTSAPATTTTASAGPKAG 500
501 RWQCCGGIGFTGPTQCEEPYTCTKLNDWYSQCL- 534

Figura 14C
CBH1 de *Scytalidium thermophilum*, incluyendo la secuencia señal

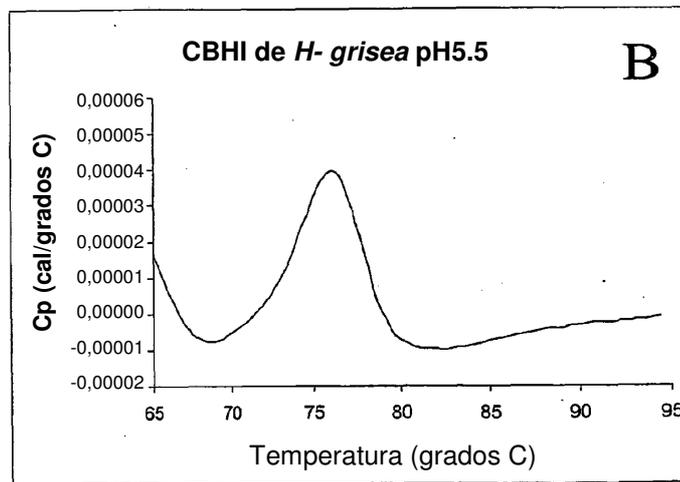
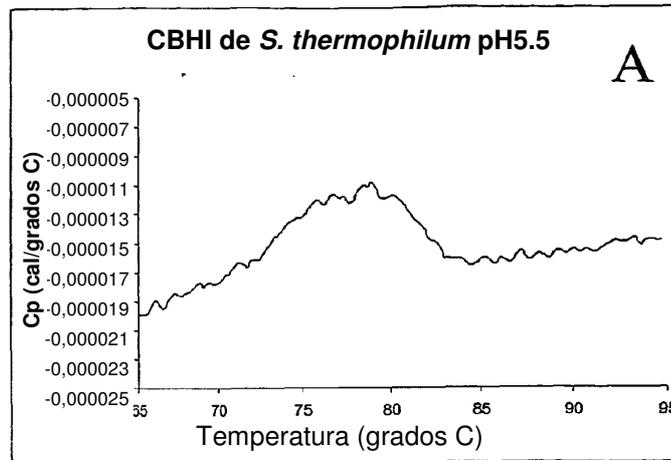


Figura 16