

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 619 902**

51 Int. Cl.:

C12N 9/42 (2006.01)

C12P 19/02 (2006.01)

C12P 19/14 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **03.06.2009 PCT/US2009/046159**

87 Fecha y número de publicación internacional: **10.12.2009 WO2009149202**

96 Fecha de presentación y número de la solicitud europea: **03.06.2009 E 09759361 (0)**

97 Fecha y número de publicación de la concesión europea: **21.12.2016 EP 2288698**

54 Título: **Composiciones y métodos que comprenden variantes de celulasa con afinidad reducida para materiales no celulósicos**

30 Prioridad:

06.06.2008 US 59506 P

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

27.06.2017

73 Titular/es:

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

72 Inventor/es:

**CASCAO-PEREIRA, LUIS, G.;
KAPER, THIJS;
KELEMEN, BRADLEY, R. y
LIU, AMY**

74 Agente/Representante:

RIZZO, Sergio

ES 2 619 902 T3

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

DESCRIPCIÓN

Composiciones y métodos que comprenden variantes de celulasa con afinidad reducida para materiales no celulósicos

III. CAMPO

- 5 **[0001]** La presente exposición se refiere a enzimas y, en concreto, a variantes de celulasa. Asimismo, se describen ácidos nucleicos que codifican las variantes de celulasa, composiciones que comprenden las variantes de celulasa, métodos de identificar variantes de celulasa útiles adicionales y métodos de utilizar las composiciones.

IV. ANTECEDENTES

- 10 **[0002]** La celulosa y la hemicelulosa son los materiales vegetales producidos por fotosíntesis más abundantes. Pueden ser degradados y utilizados como fuente de energía por numerosos microorganismos (p. ej., bacterias, levaduras y hongos) que producen enzimas extracelulares capaces de hidrolizar los sustratos poliméricos en azúcares monoméricos (Aro et al., J Biol Chem, 276: 24309-24314, 2001). Puesto que se acercan los límites de los recursos no renovables, el potencial de la celulosa para convertirse en un importante recurso de energía renovable es enorme (Krishna et al., Bioresource Tech, 77: 193-196, 2001). La utilización efectiva de la celulosa a través de procesos biológicos es un enfoque para superar la escasez de alimentos, piensos y combustibles (Ohmiya et al., Biotechnol Gen Engineer Rev, 14: 365-414, 1997).

- 20 **[0003]** Las celulasas son enzimas que hidrolizan celulosa (enlaces de glucano beta-1,4 o beta D-glucosídicos) dando lugar a la formación de glucosa, celobiosa, celooligosacáridos y similares. Tradicionalmente, las celulasas se han dividido en tres clases principales: endoglucanasas (EC 3.2.1.4) ("EG"), exoglucanasas o celobiohidrolasas (EC 3.2.1.91) ("CBH") y beta-glucosidasas ([beta]-D-glucósido glucohidrolasa; EC 3.2.1.21) ("BG"). (Knowles et al., TIBTECH 5: 255-261, 1987; y Schulein, Methods Enzymol, 160: 234-243, 1988). Las endoglucanasas actúan principalmente en las partes amorfas de la fibra de celulosa, mientras que las celobiohidrolasas son también capaces de degradar celulosa cristalina (Nevalainen y Penttila, Mycota, 303-319, 1995). Por consiguiente, es necesaria la presencia de una celobiohidrolasa en un sistema de celulasas para la solubilización eficaz de la celulosa cristalina (Suurnakki et al., Cellulose 7: 189-209, 2000). La beta-glucosidasa actúa para liberar unidades de D-glucosa de la celobiosa, celo-oligosacáridos y otros glucósidos (Freer, J Biol Chem, 268: 9337-9342, 1993).

- 30 **[0004]** Se sabe que las celulasas son producidas por un gran número de bacterias, levaduras y hongos. Algunos hongos producen un sistema de celulasas completo capaz de degradar formas cristalinas de celulosa, de manera que las celulasas se producen fácilmente en grandes cantidades a través de la fermentación. Los hongos filamentosos desempeñan un papel especial, puesto que muchas levaduras, tales como *Saccharomyces cerevisiae*, carecen de la capacidad para hidrolizar celulosa (Véase p. ej., Wood et al., Methods in Enzymology, 160: 87-116, 1988).

- 35 **[0005]** Las clasificaciones de las celulasas fúngicas de CBH, EG y BG se pueden ampliar adicionalmente para incluir múltiples componentes dentro de cada clasificación. Por ejemplo, se han aislado múltiples CBH, EG y BG de una variedad de fuentes fúngicas, incluyendo *Trichoderma reesei* (también denominada *Hypocrea jecorina*), que contiene genes conocidos para dos CBH, es decir, CBH I ("CBH1") y CBH II ("CBH2"), al menos 8 EG, es decir, EG I, EG II, EG III, EGIV, EGV, EGVI, EGVII y EGVIII, y al menos 5 BG, es decir, BG1, BG2, BG3, BG4 y BG5. EGIV, EGVII y EGVIII también tienen actividad de xiloglucanasa.

- 45 **[0006]** Con el fin de convertir de manera eficaz celulosa cristalina en glucosa, se necesita el sistema de celulasas completo que comprende componentes de cada una de las clasificaciones CBH, EG y BG, con componentes aislados menos eficaces en la hidrólisis de celulosa cristalina (Filho et al., Can J Microbiol, 42:1-5, 1996). Se ha observado una relación sinérgica entre componentes de celulasa de diferentes clasificaciones. En concreto, las celulasas de tipo EG y las celulasas de tipo CBH interactúan sinérgicamente para degradar celulosa de manera más eficiente.

- 50 **[0007]** Las celulasas se conocen en la técnica por ser útiles en el tratamiento de tejidos con el fin de mejorar la capacidad de limpieza de las composiciones detergentes, para la utilización como agente suavizante, para mejorar el tacto y la apariencia de los tejidos de algodón y similares (Kumar et al., Textile Chemist and Colorist, 29:37-42, 1997). Se han descrito composiciones detergentes que contienen celulasas con eficacia limpiadora mejorada (patente estadounidense con n.º 4,435,307; solicitudes de patente GB con n.º 2,095,275 y 2,094,826) y para su utilización en el tratamiento de tejido para mejorar el tacto y la apariencia del tejido (patentes

estadounidenses con n.º 5,648,263, 5,691,178 y 5,776,757; y solicitud de patente GB con n.º 1,358,599). Por lo tanto, las celulasas producidas en hongos y bacterias han recibido una atención significativa. En concreto, la fermentación de *Trichoderma spp.* (p. ej., *Trichoderma longibrachiatum* o *Trichoderma reesei*) produce un sistema de celulasas completo capaz de degradar formas cristalinas de la celulosa.

5 **[0008]** WO 2006/074005 describe variantes de celulasas CBH2 de *H. jecorina* que tienen termoestabilidad y actividad alterada. WO 2004/056981 describe polipéptidos que tienen actividad de celobiohidrolasa II y polinucleótidos que los codifican. EP1726644A1 describe variantes de celulasa y métodos para mejorar las propiedades de las enzimas celulolíticas mediante la sustitución, delección o inserción de aminoácidos. WO
10 2008/153925 describe métodos para mejorar múltiples propiedades de proteínas, incluyendo métodos para producir enzimas superiores para una determinada aplicación a través del cribado de un conjunto abreviado de enzimas candidatas.

[0009] Si bien se han descrito anteriormente composiciones de celulasa, se siguen necesitando composiciones de celulasa nuevas y mejoradas. Las composiciones de celulosa mejoradas se utilizan en detergentes domésticos, tratamientos textiles, conversión de biomasa y fabricación de papel. Las celulasas que presentan un
15 rendimiento mejorado son especialmente interesantes.

V. SUMARIO

[0010] La presente instrucción se refiere a variantes de celulasa modificadas para reducir la unión a materiales no celulósicos. En general, las variantes de celulasa tienen una actividad celulolítica aumentada en presencia de materiales no celulósicos en comparación con celulasas de tipo salvaje. En consecuencia, en un primer aspecto
20 la presente invención proporciona una variante de celulasa aislada, donde dicha variante es una forma madura que tiene actividad de celulasa, y donde dicha variante es una variante de una celulasa original cuya secuencia de aminoácidos es al menos un 97 % idéntica a la secuencia de aminoácidos de SEQ ID NO: 3, consistiendo la variación en una sustitución en la posición 63, numerada por correspondencia con la secuencia de aminoácidos de SEQ ID NO:3, y donde la sustitución hace que la variante de celulasa tenga una carga neta más negativa en
25 comparación con la secuencia de aminoácidos de SEQ ID NO: 3. En algunos casos, la sustitución comprende la eliminación de una carga positiva. En concreto, la sustitución puede ser R63Q numerada por correspondencia con la secuencia de aminoácidos de SEQ ID NO:3.

[0011] En algunos casos, la variante puede ser la variante de sustitución R63Y o la R63Q, R63A, R63C, R63D, R63E, R63F, R63G, R63I, R63L, R63M, R63N, R63P, R63S, R63T, R63V, R63W de la secuencia de
30 aminoácidos de SEQ ID NO:3.

[0012] En algunos casos, la carga neta más negativa puede ser -1 o -2 en comparación con la secuencia de aminoácidos de SEQ ID NO:3.

[0013] En un segundo aspecto, la presente invención proporciona un método de convertir biomasa en azúcares comprendiendo poner en contacto dicha biomasa con la variante de celulasa del primer aspecto de la invención.

35 **[0014]** En un tercer aspecto, la presente invención proporciona un método de producir un combustible que comprende:

poner en contacto una composición de biomasa con una composición enzimática que comprende la variante de celulasa del primer aspecto de la invención para producir una solución de azúcar; y
40 cultivar la solución de azúcar con un microorganismo fermentativo en condiciones suficientes para producir un combustible.

VI. BREVE DESCRIPCIÓN DE LOS DIBUJOS

[0015]

La figura 1 ilustra la sacarificación de BPA mediante preparaciones de celulasa de *Trichoderma sp.* modificadas (cuadrados) y no modificadas (círculos) en presencia de cantidades crecientes de inhibidor de lignina. La figura 1A y la figura 1B muestran resultados tras incubaciones de 24 y 48 horas, respectivamente.
45

La figura 2A ilustra la sacarificación con la comparación de celulasas modificadas y la figura 2B muestra la diferencia de sacarificación mediante la utilización de celulasas modificadas y no modificadas.

La figura 3 proporciona un alineamiento de las secuencias de aminoácidos de la forma madura de distintas celulasas: *Hypocrea jecorina* (también conocida como *T. reesei*) CBH2 (SEQ ID NO:3), *Hypocrea koningii* CBH2 (SEQ ID NO:4), *Humicola insolens* CBH2 (SEQ ID NO:5), *Acremonium cellulolyticus* CBH2 (SEQ ID NO:6), *Agaricus bisporus* CBH2 (SEQ ID NO:7), *Fusarium osyosporum* EG (SEQ ID NO:8), *Phanerochaete chrysosporium* CBH2 (SEQ ID NO:9), *Talaromyces emersonii* CBH2 (SEQ ID NO:10), *Thermobifida fusca* 6B/E3 CBH2 (SEQ ID NO:11), *Thermobifida fusca* 6A/E2 EG (SEQ ID NO:12), y *Cellulomonas fimi* CenA EG (SEQ ID NO:13).

La figura 4 proporciona un gráfico de la frecuencia relativa de los ganadores de las bibliotecas de evaluación de sitio (SEL, por sus siglas en inglés) de la variante CBH2 del ensayo de rastreo de maíz pretratado (PCS, por sus siglas en inglés) observados por encima de lo previsto, como resultado del cambio de carga. La disminución de la carga de CBH2 produce una frecuencia considerablemente mayor de ganadores PCS.

La figura 5 proporciona un mapa de plásmido de pTTTpyr-cbh2.

VII. DESCRIPCIÓN DETALLADA DE DIVERSOS MODOS DE REALIZACIÓN

[0016] La presente instrucción se refiere a variantes de celulasa modificadas para reducir la unión a materiales no celulósicos. En general, la variante de celulasa tiene una actividad celulolítica aumentada en presencia de materiales no celulósicos en comparación con la celulasa de tipo salvaje. La variante de celulasa es como se define en las reivindicaciones.

[0017] Se debe observar que tanto la descripción general anterior como la descripción detallada siguiente tienen carácter ilustrativo y explicativo únicamente y no restringen las composiciones y métodos descritos en el presente documento. A menos que se definan de otra forma en el presente documento, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que el entendido comúnmente por un experto en la materia a la que pertenece la presente exposición. En la presente solicitud, la utilización del singular incluye el plural a menos que se indique específicamente otra cosa. La utilización de "o" significa "y/o" a menos que se indique otra cosa. De forma similar, los términos "que comprende(n)", "que incluye(n)" no pretenden ser limitativos. Los encabezamientos proporcionados en el presente documento no son limitaciones de los diversos aspectos o modos de realización de la exposición que se pueden tomar como referencia a la memoria en general. Por consiguiente, los términos en el presente documento se definen de forma más completa en referencia a la memoria en general.

[0018] A menos que se definan de otra forma en el presente documento, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que el entendido comúnmente por un experto en la materia a la que pertenece la presente exposición. Singleton, et al., DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY, 2ª ED., John Wiley and Sons, New York (1994), y Hale & Marham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, NY (1991) proporcionan a un experto en la materia un diccionario general de muchos de los términos utilizados en la presente exposición. Aunque en la práctica o en el ensayo de la presente exposición se puede utilizar cualquier método y material similar o equivalente a los descritos en el presente documento, se describen los métodos y los materiales preferidos. Los intervalos numéricos incluyen los números que definen el intervalo. A menos 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. Se dirige a los profesionales en concreto a Sambrook et al., MOLECULAR CLONING: A LABORATORY MANUAL (Segunda edición), Cold Spring Harbor Press, Plainview, N.Y., 1989, y Ausubel FM et al., Current Protocols in Molecular Biology, John Wiley & Sons, New York, NY, 1993, para definiciones y términos de la técnica. Ha de entenderse que la presente exposición no se limita a la metodología, los protocolos y los reactivos concretos descritos, puesto que estos pueden variar.

I. Definiciones

[0019] Los términos que aparecen a continuación se definen de forma más completa en referencia a la memoria en general.

[0020] El término "polipéptido" como se utiliza en el presente documento se refiere a un compuesto formado por una sola cadena de residuos de aminoácidos unidos por enlaces peptídicos. El término "proteína" como se utiliza en el presente documento puede ser sinónimo del término "polipéptido".

[0021] "Variante" significa una proteína que se deriva de una proteína precursora (p. ej. la proteína nativa) mediante la adición de uno o más aminoácidos a cualquiera de los extremos C- y N-terminal o a ambos, la sustitución de uno o más aminoácidos en un sitio o en un conjunto de sitios diferentes en la secuencia de aminoácidos, o la delección de uno o más aminoácidos en cualquiera de los extremos o en ambos extremos de la proteína o en uno o más sitios en la secuencia de aminoácidos, o se modifican uno o más aminoácidos mediante el cambio de la carga (esto es, mediante la eliminación de una carga positiva, la adición de una carga negativa o mediante ambas). La preparación de una variante de celulasa puede realizarse mediante cualquier medio conocido en la técnica, incluyendo la modificación química de los aminoácidos, mediante la modificación de una secuencia de ADN que codifica la proteína nativa, la transformación de la secuencia de ADN modificada en un huésped adecuado y la expresión de la secuencia de ADN modificada para formar la variante de enzima. La variante de celulasa de la exposición incluye péptidos que comprenden secuencias de aminoácidos alteradas en comparación con una secuencia de aminoácidos de enzima precursora donde la variante de celulasa retiene la naturaleza celulolítica característica de la enzima precursora, pero puede tener propiedades alteradas en algunos aspectos específicos. Por ejemplo, una variante de celulasa puede tener un pH óptimo aumentado o temperatura aumentada o estabilidad oxidativa o afinidad disminuida o unión a materiales no celulósicos, pero mantendrá su actividad celulolítica característica. Se contempla que las variantes según la presente exposición se pueden derivar de un fragmento de ADN que codifica una variante de celulasa donde se conserva la actividad funcional de la variante de celulasa expresada. Por ejemplo, un fragmento de ADN que codifica una celulasa puede incluir además una secuencia de ADN o una parte de la misma que codifica una bisagra o conector unido a la secuencia de ADN de celulasa en cualquiera de los extremos 5' o 3' donde se conserva la actividad funcional del dominio de celulasa codificado. Los términos variante y derivado se pueden utilizar indistintamente en el presente documento.

[0022] Los "residuos equivalentes" también se pueden definir determinando la homología al nivel de la estructura terciaria para una celulasa precursora cuya estructura terciaria se ha determinado mediante cristalografía de rayos X. Los residuos equivalentes se definen como aquellos para los que las coordenadas atómicas de dos o más átomos de la cadena principal de un residuo de aminoácido concreto de una celulasa y CBH2 de *Hypocrea jecorina* (N en N, CA en CA, C en C y O en O) están dentro de 0,13 nm y preferiblemente 0,1 nm después del alineamiento. El alineamiento se consigue después de haber orientado y posicionado el mejor modelo para proporcionar el máximo solapamiento de coordenadas atómicas de átomos de proteína que no son hidrógeno de la celulasa en cuestión con respecto a la CBH2 de *H. jecorina*. El mejor modelo es el modelo cristalográfico que proporciona el factor R más bajo para datos de difracción experimental a la resolución más alta disponible. Véase para ejemplos la patente estadounidense con n.º 2006/0205042.

[0023] Los residuos equivalentes que son funcionalmente análogos a un residuo específico de CBH2 de *H. jecorina* se definen como aquellos aminoácidos de una celulasa que pueden adoptar una conformación de modo que alteren, modifiquen o contribuyan a la estructura de proteína, unión de sustrato o catálisis de una manera definida y atribuida a un residuo específico de CBH2 de *H. jecorina*. Además, son aquellos residuos de la celulasa (para los que se ha conseguido una estructura terciaria mediante cristalografía de rayos X) que ocupan una posición análoga en la medida en que, aunque los átomos de la cadena principal del residuo determinado puedan no satisfacer los criterios de equivalencia sobre la base de ocupar una posición homóloga, las coordenadas atómicas de al menos dos de los átomos de la cadena lateral del residuo se sitúen con 0,13 nm de los átomos de la cadena lateral correspondientes de CBH2 de *H. jecorina*. La estructura cristalina de CBH2 de *H. jecorina* se muestra en Zou et al. (1999) (Ref. 5, supra).

[0024] El término "molécula de ácido nucleico" incluye moléculas de ARN, ADN y ADNc. Se entenderá que, como resultado de la degeneración del código genético, se pueden producir una multitud de secuencias de nucleótidos que codifican una proteína determinada, tal como CBH2 y/o variantes de la misma. La presente exposición contempla todas las posibles variantes de secuencias de nucleótidos, que codifican una variante de celulasa, tal como CBH2, las cuales son todas posibles debido a la degeneración del código genético.

[0025] Un constructo o secuencia de ácido nucleico "heterólogo" tiene una parte de la secuencia que no es nativa a la célula en la que se expresa. Heterólogo, con respecto a una secuencia de control, se refiere a una secuencia de control (es decir, promotor o potenciador) que no funciona en la naturaleza para regular el mismo gen cuya expresión está regulando actualmente. En general, las secuencias de ácido nucleico heterólogas no son endógenas a la célula o parte del genoma en el que están presentes, y se han añadido a la célula mediante infección, transfección, transformación, microinyección, electroporación o similares. Un constructo de ácido nucleico "heterólogo" puede contener una combinación de secuencia de control/secuencia codificadora de ADN que sea la misma que una combinación de secuencia de control/secuencia codificadora de ADN encontrada en la célula nativa, o diferente de la misma.

[0026] Como se utiliza en el presente documento, el término "vector" se refiere a un constructo de ácido nucleico diseñado para la transferencia entre células huésped diferentes. Un "vector de expresión" se refiere a un vector

que tiene la capacidad de incorporar y expresar fragmentos de ADN heterólogo en una célula extraña. Muchos vectores de expresión procariotas y eucariotas están comercialmente disponibles. La selección de vectores de expresión adecuados está dentro del conocimiento de los expertos en la materia.

5 **[0027]** Por consiguiente, un "casete de expresión" o "vector de expresión" es un constructo de ácido nucleico generado de forma recombinante o sintética, con una serie de elementos de ácido nucleico especificados que permiten la transcripción de un ácido nucleico concreto en una célula diana. El casete de expresión recombinante se puede incorporar en un plásmido, cromosoma, ADN mitocondrial, ADN plastidial, virus, o fragmento de ácido nucleico. Normalmente, la parte del casete de expresión recombinante de un vector de expresión incluye, entre otras secuencias, una secuencia de ácido nucleico para ser transcrita y un promotor.

10 **[0028]** Como se utiliza en el presente documento, el término "plásmido" se refiere a un constructo de ADN bicatenario circular utilizado como un vector de clonación, y que forma un elemento genético autorreplicante extracromosómico en muchas bacterias y algunos eucariotas.

15 **[0029]** Como se utiliza en el presente documento, el término "secuencia de nucleótidos que codifica un marcador seleccionable" se refiere a una secuencia de nucleótidos que es capaz de expresión en células y donde la expresión del marcador seleccionable confiere a las células que contienen el gen expresado la capacidad de crecer en presencia de un agente selectivo correspondiente, o bajo condiciones de crecimiento selectivo correspondientes.

20 **[0030]** Como se utiliza en el presente documento, el término "promotor" se refiere a una secuencia de ácido nucleico que funciona para dirigir la transcripción de un gen en dirección 3'. El promotor será en general adecuado para la célula huésped en la que se está expresando el gen diana. El promotor, junto con otras secuencias de ácido nucleico reguladoras de la transcripción o la traducción, (también denominadas "secuencias de control") son necesarios para expresar un gen determinado. En general, las secuencias reguladoras de la transcripción y la traducción incluyen, pero sin carácter limitativo, secuencias promotoras, sitios de unión al ribosoma, secuencias de inicio y de detención de la transcripción, secuencias de inicio y detención de la traducción y secuencias potenciadoras o activadoras.

25

[0031] "Gen quimérico" o "constructo de ácido nucleico heterólogo", como se define en el presente documento, se refiere a un gen no nativo (es decir, uno que se ha introducido en un huésped) que puede estar compuesto por partes de genes diferentes, incluyendo elementos reguladores. Un constructo de gen quimérico para la transformación de una célula huésped se compone normalmente de una región reguladora de la transcripción (promotor) operativamente unida a una secuencia codificadora de proteínas heterólogas, o en un gen quimérico marcador seleccionable, a un gen marcador seleccionable que codifica una proteína que confiere, por ejemplo, resistencia a los antibióticos a las células transformadas. Un gen quimérico típico de la presente exposición, para la transformación en una célula huésped, incluye una región reguladora de la transcripción que es constitutiva o inducible, una secuencia codificadora de proteína y una secuencia de terminación. Un constructo de gen quimérico puede incluir también una segunda secuencia de ADN que codifica un péptido señal si se desea la secreción de la proteína diana.

30

35

[0032] Un ácido nucleico está "operativamente unido" cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, el ADN que codifica un líder secretor está operativamente unido a ADN para un polipéptido si se expresa como una preproteína que participa en la secreción del polipéptido; un promotor o potenciador está operativamente unido a una secuencia codificadora si afecta a la transcripción de la secuencia; o un sitio de unión al ribosoma está operativamente unido a una secuencia codificadora si se posiciona de manera que facilite la traducción. En general, "operativamente unido" significa que las secuencias de ADN que están unidas son contiguas, y, en el caso de un líder secretor, contiguas y en el marco de lectura. No obstante, los potenciadores no tienen que ser contiguos. La unión se consigue mediante ligadura en sitios de restricción convenientes. Si dichos sitios no existen, los ligadores, cebadores o adaptadores oligonucleótidos sintéticos para PCR se utilizan de conformidad con la práctica convencional.

40

45

[0033] Como se utiliza en el presente documento, el término "gen" significa el segmento de ADN que participa en la producción de una cadena de polipéptidos, que puede o no incluir regiones precedentes y posteriores a la región codificadora, por ejemplo, secuencias 5' no traducidas (5' UTR) o secuencias "líder" y secuencias 3' UTR o secuencias "tráiler", así como secuencias intercaladas (intrones) entre segmentos codificadores individuales (exones).

50

[0034] En general, las moléculas de ácido nucleico que codifican la variante de celulasa tal como CBH2 se hibridarán, bajo condiciones de astringencia moderadas a altas a la secuencia de tipo salvaje, tal como se proporciona en el presente documento como SEQ ID NO:1. No obstante, en algunos casos se emplea una

secuencia de nucleótidos que codifica CBH2 que posee una utilización de codón sustancialmente diferente, mientras que la proteína codificada por la secuencia de nucleótidos que codifica CBH2 tiene la secuencia de aminoácidos igual o sustancialmente igual que la de la proteína nativa. Por ejemplo, la secuencia codificadora se puede modificar para facilitar la expresión más rápida de CBH2 en un sistema concreto de expresión procariota o eucariota, de conformidad con la frecuencia con la que el huésped utiliza un codón concreto (Te'o et al., FEMS Microbiology Letters, 190: 13-19, 2000, por ejemplo, describe la optimización de genes para la expresión en hongos filamentosos).

[0035] Se considera que una secuencia de ácido nucleico "puede hibridarse de forma selectiva" con una secuencia de ácido nucleico de referencia si las dos secuencias de forma específica se hibridan mutuamente en condiciones de lavado e hibridación de astringencia moderada a elevada. Las condiciones de hibridación se basan en la temperatura de fusión (T_f) de la sonda o del complejo de unión a ácido nucleico. Por ejemplo, la "astringencia máxima" ocurre normalmente a aproximadamente $T_f - 5\text{ }^\circ\text{C}$ ($5\text{ }^\circ\text{C}$ por debajo de la T_f de la sonda); la "astringencia elevada" a aproximadamente $5\text{ }^\circ\text{C} - 10\text{ }^\circ\text{C}$ por debajo de la T_f ; la "astringencia moderada" o "intermedia" a aproximadamente $10\text{ }^\circ\text{C} - 20\text{ }^\circ\text{C}$ por debajo de la T_f de la sonda; y la "astringencia baja" a aproximadamente $20\text{ }^\circ\text{C} - 25\text{ }^\circ\text{C}$ por debajo de la T_f . Funcionalmente, las condiciones de astringencia máxima se pueden utilizar para identificar secuencias que tienen identidad estricta o identidad casi estricta con la sonda de hibridación; mientras que las condiciones de astringencia elevada se utilizan para identificar secuencias que tienen aproximadamente 80 % o más de identidad de secuencia con la sonda.

[0036] Las condiciones de hibridación de moderada y elevada astringencia se conocen bien en la técnica (véase, por ejemplo, Sambrook, et al, 1989, Capítulos 9 y 11, y en Ausubel, F.M., et al, 1993). Un ejemplo de condiciones de astringencia elevada incluye hibridación a aproximadamente $42\text{ }^\circ\text{C}$ en formamida al 50 %, 5x SSC, 5 veces de solución Denhardt, SDS al 0,5 % y 100 $\mu\text{g/ml}$ ADN portador desnaturalizado seguido del lavado dos veces en 2 veces SSC y SDS al 0,5 % a temperatura ambiente y dos veces adicionales en 0,1x SSC y SDS al 0,5 % a $42\text{ }^\circ\text{C}$.

[0037] El término "recombinante", cuando se utiliza con referencia a, p. ej., una célula o ácido nucleico, proteína, o vector, indica que la célula, ácido nucleico, proteína o vector se han modificado por la introducción de un ácido nucleico o proteína heterólogos o la alteración de un ácido nucleico o proteína nativos, o que la célula se deriva de una célula modificada de dicha forma. De esta manera, por ejemplo, las células recombinantes expresan genes que no se encuentran en la forma nativa (no recombinante) de la célula o expresan genes nativos que se expresan anormalmente de otro modo, se subexpresan o no se expresan en absoluto.

[0038] Como se utilizan en el presente documento, los términos "transformada", "establemente transformada" o "transgénica" con referencia a una célula significan que la célula tiene una secuencia de ácido nucleico no nativa (heteróloga) integrada en su genoma o como un plásmido episomal que se mantiene a través de múltiples generaciones.

[0039] Como se utiliza en el presente documento, el término "expresión" se refiere al proceso por el cual se produce un polipéptido basándose en la secuencia de ácido nucleico de un gen. El proceso incluye tanto transcripción como traducción.

[0040] 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 referencia a la incorporación de una secuencia de ácido nucleico en una célula eucariota o procariota en la que la secuencia de ácido nucleico puede incorporarse en el genoma de la célula (por ejemplo, cromosoma, plásmido, plástido o ADN mitocondrial), convertirse en un replicón autónomo o expresarse transitoriamente (por ejemplo, ARNm transfectado).

[0041] Se deduce que el término "expresión de CBH2" se refiere a la transcripción y traducción del gen *cbh2* o variantes del mismo, cuyos productos incluyen ARN precursor, ARNm, polipéptido, polipéptidos procesados tras la traducción y derivados de los mismos, que incluyen CBH2 de especies relacionadas, tales como *Trichoderma koningii*, *Hypocrea jecorina* (también conocida como *Trichoderma longibrachiatum*, *Trichoderma reesei* o *Trichoderma viride*) e *Hypocrea schweinitzii*. A modo de ejemplo, los ensayos para la expresión de CBH2 incluyen transferencia Western para proteína CBH2, análisis de transferencia Northern y ensayos de reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR, por sus siglas en inglés) para ARNm de *cbh2*, y ensayos de celulosa hinchada con ácido fosfórico (PASC, por sus siglas en inglés) y de ácido p-hidroxibenzoico hidrazida (PAHBAH, por sus siglas en inglés) como se describe en lo que sigue: (a) PASC: (Karlsson, J. et al. (2001), Eur. J. Biochem, 268, 6498-6507, Wood, T. (1988) en Methods in Enzymology, Vol.160. Biomass Part a Cellulose and Hemicellulose (Wood, W. & Kellog, S. Eds.), pp.19-25, Academic Press, San Diego, Calif., USA) y (b) PAHBAH: (Lever, M. (1972) Analytical Biochemistry, 47, 273, Blakeney, A.B. & Mutton, L.L. (1980) Journal of Science of Food and Agriculture, 31, 889, Henry, R.J. (1984) Journal of the Institute of Brewing, 90, 37).

- 5 **[0042]** El término “corte y empalme alternativo” se refiere al proceso por el cual se generan múltiples isoformas de polipéptido a partir de un solo gen, e implica el corte y empalme junto de exones no consecutivos durante el procesamiento de algunas, pero no todas, las transcripciones del gen. De esta manera, se puede conectar un exón concreto a cualquiera de los varios exones alternativos para formar ARNm mensajeros. Los ARNm cortados y empalmados de manera alternativa producen polipéptidos (“variantes de corte y empalme”) en los que algunas partes son comunes, mientras que otras partes son diferentes.
- [0043]** El término "secuencia señal" se refiere a una secuencia de aminoácidos en la parte N-terminal de una proteína que facilita la secreción de la forma madura de la proteína fuera de la célula. La forma madura de la proteína extracelular carece de la secuencia señal, que se escinde durante el proceso de secreción.
- 10 **[0044]** 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) del constructo de expresión. Las células huésped utilizadas en la presente exposición pueden ser células procariotas, tales como *E. coli*, o células eucariotas, tales como células de levaduras, plantas, insectos, anfibios o mamíferos. En general, las células huésped son hongos filamentosos.
- 15 **[0045]** El término "hongos filamentosos" significa cualquiera y todos los hongos filamentosos reconocidos por los expertos en la materia. Se selecciona un hongo preferido a partir del grupo que consiste en *Aspergillus*, *Trichoderma*, *Fusarium*, *Chrysosporium*, *Penicillium*, *Humicola*, *Neurospora*, o formas sexuales alternativas de los mismos tal como *Emericella*, *Hypocrea*. Ahora se ha demostrado que el hongo industrial asexual *Trichoderma reesei* es un derivado clónico del ascomiceto *Hypocrea jecorina* (Véase, Kuhls et al., PNAS, 93:7755-7760, 20 1996).
- [0046]** El término “celooligosacárido” se refiere a grupos oligosacáridos que contienen de 2 a 8 unidades de glucosa y tienen enlaces beta-1,4, por ejemplo, celobiosa.
- 25 **[0047]** El término “celulasa”, "enzimas celulolíticas" o "enzimas de celulasa" se refiere a una categoría de enzimas capaces de hidrolizar polímeros de celulosa a oligómeros de celooligosacáridos más cortos, celobiosa y/o glucosa. Se han obtenido numerosos ejemplos de celulasas, tales como exoglucanasas, exocelobiohidrolasas, endoglucanasas y glucosidasas, a partir de organismos celulolíticos, incluyendo concretamente hongos, plantas y bacterias. Las enzimas creadas por 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 (BGL o Bglu). Estos tres tipos diferentes de enzimas de celulasa 30 actúan sinérgicamente para convertir celulosa y sus derivados en glucosa.
- [0048]** Muchos microbios crean enzimas que hidrolizan celulosa, incluyendo el hongo de descomposición de la madera *Trichoderma*, las bacterias del compost *Thermomonospora*, *Bacillus*, y *Cellulomonas*; *Streptomyces*; y los hongos *Humicola*, *Aspergillus* y *Fusarium*.
- 35 **[0049]** La CBH2 de *Hypocrea jecorina* es un miembro de la Familia 6 de glicosil hidrolasa (por tanto, Cel6) y, específicamente, fue el primer miembro de esa familia identificado en *Hypocrea jecorina* (por tanto, Cel6A). La Familia 6 de glicosil hidrolasa contiene tanto endoglucanasas como celobiohidrolasas/exoglucanasas, y esa CBH2 es esta última. De esta manera, las expresiones CBH2, proteína de tipo CBH2 y celobiohidrolasas Cel6 se pueden utilizar indistintamente en el presente documento.
- 40 **[0050]** El término “dominio de unión a celulosa” como se utiliza en el presente documento se refiere a la parte de la secuencia de aminoácidos de una celulasa o de una región de la enzima que interviene en la actividad de unión a celulosa de una celulasa o un derivado de la misma. Los dominios de unión a celulosa generalmente funcionan uniéndose de manera no covalente a la celulosa a la celulosa, un derivado de celulosa u otro polisacárido equivalente de la misma. Los dominios de unión a celulosa permiten o facilitan la hidrólisis de fibras de celulosa por la región de núcleo catalítico estructuralmente distinto y funcionan normalmente independientes del núcleo 45 catalítico. De esta manera, un dominio de unión a celulosa no poseerá la actividad hidrolítica significativa atribuible a un núcleo catalítico. Dicho de otro modo, un dominio de unión a celulosa es un elemento estructural de la estructura terciaria de la proteína enzimática de celulasa que es distinto del elemento estructural que posee actividad catalítica. Dominio de unión a celulosa y módulo de unión a celulosa se pueden utilizar indistintamente en el presente documento.
- 50 **[0051]** Como se utiliza en el presente documento, el término “surfactante” se refiere a cualquier compuesto generalmente reconocido en la técnica por tener cualidades activas de superficie. De esta manera, por ejemplo, los surfactantes comprenden surfactantes aniónicos, catiónicos y no iónicos, tales como los que se encuentran habitualmente en los detergentes. Los surfactantes aniónicos incluyen alquilbenzenosulfonatos lineales o

ramificados; sulfatos de éter de alquilo o alqueno que tienen grupos alquilo o alqueno lineales o ramificados; sulfatos de alquilo o alqueno; olefinsulfonatos; y alcanosulfonatos. Los surfactantes anfólicos incluyen sulfonatos de sal de amonio cuaternario y surfactantes anfólicos de tipo betaína. Dichos surfactantes anfólicos tienen tanto los grupos de carga positiva como negativa en la misma molécula. Los surfactantes no iónicos pueden comprender éteres de polioxialquilo, así como alcanolamidas de ácidos grasos superiores o aductos de óxido de alquilo del mismo, monoésteres de glicerina de ácidos grasos y similares.

5

[0052] Como se utiliza en el presente documento, el término “tejido que contiene celulosa” se refiere a cualquier tejido cosido o no cosido, hilo o fibra de algodón o no de algodón que contienen celulosa o de algodón o no de algodón que contienen mezclas de celulosa, incluyendo celulósicos naturales y celulósicos sintéticos (tales como yute, lino, ramio, rayón y liocel).

10

[0053] Como se utiliza en el presente documento, el término “tejido que contiene algodón” se refiere a tejidos cosidos o no cosidos, hilos o fibras de algodón puro o de mezclas de algodón, incluyendo tejidos de algodón tejido, prendas de punto de algodón, vaqueros de algodón, hilos de algodón, algodón en bruto y similares.

[0054] Como se utiliza en el presente documento, el término “composición para lavado a la piedra” se refiere a una formulación para utilizarse en tejidos que contienen celulosa lavados a la piedra. Las composiciones para lavado a la piedra se utilizan para modificar 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 ropa sucia y no se utilizan durante el proceso de fabricación.

15

[0055] Como se utiliza en el presente documento, el término “composición detergente” se refiere a una mezcla que está destinada a utilizarse en un medio de lavado para el lavado de tejidos sucios que contienen celulosa. En el contexto de la presente exposición, tales composiciones pueden incluir, además de celulasas y surfactantes, enzimas hidrolíticas adicionales, mejoradores, agentes blanqueadores, activadores de blanqueo, agentes azulantes y tintes fluorescentes, inhibidores del apelmazamiento, agentes enmascarantes, activadores de celulosa, antioxidantes y solubilizantes.

20

[0056] Como se utiliza en el presente documento, el término “reducción o eliminación en la expresión del gen *cbh2*” significa que el gen *cbh2* ha sido deletado del genoma y, por tanto, no puede ser expresado por el microorganismo huésped recombinante; o que el gen *cbh2* o la transcripción ha sido modificado de tal manera que el microorganismo huésped no produce una enzima CBH2 funcional o en niveles que son considerablemente menores que el gen *cbh2* no modificado o la transcripción.

25

[0057] El término “variante del gen *cbh2*” significa que la secuencia de ácido nucleico del gen *cbh2* de *H. jecorina* se ha alterado mediante la eliminación, adición y/o manipulación de la secuencia codificadora.

30

[0058] Como se utiliza en el presente documento, el término “purificar” generalmente se refiere a someter ácidos nucleicos transgénicos o células que contienen proteínas a purificación bioquímica y/o cromatografía en columna.

[0059] Como se utiliza en el presente documento, los términos “activo” o “biológicamente activo” se refieren a una actividad biológica asociada a una determinada proteína y se utilizan indistintamente en el presente documento. Por ejemplo, la actividad enzimática asociada a una proteasa es la 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 los expertos en la materia atribuyen normalmente a esta proteína.

35

[0060] Como se utiliza en el presente documento, el término “enriquecida” significa que la celulosa tal como CBH2 se encuentra en una concentración que es mayor en relación con la concentración de CBH2 encontrada en una composición de celulosa fúngica salvaje, o de origen natural. Los términos enriquecida, elevada y mejorada se pueden utilizar indistintamente en el presente documento.

40

[0061] Una composición de celulosa fúngica salvaje es una producida por una fuente fúngica de origen natural y que comprende uno o más componentes BGL, CBH y EG, donde cada uno de estos componentes se encuentra en la proporción producida por la fuente fúngica. De esta manera, una composición de CBH enriquecida tendría CBH en una proporción alterada, donde la proporción de CBH con respecto a otros componentes de celulosa (esto es, EG, beta-glucosidasas y otras endoglucanasas) es elevada. Esta proporción se puede incrementar ya sea aumentando la CBH o disminuyendo (o eliminando) al menos algún otro componente mediante cualquier medio conocido en la técnica.

45

50

[0062] El término “aislado” o “purificado” como se utiliza en el presente documento se refiere a un ácido nucleico o aminoácido que se elimina de al menos un componente con el cual está asociado de manera natural.

[0063] De esta manera, como ilustración, un sistema de celulasas de origen natural se puede purificar en componentes sustancialmente puros mediante técnicas de separación reconocidas publicadas en la literatura, incluyendo cromatografía de intercambio iónico a un pH adecuado, cromatografía de afinidad, exclusión de tamaño y similares. Por ejemplo, en la cromatografía de intercambio iónico (normalmente cromatografía de intercambio aniónico) es posible separar los componentes de celulasa mediante elución con un gradiente de pH, o un gradiente de sal, o ambos, un gradiente de pH y de sal. La CBH purificada se puede añadir entonces a la solución enzimática dando lugar a una solución de CBH enriquecida. También es posible elevar la cantidad de CBH producida mediante un microbio utilizando métodos de genética molecular para sobreexpresar el gen que codifica CBH, posiblemente en conjunción con la delección de uno o más genes que codifican otras celulasas.

[0064] Las celulasas fúngicas pueden contener más de un componente CBH. Los componentes diferentes generalmente tienen puntos isoeléctricos diferentes que permiten su separación mediante cromatografía de intercambio iónico y similares. En una solución enzimática se puede emplear un único componente CBH o una combinación de componentes de CBH.

[0065] Cuando se emplea en soluciones enzimáticas, el homólogo o variante de componente CBH2 se añade generalmente en una cantidad suficiente para permitir el índice más alto de liberación de azúcares solubles de la biomasa. La cantidad de homólogo o variante de componente CBH2 añadida depende del tipo de biomasa que se sacarifica, que puede determinarse fácilmente por el experto en la materia cuando se utiliza, el porcentaje en peso del homólogo o la variante de componente CBH2 presente en la composición de celulasas es de preferiblemente entre 1 y 100 con ejemplos ilustrativos siendo aproximadamente 1, preferiblemente aproximadamente 5, preferiblemente aproximadamente 10, preferiblemente aproximadamente 15, o preferiblemente aproximadamente 20 por ciento en peso a preferiblemente aproximadamente 25, preferiblemente aproximadamente 30, preferiblemente aproximadamente 35, preferiblemente aproximadamente 40, preferiblemente aproximadamente 45 o preferiblemente aproximadamente 50 por ciento en peso. Además, los intervalos preferidos pueden ser aproximadamente 0,5 a aproximadamente 15 por ciento en peso, 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 5 a aproximadamente 50 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 10 a aproximadamente 50 por ciento en peso, de aproximadamente 15 a aproximadamente 60 por ciento en peso, de aproximadamente 15 a aproximadamente 65 por ciento en peso, de aproximadamente 15 a aproximadamente 70 por ciento en peso, de aproximadamente 15 a aproximadamente 75 por ciento en peso, de aproximadamente 15 a aproximadamente 80 por ciento en peso, de aproximadamente 15 a aproximadamente 85 por ciento en peso, de aproximadamente 15 a aproximadamente 95 por ciento en peso. No obstante, cuando se utiliza, el porcentaje en peso del homólogo o variante de componente CBH2 en relación con cualquier componente de tipo EG presente en la composición de celulasas es de preferiblemente aproximadamente 1, preferiblemente aproximadamente 5, preferiblemente aproximadamente 10, preferiblemente aproximadamente 15, o preferiblemente aproximadamente 20 por ciento en peso a preferiblemente aproximadamente 25, preferiblemente aproximadamente 30, preferiblemente aproximadamente 35, preferiblemente aproximadamente 40, preferiblemente aproximadamente 45 o preferiblemente aproximadamente 50 por ciento en peso. Además, los intervalos preferidos pueden ser aproximadamente 0,5 a aproximadamente 15 por ciento en peso, 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 5 a aproximadamente 50 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 10 a aproximadamente 50 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 40 por ciento en peso, de aproximadamente 15 a aproximadamente 45 por ciento en peso, de aproximadamente 15 a aproximadamente 50 por ciento en peso, de aproximadamente 20 a aproximadamente 30 por ciento en peso, de aproximadamente 20 a aproximadamente 35 por ciento en peso, de aproximadamente 20 a aproximadamente 40 por ciento en peso, de aproximadamente 20 a aproximadamente 45 por ciento en peso, de aproximadamente 20 a aproximadamente 50 por ciento en peso, de aproximadamente 25 a aproximadamente 35 por ciento en peso, de aproximadamente 25 a aproximadamente 40 por ciento en peso, de aproximadamente 25 a aproximadamente 45 por ciento en peso, de aproximadamente 25 a aproximadamente 50 por ciento en peso, de aproximadamente 30 a aproximadamente 40 por ciento en peso, de aproximadamente 30 a aproximadamente 45 por ciento en peso, de aproximadamente 30 a aproximadamente 50 por ciento en peso, de aproximadamente 35 a aproximadamente 45 por ciento en peso, de aproximadamente 35 a aproximadamente 50 por ciento en peso, de aproximadamente 40 a aproximadamente 50 por ciento en peso, de aproximadamente 45 a aproximadamente 50 por ciento en peso.

[0066] 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 5 a aproximadamente 50 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 10 a aproximadamente 50 por ciento en peso, de aproximadamente 15 a aproximadamente 60 por ciento en peso, de aproximadamente 15 a aproximadamente 65 por ciento en peso, de aproximadamente 15 a aproximadamente 70 por ciento en peso, de aproximadamente 15 a aproximadamente 75 por ciento en peso, de aproximadamente 15 a aproximadamente 80 por ciento en peso, de aproximadamente 15 a aproximadamente 85 por ciento en peso, de aproximadamente 15 a aproximadamente 95 por ciento en peso. No obstante, cuando se utiliza, el porcentaje en peso del homólogo o variante de componente CBH2 en relación con cualquier componente de tipo EG presente en la composición de celulasas es de preferiblemente aproximadamente 1, preferiblemente aproximadamente 5, preferiblemente aproximadamente 10, preferiblemente aproximadamente 15, o preferiblemente aproximadamente 20 por ciento en peso a preferiblemente aproximadamente 25, preferiblemente aproximadamente 30, preferiblemente aproximadamente 35, preferiblemente aproximadamente 40, preferiblemente aproximadamente 45 o preferiblemente aproximadamente 50 por ciento en peso. Además, los intervalos preferidos pueden ser aproximadamente 0,5 a aproximadamente 15 por ciento en peso, 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 5 a aproximadamente 50 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 10 a aproximadamente 50 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 40 por ciento en peso, de aproximadamente 15 a aproximadamente 45 por ciento en peso, de aproximadamente 15 a aproximadamente 50 por ciento en peso, de aproximadamente 20 a aproximadamente 30 por ciento en peso, de aproximadamente 20 a aproximadamente 35 por ciento en peso, de aproximadamente 20 a aproximadamente 40 por ciento en peso, de aproximadamente 20 a aproximadamente 45 por ciento en peso, de aproximadamente 20 a aproximadamente 50 por ciento en peso, de aproximadamente 25 a aproximadamente 35 por ciento en peso, de aproximadamente 25 a aproximadamente 40 por ciento en peso, de aproximadamente 25 a aproximadamente 45 por ciento en peso, de aproximadamente 25 a aproximadamente 50 por ciento en peso, de aproximadamente 30 a aproximadamente 40 por ciento en peso, de aproximadamente 30 a aproximadamente 45 por ciento en peso, de aproximadamente 30 a aproximadamente 50 por ciento en peso, de aproximadamente 35 a aproximadamente 45 por ciento en peso, de aproximadamente 35 a aproximadamente 50 por ciento en peso, de aproximadamente 40 a aproximadamente 50 por ciento en peso, de aproximadamente 45 a aproximadamente 50 por ciento en peso.

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 10 a aproximadamente 50 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 30 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. Celulasas

[0067] Las celulasas se conocen en la técnica como enzimas que hidrolizan celulosa (enlaces beta-1,4-glucano o beta D-glucosídicos) dando lugar a la formación de glucosa, celobiosa, celooligosacáridos y similares. Como se ha establecido más arriba, tradicionalmente se han dividido las celulasas en tres clases principales: endoglucanasas (EC 3.2.1.4) ("EG"), exoglucanasas o celobiohidrolasas (EC 3.2.1.91) ("CBH") y beta-glucosidasas (EC 3.2.1.21) ("BG").

[0068] 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. No obstante, a 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 celulosa de manera más eficiente. Los componentes diferentes, es decir, las diversas endoglucanasas y exocelobiohidrolasas en un sistema de celulasas completo o multicomponente, generalmente tienen propiedades diferentes, tales como punto isoeléctrico, peso molecular, grado de glicosilación, especificidad de sustrato y patrones de acción enzimática.

[0069] Se cree que las celulasas de tipo endoglucanasa hidrolizan los enlaces internos beta-1,4-glucosídicos en regiones de baja cristalinidad de la celulosa y que las celulasas de tipo exocelobiohidrolasa hidrolizan la celobiosa a partir del extremo reductor o no reductor de la celulosa. Se deduce que la acción de los componentes de endoglucanasa puede facilitar en gran medida la acción de las exocelobiohidrolasas mediante la creación de nuevos extremos de cadena que son reconocidos por los componentes de exocelobiohidrolasa. Además, se ha demostrado que las celulasas de tipo beta-glucosidasa catalizan la hidrólisis de alquil y/o aril beta-D-glucósidos, tales como metil beta-D-glucósido y p-nitrofenil glucósido, así como glicósidos que contienen únicamente residuos de carbohidratos, tales como celobiosa. Esto produce glucosa como el único producto para el microorganismo y reduce o elimina la celobiosa que inhibe las celobiohidrolasas y las endoglucanasas.

[0070] Las celulasas también encuentran una serie de usos en composiciones detergentes, incluyendo mejorar la capacidad de limpieza, como un agente suavizante y mejorar el tacto de los tejidos de algodón (Hemmpel, ITB Dyeing/Printing/Finishing 3:5-14, 1991; Tyndall, Textile Chemist and Colorist 24:23-26, 1992; y Kumar et al., Textile Chemist and Colorist, 29:37-42, 1997). Si bien el mecanismo no forma parte de la exposición, las propiedades suavizantes y de restauración de color de la celulasa se han atribuido a los componentes de endoglucanasa alcalina en composiciones de celulasa, como lo demuestran las patentes estadounidenses con n.º 5,648,263, 5,691,178, y 5,776,757, que exponen que las composiciones detergentes que contienen una composición de celulasas enriquecida en un componente de endoglucanasa alcalina especificado imparten restauración de color y suavizado mejorado a las prendas tratadas en comparación con las composiciones de celulasa no enriquecidas en dicho componente. Además, se ha demostrado que la utilización de dichos componentes de endoglucanasa alcalina en composiciones detergentes complementa los requisitos de pH de la composición detergente (p. ej., exhibiendo actividad máxima en un pH alcalino de 7,5 a 10, como se describe en las patentes estadounidenses con n.º 5,648,263, 5,691,178, y 5,776,757).

[0071] También se ha demostrado que las composiciones de celulasa degradan tejidos que contienen algodón, dando lugar a una pérdida de resistencia reducida en el tejido (patente estadounidense con n.º 4,822,516), contribuyendo a la retención a utilizar composiciones de celulasa en aplicaciones detergentes comerciales. Se ha sugerido que las composiciones de celulasa que comprenden componentes de endoglucanasa presentan una pérdida de resistencia reducida para tejidos que contienen algodón en comparación con las composiciones que comprenden un sistema de celulasas completo.

[0072] También se ha demostrado que las celulasas son útiles en la degradación de biomasa de celulasa en etanol (donde la celulasa degrada celulosa en glucosa y levadura u otros microbios y además fermenta la glucosa en etanol), en el tratamiento de pulpa mecánica (Pere et al., In Proc. Tappi Pulping Conf., Nashville, Tenn., 27-31, pp. 693-696, 1996), para su utilización como aditivo para alimentación animal (WO 91/04673) y en la molienda en húmedo de grano.

[0073] La mayoría de CBH y EG tiene una estructura multidominio que consiste en un dominio central separado de un dominio de unión a celulosa (CBD, por sus siglas en inglés) mediante un péptido de enlace (Suurnakki et al., 2000). El dominio central contiene el sitio activo mientras que el CBD interactúa con la celulosa uniendo la enzima al mismo (van Tilbeurgh et al., FEBS Lett. 204:223-227, 1986; Tomme et al., Eur. J. Biochem. 170:575-581, 1988). Los CBD 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, J. Biotechnol. 57:15-28, 1997). No obstante, 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 incrementando la concentración de enzima efectiva en la superficie de celulosa (Stahlberg et al., Bio/Technol. 9:286-290, 1991), y/o aflojando cadenas de celulosa simples de la superficie de celulosa (Tormo et al., EMBO J. vol. 15, n.º 21, pp. 5739-5751, 1996). La mayoría de estudios sobre los efectos de los dominios de celulasa en sustratos diferentes se han llevado a cabo con proteínas centrales de celobiohidrolasas, puesto que sus proteínas centrales se pueden producir fácilmente mediante proteólisis limitada con papaína (Tomme et al., 1988). Se han descrito numerosas celulasas en la literatura científica, ejemplos de las cuales incluyen: de *Trichoderma reesei*: Shoemaker, S. et al., Bio/Technology, 1:691-696, 1983, que describe CBH1; Teeri, T. et al., Gene, 51:43-52, 1987, que describe CBH2. También se han descrito celulasas de especies distintas a *Trichoderma*, por ejemplo, Ooi et al., Nucleic Acids Research, vol.18, n.º 19, 1990, que expone la secuencia de ADNc que codifica la endoglucanasa F1-CMC producida por *Aspergillus aculeatus*; Kawaguchi T et al., Gene 173(2):287-8, 1996, que expone la clonación y la secuenciación del ADNc que codifica la beta-glucosidasa 1 de *Aspergillus aculeatus*; Sakamoto et al., Curr. Genet. 27:435-439, 1995, que expone la secuencia de ADNc que codifica la endoglucanasa CMCase-1 de *Aspergillus kawachii* IFO 4308; Saarihahti et al., Gene 90:9-14, 1990, que expone una endoglucanasa de *Erwinia carotovora*; Spilliaert R, et al., Eur J Biochem. 224(3):923-30, 1994, que expone la clonación y la secuenciación de bglA, que codifica una beta-glucanasa termoestable a partir de *Rhodothermus marinus*; y Halldorsdottir S et al., Appl Microbiol Biotechnol. 49(3):277-84, 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, todavía resulta necesaria la identificación y caracterización de nuevas celulasas, con propiedades mejoradas, tales como rendimiento mejorado en condiciones de estrés térmico o en presencia de surfactantes, actividad específica aumentada, patrón de escisión de sustrato alterado, y/o alto nivel de expresión *in vitro*.

[0074] El desarrollo de nuevas y mejoradas composiciones de celulasa que comprenden cantidades diversas de celulasas de tipo CBH, de tipo EG y de tipo BG es de interés para utilizarse: (1) en composiciones para degradar pulpa de madera u otra biomasa en azúcares (p. ej., para la producción de productos bioquímicos tal como biocombustibles); (2) en composiciones detergentes que presentan capacidad de limpieza mejorada, (3) funcionan como un agente suavizante y/o mejoran el tacto de los tejidos de algodón (p. ej., "lavado a la piedra" o "biopulido"); y/o (3) en composiciones para pienso, por ejemplo.

[0075] En el presente documento también se proporcionan preparaciones de celulasa enteras que comprenden variantes de celulasa. Como se utiliza en el presente documento, la expresión "preparación de celulasa entera" se refiere tanto a las composiciones que contienen celulasa de origen natural como a las de origen no natural. Una composición "de origen natural" es una producida por una fuente de origen natural y que comprende uno o más componentes de tipo celobiohidrolasa, uno o más componentes de tipo endoglucanasa y uno o más componentes de tipo beta-glucosidasa, donde cada uno de estos componentes se encuentra en la proporción producida por la fuente. Una composición de origen natural es una producida por un organismo no modificado con respecto a las enzimas celulolíticas de forma que la proporción de las enzimas componentes no se altere de la producida por el organismo nativo. Una composición "de origen no natural" abarca las composiciones producidas mediante: (1) la combinación de enzimas celulolíticas componentes en una proporción de origen natural o en una proporción de origen no natural, esto es, alterada; o (2) la modificación de un organismo para sobreexpresar o infraexpresar una o más enzimas celulolíticas; o (3) la modificación de un organismo de forma que se deleccione al menos una enzima celulolítica. En consecuencia, en algunos modos de realización, la preparación de celulasa entera puede tener una o más de las diversas EG y/o CBH, y/o beta-glucosidasas delecionadas. Por ejemplo, EG1 puede deleccionarse sola o en combinación con otras EG y/o CBH.

[0076] En general, la preparación de celulasa entera incluye enzimas, incluyendo, pero sin carácter limitativo: (i) endoglucanasas (EG) o 1,4 β -d-glucano-4-glucanohidrolasas (EC 3.2.1.4), (ii) exoglucanasas, incluyendo 1,4- β -d-glucano glucanohidrolasas (también conocidas como celodextrinasas) (EC 3.2.1.74) y 1,4- β -glucano-celobiohidrolasas (exocelobiohidrolasas, CBH) (EC 3.2.1.91), y (iii) β -glucosidasa (BG) o β -glucosida glucosidrolasas (EC 3.2.1.21).

[0077] En la presente exposición, la preparación de celulasa entera puede ser de cualquier microorganismo que sea útil para la hidrólisis de un material celulósico. En algunos modos de realización, la preparación de celulasa entera es una celulasa entera de hongos filamentosos. Los "hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión *Eumycota* y *Oomycota*. En algunos modos de realización, la preparación de

celulasa entera es una celulasa entera de especies de *Acremonium*, *Aspergillus*, *Emericella*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Scytalidium*, *Thielavia*, *Tolyposcladium*, o *Trichoderma*. En algunos modos de realización, la preparación de celulasa entera es una celulasa entera de *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, o *Aspergillus oryzae*. En otro aspecto, la preparación de celulasa entera es una celulasa entera de *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochroum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, o *Fusarium venenatum*. En otro aspecto, la preparación de celulasa entera es una celulasa entera de *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Scytalidium thermophilum*, o *Thielavia terrestris*. En otro aspecto, la preparación de celulasa entera es una celulasa entera de *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* p. ej., RL-P37 (Sheir-Neiss et al., Appl. Microbiol. Biotechnology, 20 (1984) pp. 46-53; Montencourt B.S., Can., 1-20, 1987), QM9414 (ATCC n.º 26921), NRRL 15709, ATCC 13631, 56764, 56466, 56767, o *Trichoderma viride* p. ej., ATCC 32098 y 32086. En algunos modos de realización, la preparación de celulasa entera es una celulasa entera de *Trichoderma reesei* RutC30, que está disponible en American Type Culture Collection como *Trichoderma reesei* ATCC 56765.

[0078] Entre los ejemplos de preparaciones de celulasa comerciales adecuadas para utilizarse en la presente exposición se incluyen, por ejemplo, CELLUCLAST™ (disponible en Novozymes A/S) y las enzimas LAMINEX™, IndiAge™ y Primafast™ LAMINEX BG, ACCELLERASE™ 100 y ACCELLERASE™ 1500 (disponibles en Genencor Division, Danisco US. Inc.)

[0079] En la presente exposición, la preparación de celulasa entera puede ser de cualquier método de cultivo de microorganismos conocido en la técnica que dé lugar a la expresión de enzimas capaces de hidrolizar un material celulósico. La fermentación puede incluir cultivo en matraz de agitación, fermentación a pequeña o a gran escala, tales como fermentaciones continuas, por lotes, por lote alimentado, o en medio sólido en laboratorios o fermentadores industriales realizadas en un medio adecuado y en condiciones que permitan que la celulasa se exprese o se aíse.

[0080] Por lo general, el microorganismo se cultiva en un medio de cultivo celular adecuado para la producción de enzimas capaces de hidrolizar un material celulósico. El cultivo tiene lugar en un medio nutritivo adecuado que comprende fuentes de carbono y nitrógeno, así como sales inorgánicas, utilizando procedimientos conocidos en la técnica. En la técnica, se conocen medios de cultivo, intervalos de temperatura y otras condiciones de crecimiento y producción de celulasa adecuados. Como ejemplo no limitativo, el intervalo de temperatura normal para la producción de celulasas a partir de *Trichoderma reesei* es de 24 °C a 28 °C.

[0081] Generalmente, la preparación de celulasa entera se utiliza como se produce mediante la fermentación con recuperación y/o purificación mínima o inexistente. Por ejemplo, una vez que una célula secreta las celulasas en el medio de cultivo celular, puede utilizarse el medio de cultivo celular que contiene las celulasas. En algunos modos de realización, la preparación de celulasa entera comprende los contenidos no fraccionados de material de fermentación, incluyendo medio de cultivo celular, enzimas extracelulares y células. Alternativamente, la preparación de celulasa entera puede ser procesada mediante cualquier método conveniente, por ejemplo, mediante precipitación, centrifugación, afinidad, filtración o cualquier otro método conocido en la técnica. En algunos modos de realización, la preparación de celulasa entera puede ser concentrada, por ejemplo, y utilizarse sin purificación adicional. En algunos modos de realización, la preparación de celulasa entera comprende agentes químicos que disminuyen la viabilidad de células o matan las células. En algunos modos de realización, las células se lisan o se permeabilizan mediante la utilización de métodos conocidos en la técnica.

45 III. Biología Molecular

[0082] En un modo de realización la presente exposición proporciona la expresión de variantes de genes *cbh2* bajo el control de un promotor funcional en un hongo filamentoso. Por consiguiente, la presente exposición se basa en técnicas rutinarias en el campo de la genética recombinante (Véase, p. ej., Sambrook et al., Molecular Cloning, A Laboratory Manual, 2ª ed., 1989; Kriegl, Gene Transfer and Expression: A Laboratory Manual, 1990; y Ausubel et al., eds., CURRENT PROTOCOLS IN MOLECULAR BIOLOGY, Greene Publishing and Wiley-Interscience, New York, 1994).

Métodos de mutación de secuencias de ácido nucleico de *cbh2*

[0083] En la presente exposición, se contempla cualquier método conocido en la técnica que pueda introducir mutaciones.

[0084] La presente descripción se refiere a la expresión, la purificación y/o el aislamiento y la utilización de variante de CBH2. Estas enzimas se preparan preferiblemente mediante métodos recombinantes utilizando el gen *cbh2* de *H. jecorina*. El caldo de fermentación puede utilizarse con o sin purificación.

5 [0085] Tras el aislamiento y la clonación del gen *cbh2* de *H. jecorina*, se utilizan otros métodos conocidos en la técnica, tales como mutagénesis de sitio dirigido, para realizar las sustituciones, adiciones o deleciones que corresponden a aminoácidos sustituidos en la variante de CBH2 expresada. De nuevo, la mutagénesis de sitio dirigido y otros métodos para incorporar cambios de aminoácidos en proteínas expresadas a nivel de ADN se conocen en la técnica (Sambrook et al., supra; y Ausubel et al., supra).

10 [0086] El ADN que codifica una variante de secuencia de aminoácidos de la CBH2 de *H. jecorina* se prepara mediante diversos métodos conocidos en la técnica. Estos métodos incluyen, pero sin carácter limitativo, la preparación mediante mutagénesis de sitio dirigido (o mediada por oligonucleótidos), mutagénesis por PCR y mutagénesis por inserción de un casete de un ADN preparado anteriormente que codifica la CBH2 de *H. jecorina*.

15 [0087] La mutagénesis de sitio dirigido es un método preferido para preparar variantes de sustitución. Esta técnica se conoce bien en la técnica (véase, p. ej., Carter et al. Nucleic Acids Res. 13:4431-4443 (1985) y Kunkel et al., Proc. Natl. Acad. Sci. EE. UU 82:488 (1987)). En resumen, al llevar a cabo la mutagénesis de sitio dirigido del ADN, el ADN inicial se altera hibridando en primer lugar un oligonucleótido que codifica la mutación deseada en una cadena simple de dicho ADN inicial. Después de la hibridación, se utiliza una ADN polimerasa para sintetizar una segunda cadena completa, utilizando el oligonucleótido hibridado como cebador, y utilizando la
20 cadena simple del ADN inicial como plantilla. De esta manera, el oligonucleótido que codifica la mutación deseada se incorpora en el ADN bicatenario resultante.

[0088] La mutagénesis por PCR también es adecuada para realizar variantes de secuencias de aminoácidos del polipéptido inicial, es decir, CBH2 de *H. jecorina*. Véase Higuchi, en PCR Protocols, pp.177-183 (Academic Press, 1990); y Vallette et al., Nuc. Acids Res. 17:723-733, 1989). Véase, también, por ejemplo, Cadwell et al.,
25 PCR Methods and Applications, Vol 2, 28-33 (1992). En resumen, cuando se utilizan pequeñas cantidades de ADN plantilla como material inicial en una PCR, se pueden utilizar cebadores que difieren ligeramente en la secuencia de la región correspondiente en un ADN plantilla para generar cantidades relativamente grandes de un fragmento de ADN específico que difiere de la secuencia de la plantilla solo en las posiciones en las que los cebadores difieren de la plantilla.

30 [0089] Otro método para preparar variantes, la mutagénesis por inserción de un casete, se basa en la técnica descrita por Wells et al., Gene 34:315-323 (1985). El material inicial es el plásmido (u otro vector) que comprende el ADN del polipéptido inicial que ha de ser mutado. Se identifica el codón o los codones en el ADN inicial que ha de ser mutado. Debe haber un único sitio de endonucleasa de restricción en cada lado del sitio o de los sitios de mutación identificados. Si no existen dichos sitios de restricción, se pueden generar utilizando el método de mutagénesis mediada por oligonucleótidos descrito anteriormente para introducirlos en localizaciones apropiadas en el ADN del polipéptido inicial. El ADN plasmídico se corta en estos sitios para linealizarlo. Un oligonucleótido bicatenario que codifica la secuencia del ADN entre los sitios de restricción, pero que contiene la mutación o las mutaciones deseadas, se sintetiza utilizando procedimientos estándar, donde las dos cadenas del oligonucleótido se sintetizan por separado y después se hibridan juntas utilizando técnicas estándar. Este oligonucleótido bicatenario se denomina casete. Este casete está diseñado para que tenga extremos 5' y 3' compatibles con los extremos del plásmido linealizado, de manera que se puede ligar directamente al plásmido. Este plásmido contiene ahora la secuencia de ADN mutada.

[0090] De manera alternativa, o adicional, se puede determinar la secuencia de aminoácidos deseada que codifica una variante de CBH2, y se puede generar sintéticamente una secuencia de ácido nucleico que codifica dicha variante de secuencia de aminoácidos.
45

[0091] La variante o las variantes de CBH2 preparadas de esta manera se pueden someter a modificaciones adicionales, dependiendo a menudo del uso pretendido de la celulasa. Dichas modificaciones pueden implicar la alteración adicional de la secuencia de aminoácidos, la fusión a un polipéptido o polipéptidos heterólogos y/o modificaciones covalentes.

50 IV. Ácidos nucleicos *cbh2* y polipéptidos CBH2

A. Variantes de ácidos nucleicos de tipo *cbh2*

[0092] La secuencia de ácido nucleico para la *cbh2* de tipo salvaje se muestra en SEQ ID NO:1. La exposición abarca una molécula de ácido nucleico que codifica las variantes de celulasas descritas en el presente documento. El ácido nucleico puede ser una molécula de ADN.

5 **[0093]** Después de que se hayan clonado las secuencias de ADN que codifican las variantes de CBH2 en constructos de ADN, se utiliza el ADN para transformar microorganismos. El microorganismo que se ha de transformar con el propósito de expresar una variante de CBH2 según la presente exposición puede comprender de manera ventajosa una cepa derivada de *Trichoderma sp.* De esta manera, un modo preferido para preparar una variante de celulasas CBH2 según la presente exposición comprende transformar una célula huésped de *Trichoderma sp.* con un constructo de ADN que comprende al menos un fragmento de ADN que codifica una parte o toda la variante de CBH2. El constructo de ADN se unirá en general de manera funcional a un promotor. 10 La célula huésped transformada se cultiva entonces en condiciones para expresar la proteína deseada. Posteriormente, el producto de proteína deseada se puede purificar hasta una homogeneidad sustancial.

15 **[0094]** No obstante, de hecho, puede ser que el mejor vehículo de expresión para un ADN determinado que codifica una variante de CBH2 pueda diferir de *H. jecorina*. De este modo, puede ser que lo más ventajoso sea expresar una proteína en un huésped de transformación que porte similitud filogenética al organismo original para la variante de CBH2. En un modo de realización alternativo, se puede utilizar *Aspergillus niger* como vehículo de expresión. Para una descripción de técnicas de transformación con *A. niger*, véase WO 98/31821.

20 **[0095]** Por consiguiente, se proporciona la presente descripción de un sistema de expresión de *Aspergillus spp.* con fines ilustrativos únicamente y como una opción para expresar la variante de CBH2 de la exposición. No obstante, un experto en la materia puede inclinarse por expresar el ADN que codifica la variante de CBH2 en una célula huésped diferente si procede y debe entenderse que el origen de la variante de CBH2 debería considerarse al determinar el huésped de expresión óptimo. Adicionalmente, el experto en la materia será capaz de seleccionar el mejor sistema de expresión para un gen concreto mediante técnicas rutinarias utilizando las herramientas disponibles en la técnica.

25 **B. Polipéptidos de variantes de CBH2**

30 **[0096]** Las variantes de CBH2 de la presente exposición tienen secuencias de aminoácidos que se derivan de la secuencia de aminoácidos de un precursor de CBH2. La secuencia de aminoácidos de la variante de CBH2 difiere de la secuencia de aminoácidos del precursor de CBH2 en la sustitución, delección o inserción de uno o varios aminoácidos de la secuencia de aminoácidos del precursor, siempre que dicha variante sea como se define en las reivindicaciones. En un modo de realización preferido, el precursor de CBH2 es CBH2 de *Hypocrea jecorina*. La secuencia de aminoácidos madura de CBH2 de *H. jecorina* se muestra en SEQ ID NO:3. De esta manera, la exposición se dirige a variantes de CBH2 que contienen residuos de aminoácidos en posiciones que son equivalentes al residuo concreto identificado en CBH2 de *H. jecorina*. Un residuo (aminoácido) de un homólogo de CBH2 es equivalente a un residuo de CBH2 de *Hypocrea jecorina* si es homólogo (es decir, correspondiente en la posición en la estructura primaria o terciaria) o si es funcionalmente análogo a un residuo específico o parte de ese residuo en CBH2 de *Hypocrea jecorina* (es decir, que tiene la misma o similar capacidad funcional para combinarse, reaccionar, o interaccionar química o estructuralmente). Como se utiliza en el presente documento, la numeración pretende corresponder a la de la secuencia de aminoácidos de la CBH2 madura (SEQ ID NO:3).

40 **[0097]** El alineamiento de las secuencias de aminoácidos para determinar la homología se determina preferiblemente utilizando un "algoritmo de comparación de secuencias". El alineamiento óptimo de secuencias para la comparación se puede llevar a cabo, por ejemplo, mediante el algoritmo de homología local de Smith & Waterman, Adv. Appl. Math. 2:482 (1981), mediante el algoritmo de alineamiento por homología de Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), mediante la búsqueda de un método de similitud de Pearson & Lipman, Proc. Nat'l Acad. Sci. USA 85:2444 (1988), mediante implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), o mediante inspección visual. La inspección visual puede utilizar paquetes de gráficos tales como, por ejemplo, MOE de Chemical Computing Group, Montreal Canadá. 45

50 **[0098]** Un ejemplo de un algoritmo que es adecuado para determinar la similitud de secuencia es el algoritmo BLAST, que se describe en Altschul, et al., J. Mol. Biol. 215:403-410, 1990). El *software* para realizar análisis BLAST está disponible públicamente en el Centro Nacional de Información Biotecnológica (www.ncbi.nlm.nih.gov). Este algoritmo implica en primer lugar identificar pares de secuencias de alta puntuación (HSP, por sus siglas en inglés) mediante la identificación de palabras cortas de longitud W en la secuencia problema que coinciden con o cumplen alguna puntuación umbral T con valor positivo cuando se alinean con una palabra de la misma longitud en una secuencia de la base de datos. Estas coincidencias de palabras iniciales 55 próximas actúan como puntos de partida para encontrar pares de secuencias de alta puntuación más largas que

las contengan. Las coincidencias de palabras se expanden en ambas direcciones a lo largo de cada una de las dos secuencias que se comparan hasta donde se pueda aumentar la puntuación de alineamiento acumulativa. La extensión de las coincidencias de palabras se detiene cuando: la puntuación de alineamiento acumulativa disminuye en la cantidad X a partir de un valor máximo conseguido; la puntuación acumulativa es cero o inferior; o se alcanza el final de cualquier secuencia. Los parámetros W, T y X del algoritmo BLAST determinan la sensibilidad y la velocidad del alineamiento. El programa BLAST utiliza por defecto una longitud de palabra (W) de 11, la matriz de puntuación BLOSUM62 (véase Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989)) alineamientos (B) de 50, expectación (E) de 10, M⁵, N⁻⁴, y una comparación de ambas cadenas.

[0099] El algoritmo BLAST realiza a continuación un análisis estadístico de la similitud entre las dos secuencias (véase, por ejemplo, Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90:5873-5787 (1993)). Una medida de similitud proporcionada por el algoritmo BLAST es la probabilidad de suma más pequeña (P(N)), que proporciona una indicación de la probabilidad por la que un emparejamiento entre dos secuencias de nucleótidos o aminoácidos ocurriría al azar. Por ejemplo, una secuencia de aminoácidos se considera similar a una proteasa si la probabilidad de suma más pequeña en una comparación de la secuencia de aminoácidos de prueba con respecto a una secuencia de aminoácidos de proteasa es inferior a aproximadamente 0,1, más preferiblemente inferior a aproximadamente 0,01 y siendo lo más preferible inferior a aproximadamente 0,001.

[0100] Para fines de la presente exposición, el grado de identidad se puede determinar adecuadamente mediante programas informáticos conocidos en la técnica, tales como GAP proporcionados en el paquete de programa GCG (Program Manual for the Wisconsin Package, Version 8, August 1994, Genetics Computer Group, 575 Science Drive, Madison, Wis., USA 53711) (Needleman, S. B. y Wunsch, C. D., (1970), Journal of Molecular Biology, 48, 443-45), utilizando GAP con los siguientes ajustes para comparación de secuencias de polinucleótidos: GAP creation penalty (penalización por creación de hueco) de 5.0 y GAP extension penalty (penalización por extensión de hueco) de 0.3.

[0101] Puede utilizarse un alineamiento estructural entre un CBH2 de *T. reesei* y otras celulasas para identificar posiciones equivalentes/correspondientes en otras celulasas que tienen un grado de homología moderado a elevado, por ejemplo, aproximadamente 50 %, 55 %, 60 %, 65 %, 70 %, 75 %, 80 %, 81 %, 82 %, 83 %, 84 %, 85 %, 86 %, 87 %, 88 %, 89 %, 90 %, 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, o incluso 99 %, con CBH2 de *T. reesei* (SEQ ID NO: 3). Un método de obtener dicho alineamiento estructural es utilizar el programa Pile Up del paquete GCG utilizando valores por defecto de penalizaciones por hueco, esto es, una penalización por creación de hueco de 3.0 y una penalización por extensión de hueco de 0.1. Otros métodos de alineamiento estructural incluyen el análisis de grupos hidrofóbicos (Gaboriaud et al., FEBS Letters, 224:149-155, 1987) y plegamiento inverso (Huber and Torda, Protein Science, 7:142-149, 1998).

[0102] Un alineamiento ilustrativo de la forma madura de diversas celulasas de referencia se proporciona a modo de figura 3. Las celulasas de referencia incluyen: *Hypocrea jecorina* (también conocida como *T. reesei*) CBH2 (SEQ ID NO:3), *Hypocrea koningii* CBH2 (SEQ ID NO:4), *Humicola insolens* CBH2 (SEQ ID NO:5), *Acremonium cellulolyticus* CBH2 (SEQ ID NO:6), *Agaricus bisporus* CBH2 (SEQ ID NO:7), *Fusarium osysporum* EG (SEQ ID NO:8), *Phanerochaete chrysosporium* CBH2 (SEQ ID NO:9), *Talaromyces emersonii* CBH2 (SEQ ID NO:10), *Thermobifida fusca* 6B/E3 CBH2 (SEQ ID NO:11), *Thermobifida fusca* 6A/E2 EG (SEQ ID NO:12), y *Cellulomonas fimi* CenA EG (SEQ ID NO:13). Las secuencias se alinearon mediante los algoritmos de alineamiento de secuencias múltiples ClustalW y MUSCLE. En la tabla 1, se proporciona una matriz que muestra el porcentaje de identidad de las celulasas del alineamiento de secuencia.

Tabla 1. Matriz* de porcentaje de identidad de celulasas

Porcentaje_ID	3	4	5	6	7	8	9	10	11	12	13
3	100	95,5	62,3	64,7	59,6	63,1	55,4	63,4	31,9	13,5	27
4	95,5	100	61,6	64	59,1	63,6	54,7	63	32,9	13,5	26,8
5	62,3	61,6	100	59,1	57,6	61,3	54	58,8	31,9	15,9	26,6
6	64,7	64	59,1	100	58,6	56,4	54	72,6	32,8	13,5	29,2
7	59,6	59,1	57,6	58,6	100	55,8	69,1	58,1	34,9	17,5	27,6
8	63,1	63,6	61,3	56,4	55,8	100	48,7	54,8	31,1	13,9	25,2
9	55,4	54,7	54	54	69,1	48,7	100	52,6	32,4	15,4	25,6
10	63,4	63	58,8	72,6	58,1	54,8	52,6	100	33,9	13,2	27,3
11	31,9	32,9	31,9	32,8	34,9	31,1	32,4	33,9	100	15,9	36,3
12	13,5	13,5	15,9	13,5	17,5	13,9	15,4	13,2	15,9	100	12,8
13	27	26,8	26,6	29,2	27,6	25,2	25,6	27,3	36,3	12,8	100

*Los números de la fila superior y de la columna izquierda corresponden a los SEQ ID NOS de las secuencias

alineadas de la figura 3.

5 [0103] Las búsquedas de secuencias se llevan a cabo normalmente utilizando el programa BLASTN cuando se evalúa una secuencia de ácido nucleico determinada en relación con las secuencias de ácido nucleico de las secuencias de ADN de GenBank y otras bases de datos públicas. Se prefiere el programa BLASTX para buscar secuencias de ácido nucleico que se han traducido en todos los marcos de lectura con respecto a las secuencias de aminoácidos de las secuencias de proteínas de GenBank y otras bases de datos públicas. Tanto BLASTN como BLASTX funcionan 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.)

V. Expresión de variantes de CBH2 recombinantes

10 [0104] Los métodos de la exposición se basan en la utilización de células para expresar la variante de CBH2, sin que se requiera ningún método concreto de expresión de CBH2. La variante de CBH2 se secreta preferiblemente de las células. La exposición proporciona células huésped que se han transducido, transformado o transfectado con un vector de expresión que comprende una secuencia de ácido nucleico que codifica una variante de CBH2. Las condiciones de cultivo, tales como la temperatura, el pH y similares, son las utilizadas anteriormente para la célula huésped parental antes de la transducción, transformación o transfección y resultarán evidentes para los expertos en la materia.

20 [0105] En un enfoque, se transfecta una célula de hongo filamentoso o una célula de levadura con un vector de expresión que tiene un promotor o un fragmento de promotor biológicamente activo o uno o varios (p. ej., una serie) de los potenciadores que funciona en la línea de la célula huésped, operativamente unidos a un segmento de ADN que codifica la variante de CBH2, de manera que la variante de CBH2 se expresa en la línea celular.

A. Vectores de expresión/Constructos de ácido nucleico

25 [0106] Se pueden incorporar fragmentos de polinucleótidos naturales o sintéticos que codifican la variante de CBH2 ("secuencias de ácido nucleico que codifican CBH2") en vectores o constructos de ácido nucleico heterólogos, capaces de la introducción y la replicación en una célula de levadura o de hongo filamentoso. Los vectores y métodos expuestos en el presente documento son adecuados para utilizarse en células huésped para la expresión de la variante de CBH2. Se puede utilizar cualquier vector siempre que 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 están disponibles comercialmente. Los vectores de clonación y expresión también se describen en Sambrook et al., 1989, Ausubel F M et al., 1989, y Strathern et al., The Molecular Biology of the Yeast Saccharomyces, 1981. Se describen vectores de expresión adecuados para hongos 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, pp. 396-428. La secuencia de ADN adecuada se puede insertar en un plásmido o vector (denominados colectivamente en el presente documento como "vectores") mediante una variedad de procedimientos. En general, la secuencia de ADN se inserta en un sitio o sitios de endonucleasa de restricción adecuado(s) mediante procedimientos estándar. Se considera que dichos procedimientos y los procedimientos de subclonación relacionados están dentro del alcance del conocimiento de los expertos en la materia.

[0107] Se pueden producir hongos filamentosos recombinantes que comprenden la secuencia codificadora para la variante de CBH2 introduciendo un constructo de ácido nucleico heterólogo que comprende la secuencia codificadora de la variante de CBH2 en las células de una cepa seleccionada de los hongos filamentosos.

40 [0108] Una vez que se obtiene la forma deseada de una secuencia de ácido nucleico de la variante de cbh2, se puede modificar de diversas maneras. Cuando la secuencia implica regiones flanqueantes no codificadoras, las regiones flanqueantes se pueden someter a resección, mutagénesis, etc. De esta manera, se pueden realizar transiciones, transversiones, deleciones e inserciones en la secuencia de origen natural.

45 [0109] Puede insertarse una secuencia codificadora de la variante de cbh2 seleccionada en un vector adecuado de conformidad con técnicas recombinantes conocidas y utilizarse para transformar los hongos filamentosos capaces de la expresión de CBH2. Debido a la degeneración inherente del código genético, se pueden utilizar otras secuencias de ácido nucleico que codifican una secuencia de aminoácidos sustancialmente igual o funcionalmente equivalente para clonar y expresar la variante de CBH2. Por lo tanto, se entiende que dichas sustituciones en la región codificadora están dentro de las variantes de secuencias cubiertas por la presente exposición. Se pueden utilizar todas y cada una de estas variantes de secuencias de la misma manera que se describe en el presente documento para una secuencia de ácido nucleico parental que codifica CBH2.

[0110] La presente exposición también incluye constructos de ácido nucleico recombinante que comprenden una o varias de las secuencias de ácido nucleico que codifican variantes de CBH2 como se describe anteriormente. Los constructos comprenden un vector, tal como un plásmido o vector viral, en el que se ha insertado una secuencia de la exposición, en una orientación directa o inversa.

5 **[0111]** Los constructos de ácido nucleico heterólogos pueden incluir la secuencia codificadora para la variante de cbh2: (i) de manera aislada; (ii) en combinación con secuencias codificadoras adicionales; tales como secuencias codificadoras de proteína de fusión o de péptido señal, en las que la secuencia que codifica cbh2 es la secuencia codificadora dominante; (iii) en combinación con secuencias no codificadoras, tales como intrones y elementos de control, tales como elementos promotores y terminadores o regiones 5' y/o 3' no traducidas, eficaces para la
10 expresión de la secuencia codificadora en un huésped adecuado; y/o (iv) en un vector o medio huésped en el que la secuencia que codifica cbh2 es un gen heterólogo.

[0112] En un aspecto de la presente exposición, se emplea un constructo de ácido nucleico heterólogo para transferir una secuencia de ácido nucleico que codifica una variante de CBH2 en una célula *in vitro*, con líneas de levadura y de hongos filamentosos establecidos preferidas. Para la producción a largo plazo de la variante de
15 CBH2, se prefiere la expresión estable. Se deduce que se puede utilizar cualquier método efectivo para generar transformantes estables en la práctica de la exposición.

[0113] Los vectores adecuados están normalmente equipados con una secuencia de ácido nucleico que codifica un marcador seleccionable, sitios de inserción y elementos de control adecuados, tales como secuencias de promotor y de terminación. El vector puede comprender secuencias reguladoras, que incluyen, por ejemplo,
20 secuencias no codificadoras, tales como intrones y elementos de control, es decir, elementos promotores y terminadores o regiones 5' y/o 3' no traducidas, eficaces para la expresión de la secuencia codificadora en células huésped (y/o en un medio de célula huésped o vector en el que normalmente no se expresa una secuencia codificadora de antígeno de proteína soluble modificada), operativamente unidas a la secuencia codificadora. Los expertos en la materia conocen un gran número de vectores y promotores adecuados, muchos de los cuales están disponibles comercialmente y/o se describen en Sambrook, et al., (*supra*).

[0114] Promotores de ejemplo incluyen promotores constitutivos y promotores inducibles, ejemplos de los cuales incluyen un promotor CMV, un promotor SV40 temprano, un promotor RSV, un promotor EF-1 alfa, un promotor que contiene el elemento de respuesta *Tet* (TRE, por sus siglas en inglés) en el sistema *tet-on* o *tet-off* como se describe (ClonTech y BASF), el promotor beta actina y el promotor metalotionina que se pueden regular por
30 incremento mediante la adición de determinadas sales de metales. Una secuencia promotora es una secuencia de ADN que es reconocida por el hongo filamentosamente concreto con fines de expresión. Está operativamente unida a una secuencia de ADN que codifica un polipéptido de variante de CBH2. Dicha unión comprende el posicionamiento del promotor con respecto al codón de iniciación de la secuencia de ADN que codifica el polipéptido de variante de CBH2 en los vectores de expresión expuestos. La secuencia promotora contiene
35 secuencia de control de la transcripción y la traducción que media en la expresión del polipéptido de variante de CBH2. Ejemplos incluyen los promotores de 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 proteinasa aspártica de *A. niger* o *Rhizomucor miehei*; los genes que codifican *cbh1*, *cbh2*, *egl1*, *egl2*, u otras celulasas de *H. jecorina* (*T. reesei*).

40 **[0115]** La elección del marcador seleccionable adecuado dependerá de la célula huésped, y en la técnica se conocen bien los marcadores adecuados para diferentes huéspedes. Los genes marcadores seleccionables típicos incluyen *argB* de *A. nidulans* o *T. reesei*, *amdS* de *A. nidulans*, *pyr4* de *Neurospora crassa* o *T. reesei*, *pyrG* de *Aspergillus niger* o *A. nidulans*. Ejemplos adicionales de marcadores seleccionables incluyen, pero sin carácter limitativo, *trpc*, *trp1*, *olic31*, *niaD* o *leu2*, que se incluyen en constructos de ácido nucleico heterólogo utilizados
45 para transformar una cepa mutante tal como *trp-*, *pyr-*, *leu-* y similares.

[0116] 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 crezcan en acetamida como fuente de nitrógeno. El marcador seleccionable (p. ej., *pyrG*) puede restaurar la capacidad de una cepa mutante
50 auxotrófica para crecer en un medio mínimo selectivo o el marcador seleccionable (p. ej., *olic31*) puede conferir a los transformantes la capacidad de crecer en presencia de un fármaco inhibidor o antibiótico.

[0117] La secuencia codificadora del marcador seleccionable se clona en cualquier plásmido adecuado utilizando métodos generalmente empleados en la técnica. Ejemplos de plásmidos incluyen pUC18, pBR322, pRAX y pUC100. El plásmido pRAX contiene secuencias AMAL de *A. nidulans*, que posibilitan la replicación en *A. niger*.

[0118] La práctica de la presente exposición empleará, a menos que se indique lo contrario, técnicas convencionales de biología molecular, microbiología, ADN recombinante e inmunología, que están dentro de la técnica. Dichas técnicas se explican de forma completa en la literatura. Véase, por ejemplo, Sambrook et al., 1989; Freshney, *Animal Cell Culture*, 1987; Ausubel, et al., 1993; y Coligan et al., *Current Protocols in Immunology*, 1991.

B. Células huésped y condiciones de cultivo para la producción de CBH2

(i) Hongos filamentosos

[0119] De esta manera, la presente exposición proporciona hongos filamentosos que comprenden células que han sido modificadas, seleccionadas y cultivadas de manera efectiva para dar lugar a la producción o expresión de variante de CBH2 con respecto a los correspondientes hongos parentales no transformados.

[0120] Ejemplos de especies de hongos filamentosos parentales que se pueden tratar y/o modificar para la expresión de variante de CBH2 incluyen, pero sin carácter limitativo, *Trichoderma*, por ejemplo, *Trichoderma reesei*, *Trichoderma longibrachiatum*, *Trichoderma viride*, *Trichoderma koningii*; *Penicillium sp.*, *Hemicella sp.*, incluyendo *Hemicella insolens*; *Aspergillus sp.*, *Chrysosporium sp.*, *Fusarium sp.*, *Hypocrea sp.*, y *Emericella sp.*

[0121] Las células que expresan CBH2 se cultivan bajo condiciones normalmente empleadas para cultivar la línea fúngica parental. Generalmente, las células se cultivan en un medio estándar que contiene sales fisiológicas y nutrientes, tal como se describe en Pourquie, J. et al., *Biochemistry and Genetics of Cellulose Degradation*, eds. Aubert, J. P. et al., Academic Press, pp. 71-86, 1988 y Ilmen, M. et al., *Appl. Environ. Microbiol.* 63:1298-1306, 1997). Las condiciones de cultivo también son estándar, por ejemplo, los cultivos se incuban a 28 °C en agitadores de cultivos o fermentadores hasta que se consiguen niveles de expresión de CBH2 deseados.

[0122] Las condiciones de cultivo preferidas para un hongo filamentosos determinado se pueden encontrar en la literatura científica y/o de la fuente de los hongos, tal como la American Type Culture Collection (ATCC; www.atcc.org/). Después de establecer crecimiento fúngico, las células se exponen a condiciones efectivas para causar o permitir la expresión de la variante de CBH2.

[0123] En casos en los que una secuencia que codifica CBH2 está bajo el control de un promotor inducible, el agente inductor, por ejemplo, un azúcar, una sal de metal o antibióticos, se añade al medio en una concentración efectiva para inducir la expresión de CBH2.

[0124] 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). Se conocen otras cepas de *Aspergillus niger varawamori*, tales como GCDAP3, GCDAP4 y GAP3-4 (Ward et al., *Appl. Microbiol. Biotechnol.* 39:738-743, 1993).

[0125] 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 de celulasa. Los equivalentes funcionales de RL-P37 incluyen la cepa RUT-C30 (ATCC n.º 56765) y la cepa QM9414 (ATCC n.º 26921) de *Trichoderma reesei*. Se contempla que estas cepas también serían útiles en la sobreexpresión de la variante de CBH2.

[0126] Cuando se desea obtener la variante de CBH2 en ausencia de actividad celolítica nativa potencialmente perjudicial, es útil obtener una cepa de célula huésped de *Trichoderma* en la que se hayan delecionado uno o varios genes de celulasa antes de la introducción de un constructo de ADN o plásmido que contiene el fragmento de ADN que codifica la variante de CBH2. Dichas cepas se pueden preparar mediante el método expuesto en la patente estadounidense con n.º 5,246,853 y WO 92/06209. Mediante la expresión de una variante de celulasa CBH2 en un microorganismo huésped al que le falta uno o varios genes de celulasa, se simplifican los procedimientos de identificación y posterior purificación. Cualquier gen de *Trichoderma sp.* que se haya clonado se puede delecionar, por ejemplo, los genes *cbh1*, *cbh2*, *egl1*, y *egl2*, así como los que codifican la proteína EG III y/o EGV (véase, por ejemplo, la patente estadounidense con n.º 5,475,101 y WO 94/28117, respectivamente).

[0127] La delección de genes se puede conseguir insertando una forma del gen que se desea delecionar o alterar en un plásmido mediante métodos conocidos en la técnica. El plásmido de delección se corta a continuación en un sitio o sitios de enzimas de restricción adecuado(s), interno a la región codificadora del gen deseado, y la secuencia codificadora del gen o parte de la misma se sustituye por un marcador seleccionable. Las secuencias

de ADN flanqueantes del lugar del gen que se ha de delecionar o alterar, preferiblemente entre aproximadamente 0,5 kb y 2,0 kb, permanecen en cualquier lado del gen del marcador seleccionable. Un plásmido de delección adecuado tendrá generalmente sitios únicos de enzimas de restricción presentes en el mismo para permitir que se eliminen el fragmento que contiene el gen delecionado, incluyendo las secuencias de ADN flanqueantes, y el gen del marcador seleccionable como una única pieza lineal.

[0128] Se debe elegir un marcador seleccionable que permita la detección del microorganismo transformado. Cualquier gen del marcador seleccionable que se exprese en el microorganismo seleccionado será adecuado. Por ejemplo, con *Aspergillus sp.*, el marcador seleccionable se elige de manera que la presencia del marcador seleccionable en los transformantes no afecte significativamente a las propiedades de los mismos. Dicho marcador seleccionable puede ser un gen que codifica un producto analizable. Por ejemplo, se puede utilizar una copia funcional de un gen de *Aspergillus sp.* que, si falta en la cepa huésped, dé lugar a que la cepa huésped muestre un fenotipo auxotrófico. De manera similar, existen marcadores seleccionables para *Trichoderma sp.*

[0129] En un modo de realización, se transforma una cepa derivada de pyrG- de *Aspergillus sp.* con un gen pyrG funcional, que proporciona de esta manera 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, por sus siglas en inglés). El gen pyrG codifica la orotidina-5'-monofosfato descarboxilasa, una enzima requerida para la biosíntesis de uridina. Las cepas con un gen pyrG intacto crecen en un medio que carece de uridina, pero son sensibles al ácido fluoroorótico. Es posible seleccionar cepas derivadas de pyrG- que carecen de una enzima de orotidina monofosfato descarboxilasa funcional y requieren uridina para el crecimiento mediante la selección de resistencia al FOA. Mediante la utilización de la técnica de selección del FOA también es posible obtener cepas que requieren uridina que carecen de una orotato pirofosforribosil 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 realiza fácilmente utilizando la técnica de resistencia al FOA mencionada anteriormente y, de este modo, el gen pyrG se emplea preferiblemente como marcador seleccionable.

[0130] En un segundo modo de realización, se transforma una cepa derivada de pyr4- de *Hyprocrea sp.* (*Trichoderma sp.*) con un gen pyr4 funcional, que proporciona de esta manera un marcador seleccionable para la transformación. Se puede obtener una cepa derivada de pyr4.sup.- mediante la selección de cepas de *Hyprocrea sp.* (*Trichoderma sp.*) que son resistentes al ácido fluoroorótico (FOA, por sus siglas en inglés). El gen pyr4 codifica la orotidina-5'-monofosfato descarboxilasa, una enzima requerida para la biosíntesis de uridina. Las cepas con un gen pyr4 intacto crecen en un medio que carece de uridina, pero son sensibles al ácido fluoroorótico. Es posible seleccionar cepas derivadas de pyr4.sup.- que carecen de una enzima de orotidina monofosfato descarboxilasa funcional y requieren uridina para el crecimiento mediante la selección de resistencia al FOA. Mediante la utilización de la técnica de selección del FOA también es posible obtener cepas que requieren uridina que carecen de una orotato pirofosforribosil transferasa funcional. Es posible transformar estas células con una copia funcional del gen que codifica esta enzima (Berges & Barreau, 1991). La selección de cepas derivadas se realiza fácilmente utilizando la técnica de resistencia al FOA mencionada anteriormente y, de este modo, el gen pyr4 se emplea preferiblemente como marcador seleccionable.

[0131] Para transformar pyrG.- de *Aspergillus sp.* o pyr4- de *Hyprocrea sp.* (*Trichoderma sp.*) para que carezcan de la capacidad para expresar uno o varios genes de celulasa, un fragmento de ADN simple que comprende un gen de celulasa alterado o delecionado se aísla entonces del plásmido de delección y se utiliza para transformar un huésped pyr de *Aspergillus* o pyr- de *Trichoderma* adecuado. A continuación, se identifican los transformantes y se seleccionan basándose en su capacidad para expresar el producto génico de pyrG o pyr4, respectivamente, y complementar de esta manera la auxotrofia de uridina de la cepa huésped. A continuación, se lleva a cabo un análisis de transferencia Southern en los transformantes resultantes para identificar y confirmar un caso de integración de doble cruzamiento que sustituye parte o toda la región codificadora de la copia genómica del gen que se ha de delecionar por los marcadores seleccionables pyr adecuados.

[0132] Aunque los vectores plasmídicos específicos descritos anteriormente se refieren a la preparación de transformantes de pyr, la presente exposición no se limita a estos vectores. Se pueden delecionar y sustituir diversos genes en la cepa de *Aspergillus sp.* o *Hyprocrea sp.* (*Trichoderma sp.*) mediante la utilización de las técnicas anteriores. Además, se puede utilizar cualquier marcador seleccionable disponible, como se describe anteriormente. De hecho, cualquier gen huésped, por ejemplo, *Aspergillus sp.* o *Hyprocrea sp.*, que haya sido clonado y, de esta manera, identificado, se puede delecionar del genoma utilizando la estrategia descrita anteriormente.

[0133] Como se indica anteriormente, las cepas huésped utilizadas pueden ser derivadas de *Hyprocrea sp.* (*Trichoderma sp.*) que carecen de o tienen un gen o genes no funcional(es) correspondiente(s) al marcador

seleccionable elegido. Por ejemplo, si se elige el marcador seleccionable de *pyrG* para *Aspergillus sp.*, entonces se utiliza una cepa derivada de *pyrG*- específica como receptor en el procedimiento de transformación. Además, por ejemplo, si se elige el marcador seleccionable de *pyr4* para una *Hyprocrea sp.*, entonces se utiliza una cepa derivada de *pyr4*- específica como receptor en el procedimiento de transformación. De manera similar, se pueden utilizar los marcadores seleccionables que comprenden genes de *Hyprocrea sp.* (*Trichoderma sp.*) equivalentes a los genes *amdS*, *argB*, *trpC*, *niaD* de *Aspergillus nidulans*. Por lo tanto, la cepa receptora correspondiente debe ser una cepa derivada, tal como *argB*-, *trpC*-, *niaD*-, respectivamente.

[0134] A continuación, se prepara el ADN que codifica la variante de CBH2 para la inserción en un microorganismo adecuado. De acuerdo con la presente exposición, el ADN que codifica una variante de CBH2 comprende el ADN necesario para codificar una proteína que tiene actividad celulolítica funcional. El fragmento de ADN que codifica la variante de CBH2 puede unirse funcionalmente a una secuencia promotora fúngica, por ejemplo, el promotor del gen *glaA* en *Aspergillus* o el promotor de los genes *cbh1* o *egl1* en *Trichoderma*.

[0135] También se contempla que más de una copia de ADN que codifica una variante de CBH2 se pueda recombinar en la cepa para facilitar la sobreexpresión. El ADN que codifica la variante de CBH2 se puede preparar mediante la construcción de un vector de expresión que porta el ADN que codifica la variante. El vector de expresión que porta el fragmento de ADN insertado que codifica la variante de CBH2 puede ser cualquier vector que sea capaz de replicarse de manera autónoma 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 los genes. El primero contiene secuencias de ADN en las que el promotor, la región codificadora del gen y la secuencia de terminación se originan todas del gen que se ha de expresar. El truncamiento del gen se puede obtener cuando se desee deleccionando las secuencias de ADN no deseadas (p. ej., que codifican dominios no deseados) para dejar que el dominio se exprese bajo el control de sus propias secuencias reguladoras de la transcripción y la traducción. Un marcador seleccionable también puede contenerse en el vector que permite la selección para la integración en el huésped de múltiples copias de las nuevas secuencias de genes.

[0136] El segundo tipo de vector de expresión se preensambla y contiene secuencias requeridas para la transcripción a alto nivel y un marcador seleccionable. Se contempla que la región codificadora para un gen o parte del mismo se puede insertar en este vector de expresión de propósito general, de manera que está bajo el control transcripcional de las secuencias de promotor y terminador de los casetes de expresión.

[0137] Por ejemplo, en *Aspergillus*, pRAX es dicho vector de expresión de propósito general. Se pueden insertar genes o partes de los mismos en dirección 3' del promotor *glaa* fuerte.

[0138] Por ejemplo, en *Hyprocrea*, pTEX es dicho vector de expresión de propósito general. Pueden insertarse genes o partes de los mismos en dirección 3' del promotor *cbh1* fuerte.

[0139] En el vector, la secuencia de ADN que codifica la variante de CBH2 de la presente exposición debe estar operativamente unida a las secuencias de transcripción y traducción, es decir, una secuencia promotora y una secuencia señal adecuadas en el marco de lectura al gen estructural. El promotor puede ser cualquier secuencia de ADN que muestra actividad transcripcional en la célula huésped y puede derivarse de genes que codifican proteínas homólogas o heterólogas a la célula huésped. Un péptido señal opcional proporciona la producción extracelular de la variante de CBH2. El ADN que codifica la secuencia señal es preferiblemente el que está asociado de manera natural al gen que se ha de expresar, aunque se contempla la secuencia señal de cualquier fuente adecuada, por ejemplo, una exocelobiohidrolasa o endoglucanasa de *Trichoderma*.

[0140] Los procedimientos utilizados para ligar las secuencias de ADN que codifican la variante de CBH2 de la presente exposición con el promotor, y la inserción en vectores adecuados se conocen bien en la técnica.

[0141] El vector o constructo de ADN descrito anteriormente puede introducirse en la célula huésped según técnicas conocidas, tales como transformación, transfección, microinyección, microporación, bombardeo biolístico y similares.

[0142] En la técnica de transformación preferida, debe tenerse en cuenta que la permeabilidad de la pared celular al ADN en *Hyprocrea sp.* (*Trichoderma sp.*) es muy baja. Por consiguiente, la absorción de la secuencia de ADN deseada, el gen o el fragmento de gen es, como mucho, mínima. Existen varios métodos para aumentar la permeabilidad de la pared celular de *Hyprocrea sp.* (*Trichoderma sp.*) en la cepa derivada (es decir, que carece de un gen funcional correspondiente al marcador seleccionable utilizado) antes del proceso de transformación.

5 **[0143]** El método preferido en la presente exposición para preparar *Aspergillus sp.* o *Hyprocrea sp.* (*Trichoderma sp.*) para la transformación conlleva la preparación de protoplastos 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 vegetativas germinadas. El micelio se trata con una enzima que digiere la pared celular dando lugar a protoplastos. A continuación, los protoplastos se protegen mediante la presencia de un estabilizador osmótico en el medio de suspensión. Estos estabilizadores incluyen sorbitol, manitol, cloruro de potasio, sulfato de magnesio y similares. Normalmente, la concentración de estos estabilizadores varía entre 0,8 M y 1,2 M. Es preferible utilizar aproximadamente una solución de 1,2 M de sorbitol en el medio de suspensión.

10 **[0144]** La absorción del ADN en la cepa huésped, (*Aspergillus sp.* o *Hyprocrea sp.* (*Trichoderma sp.*), depende de la concentración de ion calcio. En general, se utilizan entre aproximadamente 10 mM de CaCl₂.sub.2 y 50 mM de CaCl₂.sub.2 en una solución de absorción. Además de la necesidad del ion calcio en la solución de absorción, otros elementos incluidos generalmente son un sistema tampón, tal como un 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).
15 Se cree que el polietilenglicol actúa para fusionar las membranas celulares, permitiendo de esta manera que se distribuya el contenido del medio en el citoplasma de la célula huésped, a modo de ejemplo la cepa de *Aspergillus sp.* o *Hyprocrea sp.* y el ADN plasmídico se transfiere al núcleo. Esta fusión deja frecuentemente múltiples copias del ADN plasmídico integradas en el cromosoma huésped.

20 **[0145]** Normalmente, en la transformación se utiliza una suspensión que contiene los protoplastos o células de *Aspergillus sp.* que han sido sometidos a un tratamiento de permeabilidad a una densidad de 10.sup.5 a 10.sup.6/mL, preferiblemente 2 veces 10.sup.5/mL. De forma similar, en la transformación se utiliza una suspensión que contiene los protoplastos o células de *Hyprocrea sp.* (*Trichoderma sp.*) que han sido sometidos a un tratamiento de permeabilidad a una densidad de 10.sump.8 a 10.sup.9/mL, preferiblemente 2 veces 10.sup.8/mL. Se mezcla un volumen de 100 .mu.L de estos protoplastos o células en una solución adecuada (p. ej., 1,2 M de sorbitol; 50 mM de CaCl₂.sub.2) con el ADN deseado. En general, se añade una concentración alta de PEG a la solución de absorción. 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. También se pueden añadir aditivos, tales como dimetilsulfóxido, heparina, espermidina, cloruro de potasio y similares, a la solución de absorción y ayudar en la transformación.

30 **[0146]** En general, la mezcla se incuba a continuación a aproximadamente 0 °C durante un periodo entre 10 y 30 minutos. A continuación se añade PEG adicional a la mezcla para mejorar de manera adicional la absorción de la secuencia de ADN o del gen deseado. El PEG 4000 al 25 % se añade generalmente en volúmenes de 5 a 15 veces el volumen de la mezcla de transformación; no obstante, volúmenes mayores o menores pueden ser adecuados. El PEG 4000 al 25 % es preferiblemente aproximadamente 10 veces el volumen de la mezcla de transformación.
35 Después de añadir el PEG, la mezcla de transformación se incuba a continuación a temperatura ambiente o en hielo antes de la adición de una solución de sorbitol y CaCl₂.sub.2. A continuación, la suspensión de protoplastos se añade a alícuotas fundidas de un medio de crecimiento. Este medio de crecimiento permite el crecimiento de transformantes únicamente. Se puede utilizar cualquier medio de crecimiento en la presente exposición que sea adecuado para el crecimiento de los transformantes deseados. No obstante, si se seleccionan los transformantes Pyr.sup.+, es preferible utilizar un medio de crecimiento que no contenga uridina. Las colonias posteriores se transfieren y se purifican en un medio de crecimiento agotado de uridina.

40 **[0147]** En esta etapa, se pueden diferenciar los transformantes estables de los transformantes inestables por su índice de crecimiento más rápido y, en *Trichoderma*, por ejemplo, la formación de colonias circulares con un contorno liso en vez de irregular en el medio de cultivo sólido que carece de uridina. De manera adicional, en algunos casos se puede realizar una prueba adicional de estabilidad mediante el crecimiento de los transformantes en un medio 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 posteriormente y crecerán en un medio selectivo que carece de uridina.

50 **[0148]** En un modo de realización concreto del método anterior, la(s) variante(s) de CBH2 se recupera(n) en forma activa de la célula huésped después del crecimiento en medios líquidos como resultado del procesamiento post traduccional adecuado de la variante de CBH2.

(ii) Levadura

55 **[0149]** La presente exposición también contempla la utilización de levadura como célula huésped para la producción de CBH2. Algunos otros genes que codifican enzimas hidrolíticas se han expresado en diversas cepas de la levadura *S.cerevisiae*. Estos incluyen secuencias que codifican dos endoglucanasas (Penttila et al., *Yeast* vol. 3, pp 175-185, 1987), dos celobiohidrolasas (Penttila et al., *Gene*, 63: 103-112, 1988) y una beta-

glucosidasa de *Trichoderma reesei* (Cummings y Fowler, Curr. Genet. 29:227-233, 1996), una xilanasa de *Aureobasidium pullulans* (Li y Ljungdahl, Appl. Environ. Microbiol. 62, n.º 1, pp. 209-213, 1996), una alfa amilasa de trigo (Rothstein et al., Gene 55:353-356, 1987), etc. Además, un casete del gen de celulasa que codifica la endo-[beta]-1,4-gluconasa (END1) de *Butyrivibrio fibrisolvens*, celobiohidrolasa (CBH1) de *Phanerochaete chrysosporium*, la celodextrinasa (CEL1) de *Ruminococcus flavefaciens* y la celobiasa de *Endomyces fibrilizer* (Bgl1) se expresó satisfactoriamente en una cepa de laboratorio de *S. cerevisiae* (Van Rensburg et al., Yeast, vol. 14, pp. 67-76, 1998).

C. Introducción de una secuencia de ácido nucleico que codifica CBH2 en células huésped.

10 [0150] La exposición proporciona, además, células y composiciones celulares que han sido modificadas genéticamente para comprender una secuencia de ácido nucleico que codifica una variante de CBH2 proporcionada de manera exógena. Se puede modificar genéticamente una célula parental o línea celular (es decir, transducirse, transformarse o transfectarse) con un vector de clonación o un vector de expresión. El vector puede estar, por ejemplo, en forma de un plásmido, una partícula viral, un fago, etc., como se describe de manera adicional anteriormente.

15 [0151] Los métodos de transformación de la presente exposición pueden dar lugar a la integración estable de todo o parte del vector de transformación en el genoma del hongo filamentoso. No obstante, también se contempla la transformación que da lugar al mantenimiento de un vector de transformación extracromosómico autorreplicante.

20 [0152] Se pueden utilizar muchos métodos de transfección estándar para producir líneas celulares de *Trichoderma reesei* que expresen grandes cantidades de la proteína heteróloga. Algunos de los métodos publicados para la introducción de constructos de ADN en cepas que producen celulasa de *Trichoderma* incluyen Lorito, Hayes, DiPietro y Harman, 1993, Curr. Genet. 24: 349-356; Goldman, VanMontagu y Herrera-Estrella, 1990, Curr. Genet. 17:169-174; Penttila, Nevalainen, Ratto, Salminen y Knowles, 1987, Gene 6: 155-164, para *Aspergillus Yelton*, Hamer y Timberlake, 1984, Proc. Natl. Acad. Sci. USA 81: 1470-1474, para *Fusarium Bajar*, Podila y Kolattukudy, 1991, Proc. Natl. Acad. Sci. USA 88: 8202-8212, para *Streptomyces Hopwood* et al., 1985, The John Innes Foundation, Norwich, UK y para *Bacillus Brigidi*, DeRossi, Bertarini, Riccardi y Matteuzzi, 1990, FEMS Microbiol. Lett. 55: 135-138).

30 [0153] Otros métodos para introducir un constructo de ácido nucleico heterólogo (vector de expresión) en hongos filamentosos (p. ej., *H. jecorina*) incluyen, pero sin carácter limitativo, la utilización de una pistola de genes o de partículas, permeabilización de las paredes celulares de hongos filamentosos antes del proceso de transformación (p. ej., mediante la utilización de altas concentraciones de álcali, por ejemplo, 0,05 M a 0,4 M de CaCl₂.sub.2 o acetato de litio), fusión de protoplastos o transformación mediada por *Agrobacterium*. Se describe un método de ejemplo para la transformación de hongos filamentosos mediante el tratamiento de protoplastos o esferoplastos con polietilenglicol y CaCl₂.sub.2 en Campbell, E.1. et al., Curr. Genet. 16:53-56, 1989 y Penttila, M. et al., Gene, 63:11-22, 1988.

40 [0154] Se puede utilizar cualquiera de los procedimientos conocidos para introducir secuencias de nucleótidos exógenas en células huésped. Estos incluyen la utilización de la transfección con fosfato de calcio, polibreno, fusión de protoplastos, electroporación, biolística, liposomas, microinyección, vectores plasma, vectores virales y cualquiera de los otros métodos conocidos para introducir ADN genómico clonado, ADNc, ADN sintético u otro material genético exógeno en una célula huésped (véase, p. ej., Sambrook et al., supra). También se utiliza el método de transfección mediada por *Agrobacterium* descrito en la patente estadounidense con n.º 6,255,115. Solo es necesario que el procedimiento de ingeniería genética concreto utilizado sea capaz de introducir de manera satisfactoria al menos un gen en la célula huésped capaz de expresar el gen heterólogo.

45 [0155] Además, los constructos de ácido nucleico heterólogos que comprenden una secuencia de ácido nucleico que codifica una variante de CBH2 se pueden transcribir *in vitro*, y el ARN resultante se puede introducir en la célula huésped mediante los métodos conocidos, por ejemplo, mediante inyección.

50 [0156] La exposición incluye, además, transformantes nuevos y útiles de hongos filamentosos, tales como *H. jecorina* y *A. niger* para utilizarse en la producción de composiciones de celulasa fúngicas. La exposición incluye transformantes de hongos filamentosos, especialmente hongos que comprenden la secuencia que codifica la variante de CBH2, o la delección de la secuencia que codifica la cbh endógena.

[0157] Después de la introducción de un constructo de ácido nucleico heterólogo que comprende la secuencia codificadora para una variante de cbh2, las células modificadas genéticamente se pueden cultivar en medios de nutrientes convencionales modificados según convenga para activar los promotores, seleccionar los

transformantes o aumentar la expresión de una secuencia de ácido nucleico que codifica la variante de CBH2. Las condiciones del cultivo, tales como la temperatura, el pH y similares, son las utilizadas anteriormente para la célula huésped seleccionada para la expresión, y resultarán evidentes para los expertos en la materia.

5 [0158] Generalmente, se considera que la progenie de células en las que se han introducido dichos constructos de ácido nucleico heterólogos comprende la secuencia de ácido nucleico que codifica la variante de CBH2 encontrada en el constructo de ácido nucleico heterólogo.

10 [0159] La exposición incluye, además, transformantes nuevos y útiles de hongos filamentosos, tales como *H. jecorina* para utilizarse en la producción de composiciones de celulasa fúngicas. También se puede utilizar *Aspergillus niger* en la producción de la variante de CBH2. La exposición incluye transformantes de hongos filamentosos, especialmente hongos que comprenden la secuencia que codifica la variante de cbh 2, o la delección de la secuencia que codifica la cbh2 endógena.

15 [0160] En general, se pueden diferenciar los transformantes estables de hongos filamentosos de los transformantes inestables por su índice de crecimiento más rápido y, en *Trichoderma*, por ejemplo, la formación de colonias circulares con un contorno liso en vez de irregular en el medio de cultivo sólido. De manera adicional, en algunos casos, se puede realizar una prueba adicional de estabilidad mediante el crecimiento de los transformantes en un medio sólido no selectivo, la recogida de esporas de este medio de cultivo y la determinación del porcentaje de estas esporas que germinarán posteriormente y crecerán en un medio selectivo.

VI. Aislamiento y purificación de proteína CBH2 recombinante

20 [0161] En general, una proteína de variante CBH2 producida en cultivo celular se secreta en el medio y se puede purificar o aislar, por ejemplo, eliminando los componentes no deseados del medio de cultivo celular. No obstante, en algunos casos, se puede producir una proteína de variante CBH2 en una forma celular que requiere la recuperación de un lisado celular. En tales casos, la proteína de variante CBH2 se purifica de las células en las que se produjo utilizando técnicas empleadas de forma rutinaria por los expertos en la materia. Los ejemplos incluyen, pero sin carácter limitativo, cromatografía de afinidad (Tilbeurgh et al., FEBS Lett. 16:215, 1984), métodos cromatográficos de intercambio iónico (Goyal et al., Bioresource Technol. 36:37-50, 1991; Fliess et al., Eur. J. Appl. Microbiol. Biotechnol. 17:314-318, 1983; Bhikhabhai et al., J. Appl. Biochem. 6:336-345, 1984; Ellouz et al., J. Chromatography 396:307-317, 1987), incluyendo intercambio iónico utilizando materiales con alto poder de resolución (Medve et al., J. Chromatography A 808:153-165, 1998), cromatografía de interacción hidrofóbica (Tomaz y Queiroz, J. Chromatography A 865:123-128, 1999), y separación de dos fases (Brumbauer, et al., Bioseparation 7:287-295, 1999).

[0162] Normalmente, la proteína de variante CBH2 se fracciona para segregar proteínas que tienen propiedades seleccionadas, tales como afinidad de unión a agentes de unión concretos, por ejemplo, anticuerpos o receptores; o que tienen un intervalo de peso molecular seleccionado, o un intervalo de puntos isoeléctricos.

35 [0163] Una vez se consigue la expresión de una proteína de variante CBH2 determinada, la proteína CBH2 producida de esta manera se purifica a partir de las células o cultivo celular. Ejemplos de procedimientos adecuados para tal purificación incluyen los siguientes: cromatografía en columna de afinidad de anticuerpos, cromatografía de intercambio iónico; precipitación etanólica; HPLC con inversión de fase; cromatografía en sílice o en una resina de intercambio de cationes como DEAE; cromatoenfoco; SDS-PAGE; precipitación con sulfato amónico y filtración por gel utilizando, por ejemplo, Sephadex G-75. Pueden emplearse diversos métodos de purificación de proteínas y dichos métodos son conocidos en la técnica y se describen, por ejemplo, en Deutscher, Methods in Enzymology, vol. 182, n.º 57, pp 779, 1990; Scopes, Methods Enzymol. 90: 479-91, 1982. La(s) etapa(s) de purificación seleccionada(s) dependerá(n), por ejemplo, de la naturaleza del proceso de producción utilizado y de la proteína concreta producida.

VII. Utilidad de cbh2 y CBH2

45 [0164] Se puede entender que los ácidos nucleicos de la variante de cbh, la proteína de variante CBH2 y las composiciones que comprenden actividad de proteína de variante CBH2 resultan útiles en una amplia variedad de aplicaciones, algunas de las cuales se describen a continuación.

50 [0165] Las nuevas y mejoradas composiciones de celulasa que comprenden cantidades distintas de celulasas de tipo BG, de tipo EG y de tipo variante de CBH son útiles en composiciones detergentes que muestran capacidad de limpieza mejorada, funcionan como un agente suavizante y/o mejoran el tacto de los tejidos de algodón (p. ej., "lavado a la piedra" o "biopulido"), en las composiciones para degradar pulpa de madera en azúcares (p. ej., para

la producción de bioetanol), y/o en composiciones para pienso. El aislamiento y caracterización de la celulasa de cada tipo proporciona la capacidad para controlar los aspectos de tales composiciones.

[0166] En un enfoque, la celulasa de la exposición es útil en composiciones detergentes o en el tratamiento de tejidos para mejorar el tacto y la apariencia.

- 5 **[0167]** Puesto que la velocidad de hidrólisis de los productos celulósicos se puede aumentar mediante la utilización de un transformante que tenga al menos una copia adicional del gen *cbh* insertado en el genoma, los productos que contienen celulosa o heteroglicanos se pueden degradar a un ritmo más rápido y en mayor medida. Los productos fabricados con celulosa, tales como papel, algodón, pañales de celulosa, y similares se pueden degradar de manera más eficiente en un vertedero. De esta manera, el producto de la fermentación
10 obtenible de los transformantes o los transformantes solos se puede utilizar en composiciones para ayudar a degradar mediante licuefacción una variedad de productos de celulosa que se añaden a los vertederos llenos.

- [0168]** La sacarificación y fermentación por separado es un proceso por el que la celulosa presente en la biomasa, por ejemplo, rastrojos de maíz, se convierte en glucosa y, posteriormente, las cepas de levadura convierten la glucosa en etanol. La sacarificación y fermentación simultánea es un proceso por el que la celulosa
15 presente en la biomasa, por ejemplo, rastrojos de maíz, se convierte en glucosa y, al mismo tiempo y en el mismo reactor, las cepas de levadura convierten la glucosa en etanol. De esta manera, en otro enfoque, la variante de celulasa de tipo CBH de la exposición resulta útil en la degradación de la biomasa en etanol. La producción de etanol a partir de fuentes de celulosa fácilmente disponibles proporciona una fuente de combustible renovable estable.

- 20 **[0169]** Las materias primas basadas en celulosa están compuestas por residuos agrícolas, pastos y maderas y otra biomasa de bajo valor, tal como residuos municipales (p. ej., papel reciclado, hierba cortada, etc.). El etanol se puede producir a partir de la fermentación de cualquiera de estas materias primas celulósicas. No obstante, la celulosa debe convertirse primero en azúcares antes de que pueda haber una conversión en etanol.

- [0170]** Se pueden utilizar una gran variedad de materias primas con la variante de CBH inventiva y la seleccionada para su uso puede depender de la región en la que se realiza la conversión. Por ejemplo, en el
25 medio oeste de Estados Unidos pueden predominar residuos agrícolas tales como paja de trigo, rastrojo de maíz y bagazo, mientras que en California puede predominar la paja de arroz. No obstante, debe entenderse que se puede utilizar cualquier biomasa celulósica disponible en cualquier región.

- [0171]** Los métodos de la presente exposición pueden utilizarse en la producción de monosacáridos, disacáridos y polisacáridos como materias primas químicas o de fermentación para microorganismos para la producción de
30 productos orgánicos, químicos y combustibles, plásticos y otros productos terminados o intermedios. En concreto, puede aumentar el valor de la transformación de residuos (residuos desecados de destilería, residuos desecados de cervecería, bagazo de caña de azúcar, etc.) mediante solubilización completa o parcial de la celulosa o hemicelulosa. Además del etanol, algunos productos químicos que pueden producirse a partir de la
35 celulosa y la hemicelulosa incluyen: acetona, acetato, glicina, lisina, ácidos orgánicos (p. ej., ácido láctico), 1,3-propanodiol, butanodiol, glicerol, etilenglicol, furfural, polihidroxialcanoatos, ácido cis, cis-mucónico, piensos animales y xilosa.

- [0172]** Una composición de celulasas que contiene una cantidad mejorada de celobiohidrolasa resulta útil en la
40 producción de etanol. El etanol de este proceso también se puede utilizar como un potenciador de octanaje o directamente como combustible en lugar de gasolina, lo cual es ventajoso puesto que el etanol como fuente de combustible es más ecológico que los productos derivados del petróleo. Se sabe que la utilización de etanol mejorará la calidad del aire y posiblemente reducirá los niveles locales de ozono y smog. Además, la utilización de etanol en lugar de gasolina puede ser de importancia estratégica en la amortiguación del impacto de los cambios repentinos en la energía no renovable y los suministros petroquímicos.

- 45 **[0173]** El etanol se puede producir a través de procesos de sacarificación y fermentación a partir de biomasa celulósica, tales como árboles, plantas herbáceas, residuos sólidos municipales y residuos agrícolas y forestales. No obstante, la proporción de enzimas de celulasa individuales en una mezcla de celulasas de origen natural producida mediante un microbio puede no ser la más eficaz para la conversión rápida de la celulosa de la biomasa en glucosa. Se sabe que las endoglucanasas actúan para producir nuevos extremos en la cadena de
50 celulosa que por sí mismos son sustratos para la acción de celobiohidrolasas y mejoran, de esta manera, la eficacia de la hidrólisis del sistema de celulasas entero. Por lo tanto, la utilización de actividad de celobiohidrolasa aumentada u optimizada puede mejorar en gran medida la producción de etanol.

[0174] De esta manera, la celobiohidrolasa inventiva resulta útil en la hidrólisis de celulosa en sus componentes de azúcar. En un modo de realización, se añade una variante de celobiohidrolasa a la biomasa antes de la adición de un organismo fermentador. En un segundo modo de realización, se añade una variante de celobiohidrolasa a la biomasa al mismo tiempo que un organismo fermentador. De manera opcional, puede haber otros componentes de celulosa presentes en cualquier modo de realización.

[0175] En otro modo de realización, la materia prima celulósica se puede pretratar. El pretratamiento puede ser mediante temperatura elevada y la adición de ácido diluido, ácido concentrado o una solución de álcali diluido. La solución de pretratamiento se añade durante un tiempo suficiente para hidrolizar al menos parcialmente los componentes de la hemicelulosa y después neutralizarlos.

[0176] El producto principal de la acción de CBH2 en la celulosa es la celobiosa, que está disponible para la conversión en glucosa mediante la actividad de BG (por ejemplo, en un producto de celulasa fúngica). Ya sea mediante el pretratamiento de la biomasa celulósica o mediante la acción enzimática en la biomasa, otros azúcares, además de la glucosa y la celobiosa, pueden estar disponibles a partir de la biomasa. El contenido de hemicelulosa de la biomasa se puede convertir (mediante hemicelulasas) en azúcares, tales como xilosa, galactosa, manosa y arabinosa. De esta manera, en un proceso de conversión de biomasa, la sacarificación enzimática puede producir azúcares que están disponibles para conversiones biológicas o químicas en otros productos intermedios o finales. Por lo tanto, los azúcares generados a partir de la biomasa son útiles en una variedad de procesos, además de la generación de etanol. Ejemplos de dichas conversiones son la fermentación de glucosa en etanol (como lo analiza M.E. Himmel et al. pp 2-45, en "Fuels and Chemicals from Biomass", ACS Symposium Series 666, ed B. C. Saha y J. Woodward, 1997) y otras conversiones biológicas de glucosa en 2,5-diceto-D-gluconato (patente estadounidense con n.º 6,599,722), ácido láctico (R. Datta y S-P. Tsai pp 224-236, *ibid*), succinato (R. R. Gokam, M. A. Eiteman y J. Sridhar pp 237-263, *ibid*), 1,3-propanodiol (A-P. Zheng, H. Biebl y W-D. Deckwer pp 264-279, *ibid*), 2,3-butanodiol (C. S. Gong, N. Cao y G. T. Tsao pp 280-293, *ibid*), y las conversiones químicas y biológicas de xilosa en xilitol (B. C. Saha y R. J. Bothast pp 307-319, *ibid*). Véase también, por ejemplo, WO 98/21339.

[0177] Las composiciones detergentes de la presente exposición pueden emplear, además de la composición de celulasas (independientemente del contenido de celobiohidrolasa, es decir, sin celobiohidrolasa, sustancialmente sin celobiohidrolasa, o con celobiohidrolasa mejorada), un surfactante, incluyendo surfactantes aniónicos, no iónicos y anfólicos, una hidrolasa, agentes mejoradores, agentes blanqueadores, agentes azulantes y tintes fluorescentes, inhibidores del apelmazamiento, solubilizantes, surfactantes catiónicos y similares. Todos estos componentes se conocen en el sector de los detergentes. La composición de celulasas como se describe anteriormente se puede añadir a la composición detergente ya sea en un diluyente líquido, en gránulos, en emulsiones, en geles, en pastas y similares. El experto en la materia conoce perfectamente dichas formas. Cuando se emplea una composición detergente sólida, la composición de celulasas se formula, preferiblemente, en forma de gránulos. Preferiblemente, los gránulos se pueden formular de manera que contengan un agente protector de celulasa. Para una exposición más completa, véase la patente estadounidense con n.º 6,162,782 titulada "Detergent compositions containing cellulase compositions deficient in CBH2 type components."

[0178] Preferiblemente, las composiciones de celulasas se emplean desde aproximadamente 0,00005 por ciento en peso hasta aproximadamente 5 por ciento en peso en relación con la composición detergente total. Más preferiblemente, las composiciones de celulasas se emplean desde aproximadamente 0,0002 por ciento en peso hasta aproximadamente 2 por ciento en peso en relación con la composición detergente total.

[0179] Además, la secuencia de ácido nucleico de la variante de CBH2 resulta útil en la identificación y caracterización de secuencias de ácido nucleico relacionadas. Varias técnicas útiles para determinar (predecir o confirmar) la función de los genes o productos génicos relacionados incluyen, pero sin carácter limitativo, (A) análisis de ADN/ARN, tal como (1) sobreexpresión, expresión ectópica, y expresión en otras especies; (2) bloqueo de genes (genética inversa, bloqueo dirigido, silenciamiento génico inducido por virus (VIGS, por sus siglas en inglés, véase Baulcombe, 100 Years of Virology, Calisher y Horzinek eds., Springer-Verlag, New York, NY 15:189-201, 1999); (3) análisis del estado de metilación del gen, especialmente las regiones reguladoras flanqueantes; y (4) hibridación *in situ*; (B) análisis de producto génico, tal como (1) expresión de proteína recombinante; (2) producción de antisueros, (3) inmunolocalización; (4) ensayos bioquímicos para actividad catalítica u otra actividad; (5) estado de fosforilación; e (6) interacción con otras proteínas a través de análisis de dos híbridos de levadura; (C) análisis de sistema, tal como colocar un gen o producto génico en el interior de un sistema bioquímico o de transducción de señales concreto basado en su fenotipo de sobreexpresión o mediante homología de 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 un sistema metabólico o de transducción de señales concreto, y ayudar a determinar la función del gen.

EXPERIMENTOS

[0180] La presente exposición se describe con más detalle en los siguientes ejemplos que, de ningún modo, pretenden limitar el alcance de la exposición, como se reivindica. Las figuras adjuntas deben considerarse partes integrales de la memoria y la descripción de la exposición. Los siguientes ejemplos se ofrecen a modo ilustrativo, pero no limitan la exposición reivindicada.

[0181] En la exposición experimental que sigue, se aplican las siguientes abreviaturas: M (molar); mM (milimolar); μ M (micromolar); nM (nanomolar); mol (moles); mmol (milimoles); μ mol (micromoles); nmol (nanomoles); g (gramos); mg (miligramos); μ g (microgramos); pg (picogramos); L (litros); ml y mL (mililitros); μ l y μ L (microlitros); cm (centímetros); mm (milímetros); μ m (micrómetros); nm (nanómetros); U (unidades); V (voltios); PM (peso molecular); s (segundos); min(s) (minuto/minutos); h (hora/horas); °C (grados centígrados); CS (cantidad suficiente); NH (no hecho); NA (no aplicable); rpm (revoluciones por minuto); H₂O (agua); dH₂O (agua desionizada); HCl (ácido clorhídrico); aa (aminoácido); pb (par de bases); kb (kilobase); kD (kilodaltons); ADNc (ADN complementario o copia); ADN (ácido desoxirribonucleico); ssDNA (ADN monocatenario); dsDNA (ADN bicatenario); dNTP (desoxirribonucleótido trifosfato); ARN (ácido ribonucleico); MgCl₂ (cloruro de magnesio); NaCl (cloruro de sodio); p/v (peso/volumen); % v (porcentaje en volumen); gr (gravedad); DO (densidad óptica); CNPG (cloro-nitrofenil-beta-D-glucósido); CNP (2-cloro-4-nitrofenol); BPA (bagazo pretratado con ácido); PASC (celulosa hinchada con ácido fosfórico) PCS (rastrero de maíz pretratado con ácido); IR (índice de rendimiento); PAGE (electroforesis en gel de poliacrilamida); PCR (reacción en cadena de polimerasa); RT-PCR (reacción en cadena de la polimerasa con transcriptasa inversa); y HPLC (cromatografía líquida de alto rendimiento).

EJEMPLO 1**Modificación química de una preparación de celulasa *Trichoderma sp.* y ensayos para analizar variantes de CBH2**

[0182] Este ejemplo describe el tratamiento de una preparación de celulasa *Trichoderma sp.* comercial (LAMINEX BG complejo enzimático (Genencor Division, Danisco US, Inc.) con anhídrido succínico para acetilar los residuos de lisina. La acetilación de los residuos de lisina del complejo enzimático LAMINEX BG altera la carga neta de las proteínas (p. ej., carga negativa aumentada). Pueden utilizarse otras modificaciones químicas similares también para convertir la carga positiva de la lisina en un grupo de carga negativa (por ejemplo, con tratamiento con anhídrido ftálico, anhídrido acetoxisuccínico, anhídrido maleico y anhídrido tartárico) o incluso en dos cargas negativas (por ejemplo, con tratamiento con anhídrido 4-nitroftálico, anhídrido trimelítico y anhídrido cis-aconítico). Pueden utilizarse otras modificaciones químicas para eliminar las cargas positivas de residuos de lisina que den lugar a un residuo no cargado (por ejemplo, tratamiento con anhídrido piválico, anhídrido acético, anhídrido butírico, anhídrido isobutírico, anhídrido hexanoico, anhídrido valérico y anhídrido isovalérico).

[0183] Los residuos de lisina en una preparación de celulasa se modificaron mediante la utilización de anhídrido succínico, utilizando una variación de métodos publicados (Lundblad, Chemical Reagents for Protein Modification, Editor: R. Lundblad, 3rd Edition CRC press, 1984). Para esta reacción, se preparó una muestra de 236 mg de complejo enzimático LAMINEX BG en 1 mL de tampón HEPES 500 mM, pH 8. Se preparó una solución de anhídrido succínico (Aldrich) mediante la disolución del polvo en dimetilsulfóxido a una concentración final de 500 mg/mL antes de la adición del complejo enzimático. Se añadió una alícuota de anhídrido succínico de modo que se consiguiera una proporción de lisina en relación con ácido succínico de > 1:100 en el tubo de reacción. Se preparó otro tubo de reacción solamente con dimetilsulfóxido y enzima, utilizando volúmenes similares, para servir como el control de proteína sin modificar. Se agitaron los tubos vorticialmente y se dejaron a temperatura ambiente durante la noche. El día siguiente, se añadió un volumen 1:10 de glicina 1M, pH 3, a cada tubo para extinguir la reacción de anhídrido succínico.

[0184] Se confirmó la modificación química mediante la comparación de proteínas modificadas y sin modificar en geles nativos. Se analizaron alícuotas de cada reacción (modificadas químicamente y sin modificar) en geles nativos en gradiente 8-25 %, con un pH 8.8 a 100 voltios (Phast System gels, GE Healthcare). Se visualizaron las proteínas tras la tinción del gel con azul de Coomassie, para confirmar que la modificación se realizó con éxito. La tinción mostró cambios en la migración de la banda proteica, y se confirmaron los cambios en la carga de los diversos componentes proteicos de la preparación de celulasa. Las muestras modificadas de preparación de celulasa *Trichoderma sp.* presentaban una carga más negativa que las muestras sin modificar.

[0185] Con el fin de aislar las proteínas modificadas y sin modificar (control), se desalaron alícuotas de 80 μ l de cada muestra mediante la utilización de columnas de centrifugación y desalinización (Pierce). La absorbancia a 280 nm de muestras desalinizadas (incluyendo el control sin modificación) se midió con un espectrofotómetro

NanoDrop™ (Thermo), en duplicados tras una dilución de muestra 1:10 para determinar la concentración total de proteínas de las muestras.

Determinaciones del potencial zeta

- [0186]** Este ejemplo describe la determinación del potencial zeta de una enzima y un sustrato. La presencia de una carga en la superficie de una partícula influye en la distribución de los iones en la región interfacial circundante. El resultado es una concentración de iones contrarios incrementada de carga opuesta a la de la partícula situada cerca de la superficie de la partícula. A medida que uno se aleja de la superficie de la partícula, la distribución heterogénea de iones será, finalmente, homogénea. La distancia a la que se obtiene una distribución homogénea se denomina longitud de Debye ($1/k$) o distancia de cribado, y depende de la fuerza iónica, como se muestra en la expresión que sigue, donde ϵ_0 es la permitividad de espacio libre (8.854×10^{-12} F m⁻¹), ϵ_r es la permitividad del líquido, k es la constante de Boltzmann (1.38×10^{-23} J K⁻¹), T es la temperatura en Kelvin, e es la carga electrónica (1.6022×10^{-19} C), I es la fuerza iónica molar, y NA es la constante de Avogadro (6.022×10^{23} mol⁻¹).

$$\text{Longitud de Debye} = \frac{1}{k} = \sqrt{\frac{\epsilon_0 \epsilon_r k T}{NA e^2 2I}} \quad (0.1)$$

- 15 La fuerza iónica molar puede calcularse a partir de la siguiente ecuación, donde C_i es la concentración de especies iónicas y Z_i es la valencia.

$$I = \frac{1}{2} \sum C_i Z_i^2 \quad (0.2)$$

Para agua a 298 K, la expresión de longitud de Debye se reduce a la forma siguiente.

$$k^{-1} = 0.304(I^{-0.5}) \quad (0.3)$$

- 20 La capa líquida alrededor de la partícula tiene dos partes; una región interior (capa de Stern) donde los iones están fuertemente unidos y una región exterior (difusa) donde están asociados con menos firmeza. En la capa difusa, existe un límite en el interior del cual los iones y las partículas forman una entidad estable. Cuando se mueve una partícula, los iones que se encuentran en este límite se mueven con ella. Los iones que se encuentran más allá del límite no se desplazan con la partícula. El potencial eléctrico en este límite, también denominado superficie de cizalla hidrodinámica, se define como potencial zeta.

[0187] En la dispersión de luz electroforética, el potencial zeta ζ se calcula a partir de la movilidad electroforética medida u mediante la utilización de la ecuación de Henry que se muestra a continuación, donde ϵ es la constante dieléctrica, h es la viscosidad de la solución, k es la longitud de Debye inversa, a es el radio de la partícula y $f(ka)$ es la función de Henry.

$$u = \frac{2e\zeta}{3h} f(ka) \quad (0.4)$$

- [0188]** Las unidades de k son longitud recíproca, siendo $1/k$ el "grosor" de la doble capa eléctrica (longitud de Debye). El parámetro a se refiere al radio de la partícula y, por consiguiente, ka es la proporción del radio de la partícula en relación con el grosor de la doble capa eléctrica. La función de Henry, $f(ka)$ depende de la forma de la partícula, pero se conoce para una esfera. En la expresión anterior, varía entre $f(0) = 1$ (límite de Hückel) y $f(\infty) = 1.5$ (límite de Smoluchowski). Para las partículas pequeñas como las proteínas en un medio dieléctrico bajo (o fuerza iónica baja), el límite de Hückel de $f(ka) = 1$ es el modelo más adecuado.

- [0189]** Los potenciales zeta de las proteínas se midieron con el Zetasizer NS (Malvern Instruments, Reino Unido) según el principio descrito anteriormente. Los potenciales zeta de tejidos teñidos con bismaleimida se midieron con el SurPass (Anton-Paar, Austria) utilizando la implementación del potencial de flujo del principio anterior. A partir de la definición de carga superficial, normalmente expresada en Culombios:

$$q_s = 4\pi e_0 \epsilon_r a (1 + ka) \zeta \quad (0.5)$$

También puede expresarse como una carga neta z multiplicada por la carga elemental e 1.6×10^{-19} C:

$$q_s = ze \quad (0.6)$$

Por consiguiente, el cambio esperado en el potencial zeta debido a un incremento de la carga neta es proporcionado por:

$$\frac{Dz}{Dz} = \frac{e}{4pe_o e_p a (1 + ka)} \quad (0.7)$$

- 5 También es posible medir los potenciales zeta con una técnica de gel nativo (Sparks et al., Journal of Lipid Research, 33:123-130, 1992) como se describe en el ejemplo 1. La movilidad electroforética medida con geles nativos es normalmente menor que en solución debido al retardo provocado por la matriz de gel. Los potenciales zeta calculados de esta manera son normalmente más bajos en comparación con los métodos basados en soluciones. Por lo tanto, nos referimos a ellos como potenciales zeta aparentes cuando se obtienen a través de
10 la técnica con gel nativo.

[0190] La carga efectiva en una formulación determinada se enumera como su potencial zeta. La utilización del potencial zeta como una escala de carga común posibilita la comparación de variantes de enzimas que tienen distintos pliegues (p. ej., serina proteasas, metaloproteasas, etc.), así como interacciones con distintos sustratos (p. ej., micromuestra de bismaleimida) en las condiciones de interés (p. ej., detergente American Association of
15 Textile Chemists and Colorists [AATCC] HDL). Si bien se prefiere el potencial zeta para comparar distintos pliegues proteicos, las movilidades electroforéticas o las cargas medidas también proporcionan una escala absoluta y son adecuadas para las comparaciones. El rendimiento de la bismaleimida como una función del potencial zeta de las enzimas se describe mediante una distribución normal estándar, indicada por la línea sólida, con una media μ igual a -9,68 mV, desviación típica σ de 11,39 mV y valor pico de 0,4056 [A600-fondo]. Esta
20 distribución se indica en coordenadas reducidas estándar por la actividad de la bismaleimida divididas por el valor pico en el eje Y derecho como una función de la puntuación Z en el eje X superior. La puntuación Z se define, como es habitual, como $(X-\mu)/\sigma$, donde X es, en este caso, el potencial zeta.

Tabla 1-1 Actividad en la micromuestra de bismaleimida de las proteasas		
Nivel de rendimiento	Puntuación Z	Potencial Z ζ Ventana* para la actividad de micromuestra de bismaleimida
90 %	$\pm 0,46$	$-14,92 < \zeta < -4,44$
80 %	$\pm 0,65$	$-17,08 < \zeta < -2,28$
70 %	$\pm 0,84$	$-19,25 < \zeta < -0,11$
60 %	$\pm 1,00$	$-21,07 < \zeta < +1,71$
50 %	$\pm 1,18$	$-23,12 < \zeta < +3,76$

Media $\mu = -9,68$ mV, desviación típica $\sigma = 11,39$ mV, potencial zeta $\zeta = Z^ \sigma + \mu$ Tampón de referencia: HEPES 5mM, pH 8.0, 2,5mM NaCl

[0191] La distribución normal es única para cada tintado de sustrato en condiciones de reacción determinadas (pH, conductividad, tipo de sal, quelantes detergentes, etc.). Distintas ventajas o resultados favorables siguen a una distribución normal con una propiedad física que se mantiene en las enzimas de diversos pliegues, como por ejemplo, el caso de niveles de expresión y potenciales zeta para variantes de carga en escalera ASP y NprE. En una distribución normal el valor pico se da en la media. La comparación de cargas de enzima y de sustrato en una escala de potencial zeta común revela que el rendimiento de la bismaleimida óptimo se da cuando el
30 potencial zeta de enzima medio, en este caso -9,68 mV, básicamente se empareja con el potencial zeta del tintado de sustrato, en este caso -8,97 mV, medido en las mismas condiciones.

[0192] Los niveles de rendimiento de las distribuciones normales estándar se describen convenientemente por lo que respecta a sus puntuaciones z como se indica en la tabla 1-1 (Véase, Abramowitz y Stegun, Handbook of
35 Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover, New York, 9th Ed., 1964). La conversión a potenciales zeta es conocimiento claro proporcionado de la media y la desviación típica definiendo la distribución para una aplicación determinada. En este ejemplo, el rendimiento de limpieza medido para un pliegue proteico se limita a valores de potencial zeta entre -40 mV y + 20 mV. Las variantes con un rendimiento de limpieza por encima de un 80 % de su pliegue óptimo (esto es $z = \pm 0,65$), se limitan a valores de potencial zeta entre -17,08 mV y -2,28 mV. Las variantes con un rendimiento de limpieza por encima de un 90 % de su
40 pliegue óptimo (esto es $z = \pm 0,46$), se limitan a valores de potencial zeta entre -14,92 mV y -4,44 mV.

[0193] Distintas manchas de sustrato (p. ej., hierba, manchas derivadas del cuerpo, tomate) tienen distintos potenciales zeta en la misma formulación y la misma mancha de sustrato tiene distintos potenciales zeta en

distintas formulaciones (p. ej., HDL norteamericano, detergente lavavajillas en polvo europeo). En cualquier caso, mientras que la carga de la mancha de sustrato varía, se espera que la desviación típica de la distribución normal siga siendo constante. El conocimiento de los potenciales zeta de la enzima y del sustrato en una formulación detergente determinada permite la identificación rápida del nivel de rendimiento esperado para esa variante, así como la dirección y la magnitud del cambio de carga necesarios para conseguir niveles de rendimiento óptimos. La medición del potencial zeta del sustrato en el medio de reacción deseado permite la optimización de la reacción enzimática en el sustrato particulado en dicho medio. Cualquier reacción enzimática en cualquier medio puede optimizarse utilizando un proceso similar.

Optimización de reacciones en sustratos que muestran carga variable

[0194] La conversión de celulosa se analizó con técnicas conocidas en la técnica (Véase, p. ej., Baker et al., Appl Biochem Biotechnol, 70-72:395-403, 1998). En los experimentos, se utilizó un ensayo de conversión celulósica estándar. En este ensayo, la enzima y el sustrato tamponado se colocaron en recipientes y se incubaron a una temperatura determinada durante un tiempo. La reacción se extinguió con suficiente glicina 100 mM, pH 11, para convertir el pH de la mezcla de reacción en al menos pH 10. Una vez se extinguió la reacción, se filtró una alícuota de la mezcla de reacción a través de una membrana de 0,2 micras para eliminar los sólidos. A continuación, se sometió a ensayo la solución filtrada para azúcares solubles mediante HPLC según los métodos descritos anteriormente en Baker et al.

Determinación del potencial zeta del rastrojo de maíz pretratado (PCS, por sus siglas en inglés)

[0195] Se pretrató el rastrojo de maíz con un 2 % de H₂SO₄ p/p, como se describe (Schell et al., J Appl Biochem Biotechnol, 105:69-86 [2003]) y, a continuación, se lavó varias veces con agua desionizada para obtener un pH de 4.5. Se añadió acetato de sodio para crear una concentración final de 50mM y se tituló la solución a pH 5.0. La concentración de celulosa en la mezcla de reacción era aproximadamente 7 %.

[0196] Las alícuotas de PCS antes y después de la sacarificación mediante una mezcla de celulasa comercial, Spezyme CP e Indiage 44L, se dosificaron en tubos de centrifugación Eppendorf de 1,5 mL y ocuparon aproximadamente un tercio del volumen. Las muestras se centrifugaron a 6,000 rpm durante 5 min, se cambió el sobrenadante por agua Milli-Q™ y se repitió el proceso 5 veces. Se preparó una solución madre de 100 mg/mL en agua Milli-Q™ a partir del rastrojo de maíz lavado. Se diluyó esta solución madre a 1 mg/mL en un tampón de acetato de sodio 50 mM, pH 5.0 para las mediciones de potencial zeta. Se transfirió una alícuota de 1 mL de cada muestra de sustrato a una cubeta desechable limpia ZetasizerNS™, de Malvern Instruments (Reino Unido).

[0197] La tabla 1-2 indica que durante la reacción de sacarificación, la carga del sustrato de PCS, expresada como potencial zeta, casi se convierte en el doble de negativa. Sin ceñirse a la teoría, existen muchas explicaciones para un aumento de la carga negativa neta incluyendo, pero sin carácter limitativo, enriquecimiento de la lignina, la parte no reactiva de este sustrato, así como la unión o inhibición no productiva de celulosas enteras y otras proteínas. Existe un potencial zeta enzimático óptimo para el rendimiento (p. ej., alcance de reacción y velocidad de reacción), que empareja el potencial zeta del sustrato en condiciones de medio de reacción. Distintos pretratamientos de biomasa pueden influenciar radicalmente la carga de sustrato inicial. En caso de que el potencial zeta de la enzima o del sustrato se desemparejen durante la reacción, la interacción enzima-sustrato ya no será óptima. Este efecto será radical para cambios de casi 10 mV, que son el caso para la conversión de biomasa.

[0198] Las estrategias para remediar esta situación incluyen, pero sin carácter limitativo, el suministro de una mezcla de enzimas que abarca varias cargas; un enfoque de proceso de lote alimentado donde las enzimas que poseen diferentes cargas en el nuevo punto óptimo se suministran en diferentes tiempos de reacción y/o alcances de conversión; control de la carga superficial de sustrato mediante adición de agentes de formulación, concretamente surfactantes (iónicos y no iónicos) u otras proteínas; control de la carga superficial de sustrato mediante ajustes del pH; ajustes de la fuerza iónica durante la reacción para cambiar los óptimos de carga de la enzima; filtración a través de membrana, concretamente ósmosis inversa y nanofiltración, para controlar la fuerza iónica durante la reacción; adición de quelantes para controlar la fuerza iónica a través de la eliminación de sales; y control de la carga del sustrato de biomasa a través de procesos de pretratamiento.

Tabla 1-2. Potencial zeta del rastrojo de maíz pretratado con ácido	
Condición PCS	Potencial zeta
Inicial (antes de la sacarificación)	-12,0 ± 7,00 mV
Tras la sacarificación mediante Spezyme CP	-22,2 ± 8,67 mV
Tras la sacarificación mediante Indiage 44L	-22,7 ± 6,84 mV

[0199] Los siguientes ensayos se utilizaron en los ejemplos descritos a continuación. Cualquier desviación de los protocolos proporcionados a continuación se indica en los ejemplos. En estos experimentos, se utilizó un espectrofotómetro para medir la absorbancia de los productos formados al término de las reacciones.

5 **Ensayo con hexoquinasa para la medición de glucosa residual**

[0200] Se midió la glucosa residual de sobrenadantes de cultivos de *H. jecorina* que expresan variantes de CBH2 mediante la utilización de un ensayo con hexoquinasa. Se añadió un volumen de 5 µl de sobrenadante al ensayo con hexoquinasa glucosa de 195 µl (Instrumentation Laboratory, Breda, Países Bajos) en una placa de microtitulación de 96 pocillos (Costar Flat Bottom PS). Las placas se incubaron a temperatura ambiente durante 10 15 min. A continuación de la incubación, se midió la absorbancia a 340 nm DO. Se excluyeron los sobrenadantes de cultivos que expresan glucosa residual del agrupamiento para estudios adicionales.

Ensayo HPLC para la determinación del contenido de proteína

[0201] La concentración de proteínas de variante de CBH2 de sobrenadantes de cultivos combinados se determinó mediante la utilización de un HPLC Agilent 1100 (Hewlett Packard) equipada con una columna Proswift RP 2H (Dionex). Se inyectaron diez microlitros de muestra, mezclados con 50 µl de 10 % de acetonitrilo en agua desmineralizada filtrada tras el equilibrio de la columna de HPLC con 10 % de acetonitrilo que contiene 0,01 % de ácido trifluoroacético. Se eluyeron los compuestos mediante la utilización de un gradiente de 10 % a 30 % de acetonitrilo desde 0,3 min hasta 1 min, seguido de un gradiente de 30 % a 65 % desde 1 min hasta 4 mins. Las concentraciones de proteína de las variantes de CBH2 se determinaron a partir de una curva de calibración generada utilizando CBH2 de tipo salvaje purificada (6,25; 12,5; 25; 50 µg/ml). Para calcular el índice de rendimiento (IR), se promedió la relación de la proteína total (media) producida por una variante y la proteína total (media) producida por el tipo salvaje en la misma dosis.

Determinación de la actividad específica mediante ensayo de hidrólisis de celulosa hinchada con ácido fosfórico (PASC, por sus siglas en inglés)

[0202] *Hidrólisis de celulosa:* La celulosa hinchada con ácido fosfórico (PASC) se preparó a partir de Avicel, según un método publicado (Walseth, Tappi 35:228, 1971; y Wood, Biochem J, 121:353-362, 1971). Este material se diluyó con tampón y agua para conseguir una mezcla de 1 % p/v de forma que la concentración final de acetato de sodio fuera 50mM, pH 5.0. Se dispersaron cien microlitros de una suspensión de PASC 1 % en tampón de acetato de sodio 50 mM (pH 5.0) en una placa de microtitulación de 96 pocillos (Costar Flat Bottom PS). Se añadieron diez microlitros de un sobrenadante de cultivo de 5 mg/ml de una cepa delecionada de CBH2 a la PASC, y a esto se le añadieron 5, 10, 15 o 20 µl de sobrenadantes de cultivo combinados de células de *H. Jecorina* que expresan CBH2 de tipo salvaje o variantes de CBH2. La deleción del gen CBH2 de *Hypocrea jecorina* (también denominada *Trichoderma reesei*) se describe en la patente estadounidense con n.º 5,861,271 y 5,650,322. Se añadieron volúmenes compensatorios de tampón de acetato para compensar las diferencias en el volumen total. La placa de microtitulación se selló y se incubó en una incubadora termostata a 50 °C en agitación constante a 900 rpm. Tras dos horas, la reacción de hidrólisis se detuvo mediante la adición de tampón de glicina de 100 µl, pH 10 a cada pocillo. Los productos de la reacción de hidrólisis se analizaron con el ensayo PAHBAH.

[0203] *Ensayo PAHBAH:* Se añadieron alícuotas de 150 µl de reactivo de azúcar reductor PAHBAH (5 % p/v de ácido p-hidroxibenzoico hidrazida (PAHBAH, Sigma # H9882, disuelto en 0,5 N HCl), (Lever, Anal Biochem, 47:273-279, 1972) a todos los pocillos de una placa de microtitulación vacía. Se añadieron diez microlitros de los sobrenadantes de reacción de hidrólisis a la placa de reacción PABA. Todas las placas se sellaron y se incubaron a 69 °C en agitación constante a 900 rpm. Después de una hora, se pusieron las placas en hielo durante cinco minutos y se centrifugaron a 720 x g a temperatura ambiente durante cinco minutos. Se transfirieron muestras de 80 µL de las mezclas de reacción PAHBAH desarrolladas a una nueva placa (de lectura) y se midió la absorbancia a 410 nm en un espectrofotómetro. Se incluyó un estándar de celobiosa como control. Se generó una curva de respuesta a la dosis para la proteína de tipo salvaje CBH2. Para calcular el índice de rendimiento (IR), se promedió la relación del azúcar total (medio) producido por una variante y el azúcar total (medio) producido por el tipo salvaje en la misma dosis.

50 **Determinación de la actividad específica mediante hidrólisis de rastrojo de maíz pretratado (PCS) con ácido diluido**

[0204] *Rastrojo de maíz pretratado (PCS):* Se pretrató el rastrojo de maíz con un 2 % de H₂SO₄ p/p, como se describe (Schell et al., J Appl Biochem Biotechnol, 105:69-86, 2003) y, a continuación, se lavó varias veces con agua desionizada para obtener una pasta que tiene un pH de 4.5. A continuación, se añadió tampón de acetato

de sodio (pH 5.0) a una concentración final de 50mM de acetato de sodio y, cuando fuera preciso, se tituló esta mezcla a pH 5.0 utilizando 1 N de NaOH. La concentración de celulosa en la mezcla de reacción era aproximadamente 7 %. Se añadieron sesenta y cinco microlitros de esta suspensión de celulosa por pocillo a una placa de microtitulación de 96 pocillos (Nunc Flat Bottom PS). Se añadieron diez microlitros de un sobrenadante de cultivo de 5 mg/ml de una cepa delecionada de CBH2 al PCS, y a esto se le añadieron 5, 10, 15 o 20 µl de sobrenadantes de cultivo combinados de células de *H. jecorina* que expresan CBH2 de tipo salvaje o variantes de CBH2. Se añadieron volúmenes compensatorios de tampón de acetato para compensar las diferencias en el volumen total. Tras sellar la placa, se colocaron las placas en una incubadora termostatizada a 50 °C en agitación constante a 1300 rpm durante cinco minutos. A continuación, se incubaron las placas a 50 °C mientras se agitaban a 220 rpm con un 80 % de humedad para evitar el secado. Tras siete días, se pusieron las placas en hielo durante cinco minutos y la reacción de hidrólisis se detuvo mediante la adición de tampón de glicina de 100 µl, pH 10 a cada pocillo. Los productos de la reacción de hidrólisis se analizaron con el ensayo PAHBAH.

[0205] Ensayo PAHBAH: Se añadieron alícuotas de 150 µl de reactivo de azúcar reductor PAHBAH (5 % p/v de ácido p-hidroxibenzoico hidrazida (PAHBAH, Sigma # H9882, disuelto en 0,5 N HCl), (Lever, Anal Biochem, 47:273-279, 1972) a todos los pocillos de una placa de microtitulación vacía. Se añadieron diez microlitros de los sobrenadantes de reacción de hidrólisis a la placa de reacción PABA. Todas las placas se sellaron y se incubaron a 69 °C en agitación constante a 900 rpm. Después de una hora, se pusieron las placas en hielo durante cinco minutos y se centrifugaron a 720 x g a temperatura ambiente durante cinco minutos. Se transfirieron muestras de 80 µL de las mezclas de reacción PAHBAH desarrolladas a una nueva placa (de lectura) y se midió la absorbancia a 410 nm en un espectrofotómetro. Se incluyó un estándar de celobiosa como control. Se generó una curva de respuesta a la dosis para la proteína de tipo salvaje CBH2. Para calcular el índice de rendimiento (IR), se promedió la relación del azúcar total (medio) producido por una variante y el azúcar total (medio) producido por el tipo salvaje en la misma dosis.

Estabilidad de las variantes de CBH2 en presencia de etanol

[0206] Se sometió a ensayo la estabilidad de las CBH2 de tipo salvaje y de las variantes de CBH2 en presencia de 4,5 % de etanol (EtOH) a 49 °C. Se añadieron sobrenadantes de cultivo combinados (80 µL) de células de *H. jecorina* que expresan variantes de CBH2 a una placa de 96 pocillos (Greiner V-bottom PS) que contiene 10 µl de 40,5 % de EtOH por pocillo. Las placas se sellaron y se incubaron en una incubadora termostatizada a 49 °C durante 16 horas en agitación a 900 rpm. A continuación de la incubación, se pusieron las placas en hielo durante 5 minutos. Se determinó la actividad de CBH2 residual utilizando el ensayo de hidrólisis con celulosa hinchada con ácido fosfórico (PASC) como se describe anteriormente.

[0207] Para calcular la actividad residual, se dividió el valor del producto formado por la adición de 5, 10, 15 y 20 µl de CBH2 incubado con EtOH al ensayo PASC de actividad residual por el valor del producto formado por la adición de 5, 10, 15 y 20 µl de CBH2 sin EtOH al ensayo PASC. A continuación, los valores individuales de estas cuatro proporciones se promediaron para proporcionar la actividad residual media. Para determinar el valor IR para la variante, se dividió entonces el valor de la actividad residual media para las variantes por la media de los valores de actividad residual de los controles de CBH2 de tipo salvaje.

Termoestabilidad de variantes de CBH2

[0208] Se sometió a ensayo la termoestabilidad de las CBH2 de tipo salvaje y de las variantes de CBH2 a 53 °C. Se añadió sobrenadante de cultivo combinado (80 µL) de células de *H. jecorina* que expresan variantes de CBH2 a una placa de 96 pocillos (Greiner V-bottom PS). Las placas se sellaron y se incubaron en una incubadora termostatizada a 53 °C durante 16 horas en agitación a 900 rpm. A continuación de la incubación, se pusieron las placas en hielo durante 5 minutos. Se determinó la actividad de CBH2 residual utilizando el ensayo de hidrólisis con celulosa hinchada con ácido fosfórico (PASC) como se describe anteriormente.

[0209] Para calcular la actividad residual, se dividió el valor del producto formado por la adición de 5, 10, 15 y 20 µl de CBH2 tratado térmicamente al ensayo PASC de actividad residual por el valor del producto formado por la adición de 5, 10, 15 y 20 µl de CBH2 sin tratar térmicamente al ensayo PASC. A continuación, los valores individuales de estas cuatro proporciones se promediaron para proporcionar la actividad residual media. Para determinar el valor IR para la variante, se dividió entonces el valor de la actividad residual media para las variantes por la media de los valores de actividad residual de los controles de CBH2 de tipo salvaje.

EJEMPLO 2

Evaluación de la unión de lignina

[0210] La lignina, un biopolímero complejo de fenilpropanoide, es el constituyente principal diferente a los carbohidratos de madera que se une a las fibras de celulosa para endurecer y fortalecer las paredes celulares de las plantas. Al estar reticulado a otros componentes de pared celular, la lignina minimiza la accesibilidad de la celulosa y la hemicelulosa a las enzimas de degradación de celulosa. Por lo tanto, generalmente se asocia la lignina con digestibilidad reducida de toda la biomasa vegetal. En concreto, la unión de celulasas a la lignina reduce la degradación de celulosa mediante celulasas. La lignina es hidrofóbica y, al parecer, tiene carga negativa. En consecuencia, se contempla la adición de cargas negativas a las celulasas para reducir su unión a la lignina.

[0211] Como se describe en el presente documento, se creó una reacción para medir el efecto de la modificación química sobre la capacidad de una preparación de celulosa *Trichoderma sp.* para unir un componente de polímeros vegetales, concretamente la lignina. Se recuperó la lignina mediante la digestión amplia de bagazo de caña de azúcar pretratado con ácido con celulasas (100 mg Laminex BG/ g de celulosa) seguido de la hidrólisis de las celulasas mediante serina proteasa no específica, como se describe precisamente en Berlin et al. (Applied Biochemistry and Biotechnology, 121:163-170, 2005), salvo que la sonicación, el secado, la molienda y el cribado no se realizaron y se añadió al procedimiento un lavado con ácido (0,1 N HCl) para eliminar la proteasa seguido de lavados constantes con tampón de acetato (50 mM acetato de sodio, pH 5) para devolver la muestra a un pH 5. Brevemente, se combinó 50 µL de lignina 1,16 % (recuperada a partir de la sacarificación completa de bagazo) preparada en tampón de acetato de sodio 50 mM a pH 5, con µl de una preparación de celulosa *Trichoderma sp.* desalinizada modificada o sin modificar. Se incubaron tubos para microcentrifugadora que contenían la mezcla de reacción a temperatura ambiente durante una hora y, a continuación, se centrifugaron a alta velocidad para separar los materiales solubles de los insolubles. Se recogieron diez µl del sobrenadante de cada tubo. Se volvieron a mezclar los tubos de reacción y se incubaron durante dos horas adicionales y, a continuación, se recogieron unas segundas alícuotas de 10 µl del sobrenadante de cada tubo. Se analizaron las muestras de sobrenadante mediante SDS-PAGE. La reducción de la intensidad de banda en las preparaciones de celulosa *Trichoderma sp.* modificadas era indicativa de una reducción en la unión de lignina.

EJEMPLO 3

Evaluación de la unión de bagazo

[0212] El bagazo es la biomasa que permanece después de que se haya machacado la caña de azúcar para extraer su jugo. Se preparó una solución que contiene 2 % de celulosa de bagazo (tratada con ácido, 28 % sólido, 57 % glicano) en acetato de sodio 50 mM a pH 5. Se diluyeron muestras de preparaciones de celulosa *Trichoderma sp.* sin modificar o modificadas químicamente diez veces en el mismo tampón de acetato de sodio. Se mezclaron alícuotas de las enzimas diluidas con solución de bagazo o tampón únicamente y se incubaron durante 1 h a temperatura ambiente. Se recogió el sobrenadante y se sometió a ensayo la actividad de un componente de celulosa, concretamente beta-glucosidasa.

[0213] Se midió la actividad de beta-glucosidasa utilizando el ensayo de cloro-nitro-fenil-beta-D-glucósido (CNPG). El ensayo CNPG es un ensayo cinético en el que la β-glucosidasa convierte CNPG en el producto coloreado 2-cloro-4-nitrofenol (CNP). La DO se mide a 405 nm durante un periodo de 10 minutos a 37 °C. Los índices se obtienen como Vmax mediante el empleo del *software* SpectraMax y, a continuación, se convierten en actividad específica (µM CNP/s/mg de Proteína). Brevemente, se añadieron 200 µl de tampón de acetato de sodio 50 mM, pH 5.0 a cada pocillo de una placa de microtitulación de 96 pocillos. Se cubrió la placa y se colocó en un Eppendorf Thermomixer a 37 °C durante 15 minutos para permitir que se equilibrara a la temperatura. Se añadieron cinco µl de las muestras de enzima, diluidas en serie en tampón de acetato de sodio 50 mM, pH 5.0 a cada pocillo después del equilibrio. Se diluyó una solución madre de CNPG 10 mM 1:5 utilizando tampón de acetato de sodio 50 mM, pH 5.0, y a continuación, se añadieron 20 µl de la solución de CNPG diluida (2 mM) a cada pocillo que contenía muestras de enzima. Se transfirió la placa de microtitulación a un espectrofotómetro (SpectraMAX, tipo 340; Molecular Devices) a 37 °C y se leyó la DO a 405 nm durante 0-15 min, con una lectura a intervalos de ≥ 9 s.

[0214] La cantidad de actividad de beta-glucosidasa de las muestras de enzimas de celulosa que no se unió al sustrato de bagazo era considerablemente mayor en el caso de la preparación de celulosa *Trichoderma sp.* modificada químicamente. En concreto, como determina el ensayo CNPG, menos del 50 % de la beta-glucosidasa sin modificar no se unió (p. ej., 50 % unido) al sustrato de bagazo, mientras que aproximadamente el 80 % de la beta-glucosidasa modificada no se unió (p. ej., 20 % unido) al sustrato de bagazo. Tomados conjuntamente, los datos de unión de celulosa modificada indican que la reducción de las cargas positivas en la celulosa conlleva la reducción de la unión a un sustrato de polímero vegetal con carga más negativa. En este caso, el sustrato de polímero vegetal era lignina que permanecía en biomasa tratada con ácido. Se demostró que la biomasa tratada con ácido a partir de rastrojo de maíz, un biopolímero vegetal de composición química similar,

adopta una carga cada vez más negativa durante la sacarificación, como determina la medición del potencial zeta (Véase la tabla 1-2).

EJEMPLO 4

Sacarificación de bagazo pretratado con ácido

5 **[0215]** La sacarificación de la celulosa presente en el bagazo pretratado con ácido (BPA) que contiene cantidades distintas de lignina adicional se evaluó utilizando preparaciones de celulasas *Trichoderma sp.* químicamente modificadas y sin modificar y se sometió a ensayo mediante HPLC para controlar la liberación de azúcares, DP1 a DP7. Los resultados se muestran en la figura 1 como conversión en porcentaje del sustrato de polímero. En una placa de microtitulación, se prepararon 200 µL de BPA (glucano 3,5 %) en tampón de acetato de sodio 50 mM a pH 5, y se ajustaron a distintas cantidades de lignina. Se añadieron veinte microlitros de solución de enzima de celulasa (LAMINEX BG modificado o sin modificar) a los pocillos. Se cubrieron las placas con selladores de placa de aluminio y se colocaron en incubadoras a 50 °C, en agitación, durante 24 h o 48 h. Se terminó la reacción mediante la adición de 100 µl de glicina 100 mM pH 10 a cada pocillo. Después de la mezcla exhaustiva, el contenido de los pocillos de placa de microtitulación se filtró a través de una placa de filtrado de 96 pocillos Millipore (0,45 µm, PES). Se diluyó el filtrado en una placa que contenía 100 µl de glicina 10 mM pH 10 y se midió la cantidad de azúcares solubles (DP1 hasta DP7) producidos mediante HPLC. La serie Agilent 1100 HPLC se equipó con una columna de descalcificación/precolumna (Catálogo Biorad n.º 125-0118) y una columna de carbohidratos Aminex a base de plomo (Catálogo Aminex n.º HPX-87P). La fase móvil utilizada era agua con un caudal de 0,6 ml/min. Se diluyeron todos los estándares de azúcares solubles (DP1-DP7) obtenidos de Sigma en agua Milli-Q a 100 mg/mL y se utilizaron para convertir el área de pico para los azúcares individuales en concentraciones de azúcares reales. El porcentaje de conversión se calculó dividiendo los azúcares medidos mediante HPLC entre el 100 % de la conversión de celulosa en glucosa.

15 **[0216]** La unión de celulasa a lignina disminuirá su eficacia en la degradación de celulosa. Esto se demuestra como una reducción en la conversión de celulosa en presencia de cantidades cada vez mayores de lignina presentes en las reacciones de sacarificación. Esta tendencia persiste en las preparaciones de celulasa modificadas. Sin embargo, se da un aumento del 10 % en la conversión de celulosa en las muestras de celulasa modificadas en comparación con las muestras de celulasa sin modificar. Este resultado indica que la carga negativa creciente de la celulasa reduce la unión no productiva de la celulasa a la lignina.

EJEMPLO 5

30 Sacarificación aumentada de BPA con CBH2 modificada químicamente

[0217] Se modificaron químicamente la *Trichoderma* CBH1, la variante de CBH2, la EG1, EG2 y la beta-glucosidasa purificadas como se describe en el ejemplo 1. La variante de CBH2 utilizada en el presente experimento tiene múltiples sustituciones (P98L/M134V/T154A/I2112V/S316P/S413Y con los números correspondiendo a la celulasa CBH2 madura de tipo salvaje) como se describe en la Pub. estadounidense con n.º 2006/0205042: La secuencia de aminoácidos de la variante de CBH2 madura es la siguiente:

QACSSVWGQCGGQNWSGPTCCASGSTCVYSNDYYSQCLPGAASSSSSTRAA
 STTSRVSPPTSRSSSATPPPGSTTRVPPVGSATYSGNPFVGVTLWANAYYASEVS
 SLAIPSLTGAMATAAAA VAKVPSFVWLDTLDKTPLMEQTLADIRAANKNGGNYAG
 QFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYS DVRTLLVIEPDS
 LANLVTNLGTPKCANAQSAYLECYAVTQLNLPNVAMYLDAGHAGWLGWPANQ
 DPAAQLFANVYKNASSPRALRGLATNVANYNGWNITSPPPYTQGNNAVYNEKLYIHA
 IGPLLANHGWSNAFFITDQGRSGKQPTGQQWGDWCNVIGTGFGIRPSANTGDSLLD
 SFVWVKPGGEDGTSDSSAPRFDYHCALPDALQPAPQAGAWFQAYFVQLLTNANPS
 FL (SEQ ID NO:14).

[0218] Se verificaron las modificaciones químicas de CBH1, variante de CBH2, EG1, EG2 y beta-glucosidasa mediante su movilidad cambiada en el gel nativo en comparación con las proteínas sin modificar. La CBH1, la variante de CBH2, la EG1, EG2 y la beta-glucosidasa modificadas tienen cargas más negativas. Se midieron

5 todas las concentraciones de proteínas con un espectrofotómetro (Thermo) NanoDrop™. Se creó una reacción de sacarificación en placa de microtitulación, en cada pocillo de una placa de microtitulación, se prepararon 150 uL de BPA (7 % de glucano preparado como se describe anteriormente para PCS) en tampón de acetato de sodio 50 mM a pH 5, se añadieron 20 µl de mezcla de enzima de 21 µg de proteína total de modo que la

10 proporción final de proteína y celulosa en cada pocillo fuera de 20 mg/g. Se elaboraron seis mezclas de enzimas mediante la adición de beta-glucosidasa, variante de CBH2, EG1 o EG2 purificadas modificadas o sin modificar a un fondo de *T.reesei* en el que los genes que codifican celobiohidrolasa I (CBHI, Cel7a), celobiohidrolasa II (CBHII, Cel6a), endoglucanasa I (EGI, Cel7b) y endoglucanasa II (EGII, Cel5a) se han inactivado (Véase, US 2007/0128690). En cada mezcla, se añadió 72,5 % del fondo de *T. reesei*, 2,5 % de beta-glucosidasa, 15 % de

15 variante de CBH2, 5 % de EG1, 5 % de EG2; las primeras cuatro mezclas tienen una proteína que no está modificada, la quinta mezcla tiene todas las proteínas sin modificar y la sexta mezcla tiene todas las proteínas modificadas. Se incubó la placa a 50 °C durante 72 horas. Se terminó la reacción mediante la adición de 100 µl de glicina 100 mM pH 10 a cada pocillo. Después de la mezcla exhaustiva, el contenido de los pocillos de placa de microtitulación se filtró a través de una placa de filtrado de 96 pocillos Millipore (0,45 µm, PES). Se diluyó el

20 filtrado en una placa que contenía 100 µl de glicina 10 mM pH 10 y se midió la cantidad de azúcares solubles (DP1 hasta DP7) producidos mediante HPLC. La serie Agilent 1100 HPLC se equipó con una columna de descalcificación/precolumna (Catálogo Biorad n.º 125-0118) y una columna de carbohidratos Aminex a base de plomo (Catálogo Aminex n.º HPX-87P). La fase móvil utilizada era agua con un caudal de 0,6 ml/min. Se diluyeron todos los estándares de azúcares solubles (DP1-DP7) obtenidos de Sigma en agua Milli-Q a 100 mg/mL y se utilizaron para convertir el área de pico para los azúcares individuales en concentraciones de azúcares reales. El porcentaje de conversión se calculó dividiendo los azúcares medidos mediante HPLC entre el 100 % de la conversión de celulosa en glucosa.

25 **[0219]** La figura 2A muestra que la sexta mezcla de enzimas (EG2, EG1, variante de CBH2 y beta-glucosidasa modificadas) con todas las proteínas modificadas, tiene la conversión de celulosa más elevada, y la quinta mezcla de enzimas, con todas las proteínas sin modificar, tiene la conversión más baja. Si se comparan las primeras cuatro mezclas de enzimas, la segunda mezcla de enzimas con la CBH2 sin modificar proporcionó la siguiente conversión más baja. La figura 2B muestra las ventajas de las proteínas modificadas sobre las proteínas sin modificar en la conversión de celulosa.

EJEMPLO 6

30 **Preparación de variantes de carga en escalera de CBH2 de *T.reesei***

35 **[0220]** Como se determina durante el desarrollo de la presente exposición, la succinilación de los residuos de lisina superficiales de CBH2 mejoró el rendimiento sobre BPA y sobre rastrojo de maíz pretratado. La carga de la variante de CBH2 modificada era de aproximadamente -17 en comparación con la variante de CBH2 sin modificar. Teniendo esto en cuenta, se designó una carga en escalera de CBH2 para la determinación de la carga superficial óptima en aplicaciones de rendimiento de celulosa.

SEQ ID NO:1 expone la secuencia de ADN que codifica CBH2 de *Hypocrea jecorina* de referencia:

atgattgtcggcattctcaccacgctggctacgctggccacactcgagctagtgctccttagaggagcggcaagcttctcaagcgt
ctggggccaatgtggtggccagaattggtcgggtccgacttctgtgcttccggaagcacatcgcttactccaacgactattactccc
agtgtcttccggcgctgaagctcaagctcgtccacgcgcccgcgtcgacgacttctcagatccccacacatcccggctga
gctccgcgacgctccacctggttctactactaccagagctaccctcagctcggatcgggaaccgctactgattcaggcaaccctttgtg
gggtcactccttggccaatgcatattacgctctgaagttagcagcctcgtattcctagcttgactggagccatggccactgctcgag
cagctgtcgaaggttccctctttatgtggctagatactctgacaagaccctctcatggagcaaacctggccgacatccgcaccg
ccaacaagaatggcggtaactatgccggacagtttgggtgatgacttggcgatcggattgcgctgacctgctcgaatggcgaa
tactctattgccgatggtggcgtcggcaataaagaactatatacgacaccattcgtcaaatgtcgtggaatattccgatatccggacct
cctggttattgagcctgactcttccaacctggtgaccaacctggactcctcaaggtgccaatgctcagtcagcctactctgagtg
atcaactacgccgtcacacagctgaaccttccaatgttgcgatgtattggacgctggccatgcaggatggcttggctggccggcaaa
ccaagaccggcgcctcagctatttgaatgttacaagaatgcatcgtctccgagagctcttcgggattggcaaccaatgtcgcca
actacaacgggtggaacattaccagcccccatctacacgaagcaacgctgtctacaacgagaagctgtacatccacgctattg
gacctcttccaatcacggctggtccaacgccttctcatcactgatcaaggtcgtcgggaaagcagcctaccggacagcaaca
gtggggagactggtgcaatgtgatcggcaccggatttggattcggccatccgcaaacactggggactcgttctgattcgttctg
ggtcaagccagggcggcagtgacggcaccagcagcagcagtgccacagattgactcccactgtcgcctccagatgcttgc
aaccggcgcctcaagctggtgcttgggtccaagcctacttctcagccttctcacaacgcaaacccatcgttctgtaa*

SEQ ID NO:2 expone la secuencia de proteína de longitud completa de CBH2 de *Hypocrea jecorina*:

MIVGILTTLATLTLAASVPLEERQACSSVWGQCGGNWSGPTCCASGSTCVYSNDYYSQCL
PGAASSSSSTRAASTTSRVSPPTTSRSSSATPPPGSTTTRVPPVSGTATYSGNPFVGVTPWA
NAYYASEVSSLAIPSLTGAMATAAAVAKVPSFMWLDTLDKTPLMEQTLADIRRTANKNGNY
AGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTLLVIEPDSL
NLVTNLGTPKCANAQSAYLECIYAVTQLNLPNVAMYLDAGHAGWLGWPANQDPAAQLFANV
YKNASSPRALRGLATNVANYNGWNITSPPSYTQGNVYNEKLYIHAIGPLLANHGWSNAFFI
TDQGRSGKQPTGQQQWGDWCNVIGTGFIRPSANTGDSLLDSFVWVKPGGECGDTSDSSAPR
FDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL*

SEQ ID NO:3 expone la secuencia de proteína madura de CBH2 de *Hypocrea jecorina*:

QACSSVWGQCGGNWSGPTCCASGSTCVYSNDYYSQCLPGAASSSSSTRAASTTSRVSPPTTS
RSSSATPPPGSTTTRVPPVSGTATYSGNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAA
AAVAKVPSFMWLDTLDKTPLMEQTLADIRRTANKNGNYAGQFVVYDLPDRDCAALASNGEYS
IADGGVAKYKNYIDTIRQIVVEYSDIRTLLVIEPDSLNLVTNLGTPKCANAQSAYLECIY
AVTQLNLPNVAMYLDAGHAGWLGWPANQDPAAQLFANVYKNASSPRALRGLATNVANYNGWN
ITSPPSYTQGNVYNEKLYIHAIGPLLANHGWSNAFFITDQGRSGKQPTGQQQWGDWCNVIG
TGFIRPSANTGDSLLDSFVWVKPGGECGDTSDSSAPRFDSHCALPDALQPAPQAGAWFQAY
FVQLLTNANPSFL*

5

[0221] Los residuos seleccionados para su mutagénesis incluían residuos no conservados, de lisina expuesta, de arginina, de asparagina y de glutamina, que se seleccionaron para sustitución para introducir cargas negativas. Las lisinas succiniladas en CBH2 modificada se identificaron mediante espectrometría de masas y se seleccionaron para mutagénesis en glutamato, dando lugar a una diferencia de carga -2 por sustitución. Se seleccionaron otros residuos para sustitución mediante análisis de la estructura tridimensional de CBH2 combinada con alineamiento de aminoácidos de secuencias de CBH2 homólogas (Véase, p. ej., la Figura 3 de la

10

Pub. estadounidense con n.º 2006/0205042). Los residuos superficiales que eran extremadamente variables en el alineamiento de secuencia de aminoácidos de CBH2 eran candidatos para mutagénesis. Sin embargo, se evitó la acumulación de sustituciones muy cerca entre sí. Se sustituyó la arginina por glutamina (carga -1), y se sustituyó la glutamina y la asparagina por las variantes carboxílicas respectivas (carga -1). Además, se seleccionaron residuos de aspartato y glutamato para la sustitución de los respectivos residuos de amino para la terminación de la carga en escalera (carga +1). En la tabla 6-1, se muestran sustituciones de CBH2 específicas, con todas las posiciones mostradas excepto R63 y R77, ubicadas en el dominio catalítico de CBH2. Puede crearse una carga positiva neta mediante la eliminación de un residuo cargado negativamente o mediante la introducción de un residuo cargado positivamente. De la misma manera, puede crearse una carga negativa neta mediante la eliminación de un residuo cargado positivamente o mediante la introducción de un residuo cargado negativamente.

Tabla 6-1. Sustituciones de CBH2

Introducción de carga negativa y eliminación de carga positiva	Eliminación de carga positiva	Introducción de carga negativa	Introducción de carga negativa	Eliminación de carga negativa	Eliminación de carga negativa
Lisina	Arginina	Asparagina	Glutamina	Aspartato	Glutamato
K157E	R153Q	N382D	Q204E	D189N	E208Q
K129E	R294Q	N344D	Q147E	D211N	E244Q
K288E	R203Q	N237D	Q239E	D405N	E146Q
K194E	R378Q	N339D	Q281E	D277N	
K356E	R63Q	N289D		D151N	
K327E	R77Q	N161D			
		N285D			
		N197D			
		N254D			
		N247D			

[0222] Para la preparación de una carga en escalera de CBH2, se designaron diez variantes de carga de CBH2 (C-1 a C-10) abarcando un intervalo de carga de +8 a -32, en comparación con la CBH2 de tipo salvaje en tramos de 4, como se muestra en la tabla 6-2.

Tabla 6-2. Carga en escalera de CBH2

C-1	C-2	CBH2	C-3	C-4	C-5	C-6	C-7	C-8	C-9	C-10
8	4	0	-4	-8	-12	-16	-20	-24	-28	-32
D189 N	D189 N	WT	K157 E	K157 E	K157 E	K157 E	K157 E	K157 E	K157 E	K157 E
E208 Q	E208 Q		K129 E	K129 E	K129 E	K129 E	K129 E	K129 E	K129 E	K129 E
D211 N	D211 N			K288 E	K288 E	K288 E	K288 E	K288 E	K288 E	K288 E
D405 N	D405 N			K194 E	K194 E	K194 E	K194 E	K194 E	K194 E	K194 E
E244 Q					K356 E	K356 E	K356 E	K356 E	K356 E	K356 E
D277 N					K327 E	K327 E	K327 E	K327 E	K327 E	K327 E
D151 N						R153 Q	R153 Q	R153 Q	R153 Q	R153 Q
E146 Q						R294 Q	R294 Q	R294 Q	R294 Q	R294 Q
						R203 Q	R203 Q	R203 Q	R203 Q	R203 Q
						R378 Q	R378 Q	R378 Q	R378 Q	R378 Q
						N382 D	N382 D	N382 D	N382 D	N382 D

Tabla 6-2. Carga en escalera de CBH2

C-1	C-2	CBH2	C-3	C-4	C-5	C-6	C-7	C-8	C-9	C-10
							N344 D	N344 D	N344 D	N344 D
							N237 D	N237 D	N237 D	N237 D
							N339 D	N339 D	N339 D	N339 D
								N289 D	N289 D	N289 D
								N161 D	N161 D	N161 D
								Q204 E	Q204 E	Q204 E
								Q147 E	Q147 E	Q147 E
									N285 D	N285 D
									N197 D	N197 D
									N254 D	N254 D
									N247 D	N247 D
										Q239 E
										Q281 E
										R63Q
										R77Q

[0223] Las secuencias de aminoácidos de las variantes se volvieron a traducir en ADN y el codón se optimizó para la expresión en *Trichoderma reesei* mediante la utilización del *software* GeneDesigner (ADN2.0). Los genes de variante de *cbh2* con codón optimizado se sintetizaron y el ADN de las variantes de carga superficial de CBH2 (SCV, por sus siglas en inglés) se amplificó a partir de los constructos de ADN2.0 mediante PCR utilizando cebadores: GGHTK22 directo 5'-CACCATGATCGTGGGAATTCTTACTACTC-3' (SEQ ID NO:15); y GGTHK23 inverso 5'-CTACAAAAACGAAGGGTTCGCATT-3' (SEQ ID NO:16). En un experimento, se utilizó la mutagénesis de sitio dirigido para introducir las mutaciones K129E y K157E (variante de carga C-3 de *cbh2*) en el ADN genómico de la CBH2 de tipo salvaje. La variante de carga C-3 de CBH2 se clonó en pTrex3GM y se expresó como se describe a continuación.

[0224] Los productos de la PCR se purificaron y se clonaron en pENTR/TOPO para la transformación de las células *E. coli* TOP10. El ADN plasmídico se aisló de colonias individuales y se verificó la secuencia correcta. Las SCV de CBH2 se clonaron en pTrex3GM, pTTTpyr(*P_{cbh1}*), y pTTTpyr(*P_{stp1}*), como se muestra en la tabla 6-3.

Tabla 6-3. Clones de expresión de variantes de carga superficial de CBH2

Variante de CBH2	Vectores de destino		
	pTrex3gM	pTTTpyr (<i>P_{cbh1}</i>)	pTTTpyr(<i>P_{stp1}</i>)
C-1 (pTK354a)	1	11	21
C-2 (pTK355a)	2	12	22
C-3 (pTK356a)	3	13	23
C-4 (pTK357a)	4	14	24
C-5 (pTK358a)	5	15	25
C-8 (pTK361a)	6	16	26
C-6 (pTK359a)	(7)	(17)	(27)
C-7 (pTK360a)	8	18	28
C-9 (pTK362a)	9	19	29
C-10 (pTK363b)	10	20	30

[0225] Variantes de carga superficial de CBH2 (SCV) en *T. reesei*. La transformación biolística de *T. reesei* se llevó a cabo con el vector de expresión pTrex3gM que contiene el marco abierto de lectura de la variante de carga C3 de *cbh2* (con mutaciones K129E y K157E) mediante la utilización del siguiente protocolo. Se utilizó *T. reesei*, en la que los genes que codifican celobiohidrolasa I (CBHI, Cel7a), celobiohidrolasa II (CBHII, Cel6a), endoglucanasa I (EGI, Cel7b), y endoglucanasa II (EGII, Cel5a) han sido inactivados. La transformación de la cepa de *Trichoderma reesei* mediante el método de transformación biolística se consiguió utilizando un sistema Biolistic® PDS-1000/he Particle Delivery System de Bio-Rad (Hercules, CA) siguiendo las instrucciones del fabricante (Véase WO 05/001036 y US 2006/0003408). Los transformantes se transfirieron a nuevas placas de selección de acetamida. Los transformantes estables se inocularon en placas de microtitulación de filtrado (Millipore), que contenían 200ul/pocillo de medio mínimo de glicina (6,0 g/L glicina; 4,7 g/L (NH₄)₂SO₄; 5,0 g/L

5 KH₂PO₄; 1,0 g/L MgSO₄·7H₂O; 33,0 g/L PIPPS; p.H. 5.5) con adición posestéril de ~2 % de mezcla de glucosa y soforosa como fuente de carbono, 10 ml/L de 100g/L de CaCl₂, 2,5 ml/L de oligoelementos de *T. reesei* (400X): 175g/L de ácido cítrico anhidro; 200g/L FeSO₄·7H₂O; 16g/L ZnSO₄·7H₂O; 3,2 g/L CuSO₄·5H₂O; 1,4 g/L MnSO₄·H₂O; 0,8 g/L H₃BO₃. Los transformantes se cultivaron en cultivo líquido durante 5 días en una cámara rica en O₂ alojada en una incubadora a 28 °C. Las muestras de sobrenadante de la placa de microtitulación de filtrado se obtuvieron mediante la utilización de un colector de vacío. Se pusieron las muestras en geles NuPAGE 4-12 % (Invitrogen) según las instrucciones del fabricante. Se tiñó el gel con tinte Simply Blue stain (Invitrogen). La expresión de variantes de carga superficial de CBH2 adicionales puede obtenerse utilizando este método.

EJEMPLO 7

10 **Generación de variantes de carga de CBH2 en *T. reesei***

15 **[0226]** Se envió el plásmido pTTTpyr-cbh2 que contiene la secuencia codificadora de proteína de CBH2 de *Hypocrea jecorina* (SEQ ID NO:1) a BASEClear (Leiden, Países Bajos), GeneArt AG (Regensburg, Alemania) y Sloning BioTechnology GmbH (Puchheim, Alemania) para la generación de bibliotecas de evaluación de sitio (SEL, por sus siglas en inglés). El mapa de plásmido de pTTTpyr-cbh2 se muestra en la figura 5. Se realizó una petición a los proveedores para la generación de bibliotecas posicionales en cada uno de los sitios en la proteína madura de CBH2 de *Hypocrea jecorina* (SEQ ID NO:3). La secuencia de aminoácidos de proteína de longitud completa de CBH2 se muestra en SEQ ID NO:2.

SEQ ID NO:1 expone la secuencia de ADN que codifica CBH2 de *Hypocrea jecorina* de referencia:

```
atgattgtcggcattctcaccacgctggctacgctggccacactcgcagctagtgtgcctctagaggagcggcaagcttctcaagcgt
ctggggccaatgtggtggccaagaattggtcgggtccgactgctgtgcttccggaagcacatgcgtctactccaacgactattactccc
agtgtcttccggcgctgcaagctcaagctcgtccacgcgcgccgctgcagcacttctcgagtatccccacaacatcccggctcga
gctccgcgacgcctccacctggttctactactaccagagtacctccagtcggatcgggaaccgctacgtattcaggcaaccctttgtg
gggtcactcctgggccaatgcatattacgcctctgaagtttagcagcctcgtattcctagcttgactggagccatggcactgctgcag
cagctgtcgaaggttccctctttatgtggctagatactctgacaagaccctctcatggagcaaaccttggccgacatccgcaccg
ccaacaagaatggcggtaactatgcccggacagtttgggtgatgactgcccgatcgcgattgcgctgcccttgcctcgaatggcgaa
tactctattgccgatggtggcgtcgccaataaagaactatatacgacaccattcgtcaattgctggaatattccgatatccggaccct
cctggttattgagcctgactctcttccaacctggtagcaacctcggtactccaaagtgtgccaatgctcagtcagcctacttgagtgc
atcaactacgccgtcacacagctgaacctccaatgttgcgatgtattggacgctggccatgcaggatggcttggctggccggcaaa
ccaagaccggcgctcagctatttgaatgtttacaagaatgcatcgtctccgagagctcttcgaggattggcaaccaatgtcgcca
actacaacgggtggaacattaccagccccatcgtacacgcaaggcaacgctgtctacaacgagaagctgtacatccacgctattg
gacctcttcttccaatcacggctggtccaacgccttctcatcactgatcaaggtcgtatcgggaaagcagcctaccggacagcaaca
gtgggggagactggtgcaatgtgatcggcaccggttgggtattgccatccgcaaacactggggactcgttgcgtggattcgtttgtctg
ggtcaagccagggcggcagtgtagcggcaccagcgacagcagtgcgccacgatttgactcccactgtgcgctcccagatgccttgc
aaccggcgcctcaagctggtgcttgggtccaagcctacttgtgcagcttctcacaacgcaaacccatcgttctctgtaa*
```

20 SEQ ID NO:2 expone la secuencia de proteína de longitud completa de CBH2 de *Hypocrea jecorina*:

MIVGILTTLATLTLAASVPLEERQACSSVWGQCGGQNWSGPTCCASGSTCVYSNDYYSQCL
 PGAASSSSSTRAASTTSRVSPPTTSRSSSATPPPGSTTTTRVPPVGSATATYSGNPFVGVTPWA
 NAYYASEVSSLAIPSLTGAMATAAAAVAKVPSFMWLDTLDKTPLMEQTLADIRRTANKNGGNY
 AGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTLVIEPDSL
 NLVTNLGTPKCANAQSAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQDPAAQLFANV
 YKNASSPRALRGLATNVANYNGWNITSPPSYTQGNVAVYNEKLYIHAIGPLLANHGWSNAFFI
 TDQGRSGKQPTGQQQWGDWCNVIGTGFIRPSANTGDSLLDSFVWVKPGGECGTSDSSAPR
 FDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL*

SEQ ID NO:3 expone la secuencia de proteína madura de CBH2 de *Hypocrea jecorina*:

QACSSVWGQCGGQNWSGPTCCASGSTCVYSNDYYSQCLPGAASSSSSTRAASTTSRVSPPTTS
 RSSSATPPPGSTTTTRVPPVGSATATYSGNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAA
 AAVAKVPSFMWLDTLDKTPLMEQTLADIRRTANKNGGNYAGQFVVYDLPDRDCAALASNGEYS
 IADGGVAKYKNYIDTIRQIVVEYSDIRTLVIEPDSLNLVTNLGTPKCANAQSAYLECIN
 AVTQLNLPNVAMYLDAGHAGWLGWPANQDPAAQLFANVYKNASSPRALRGLATNVANYNGWN
 ITSPPSYTQGNVAVYNEKLYIHAIGPLLANHGWSNAFFITDQGRSGKQPTGQQQWGDWCNVIG
 TGFIRPSANTGDSLLDSFVWVKPGGECGTSDSSAPRFDSHCALPDALQPAPQAGAWFQAY
 FVQLLTNANPSFL*

- 5 **[0227]** Se obtuvieron plásmidos de pTTTpyr-cbh2 purificados (p_{cbh1} , Amp^R, Acetamida^R) que contienen marcos
 abiertos de lectura que codifican secuencias de variante de CBH2 de los proveedores especificados
 anteriormente. Se transformaron protoplastos de cepa de *H. jecorina* ($\Delta eg1$, $\Delta eg2$, $\Delta cbh1$ $\Delta cbh2$) con los
 constructos pTTTpyr y se cultivaron en agar selectivo que contenía acetamida a 28 °C durante 7 días.
 Brevemente, se realizó la transformación biolística de *H. jecorina* utilizando el protocolo siguiente y una cepa en
 la que los genes que codifican celobiohidrolasa I (CBHI, Cel7a), celobiohidrolasa II (CBHII, Cel6a),
 10 endoglucanasa I (EGI, Cel7b), y endoglucanasa II (EGII, Cel5a) han sido inactivados. La transformación de la *H.*
jecorina mediante el método de transformación biolística se consiguió utilizando un sistema Biolistic® PDS-
 1000/he Particle Delivery System de Bio-Rad (Hercules, CA) siguiendo las instrucciones del fabricante (Véase
 WO 05/001036 y US 2006/0003408). Las esporas fueron cosechadas, colocadas de nuevo en placas de agar de
 acetamida e incubadas a 28 °C durante 7 días. Las esporas fueron cosechadas en 15 % de glicerol y
 15 almacenadas a -20 °C para su utilización adicional. Para la producción de proteína de variante CBH2, se añadió
 un volumen de 10 μ l de suspensión de spora a 200 μ l de medio mínimo de glicina complementado con 2 % de
 mezcla de glucosa/soforosa en una placa de filtrado PVDF: 6,0 g/L de glicina, 4,7 g/L (NH₄)₂SO₄; 5,0 g/L
 KH₂PO₄; 1,0 g/L MgSO₄·7H₂O; 33,0 g/L PIPPS; pH 5,5; con adición posestéril de ~2 % de mezcla de
 20 glucosa/soforosa como fuente de carbono, 10 ml/L de 100g/L de CaCl₂, 2,5 ml/L de oligoelementos de *T. reesei*
 (400X): 175g/L de ácido cítrico anhidro; 200g/L FeSO₄·7H₂O; 16g/L ZnSO₄·7H₂O; 3,2 g/L CuSO₄·5H₂O; 1,4 g/L
 MnSO₄·H₂O; 0,8 g/L H₃BO₃. Cada variante de CBH2 se cultivó por cuadruplicado. Después de sellar la placa con
 una membrana permeable de oxígeno, se incubaron las placas a 28 °C durante 6 días, con agitación a 220 rpm.
 Se cosechó el sobrenadante mediante la transferencia del medio de cultivo a una placa de microtitulación a baja
 25 presión y se sometió a ensayo para la glucosa residual mediante la utilización del ensayo con hexoquinasa,
 como se describe en el ejemplo 1.

EJEMPLO 8

Expresión, actividad y rendimiento de variantes de CBH2

- 30 **[0228]** Se sometieron a prueba las variantes de carga de CBH2 de *H. jecorina* para diversas propiedades de
 interés. En concreto, se sometieron a prueba las variantes de celulasa para la expresión de proteína utilizando el
 ensayo HPLC (HPLC), actividad específica utilizando el ensayo de hidrólisis de PASC (Act. PASC) y el ensayo
 de hidrólisis de PCS (Act. PCS), estabilidad en presencia de etanol (proporción EtOH) y termoestabilidad
 (proporción de calor) como se describe en el ejemplo 1. En la tabla 8-1, se muestran los datos de rendimiento
 para las variantes de carga de CBH2. El índice de rendimiento (IR) es la proporción de rendimiento de la variante
 de celulasa en relación con la celulasa original o de referencia. Diversos términos que se exponen a continuación
 35 se utilizan para describir la mutación: las mutaciones aceleradoras tienen un IR > 1; las mutaciones neutras
 tienen un IR \leq 0,5; las mutaciones no deletéreas tienen un IR > 0,05; las mutaciones deletéreas tienen un IR =

0,05; las mutaciones combinables son aquellas mutaciones para las cuales la variante tiene valores de índice de rendimiento $\leq 0,5$ para al menos una propiedad. Las mutaciones combinables son mutaciones que pueden combinarse para suministrar proteínas con índices de rendimiento adecuados para una o más propiedades deseadas. Las posiciones en las que se dan las mutaciones se clasifican de la manera siguiente: Las posiciones no restrictivas tienen ≤ 20 % de mutaciones neutras para al menos una propiedad; y las posiciones restrictivas tienen < 20 % de mutaciones neutras para la actividad y la estabilidad. Las posiciones completamente restrictivas no tienen mutaciones neutras para la actividad o la estabilidad.

5

[0229] Estos datos pueden utilizarse para modificar cualquier CBH2. Aunque la CBH2 que se ha de modificar tenga un aminoácido diferente al de la CBH2 de *Hypocrea jecorina* en una posición concreta, estos datos pueden utilizarse para encontrar una sustitución que altere las propiedades deseadas mediante la identificación de la mejor sustitución, incluyendo la sustitución del aminoácido de tipo salvaje de CBH2 de *Hypocrea jecorina*.

10

[0230] La tabla 8-11 muestra valores de índice de rendimiento (IR) para las variantes de CBH2 de *Hypocrea jecorina*. Los índices de rendimiento menores o iguales que 0,05 se establecieron en 0,05 y se indican en cursiva y negrita en la tabla.

Tabla 8-1: Índices de rendimiento de variantes de carga de CBH2					
Variante	HPLC	Act. esp. PASC	Act. esp. PCS	Res. EtOH	Res. Calor
R63A	1,21	1,03	0,94	0,97	1,00
R63C	1,47	1,02	1,05	0,96	1,06
R63D	0,12	0,11	0,19	0,05	1,02
R63E	0,88	0,94	0,98	0,96	1,05
R63F	0,91	0,95	0,85	0,94	1,09
R63G	0,71	0,96	0,95	1,08	1,20
R63I	0,81	0,93	1,11	1,03	1,09
R63L	0,94	0,96	0,91	1,06	1,29
R63M	1,22	0,99	1,10	1,05	1,09
R63N	1,23	1,02	1,11	1,03	1,07
R63P	0,96	0,99	1,17	1,04	1,09
R63Q	0,87	0,99	1,01	0,99	1,13
R63S	0,59	0,92	1,09	1,00	0,94
R63T	0,43	0,87	0,94	0,91	0,84
R63V	1,00	0,99	0,89	1,03	1,06
R63W	0,98	0,97	0,91	1,08	1,10
R63Y	1,29	1,03	0,99	1,06	1,07
R77F	0,25	0,81	0,58	0,92	0,81
R77G	0,41	1,02	0,75	1,00	0,93
R77L	0,14	0,53	0,36	0,94	0,79
R77N	0,83	0,99	0,87	1,00	1,02
K129A	0,23	1,03	0,81	0,74	0,57
K129L	0,80	0,94	0,80	0,82	0,71
K129N	0,15	0,93	0,65	0,73	0,61
K129Q	0,41	1,30	0,89	1,01	0,99
K129S	0,37	1,05	0,92	0,84	0,84
K129T	0,16	0,87	0,37	0,81	0,69
K129V	0,17	0,93	0,91	0,63	0,45
K129Y	0,35	0,95	0,82	0,68	0,46
Q147E	0,86	1,00	0,88	1,07	1,05
R153A	0,23	0,50	0,72	0,63	0,64
R153C	0,26	0,51	0,86	0,94	0,88
R153D	0,35	0,35	0,54	0,75	0,72
R153E	0,27	0,69	0,75	0,79	0,80
R153F	0,21	0,38	0,69	0,68	0,58

ES 2 619 902 T3

Tabla 8-1: Índices de rendimiento de variantes de carga de CBH2					
Variante	HPLC	Act. esp. PASC	Act. esp. PCS	Res. EtOH	Res. Calor
R153G	0,21	0,31	0,59	0,62	0,92
R153I	0,22	0,28	0,37	0,63	1,09
R153L	0,19	0,30	0,40	0,94	0,44
R153M	0,31	0,70	0,86	0,92	1,05
R153N	0,22	0,20	0,42	1,12	1,23
R153P	0,23	0,14	0,10	0,39	2,17
R153Q	0,24	0,52	0,99	0,77	1,05
R153S	0,41	0,78	0,99	0,93	0,93
R153T	0,31	0,63	0,74	0,84	0,94
R153V	0,23	0,28	0,59	0,63	0,81
R153W	0,42	0,77	1,07	0,90	0,87
R153Y	0,32	0,60	0,73	0,84	0,90
K157A	0,68	1,04	0,93	1,01	0,91
K157D	0,69	0,96	1,01	1,12	0,90
K157E	0,10	0,92	0,75	1,03	0,73
K157F	0,40	1,01	0,94	1,07	0,95
K157G	0,50	0,98	0,76	1,04	0,93
K157I	0,71	1,00	0,81	1,05	0,94
K157L	0,18	0,88	0,60	1,04	0,75
K157M	0,33	0,99	0,77	0,99	0,88
K157P	0,11	0,57	0,63	3,04	1,86
K157Q	0,50	1,00	0,70	1,00	0,88
K157T	0,14	0,79	0,47	1,02	0,83
K157V	0,46	0,99	0,86	0,99	0,91
K157W	1,06	0,95	0,84	1,04	0,93
K157Y	0,77	0,98	0,92	1,00	0,84
N161E	0,76	1,03	1,07	1,12	1,12
D189A	0,29	0,77	0,73	0,79	0,69
D189C	0,38	0,89	0,82	0,90	0,76
D189E	0,44	0,98	0,84	0,94	0,89
D189F	0,39	0,82	0,64	0,94	0,99
D189G	0,27	0,87	0,86	0,88	0,71
D189H	0,39	0,91	0,52	0,91	0,83
D189I	0,16	0,52	0,39	0,63	0,61
D189K	0,37	0,96	0,78	0,89	0,73
D189L	0,33	1,04	0,82	0,89	0,85
D189N	0,55	0,95	1,05	1,00	0,96
D189P	0,12	0,30	0,22	0,72	0,76
D189Q	0,73	0,98	0,72	0,93	0,86
D189R	0,38	0,96	0,82	0,94	0,88
D189S	0,14	0,60	0,32	0,86	1,22
D189T	0,21	0,84	0,84	0,85	0,76
D189V	0,36	1,03	0,99	1,03	0,91
D189W	0,13	0,18	0,34	0,45	1,06
D189Y	0,32	0,52	0,46	1,30	1,34
K194A	0,30	0,67	0,86	0,95	1,00
K194C	0,22	1,08	1,72	0,88	0,88
K194D	0,61	0,98	1,07	1,00	1,03
K194E	0,44	1,00	1,14	1,02	1,14
K194F	0,67	1,03	0,94	1,04	1,13

ES 2 619 902 T3

Tabla 8-1: Índices de rendimiento de variantes de carga de CBH2					
Variante	HPLC	Act. esp. PASC	Act. esp. PCS	Res. EtOH	Res. Calor
K194G	0,43	0,92	0,65	0,92	0,96
K194I	0,24	0,38	0,49	0,50	0,59
K194L	0,61	0,98	0,90	1,01	1,22
K194M	0,61	1,02	0,86	1,03	1,17
K194N	0,54	0,99	1,02	1,04	1,17
K194P	0,32	0,39	0,55	0,46	0,38
K194Q	0,64	1,00	1,00	1,03	1,17
K194S	0,66	1,00	0,83	1,12	1,03
K194T	0,08	0,11	1,09	1,17	9,80
K194V	0,23	0,30	0,71	0,55	0,62
K194W	0,43	0,93	0,98	0,98	1,01
K194Y	0,08	0,11	0,36	0,60	2,65
N197D	1,50	1,12	1,08	1,06	1,07
R203A	0,71	0,93	0,75	0,69	0,88
R203F	0,85	0,92	0,99	0,67	0,87
R203G	0,60	1,07	1,13	0,66	1,39
R203I	0,39	0,70	0,64	1,46	3,02
R203L	0,46	1,14	0,54	1,08	1,24
R203M	0,58	0,88	0,91	0,99	1,25
R203N	0,36	1,08	0,82	1,06	1,20
R203P	0,29	0,49	0,28	2,85	4,58
R203Q	0,80	0,99	0,78	0,84	0,86
R203S	0,54	1,19	0,53	1,99	1,96
R203T	0,66	1,04	0,82	0,99	0,91
R203V	0,85	0,94	0,89	0,73	0,78
R203W	0,39	1,10	0,63	2,28	2,42
R203Y	0,57	1,09	1,03	1,03	1,18
Q204D	1,23	1,04	0,73	1,00	1,00
Q204E	1,08	1,12	1,05	0,93	1,04
Q239D	0,59	0,98	0,91	0,92	0,92
Q239E	0,70	1,10	1,16	1,00	1,04
N247D	1,18	0,76	0,67	0,58	0,96
N254E	0,20	0,84	0,56	0,76	0,91
Q281D	1,10	0,60	0,43	1,33	0,95
Q281E	2,50	0,89	0,85	1,10	0,63
N285D	0,35	0,70	0,59	1,99	2,24
K288A	0,41	0,61	0,68	1,52	2,84
K288C	0,41	0,95	0,63	0,72	1,49
K288D	0,23	0,23	0,05	5,77	5,10
K288E	0,38	0,89	0,81	0,65	1,60
K288F	0,36	0,75	0,29	1,53	2,74
K288G	0,23	0,40	0,12	2,22	4,09
K288H	0,32	0,90	0,48	1,09	1,54
K288I	0,40	0,75	0,37	1,51	2,48
K288L	0,80	0,96	0,81	0,60	0,82
K288N	0,71	0,98	0,84	1,08	1,68
K288P	0,26	0,34	0,05	5,00	5,39
K288Q	0,23	0,80	0,18	1,23	2,51
K288S	0,19	0,71	0,21	2,02	2,72
K288T	0,25	0,71	0,35	1,27	2,96

ES 2 619 902 T3

Tabla 8-1: Índices de rendimiento de variantes de carga de CBH2					
Variante	HPLC	Act. esp. PASC	Act. esp. PCS	Res. EtOH	Res. Calor
K288V	0,31	0,84	0,38	1,67	1,87
N289D	0,84	0,96	0,87	0,93	0,92
N289E	0,13	0,26	0,63	1,03	1,48
R294A	0,97	1,06	0,89	0,92	0,77
R294C	0,70	1,15	0,74	1,01	1,16
R294D	0,75	0,94	0,85	0,94	1,09
R294E	0,73	0,82	0,66	1,02	1,20
R294F	0,37	0,52	0,54	0,90	1,56
R294G	1,25	0,93	0,84	0,71	0,87
R294I	0,77	0,84	0,72	0,98	1,01
R294L	0,54	1,06	1,03	1,05	1,22
R294M	0,34	0,76	0,78	0,81	1,78
R294N	0,50	0,80	0,64	1,30	1,06
R294P	0,48	0,86	0,74	0,74	0,99
R294Q	0,50	0,79	0,75	1,23	1,30
R294S	0,64	1,00	0,57	1,32	1,47
R294T	1,12	1,03	0,97	1,02	1,10
R294V	0,83	0,96	0,83	0,82	1,03
R294W	1,08	1,01	0,57	0,56	0,58
R294Y	0,26	0,43	0,31	2,39	3,19
K327A	0,86	1,03	0,98	0,76	0,76
K327C	0,42	0,93	1,25	1,28	1,02
K327D	0,31	0,73	0,76	1,13	1,81
K327E	0,73	0,84	0,85	1,05	0,87
K327F	0,24	0,61	0,44	1,91	2,14
K327G	0,24	0,83	0,16	1,65	1,30
K327I	0,31	0,74	0,23	1,44	1,37
K327L	0,26	0,94	0,60	0,89	1,30
K327M	0,41	0,90	0,90	0,87	1,44
K327N	0,41	0,90	1,04	0,74	1,33
K327P	0,21	0,58	0,05	1,65	2,77
K327Q	1,32	1,06	1,42	1,15	0,77
K327S	0,20	0,63	0,58	2,39	2,92
K327T	0,21	0,29	0,09	3,69	7,57
K327V	0,28	0,86	0,58	0,66	1,26
K327W	0,28	1,05	0,74	0,93	1,57
K327Y	0,93	1,04	1,08	0,49	0,60
N339D	0,92	1,04	1,05	1,04	1,10
N339E	0,94	1,03	0,83	1,08	1,13
N344D	0,29	0,85	1,10	1,43	1,31
K356A	1,22	1,03	0,88	0,23	0,25
K356C	0,62	0,92	0,98	0,29	0,45
K356D	0,42	0,87	0,57	0,22	0,42
K356E	0,43	0,98	0,71	0,28	0,43
K356F	0,69	0,99	0,45	0,15	0,39
K356G	0,97	0,98	0,20	0,20	0,40
K356I	0,76	0,96	0,60	0,26	0,46
K356L	0,67	1,00	0,68	0,45	0,39
K356M	0,68	0,99	0,88	0,54	0,53
K356N	0,98	1,04	0,72	0,19	0,37

Tabla 8-1: Índices de rendimiento de variantes de carga de CBH2					
Variante	HPLC	Act. esp. PASC	Act. esp. PCS	Res. EtOH	Res. Calor
K356P	0,23	0,52	0,33	1,09	0,99
K356Q	0,77	0,99	0,77	0,41	0,43
K356S	0,58	0,88	0,48	0,22	0,41
K356T	0,63	1,00	0,49	0,28	0,43
K356V	0,91	0,90	0,75	0,35	0,31
K356W	0,39	0,90	0,42	0,24	0,47
K356Y	0,52	0,96	0,47	0,20	0,51
R378A	0,21	0,76	0,79	0,25	0,54
R378C	0,25	0,34	0,42	0,74	1,01
R378D	0,19	0,23	0,33	0,49	0,58
R378E	0,17	0,18	0,30	0,92	0,94
R378F	0,13	0,39	1,34	0,12	0,38
R378G	0,08	0,22	1,48	0,19	0,82
R378I	0,19	0,42	0,33	0,38	0,69
R378L	0,21	0,75	0,79	0,33	0,53
R378M	0,14	0,52	0,77	0,37	0,47
R378N	0,12	0,19	0,48	0,38	0,81
R378P	0,17	0,65	0,54	0,48	0,62
R378Q	0,20	0,82	0,94	0,56	0,59
R378S	0,18	0,65	0,47	0,45	0,64
R378T	0,18	0,52	1,03	0,25	0,43
R378V	0,36	1,01	0,98	1,06	1,19
R378W	0,09	0,10	0,64	0,34	0,30
R378Y	0,24	0,60	0,62	0,23	0,53
N382D	0,71	0,81	N.H.	1,17	1,15
N382E	2,21	0,16	N.H.	6,19	5,56
D405A	1,75	0,82	0,33	0,32	0,60
D405C	0,45	0,63	0,57	0,72	0,97
D405E	0,85	0,81	0,63	0,50	0,65
D405F	0,79	0,95	0,24	0,56	0,82
D405G	0,77	0,87	0,16	0,56	0,86
D405H	0,27	0,38	0,25	3,35	3,59
D405I	0,62	1,04	0,49	0,83	1,02
D405K	0,63	0,65	0,24	0,43	0,72
D405L	0,41	0,95	0,76	0,38	0,43
D405M	0,52	1,31	0,58	0,59	0,69
D405N	0,74	1,04	0,78	1,34	1,32
D405P	0,77	0,80	0,19	0,99	1,02
D405Q	0,71	1,12	0,39	1,10	1,21
D405R	0,93	1,20	0,36	0,56	0,61
D405S	0,95	1,24	0,40	0,75	0,96
D405T	0,74	1,17	0,32	0,64	0,83
D405V	0,39	1,05	0,60	1,32	1,24
D405W	1,12	1,17	0,40	0,42	0,48
D405Y	1,56	1,07	0,39	0,31	0,07

EJEMPLO 9

Efecto de cambio de carga en la actividad de variantes de CBH2

5 **[0231]** En este ejemplo, se analizó el efecto de cambio de carga en la actividad de CBH2 en un ensayo de rastreo de maíz pretratado (PCS, por sus siglas en inglés) de actividad de celulasa. Brevemente, el número de ganadores de PCS en las bibliotecas de evaluación de sitio (SEL, por sus siglas en inglés) de CBH2 se determinó como una propiedad de cambio de carga neta. En la tabla 9-1, la proporción de ganadores observados por encima de lo previsto (o/e, por sus siglas en inglés), se determinó en el ensayo PCS. Los valores en cursiva y negrita son considerablemente diferentes del promedio de 10 distribuciones aleatorias más o menos el número de desviaciones típicas (dt) indicadas en las respectivas columnas.

Tabla 9-1: Efecto de carga en la actividad de variantes de CBH2 en un ensayo PCS

PCS	o/e 1 dt	o/e 2 dt	o/e 3 dt	Resultados
Cambio de carga				
-2,00	1,20	1,20	1,20	>90 % seguro más de lo previsto
-1,00	1,40	1,40	1,40	>99 % seguro más de lo previsto
0,00	1,03	1,03	1,03	como se preveía
1,00	0,46	0,46	0,46	>95 % seguro menos de lo previsto
2,00	0,00	0,00	0,00	como se preveía

10 **[0232]** Como se muestra en la tabla 9-1 y en la figura 4, la disminución de la carga (p. ej., -1, -2) da lugar a frecuencia considerablemente superior de ganadores de CBH2 en el ensayo PCS, mientras que el incremento de la carga (p. ej., +1) da lugar a una frecuencia considerablemente menor de ganadores de CBH2 en el ensayo PCS. En resumen, la actividad de CBH2 en PCS tiene relación con la disminución de la carga.

REIVINDICACIONES

- 5 **1.** Una variante de celulasa aislada, donde dicha variante es una forma madura que tiene actividad de celulasa, y donde dicha variante es una variante de una celulasa original cuya secuencia de aminoácidos es al menos un 97 % idéntica a la secuencia de aminoácidos de SEQ ID NO: 3, consistiendo la variación en una sustitución en la posición 63 numerada por correspondencia con la secuencia de aminoácidos de SEQ ID NO:3, y donde la sustitución hace que la variante de celulasa tenga una carga neta más negativa en comparación con la secuencia de aminoácidos de SEQ ID NO: 3.
- 2.** La variante de celulasa aislada de la reivindicación 1, donde la sustitución comprende la eliminación de una carga positiva.
- 10 **3.** La variante de celulasa aislada de la reivindicación 1, donde la sustitución es R63Q numerada por correspondencia con la secuencia de aminoácidos de SEQ ID NO:3.
- 4.** La variante de celulasa aislada de la reivindicación 1, donde la variante es la variante de sustitución R63Y o la R63Q, R63A, R63C, R63D, R63E, R63F, R63G, R63I, R63L, R63M, R63N, R63P, R63S, R63T, R63V, R63W de la secuencia de aminoácidos de SEQ ID NO:3.
- 15 **5.** La variante de celulasa aislada de la reivindicación 1, donde la carga neta más negativa es un -1 o -2 en comparación con la secuencia de aminoácidos de SEQ ID NO:3.
- 6.** Un método de convertir biomasa en azúcares que comprende poner en contacto dicha biomasa con la variante de celulasa de la reivindicación 1.
- 7.** Un método de producir un combustible que comprende:
- 20 poner en contacto una composición de biomasa con una composición enzimática que comprende la variante de celulasa de la reivindicación 1 para producir una solución de azúcar; y cultivar la solución de azúcar con un microorganismo fermentativo en condiciones suficientes para producir un combustible.

FIG. 1A

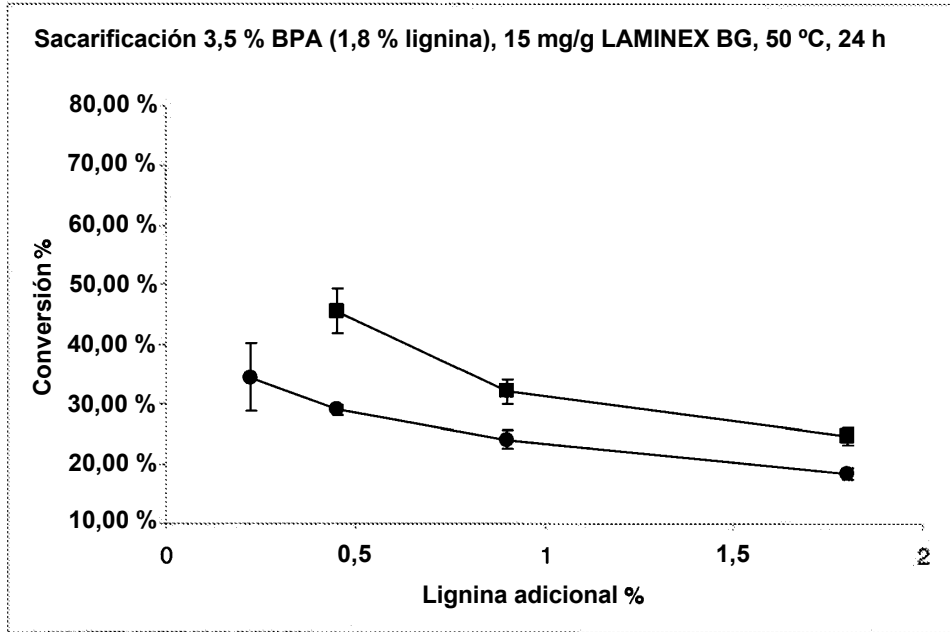


FIG. 1B

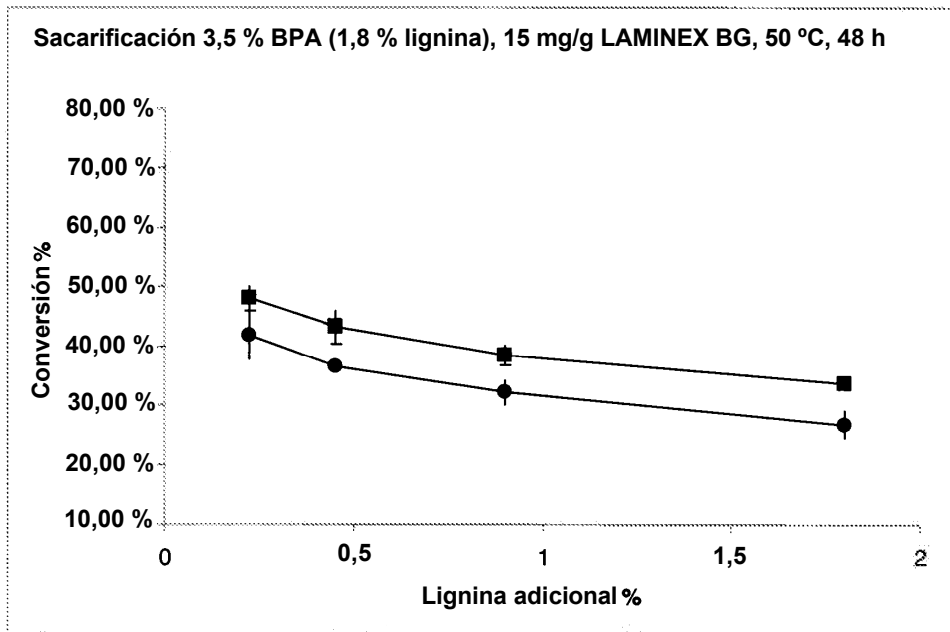


FIG. 2A

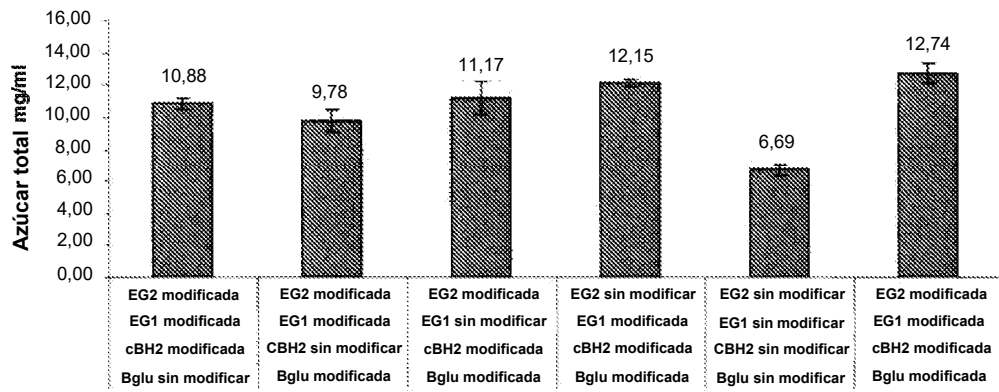


FIG. 2B

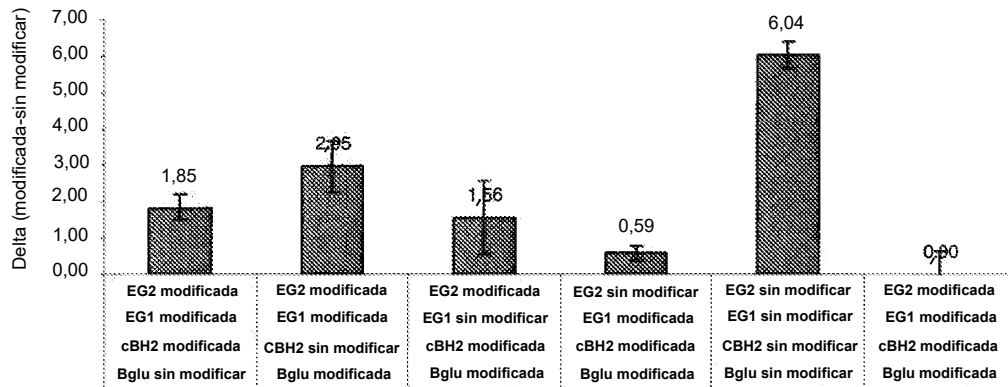


FIG. 3

```

T. reesei CBH #3          -----QACSSVWGQCGGQN----W
H. koningii CBH #4       -----QACSSVWGQCGGQN----W
H. insolens CBH #5       -----APVVEERQNCAPTWGQCGGIG----F
A. cellulolyticus CBH #6 -----MLRYLSIVAATAILTGVEAQQSVWGQCGGQ----W
A. bisporus CBH #7       -----QSPV-----WGQCGGNG----W
F. oxysporum EG #8       -----APVEERQSCSNGVWAQCGGQN----W
P. chrysosporium CBH #9 -----MKSTAFFAALVTLLPAYVAGQASE--WGQCGGIG----W
T. emersonii CBH #10     -----MRNLLALAPAALLVGAEEAQQLWQCGGSS----W
T. fusc CBH #11          -AGCSVDYTVN-SWGTGFTANVTITNLGSAINGWLEWDFPGNQVTLNW
T. fusc EG #12           -----NDSPFYVNPMSAEWVRNPNND-----
C. fimi EG #13           APGCRVDYAVTNQWPGGFANVTITNLGDPVSSWKLDWTYTAGQRIQQLW
                                *

T. reesei CBH #3          SGPTCCASGSTCVYSNDYYSQCLP--GA---ASSSSSTRAAST----TSR
H. koningii CBH #4       SGPTCCASGSTCVYSNDYYSQCLP--GA-----ASSSSSTRASSTTA
H. insolens CBH #5       NGPTCCQSGSTCVKQNDWYSQCLP--GSQVTTTSTTSTSSST----TSR
A. cellulolyticus CBH #6 SGATSCAAGSTCSTLNPPYQACIP--GT-----ATSTTLV----KTT
A. bisporus CBH #7       TGPTTCASGSTCVKQNDFYQCLP--NN-----Q
F. oxysporum EG #8       SGTPCCTSGNKCKVKNDFYSQCP--GS-----AEPS----STA
P. chrysosporium CBH #9  TGPTTCVSGTTCVNLNPPYQCLP--GS-AVTTTTSVITSHSSS----VSS
T. emersonii CBH #10     TGATSCAAGATCSTINPPYQCV--AT-----ATPTTLT----TTT
T. fusc CBH #11          NGTYTQSGQHVSVNAPYNASI-PANGTVEFGFNGSYSGSNDIPSSFKLN
T. fusc EG #12           PRTPVIRDRIASVPQGTWFHHP--GQ-----ITQVDA----LMS
C. fimi EG #13           NGTASTNGGQVSVTSLPWNGSI-PTGGTASFGFNGSWAGSNPTPASFSLN
                                :
                                *

T. reesei CBH #3          VSPTTSR---SSSATPPPGSTTRVPP-VGSGTATYSGNPFVGVTPWANA
H. koningii CBH #4       RASSTTS---RSSATPPPGSSTTRVPP-VGSGTATYSGNPFVGVTPWANA
H. insolens CBH #5       ATSTTRT---GGVTSITTAPTRTVTIPGGATTTASYNGNPFEGVQLWANN
A. cellulolyticus CBH #6 SSTS VGT-----TSPPTTTTKAST-TATTTAAASGNPFSGYQLYANP
A. bisporus CBH #7       APPSTTT---QPGTTPPATTTSGGTGP-TSGA-----GNPYTKTVWLSP
F. oxysporum EG #8       AGPSSTT---ATKTATGGSSTTAGGS-VTSAPPAASDNPNYAGVDLWANN
P. chrysosporium CBH #9  VSSHSGS---STSTSSPTGPTGTNPPP-PPSA-----NNPWTGFQIFLSP
T. emersonii CBH #10     KPTSTGG---AAPTTPPTTTGTTTTSP-VVTRPASASGNPFEQYQLYANP
T. fusc CBH #11          GVTCDGSDDPDPEPSPSPSPSPSPTDPDEPGGPTNPTNPGEKVD---NP
T. fusc EG #12           AAQAAGK---IPILVVYNAPGRDCGNH-SSGGAPSHS-----AYRSWIDE
C. fimi EG #13           GTTCTGT---VPTTSPTPTPTPTPTPT-TPTPTPTPT--PTVTPQPTSGF
                                :

T. reesei CBH #3          YYASEVSSLAIPSLTG-AMATAAAVAVKVPFVWLDLTL-----DKTP-LM
H. koningii CBH #4       YYASEVSSLAIPSLTG-AMATAAAVAVKVPFVWLDLTF-----DKTP-LM
H. insolens CBH #5       YYRSEVHTLAIPIQITDPALRAAASAVAEVPSFQWLDRN-----VTVDLTL
A. cellulolyticus CBH #6 YYSSEVHTLAIPLSLTG-SLAAAATKAAEIPSFVWLDTA-----AKVP-TM
A. bisporus CBH #7       FYADEVAQ-AAADISNPSLATKAASVAKIPTFVWFDTV-----AKVP-DL
F. oxysporum EG #8       YYRSEVMNLAVPKLSG-AKATAAAKVADVPSFQWMDTY-----DHIS-LM
P. chrysosporium CBH #9  YYANEVAA-AAKQITDPTLSKAASVANIPFTWLDV-----AKIP-DL
T. emersonii CBH #10     YYASEVSLAIPSLSS-ELVPKASEVAVKVPFVWLDQA-----AKVP-SM
T. fusc CBH #11          FEGAKLYVNPVWSAKA-AAEPGGSANESTAVWLDLDRIGAIEGNDSPTTG
T. fusc EG #12           F-AAGLKNRPAYIIVEPDLISLMSSCMQHVVQVELETM-----AYAGKAL
C. fimi EG #13           YVDPTTQGYRAWQAASGTDKALLEKIALTPQAYWVGW-----ADAS---
                                :
    
```

Figura 3 continuada

```
T. reesei CBH #3      EQTLAD-IRTANKNGGN--YAGQFVVYDLPDRDCAALASNGEYSIADGGV
H. koningii CBH #4   EQTLAD-IRTANKNGGN--YAGQFVVYDLPDRDCAALASNGEYSIADGGV
H. insolens CBH #5   VETLSE-IRAANQAGANPPYAAQIVVYDLPDRDCAAAAASNGEWAIANNGA
A. cellulolyticus CBH #6  GTYLAN-IEAANKAGASPPIAGIFVVYDLPDRDCAAAAASNGEYTVANNGV
A. bisporus CBH #7    GGYLAD-ARSKNQ-----LVQIVVYDLPDRDCAALASNGEFSLANDGL
F. oxysporum EG #8    EDTLAD-IRKANKAGGK--YAGQFVVYDLPDRDCAAAAASNGEYSLDKDGA
P. chrysosporium CBH #9  GTYLAS--ASALGKSTGTRKQLVQIVYDLPDRDCAAKASNGEFSIANNGQ
T. emersonii CBH #10   GDYDKD-IQSQNAAGADPPIAGIFVVYDLPDRDCAAAAASNGEFSIANNGV
T. fusc CBH #11       T. fusc EG #12       KAGSSQ-ARIYFDAGHSAWHSPAQMASWL----QQADISNSAHGIA-TNT
C. fimi EG #13        -HAQAE-VADYTGRAVAAGKTPMLVVYAI PGRDCGSHSGGG-----VSE
                                .          :          :          :          :          :
T. reesei CBH #3      AKYKN-YIDTIRQIVVEYS---IRTLVIEPDSLNLVNLGTPK----
H. koningii CBH #4   DKYKN-YIDTIRQIVVEYS---IRTLVIEPDSLNLVNLGTPK----
H. insolens CBH #5   NNYKG-YINRIREILISFSD---VRTILVIEPDSLNMVFNMNVAK----
A. cellulolyticus CBH #6  ANYKA-YIDSIVAQLKAYPD---VHTILIEPDSLNMVFNLNSTAK----
A. bisporus CBH #7    NKYKN-YVDQIAAQIKQFPD---VSVVAVIEPDSLNLVNLNVQK----
F. oxysporum EG #8    NKYKA-YIAKIKGILQNYSD---TKVILVIEPDSLNLVNLNVQK----
P. chrysosporium CBH #9  ANYEN-YIDQIVAQIQQFPD---VRVVAVIEPDSLNLVNLNVQK----
T. emersonii CBH #10   ALYKQ-YIDSIREQLTTSYSD---VHTILVIEPDSLNLVNLNVPK----
T. fusc CBH #11       DRYKSEYIDPIADIMWDFADYENLRIVAIIEIDSLPNLVTNVGNGGTEL
T. fusc EG #12       SNYRW-TADEVA-----YAK---AVLSAIGNPSLRAVIDT SRNGNG----
C. fimi EG #13        SEYAR-WVDTVAQGIK-----GNPIVILEPDALAQLGD-----
                                *          :          :          :          :
T. reesei CBH #3      CA--NAQSAYLECIINYAVTQL-NLPNVAMYLDAGHAGWLGWPANQDPAAQ
H. koningii CBH #4   CA--NAQSAYLECIINYAVTQL-NLPNVAMYLDAGHAGWLGWPANQDPAAQ
H. insolens CBH #5   CS--GAASTYRELTIIYALKQL-DLPHVAMYMDAGHAGWLGWPANIQPAE
A. cellulolyticus CBH #6  CA--EAQSAYYECVNIALINL-NLANVAMYIDAGHAGWLGWSANLSPAAQ
A. bisporus CBH #7    CA--NAQSAYKEGVYIYAVQKL-NAVGVTMYIDAGHAGWLGWPANLSPAAQ
F. oxysporum EG #8    CA--KAESAYKELTVYAIKEL-NLPNVSMYLDAGHGGWLGWPANIGPAK
P. chrysosporium CBH #9  CA--NAKTTYLACVNYALTNL-AKVGVMYMDAGHAGWLGWPANLSPAAQ
T. emersonii CBH #10   CA--NAQDAYLECIINYAITQL-DLNPVAMYLDAGHAGWLGWQANLAPAAQ
T. fusc CBH #11       CAYMKQNGGYVNGVGYALRKLGEIPNVVNYIDAHHGWIGWDSNFGPSVD
T. fusc EG #12       ----PAGNEWCDPSGRAIGTPSTTNTGDPMIDAFI--WIKLPEADGCIA
C. fimi EG #13        CS---GQGDRVGF LKYAAKSL-TLKGARVYIDAGHAKWLSVDTTPVNRINQ
                                *          :          :          :          :
T. reesei CBH #3      LFANVYKNAS-SPRALRGLATNVANYNGWNIT-----SPPSY
H. koningii CBH #4   LFANVYKNAS-SPRALRGLATNVANYNGWNIT-----SPPSY
H. insolens CBH #5   LFAKIYEDAG-KPRAVRGLATNVANYNAWSIS-----SPPPY
A. cellulolyticus CBH #6  LFATVYKNAS-APASLRGLATNVANYNAWSIS-----SPPSY
A. bisporus CBH #7    LFAQIYRDAG-SPRNLRGITATNVANFNALRAS-----SPDPI
F. oxysporum EG #8    LYAQIYKDAG-KPSRVRLVTVNVSNYNGWKL-----TKPDY
P. chrysosporium CBH #9  LFTQVWQAG-KSPFIKGLATNVANYNALQAA-----SPDPI
T. emersonii CBH #10   LFASVYKNAS-SPASVRGLATNVANYNAWSIS-----RCPSY
T. fusc CBH #11       IFYEAANASGSTVDYVHGFISNTANYSATVEPYLDVNGTVNGQLIRQSKW
T. fusc EG #12       GAGQFVPQAAAYEMAIAGGTNPNPNPNP--TP-----TPTPT
C. fimi EG #13        V-----GFYAVGFALNTSNYQT-----
                                *          *          *
```

Figura 3 continuada

```

T. reesei CBH #3      TQGNVAVYNEKLYIHAIGPLLANHGWSN-AFFITDQGRSGK----QPTGQQ
H. koningii CBH #4   TQGNVAVYNEQLYIHAIGPLLANHGWSN-AFFITDQGRSGK----QPTGQQ
H. insolens CBH #5   TSPNPNYDEKHYIEAFRPLLEARGFP--AQFIVDQGRSGK----QPTGQQ
A. cellulolyticus CBH #6 TSGDSNYDEKLYINALSPLLTNSNGWPN-AHFIMDTSRNGV----QPTKQQ
A. bisporus CBH #7   TQGNSNYDEIHYIEALAPMLSNAGFP--AHFIVDQGRSGV----QNIIRD
F. oxysporum EG #8   TESNPNYDEQRYTNAFAPLLAQEGWSN-VKFIVDQGRSGK----QPTGQK
P. chrysosporium CBH #9 TQGNPNYDEIHYINALAPLLQQAGWD--ATFIVDQGRSGV----QNIIRD
T. emersonii CBH #10 TQGDANCDDEEDYVNALGPLFQEQGFP--AYFIIDTSRNGV----RPTKQS
T. fusc CBH #11      VDWNQYVDELSFVQDLRQALIAKGFERSDIGMLIDTSRNGWGGPNRPTGFS
T. fusc EG #12       PTPPPGSSGACTATYTIANEWNDGFQATVTVTANQNITGW-----TV
C. fimi EG #13       -----TADSKAYGQQISQRLGG-----KKFVIDTSRNGN-----G
          .
          :
          *

T. reesei CBH #3      QWGD-----WCNVIGTGFGIRPSANTGDSLDSFVWV
H. koningii CBH #4   QWGD-----WCNVIGTGFGIRPSANTGDSLDSFVWI
H. insolens CBH #5   EWGH-----WCNAIGTGFGMRPTANTGHQYVDAFVWV
A. cellulolyticus CBH #6 AWGD-----WCNVIGTGFGVQPTTNTGDFLEDAFVWV
A. bisporus CBH #7   QWGD-----WCNVKGAGFGQRPTTNTGSSLIDAVWV
F. oxysporum EG #8   AQGD-----WCNAKGTGFGLRPSTNTGDALADAFVWV
P. chrysosporium CBH #9 QWGD-----WCNIKGAGFGTRPTTNTGSQFIDSIVWV
T. emersonii CBH #10 QWGD-----WCNVIGTGFGVRPTTDTGNPLEDAFVWV
T. fusc CBH #11      SSTDNLNTYVDESRIIDRRIHPGNWCNQAGAGLGERPTVNPA-PGVDAVWV
T. fusc EG #12       TWT-----FTDGTITNAWNADVSTSGSSVTARNVG
C. fimi EG #13       SNGE-----WCNPRGRALGERPVAVNDGSGLDLWV
          :
          :
          .
          .
          *

T. reesei CBH #3      KPGGEC DGTSDSS----APRFDSHCA-----LPDALQPAPQAG
H. koningii CBH #4   KPGGEC DGTSDSS----APRFDSHCA-----LPDALQPAPQAG
H. insolens CBH #5   KPGGEC DGTSDTT----AARYDYHCG-----LEDALKPAPEAG
A. cellulolyticus CBH #6 KPGGES DGTSNSS----ATRYDFHCG-----YSDALQPAPQAG
A. bisporus CBH #7   KPGGEC DGTSDNS----SPRFDSHCS-----LSDAHQPAPEAG
F. oxysporum EG #8   KPGGES DGTSDTS----AARYDYHCG-----LDDALKPAPEAG
P. chrysosporium CBH #9 KPGGEC DGTSNSS----SPRYDSTCS-----LPDAAQPAPEAG
T. emersonii CBH #10 KPGGES DGTSNTT----SPRYDYHCG-----LSDALQPAPQAG
T. fusc CBH #11      KPPGES DGASEEIPNDEGKGFDRMCDPTYQGNARNGMNP SGALPNAPISG
T. fusc EG #12       HNGTLSQGASTEF-----GFVG
C. fimi EG #13       KLPGES DGACNGG-----PAAG
          :
          .
          *

T. reesei CBH #3      AWFQAYFVQLLTNANPSFL
H. koningii CBH #4   AWFQAYFVQLLTNANPSFL
H. insolens CBH #5   QWFQAYFEQLLRNANPPF-
A. cellulolyticus CBH #6 TWFAQAYFVQLLTNANPALV
A. bisporus CBH #7   TWFAQAYFETLVANANPAL-
F. oxysporum EG #8   TWFAQAYFEQLLDNANPSFL
P. chrysosporium CBH #9 TWFAQAYFQTLVSAANPPL-
T. emersonii CBH #10 TWFAQAYFEQLLTNANPLF-
T. fusc CBH #11      HWFSAQFRELLANAYPPL-
T. fusc EG #12       SKGNSNSVPTLTCAAS---
C. fimi EG #13       QWQEI ALEMARNARW---
          .
          *
    
```

Figura 4

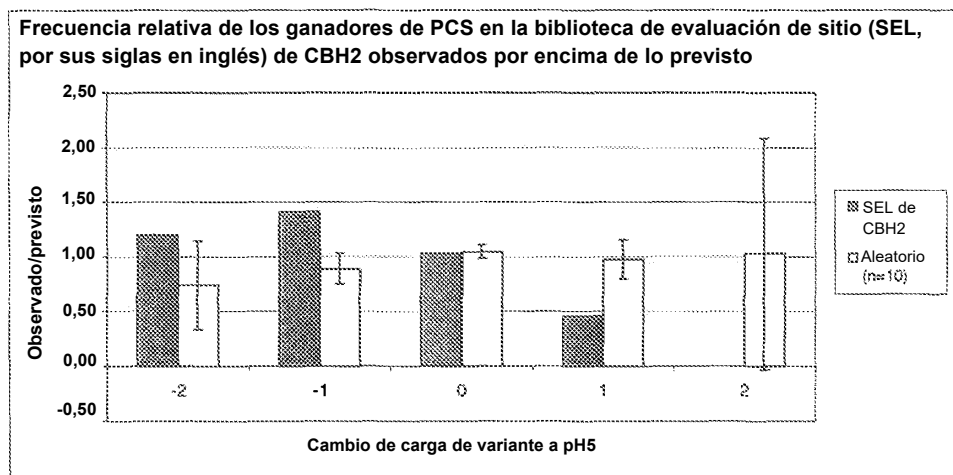


Figura 5

