



No. 13-110810-00000-0000

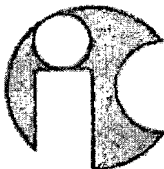
Fecha: 2013-05-02 16:48:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 90

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 10-151612-00008-0000

Fecha: 2013-05-02 16:48:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 851 DIVISIONAL Folios: 90



Industria y Comercio

SUPERINTENDENCIA

División de Nuevas Creaciones

SOLICITUD PATENTE DE INVENCION

DIVISIONAL DEL EXPEDIENTE No. 10-151612

54. TITULO COMPOSICIONES Y METODOS QUE COMPRENDEN
VARIANTES DE CELULASA CON AFINIDAD REDUCIDA PARA
MATERIALES NO CELULOSICOS.

51. CLASIFICACION INTERNACIONAL

71. SOLICITANTE DANISCO US INC

DOMICILIO 925 Page Mil Road, Palo Alto, California 94304,
Estados Unidos de America

74. APODERADO JERONIMO CASTILLO BONILLA

22. BOGOTA D.C., _____



No. 13-110810- -00000-0000

Fecha: 2013-05-02 16:48:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 90

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO



No. 10-151612- -00008-0000

Fecha: 2013-05-02 16:48:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 851 DIVISIONAL Folios: 90

DIRECCIÓN DE
CONVERSIÓN, DIVISIÓN

1. Identificación del Trámite

☐

Conversión

☒

División

☐

Fusión

☒

Patente de Invención

☐

Patente de Modelo de Utilidad

☐

Registro de Diseño Industrial

☐

Esquema de Trazado de Circuitos Integrados

2. Datos del solicitante

☐

Persona Natural

☒

Persona Jurídica

DANISCO US INC

Nombre o denominación / Nombre o razón social

Nombre del representante legal

Tipo de empresa

☐

Micro

☐

Pequeña

☐

Mediana

☐

Otra

Documento de Identificación

☐

C.C.

☐

C.E.

☐

NIT

☐

Otro

Número

Nacionalidad/País de constitución	Dirección del titular	Ciudad
Estados Unidos de America	925 Page Mill Road	Palo Alto, California
Dirección electrónica	No. Fax	Número telefónico

☒

Autorizo expresamente a la Superintendencia de Industria y Comercio para que todas las comunicaciones sean enviadas a través de la dirección de correo electrónico

3. Datos del apoderado:

CASTILLO BONILLA JERONIMO

Apellidos y Nombre

79.462.057 BOGOTA

Documento de identificación

104167

Tarjeta profesional

Dirección y domicilio		Ciudad
Calle 93 A No. 14 - 17, Oficina 311,		Bogotá, D.C.
Dirección electrónica	No. Fax	Número telefónico
info@pcastillopineda.com	218 23 51 – 257 44 60	236 34 93 – 236 34 09 – 635 12 28

En caso de haber presentado poder general para asuntos que se adelanten ante la Delegatura de Propiedad Industrial, sírvase indicar el número de su radicación o protocolo de poder general

4. Datos del trámite:

CONVERSION

Expediente No.

De:

A:

DIVISION

Expediente No.

10-151612

No. de divisionales

FUSION

Expediente No.

Expediente No.

5. Comprobante de Pago No

Fecha

DD

MM

AA

6. Anexos

X

Comprobante de pago de la tasa para la presentación de la corrección y/o modificación

Poder, si fuere el caso con el que se acredita la representación

Número de documento que ordena la comisión o fusión o división

Documento que contiene la conversión, división o fusión de la solicitud

Copia de la solicitud y sus anexos en formato magnético

7. Firma

JERONIMO CASTILLO BONILLA

Nombre y apellidos

79.462.057 BOGOTA

C.C.

Firma

104167

Tarjeta Profesional

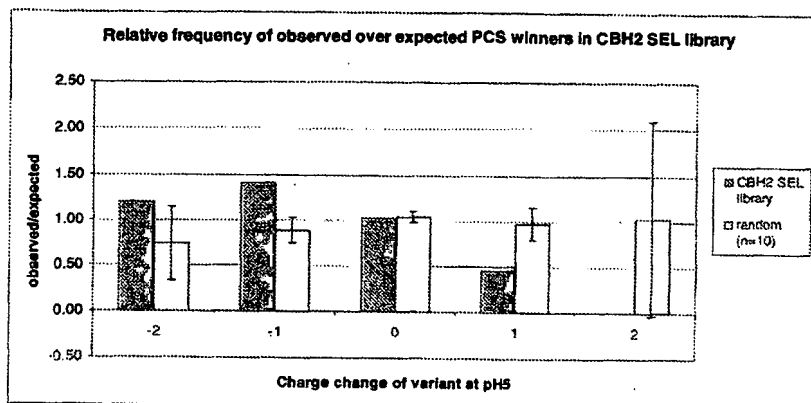
PETITORIO																
 Industria y Comercio SUPERINTENDENCIA	Espacio reservado para el adhesivo de radicación															
FORMULARIO ÚNICO DE SOLICITUD DE PATENTES PCT ENTRADA FASE NACIONAL																
SOLICITUD DE REGISTRO DE:																
☒ Patente de Invención ☐ Patente Modelo de Utilidad																
☒ Capítulo I ☐ Capítulo II																
SOLICITANTE	Nombre: DANISCO US INC Dirección: 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América Domicilio: Estadounidense Lugar de Constitución: Delaware Teléfono: Fax: E - mail:															
	IDENTIFICACIÓN ☐ C.C. ☐ NIT ☐ C.E. ☐ Otro Cual: _____ Número: _____															
REPRESENTANTE O APODERADO	Nombre: JERÓNIMO CASTILLO BONILLA Dirección: Calle 93A N° 14-17, Ofc 311 Teléfono: 2363493 Fax: 2182351 E-mail: info@pcastillopineda.com Número: 79.462.057 TP: 104167 En caso de haber presentado poder general para asuntos que se adelanten ante la Delegatura de Propiedad Industrial, sírvase indicar el número de radicación (del poder) _____															
INVENTOR (ES) (72)	Nombre: LUIS G. CASCAO - PEREIRA Dirección: 925 Page Mill Road, Palo Alto, California 94304, Estados Unidos de América. Nacionalidad o Domicilio: Brasileño															
PCT (86) Solicitud Internacional No.: <u>PCT/US2010/049849</u> Fecha: <u>Septiembre 22, 2010</u> Publicación Internacional No.: <u>WO 2011/038019 A2</u> Fecha: <u>Marzo 31, 2011</u>																
Título (54) COMPOSICIONES Y METODOS QUE COMPRENDEN VARIANTES DE CELULASA CON AFINIDAD REDUCIDA PARA MATERIALES NO CELULOSICOS.																
Clasificación Internacional (51)																
Prioridad	<table border="1"><thead><tr><th>(33) País de Origen</th><th>(32) Fecha</th><th>(31) Número de Solicitud</th></tr></thead><tbody><tr><td>US</td><td>Septiembre 23, 2009</td><td>61/245.273</td></tr><tr><td>US</td><td>Diciembre 23, 2009</td><td>61/289.886</td></tr><tr><td></td><td></td><td></td></tr><tr><td></td><td></td><td></td></tr></tbody></table>	(33) País de Origen	(32) Fecha	(31) Número de Solicitud	US	Septiembre 23, 2009	61/245.273	US	Diciembre 23, 2009	61/289.886						
(33) País de Origen	(32) Fecha	(31) Número de Solicitud														
US	Septiembre 23, 2009	61/245.273														
US	Diciembre 23, 2009	61/289.886														
☒ SI ☐ NO																
Comprobante de pago No: Fecha:																

Anexos

- ☒ Comprobante de pago de la tasa de presentación de la solicitud.
- ☒ Documento que acredita la existencia y representación legal de la persona jurídica peticionaria, en cuanto no estén inscritas en alguna Cámara de Comercio Colombiana.
- ☒ Poderes, si fuere el caso.
- ☐ Documento de cesión del inventor al solicitante o a su causante.
- ☒ Traducción de la solicitud internacional al castellano (Resumen, descripción, reivindicaciones, dibujos)
- ☐ Copia del examen preliminar internacional efectuado por la Autoridad Internacional de Examen preliminar (IPEA)
- ☐ Traducción de examen preliminar internacional efectuado por la IPEA
- ☐ De ser el caso, copia del contrato de acceso.
- ☐ De ser el caso, documento que acredite la licencia o autorización de uso de conocimientos tradicionales de las comunidades indígenas.
- ☐ De ser el caso, certificado de depósito del material biológico
- ☒ Arte final 12x12.
- ☐ De ser el caso, información sobre otras solicitudes de patente o títulos obtenidos en el extranjero por el mismo titular o su causante, relacionadas parcial o totalmente con la invención de esta solicitud.
- ☐ De ser el caso, modificaciones efectuadas a la solicitud internacional.
- ☐ Otros Publicación Internacional **Notificación de la Oficina Internacional mostrando que los solicitantes cambiaron para DANISCO US INC.**

11. FIGURA CARACTERÍSTICA:

Figure 4



12. Solicito la concesión de la patente

NOMBRE: JERÓNIMO CASTILLO B.

FIRMA: [Signature]

C.C. 79.462.057 T.P. 104167

Instrucciones para el diligenciamiento del presente formulario, ver reverso de esta página

(Continuación Inventores)

Nombre: Thijs KAPER
Dirección: 925 Page Mill Road, Palo Alto, California 94304,
Estados Unidos de América.
Nacionalidad: Holandesa

Nombre: Bradley R. KELEMEN
Dirección: 925 Page Mill Road, Palo Alto, California 94304,
Estados Unidos de América.
Nacionalidad: Estadounidense

Nombre: Amy LIU
Dirección: 925 Page Mill Road, Palo Alto, California 94304,
Estados Unidos de América.
Nacionalidad: Canadiense

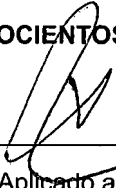
SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO
NIT : 800.176.089-2
- / -

 Industria y Comercio SUPERINTENDENCIA	RECIBO DE CAJA	No. 13 - 0041456
		Bogotá D.C., Mayo 02 de 2013 - 16:28:40

RECIBIDO DE : CASTILLO PINEDA SAS				NI 860.046.502	
*** Soporte del Pago ***					
TIPO PAGO	BANCO	CUENTA	No. PAGO	FECHA PAGO	VR. PAGO
CONSIGNACION	BANCO DE BOGOTA	062754387	522530610	02/05/2013	2.310.000.00

*** Conceptos Pagados ***				
CANT.	RENTISTICO	CONCEPTO	Vr. UNDITARIO	Vr. CONCEPTO
1	50005-01-01 SOLICITUDES	2205 DIVISIONAL SOLICITUD DE PATENTE DE INVENCIÓN CONT. DERE	450.000.00	450.000.00
				\$450.000.00
=====				

SON: **CUATROCIENTOS CINCUENTA MIL PESOS MONEDA CORRIENTE**

Responsable: 
--
Recibo de Caja Aplicado al Expediente No. _____

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO5

No. 13-110810- -00000-0000
Fecha: 2013-05-02 16:48:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 411 PRESENTACION Folios: 90

SUPERINTENDENCIA DE INDUSTRIA Y COMERCIO

No. 10-151612- -00008-0000
Fecha: 2013-05-02 16:48:23 Dep. 2020 DIR.NUEVASCR
Tra. 11 PATENTEIYII Eve: 378 FASENACIONALI
Act. 851 DIVISIONAL Folios: 90

COMPOSICIONES Y MÉTODOS QUE COMPRENDEN VARIANTES DE CELULASA CON AFINIDAD REDUCIDA PARA MATERIALES NO CELULÓSICOS

I. DECLARACIÓN RELATIVA A LOS DERECHOS EN INVENCIONES REALIZADAS MEDIANTE INVESTIGACIÓN Y DESARROLLO FINANCIADOS CON RECURSOS FEDERALES

[01] Esta invención se realizó con apoyo del Gobierno bajo subvención condicional No: DE-FC36-08GO18078, otorgada por el Departamento de Energía. El Gobierno tiene ciertos derechos en esta invención.

II. REFERENCIA CRUZADA A UNA SOLICITUD RELACIONADA

[02] Esta solicitud reclama el beneficio de la Solicitud Provisional de los Estados Unidos No. 61/059,506, presentada el 6 de junio de 2008, la cual se incorpora en esta especificación por referencia.

III. CAMPO

[03] La presente divulgación se refiere a enzimas y, en particular, a variantes de celulasa. Se describen asimismo ácidos nucleicos que codifican las variantes de celulasa, composiciones que contienen las variantes de celulasa, métodos útiles adicionales de identificar variantes de celulasa y métodos de usar las composiciones.

IV. ARTE PREVIO

[04] La celulosa y la hemicelulosa son los materiales vegetales más abundantes producidos por fotosíntesis. Pueden degradarse y usarse como una fuente de energía por numerosos microorganismos (v.g., bacterias, levadura y hongos), que producen enzimas extracelulares capaces de hidrolizar los sustratos poliméricos en azúcares monoméricos (Aro y col., J Biol Chem, 276: 24309-24314, 2001). A medida que se aproximan los límites de los recursos no renovables, el potencial de la celulosa de convertirse en un importante recurso energético renovable es enorme (Krishna y col., Bioresource Tech, 77: 193-196, 2001). La utilización efectiva de celulosa por procesos biológicos es un enfoque para superar la escasez de alimentos, forrajes, y combustibles (Ohmiya y col., Biotechnol Gen Engineer Rev, 14: 365-414, 1997).

[05] Las celulasas son enzimas que hidrolizan celulosa (enlaces beta-1,4-glucano o beta D-glucosídicos), lo cual resulta en 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 y col., TIBTECH 5: 255-261, 1987; y Schulein, Methods Enzymol, 160: 234-243, 1988). Las endoglucanasas actúan principalmente sobre las partes amorfas de la fibra de celulosa, mientras que las celobiohidrolasas son también capaces de degradar la celulosa cristalina (Nevalainen y Penttila, Mycota, 303-319, 1995). Por consiguiente, la presencia de una celobiohidrolasa en un sistema de celulasas se requiere para una eficiente solubilización de la celulosa cristalina (Suurnakki y col., 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).

[06] Se sabe que las celulasas son producidas por un gran número de bacterias, levaduras y hongos. Ciertos hongos producen un sistema completo de celulasas capaz de degradar formas cristalinas de celulosa, de modo que las celulasas se produzcan fácilmente en grandes cantidades mediante fermentación. Los hongos filamentosos juegan un papel especial puesto que muchas levaduras, tales como *Saccharomyces cerevisiae*, carecen de la capacidad de hidrolizar celulosa (véase, v.g., Wood y col., Methods in Enzymology, 160: 87-116, 1988).

[07] Las clasificaciones de CBH, EG y BG de las celulasas fúngicas pueden ampliarse además para incluir múltiples componentes dentro de cada clasificación. Por ejemplo, múltiples CBHs, EGs y BGs se han aislado de una variedad de fuentes fúngicas incluyendo *Trichoderma reesei* (también denominado *Hypocrea jecorina*), el cual contiene genes conocidos para dos CBHs, es decir, CBH I ("CBH1") y CBH II ("CBH2"), por lo menos 8 EGs, es decir, EG I, EG II, EG III, EGIV, EGV, EGVI, EGVII y EGVIII, y por lo menos 5 BGs, es decir, BG1, BG2, BG3, BG4 y BG5. EGIV, EGVI y EGVIII tienen igualmente actividad de xiloglucanasa.

[08] Con el fin de convertir eficientemente celulosa cristalina en glucosa, se requiere el sistema completo de celulasas constituido por componentes de cada una de las clasificaciones CBH, EG y BG, con componentes aislados menos efectivos en hidrolizar celulosa cristalina (Filho y col., Can J Microbiol, 42:1-5,

1996). Se ha observado una relación sinérgica entre componentes de celulasa de diferentes clasificaciones. En particular, las celulasas tipo EG y las celulasas tipo CBH interactúan sinérgicamente para degradar celulosa de manera más eficiente.

[09] Las celulasas son conocidas en el arte por ser útiles en el tratamiento de textiles para los propósitos de potenciar la capacidad limpiadora de composiciones detergentes, para uso como un agente suavizante, para mejorar el tacto y apariencia de las telas de algodón, y similares (Kumar y col., Textile Chemist and Colorist, 29:37-42, 1997). Se han descrito composiciones detergentes que contienen celulasa con rendimiento limpiador mejorado (Patente de los Estados Unidos No. 4,435,307; Solicitudes de Gran Bretaña Nos. 2,095,275 y 2,094,826) y para uso en el tratamiento de telas para mejorar el tacto y apariencia de textiles (Patentes de los Estados Unidos Nos. 5,648,263, 5,691,178, y 5,776,757; y Solicitud de la Gran Bretaña No. 1,358,599). Por consiguiente, celulasas producidas en hongos y bacterias han recibido una atención significativa. En particular, la fermentación de *Trichoderma spp.* (v.g., *Trichoderma longibrachiatum* o *Trichoderma reesei*) ha demostrado producir un sistema completo de celulasas capaz de degradar formas cristalinas de celulosa.

[10] Aunque previamente se han descrito composiciones de celulasa, subsiste una necesidad de composiciones de celulasa nuevas y mejoradas. Composiciones de celulasa mejoradas se usan en detergentes de uso en el hogar, tratamientos para textiles, conversión de biomasa y fabricación de papel. Son de particular interés las celulasas que exhiben un rendimiento mejorado.

V. SUMARIO

[11] La presente enseñanza 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 incrementada en la presencia de materiales no celulósicos, en comparación con celulasas naturales ("**wild type cellulases**"). En algunas incorporaciones las variantes de celulasa tienen una carga neta disminuída (es decir, es más negativa) en comparación con celulasas naturales. En algunas incorporaciones, las variantes de celulasa están cargadas menos positivamente que las celulasas naturales. En algunas incorporaciones, una celulasa se modifica eliminando una o más cargas positivas. En algunas incorporaciones, una celulasa se modifica agregando una o más cargas negativas. En algunas incorporaciones, una celulasa se modifica eliminando una o más cargas positivas y agregando una o más cargas negativas.

[12] En algunas incorporaciones, las presentes enseñanzas se refieren a variantes de celobiohidrolasa I (CBH1) o celobiohidrolasa II (CBH2). En algunas incorporaciones la variante de celulasa es una forma madura que tiene actividad de celulasa y una sustitución en una o más posiciones seleccionadas del grupo conformado por 63, 77, 129, 147, 153, 157, 161, 194, 197, 203, 237, 239, 247, 254, 281, 285, 288, 289, 294, 327, 339, 344, 356, 378, y 382, en donde las posiciones se numeran por correspondencia a una celulasa de referencia (v.g., *Hypocrea jecorina* CBH2 natural) que tiene la secuencia de aminoácidos de SEQ ID NO:3, y en donde la sustitución en una o más posiciones hace que la variante de celulasa tenga una carga neta más negativa en comparación con la celulasa de referencia. En algunas incorporaciones, CBH2 se modifica eliminando una o más cargas positivas, lo cual en algunas incorporaciones implica un reemplazo de una lisina o una arginina con un aminoácido neutro (v.g., K o R reemplazado por N o Q u otro residuo neutro). En algunas incorporaciones, CBH2 se modifica agregando una o más cargas negativas, lo cual en algunas incorporaciones implica un reemplazo de un aminoácido neutro con un aminoácido cargado negativamente (v.g., N o Q u otro residuo neutro reemplazado por D o E). En algunas incorporaciones, CBH2 se modifica eliminando una o más cargas positivas y agregando una o más cargas negativas, lo cual en algunas incorporaciones implica un reemplazo de una lisina o una arginina con un aminoácido cargado negativamente (v.g., K o R reemplazado con D o E). En general, la variante CBH2 tiene una actividad celulolítica incrementada en la presencia de lignina en comparación con la *Hypocrea jecorina* CBH2 natural que tiene la secuencia de aminoácidos de SEQ ID NO:3. Las presentes enseñanzas proporcionan además variantes CBH2 que incluyen una o más sustituciones seleccionadas del grupo conformado por K129E, K157E, K194E, K288E, K327E, K356E, R63Q, R77Q, R153Q, R203Q, R294Q, R378Q, N161D, N197D, N237D, N247D, N254D, N285D, N289D, N339D, N344D, N382D, Q147E, Q204E, Q239E, Q281E, D151N, D189N, D211N, D277N, D405N, E146Q, E208Q, y E244Q, en la forma madura de CBH2, en donde las citadas sustituciones se numeran de acuerdo con la forma madura de *Hypocrea jecorina* CBH2 de SEQ ID NO:3. En algunas incorporaciones, la variante incluye una sustitución adicional en una o más posiciones adicionales seleccionadas del grupo conformado por 146, 151, 189, 208, 211, 244, 277 y 405, en donde las posiciones adicionales se numeran por correspondencia con la secuencia de aminoácidos de la celobiohidrolasa II (CBH2) de referencia divulgada como SEQ ID NO:3. En algunas incorporaciones, la sustitución adicional en una o más posiciones adicionales incluye un reemplazo de ácido aspártico o ácido glutámico con un aminoácido neutro (v.g., D o E reemplazado por N o Q u

otro residuo neutro). En algunas incorporaciones, la sustitución adicional en una o más posiciones adicionales incluye una o más del grupo conformado por D151N, D189N, D211N, D277N, D405N, E146Q, E208Q, y E244Q, en donde las posiciones se numeran por correspondencia con la secuencia de aminoácidos de la celobiohidrolasa II (CBH2) de referencia divulgada como SEQ ID NO:3. En algunas incorporaciones preferidas, la sustitución en una o más posiciones se selecciona del grupo conformado por las posiciones 1, 2, 3, 4, 5, 6, 7, 8, 9 y 10. En algunas incorporaciones preferidas, la variante de celulasa se deriva de una celulasa progenitora seleccionada del grupo conformado por *Hypocrea jecorina* CBH2, *Hypocrea koningii* CBH2, *Humicola insolens* CBH2, *Acremonium cellulolyticus* CBH2, *Agaricus bisporus* CBH2, *Fusarium osysporum* EG, *Phanerochaete chrysosporium* CBH2, *Talaromyces emersonii* CBH2, *Thermobifida fusca* 6B/E3 CBH2, *Thermobifida fusca* 6A/E2 EG, y *Cellulomonas fimi* CenA EG. En algunas incorporaciones preferidas, la variante de celulasa se deriva de una celulasa progenitora cuya secuencia de aminoácidos es por lo menos 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% idéntica a un miembro del grupo conformado por SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO: 10, SEQ ID NO:11, SEQ ID NO:12, y SEQ ID NO:13. En algunas incorporaciones, la carga neta más negativa es -1 o -2 en comparación con la CBH2 de referencia.

[13] La presente divulgación provee también variantes de celulasa, en donde la variante es una forma madura que tiene actividad de celulasa e incluye una modificación química de un residuo lisina para eliminar la carga positiva del residuo lisina. En algunas incorporaciones preferidas, la modificación química comprende un tratamiento con un compuesto seleccionado del grupo conformado por anhídrido succínico, anhídrido acetoxisuccínico, anhídrido maleico, anhídrido tartárico, anhídrido ftálico, anhídrido trimelítico, anhídrido cis-aconítico, anhídrido t-nitroftálico, anhídrido acético, anhídrido butírico, anhídrido isobutírico, anhídrido hexanoico, anhídrido valérico, anhídrido isovalérico y anhídrido piválico. En algunas incorporaciones preferidas, la variante de celulasa se deriva de una celulasa progenitora seleccionada del grupo conformado por una *Hypocrea jecorina* celobiohidrolasa I, *Hypocrea jecorina* celobiohidrolasa II, *Hypocrea jecorina* endoglucanasa I, *Hypocrea jecorina* endoglucanasa II, e *Hypocrea jecorina* beta-glucosidasa. En algunas incorporaciones preferidas, la variante de celulasa se deriva de una celulasa progenitora seleccionada del grupo conformado por *Hypocrea jecorina* CBH2, *Hypocrea koningii* CBH2, *Humicola insolens* CBH2, *Acremonium cellulolyticus* CBH2, *Agaricus bisporus* CBH2, *Fusarium osysporum* EG, *Phanerochaete chrysosporium* CBH2, *Talaromyces emersonii* CBH2,

Thermobifida fusca 6B/E3 CBH2, *Thermobifida fusca* 6A/E2 EG, y *Cellulomonas fimi* CenA EG. De igual manera, se proveen variantes de celulasa derivadas de una celulasa progenitora cuya secuencia de aminoácidos es por lo menos 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98% o 99% idéntica a un miembro del grupo conformado por SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, y SEQ ID NO: 13. En algunas incorporaciones, la variante de celulasa comprende una sustitución en una o más posiciones seleccionadas del grupo conformado por 63, 77, 129, 147, 153, 157, 161, 194, 197, 203, 237, 239, 247, 254, 281, 285, 288, 289, 294, 327, 339, 344, 356, 378, y 382, en donde las posiciones se numeran por correspondencia con la secuencia de aminoácidos de una celobiohidrolasa II (CBH2) de referencia divulgada como SEQ ID NO:3.

[14] Las presentes enseñanzas se refieren además a una variante CBH2 que incluye de una a veintiséis sustituciones seleccionadas del grupo conformado por K129E, K157E, K194E, K288E, K327E, K356E, R63Q, R77Q, R153Q, R203Q, R294Q, R378Q, N161D, N197D, N237D, N247D, N254D, N285D, N289D, N339D, N344D, N382D, Q147E, Q204E, Q239E, y Q281E. En algunas incorporaciones, la variante CBH2 comprende una combinación de sustituciones seleccionadas del grupo conformado por: i) K157E / K129E; ii) K157E / K129E / K288E / K194E; iii) K157E / K129E / K288E / K194E / K356E / K327E; iv) K157E / K129E / K288E / K194E / K356E / K327E / R153Q / R294Q / R203Q / R378Q; v) K157E / K129E / K288E / K194E / K356E / K327E / R153Q / R294Q / R203Q / R378Q / N382D / N344D / N327D / N339D; vi) K157E / K129E / K288E / K194E / K356E / K327E / R153Q / R294Q / R203Q / R378Q / N382D / N344D / N327D / N339D / N289D / N161D / Q204E / Q147E; vii) K157E / K129E / K288E / K194E / K356E / K327E / R153Q / R294Q / R203Q / R378Q / N382D / N344D / N327D / N339D / N289D / N161D / Q204E / Q147E / N285D / N197D / N254D / N247D; y viii) K157E / K129E / K288E / K194E / K356E / K327E / R153Q / R294Q / R203Q / R378Q / N382D / N344D / N327D / N339D / N289D / N161D / Q204E / Q147E / N285D / N197D / N254D / N247D / Q239E / Q281E / R63Q / R77Q.

[15] En algunas incorporaciones, la variante CBH2 incluye de una a ocho sustituciones seleccionadas del grupo conformado por D151N, D189N, D211N, D277N, D405N, E146Q, E208Q, y E244Q. En algunas incorporaciones, la variante CBH2 incluye una combinación de sustituciones seleccionadas del grupo constituido por: i) D189N / E208Q / D211N / D405; y ii) D189N / E208Q / D211N / D405 / E244Q / D277N / D151 / E146Q.

[16] Se describen asimismo ácidos nucleicos aislados que codifican una variante CBH2 que posee actividad de celobiohidrolasa, como se describe en los párrafos precedentes. En un primer aspecto, la divulgación abarca un ácido nucleico aislado que codifica un polipéptido que posee actividad de celobiohidrolasa, el cual polipéptido es una variante de una glicosil hidrolasa de la familia 6, y en donde el citado ácido nucleico codifica una sustitución en un residuo que disminuye la carga neta en comparación con la *Hypocrea jecorina* CBH2 natural.

[17] En otro aspecto, la divulgación se dirige a un ácido nucleico aislado que codifica una variante CBH2, en donde la citada variante comprende una sustitución en una posición seleccionada del grupo conformado por K129E, K157E, K194E, K288E, K327E, K356E, R63Q, R77Q, R153Q, R203Q, R294Q, R378Q, N161D, N197D, N237D, N247D, N254D, N285D, N289D, N339D, N344D, N382D, Q147E, Q204E, Q239E, Q281E, D151N, D189N, D211N, D277N, D405N, E146Q, E208Q, y E244Q, en la forma madura de CBH2, en donde las citadas sustituciones se numeran de acuerdo con la forma madura de *Hypocrea jecorina* CBH2 de SEQ ID NO:3.

[18] En algunas incorporaciones, la divulgación se dirige a un casete de expresión que comprende un ácido nucleico que codifica una variante CBH2, un constructo que comprende el ácido nucleico que codifica la variante CBH2 enlazado operacionalmente a una secuencia reguladora, un vector que comprende un ácido nucleico que codifica una variante CBH2, y una célula hospedera transformada con el vector que comprende un ácido nucleico que codifica una variante CBH2. Las presentes enseñanzas ofrecen además métodos para producir una variante CBH2 cultivando las células hospederas que expresan una variante CBH2 en un medio de cultivo bajo condiciones apropiadas para producir la variante CBH2.

[19] Se proveen igualmente composiciones que comprenden la variante de celulasa de los párrafos precedentes. En algunas incorporaciones preferidas, la composición comprende además por lo menos una enzima adicional seleccionada del grupo conformado por una subtilisina una metaloproteasa neutra, una lipasa, una cutinasa, una amilasa, una carbohidrasa, una pectinasa, una mannanasa, una arabinasa, una galactanasa, una xilanasa, una oxidasa, y una peroxidasa.

[20] En esta especificación se ofrecen métodos de convertir biomasa en azúcares, métodos que comprenden contactar la citada biomasa con una variante de celulasa. Se proveen además métodos de producir un combustible contactando una composición de biomasa con una composición enzimática que comprende la

variante de celulasa para producir una solución de azúcar y cultivar con un microorganismo fermentador bajo condiciones suficientes para producir un combustible.

[21] Se proporcionan igualmente composiciones que comprenden variantes de celulasa incluyendo composiciones detergentes, aditivos de forrajes por ejemplo, y métodos de limpieza o cuidado de telas contactando una superficie y/o un artículo, que comprende una tela, con la composición detergente. Asimismo, se facilitan métodos de tratamiento para el cuidado de telas, incluyendo la eliminación de bolitas formadas por agrupación de fibras en los tejidos ("*depilling*") y acabado superficial, contactando una superficie y/o un artículo, que comprende una tela, con una variante de celulasa.

[22] Otros objetos, características y ventajas de la presente divulgación se harán evidentes a partir de la siguiente descripción detallada. Debe entenderse, sin embargo, que la descripción detallada y los ejemplos específicos, aunque indican incorporaciones preferidas de la divulgación, se dan a manera de ilustración únicamente, puesto que varios cambios y modificaciones dentro del ámbito y espíritu de la divulgación se harán evidentes a un experto en el arte a partir de esta descripción detallada.

VI. BREVE DESCRIPCIÓN DE LSO DIBUJOS

[23] FIG. 1 ilustra la sacarificación de APB (*acid pretreated bagasse* – bagazo pretratado con ácido) por preparaciones de celulasa de *Trichoderma sp.* modificada (cuadrados) y no modificada (círculos) en la presencia de cantidades crecientes de inhibidor de lignina. FIG. 1A y FIG. 1B muestran los resultados después de incubaciones de 24 y 48 horas, respectivamente.

[24] FIG. 2A ilustra la sacarificación comparando celulasas modificadas y FIG. 2B muestra la diferencia de la sacarificación usando celulasas modificadas y no modificadas.

[25] FIG. 3 presenta un alineamiento de las secuencias de aminoácidos de la forma madura de varias 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 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).

[26] FIG. 4 ilustra un gráfico de la frecuencia relativa de los ganadores observados sobre los esperados en un análisis con rastrojo de maíz pretratado (PCS - *pretreated corn stover*) de la variante CBH2 SELs como un producto del cambio de carga. Disminuir la carga de CBH2 resulta en una frecuencia significativamente mayor de ganadores PCS.

[27] FIG. 5 presenta un mapa de un plásmido de pTTTpyr-cbh2.

VII. DESCRIPCIÓN DETALLADA DE VARIAS INCORPORACIONES

[28] Las presentes enseñanzas se refieren a variantes de celulasa modificada para reducir la unión a materiales no celulósicos. En general, la variante de celulasa tiene una mayor actividad celolítica en la presencia de materiales no celulósicos, en comparación con la celulasa natural. En algunas incorporaciones la celulasa variante tiene una carga neta disminuída (*es decir*, es más negativa) en comparación con la celulasa natural. En algunas incorporaciones, las variantes de celulasa están cargadas menos positivamente que la celulasa natural. En algunas incorporaciones, una celulasa se modifica eliminando una o más cargas positivas. En algunas incorporaciones, una celulasa se modifica agregando una o más cargas negativas. En algunas incorporaciones, una celulasa se modifica eliminando una o más cargas positivas y agregando una o más cargas negativas.

[29] Hay que entender que tanto la anterior descripción general como la siguiente descripción detallada son ejemplares y explicativas únicamente y no son restrictivas de las composiciones y métodos descritos en esta especificación. A menos que se defina lo contrario en esta especificación, todos los términos técnicos y científicos empleados aquí tienen el mismo significado que se entiende comúnmente por una persona de habilidad ordinaria en el arte al cual pertenece esta divulgación. En esta solicitud, el uso del singular incluye el plural a menos que específicamente se manifieste lo contrario. El uso de "o" significa "y/o", a menos que se indique otra cosa. Los términos "comprenden," "que comprende," "comprende," "incluyen," "incluyendo," e "incluye" no pretenden ser limitantes. Todas las patentes y publicaciones, incluyendo todas las secuencias de aminoácidos y las secuencias nucleotídicas divulgadas en tales patentes y publicaciones, mencionadas en esta especificación, se incorporan expresamente

por referencia. Los encabezamientos provistos en esta especificación no constituyen limitaciones de los diferentes aspectos o incorporaciones de la divulgación que puedan ser tenidas por referencia a la especificación como un todo. Por consiguiente, los términos contenidos aquí se definen más completamente por referencia a la especificación como un todo.

[30] A menos que se defina lo contrario en esta especificación, todos los términos técnicos y científicos empleados aquí tienen el mismo significado que se entiende comúnmente por una persona de habilidad ordinaria en el arte al cual pertenece esta divulgación. Singleton, y col., *DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY*, 2D ED., John Wiley and Sons, New York (1994), y Hale & Marham, *THE HARPER COLLINS DICTIONARY OF BIOLOGY*, Harper Perennial, NY (1991) ofrecen a una persona experta un diccionario general de muchos de los términos usados en la divulgación. Aunque algunos métodos y materiales similares o equivalentes a los descritos en esta especificación pueden usarse en la práctica o experimentación de la presente divulgación, se describen los métodos y materiales preferidos. Los intervalos numéricos son inclusivos de los números que definen el intervalo. A menos que se indique otra cosa, los ácidos nucleicos se escriben de izquierda a derecha en orientación 5' a 3'; las secuencias de aminoácidos se escriben de izquierda a derecha en orientación amino a carboxilo, respectivamente. A los profesionales se los remite particularmente a Sambrook y col., *MOLECULAR CLONING: A LABORATORY MANUAL* (Second Edition), Cold Spring Harbor Press, Plainview, N.Y., 1989, y Ausubel FM y col., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y., 1993, en lo que respecta a las definiciones y términos del arte. Hay que entender que esta divulgación no se limita a la metodología, protocolos, y reactivos particulares descritos, ya que éstos pueden variar.

I. Definiciones

[31] Los siguientes términos se definen más completamente por referencia a la especificación como un todo.

[32] El término "polipéptido", tal como se emplea en esta especificación, se refiere a un compuesto constituido por una cadena única de residuos aminoácidos conectados por enlaces peptídicos. El término "proteína", tal como se emplea en esta especificación, puede ser sinónimo con el término "polipéptido".

[33] "Variante" significa una proteína que se deriva de una proteína precursora

(v.g., la proteína nativa) por adición de uno o más aminoácidos a uno cualquiera o a ambos extremos C- y N-terminal, sustitución de uno o más aminoácidos en uno o un número de diferentes sitios en la secuencia de aminoácidos, o eliminación de uno o más aminoácidos en uno cualquiera o en ambos extremos de la proteína o en uno o más sitios en la secuencia de aminoácidos, o uno o más aminoácidos se modifican cambiando la carga (es decir, suprimiendo una carga positiva, agregando una carga negativa, o ambos, suprimiendo una carga positiva y agregando una carga negativa). La preparación de una variante de celulasa puede llevarse a cabo por cualquier medio conocido en el arte, incluyendo modificación química de aminoácidos, modificando una secuencia de DNA que codifica la proteína nativa, transformación de la secuencia de DNA modificado en un hospedero apropiado, y expresión de la secuencia de DNA modificado para formar una enzima variante. La celulasa variante de la divulgación incluye péptidos que comprenden secuencias de aminoácidos alteradas en comparación con una secuencia de aminoácidos de la enzima precursora en donde la celulasa variante retiene la naturaleza celulolítica característica de la enzima precursora, pero que puede tener propiedades alteradas en algún aspecto específico. Por ejemplo, una celulasa variante puede tener un pH óptimo o una mayor temperatura o estabilidad oxidativa o afinidad disminuída o unión a materiales no celulósicos, pero conservará su actividad celulolítica característica. Se contempla que las variantes de acuerdo con la presente divulgación pueden derivarse de un fragmento de DNA que codifica una variante de celulasa en donde se conserva la actividad funcional de la variante de celulasa expresada. Por ejemplo, un fragmento de DNA que codifica una celulasa puede incluir además una secuencia de DNA o porción de la misma que codifica una región flexible o conector adherido a la secuencia de DNA de la celulasa en uno cualquiera de los extremos 5' o 3' en donde se conserva la actividad funcional del dominio de la celulasa codificada. Los términos variante y derivado pueden usarse indistintamente en esta especificación.

[34] "Residuos equivalentes" pueden definirse igualmente determinando homología en el nivel de la estructura terciaria para una celulasa precursora cuya estructura terciaria se ha determinado por cristalografía de rayos X. Residuos equivalentes se definen como aquellos para los cuales las coordenadas atómicas de dos o más de los átomos de la cadena principal de un residuo aminoácido particular de una celulasa e *Hypocrea jecorina* CBH2 (N en N, CA en CA, C en C y O en O) están en 0.13 nm y, preferiblemente, 0.1 nm después del alineamiento. El alineamiento se obtiene después de que se ha orientado y posicionado el mejor modelo para producir la superposición máxima de las coordenadas atómicas de

átomos de proteínas distintos de hidrógeno de la celulasa en cuestión con respecto a la *H. jecorina* CBH2. El mejor modelo es el modelo cristalográfico que produce el factor R mínimo para datos de difracción experimentales con la máxima resolución disponible, véase por ejemplo US 2006/0205042.

[35] Residuos equivalentes, que son funcionalmente análogos a un residuo específico de *H. jecorina* CBH2, se definen como aquellos aminoácidos de una celulasa que pueden adoptar una conformación tal que puedan alterar, modificar o contribuir a la estructura de la proteína, unión al sustrato o catálisis en una manera definida y atribuida a un residuo específico de la *H. jecorina* CBH2. Además, son aquellos residuos de la celulasa (para la cual se ha obtenido una estructura terciaria por cristalografía de rayos X) los cuales ocupan una posición análoga en la medida que, aunque los átomos de la cadena principal de un residuo dado pueden no satisfacer los criterios de equivalencia sobre la base de ocupar una posición homóloga, las coordenadas atómicas de por lo menos dos de los átomos de la cadena lateral del residuo están a 0.13 nm de los átomos de la cadena lateral correspondiente de *H. jecorina* CBH2. La estructura cristalina de *H. jecorina* CBH2 se muestra en Zou y col. (1999) (Ref. 5, supra).

[36] El término "molécula de ácido nucleico" incluye moléculas de RNA, DNA y cDNA. Se entenderá que, como resultado de la degeneración del código genético, puede producirse una multitud de secuencias nucleotídicas que codifican una proteína dada tal como CBH2 y/o variantes de la misma.

La presente divulgación contempla toda secuencia nucleotídica de una variante posible, que codifica la celulasa variante tal como CBH2, todas las cuales son posibles, dada la degeneración del código genético.

[37] Un constructo de ácido nucleico "heterólogo" o secuencia tiene una porción de la secuencia que no es nativa a la célula en la cual se expresa. Heterólogo, con respecto a una secuencia de control se refiere a una secuencia de control (es decir, promotora o potenciadora) que no funciona en la naturaleza para regular el mismo gen cuya expresión está regulando actualmente. Por lo general, secuencias de ácido nucleico heterólogo son no endógenas a la célula o parte del genoma en el cual están presentes, y se han añadido a la célula, por infección, transfección, transformación, microinyección, electroporación, o similares. Un constructo de ácido nucleico "heterólogo" puede contener una combinación de secuencia de control/secuencia codificante de DNA que es la misma que, o diferente de, una combinación de secuencia de control/secuencia codificante de

DNA hallada en la célula nativa.

[38] Tal como se emplea en esta especificación, el término "vector" se refiere a un constructo de ácido nucleico diseñado para transferencia entre diferentes células hospederas. Un "vector de expresión" se refiere a un vector que tiene la capacidad de incorporar y expresar fragmentos de DNA heterólogo en una célula extraña. En el comercio se dispone de muchos vectores de expresión procarióticos y eucarióticos. La selección de los vectores de expresión apropiados está dentro del conocimiento de los expertos en el arte.

[39] Por consiguiente, un "casete de expresión" o "vector de expresión" es un constructo de ácido nucleico generado recombinante o sintéticamente, con una serie de elementos especificados de ácido nucleico que permiten la transcripción de un ácido nucleico particular en una célula diana. El casete de expresión recombinante puede incorporarse en un DNA plasmídico, cromosómico, mitocondrial, plastídico, virus, o fragmento de ácido nucleico. Normalmente, la porción del casete de expresión recombinante de un vector de expresión incluye, entre otras secuencias, una secuencia de ácido nucleico que se va a transcribir y un promotor.

[40] Tal como se emplea en esta especificación, el término "plásmido" se refiere a un constructo de DNA de hebra doble (ds - *double-stranded*) usado como un vector de clonación, y el cual forma un elemento genético auto-replicante en muchas bacterias y algunas eucariotas.

[41] Tal como se emplea en esta especificación, el término "secuencia nucleotídica que codifica un marcador seleccionable" se refiere a una secuencia nucleotídica 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 la presencia de un agente selectivo correspondiente, o bajo condiciones correspondientes de crecimiento selectivo.

[42] Tal como se emplea en esta especificación, el término "promotor" se refiere a una secuencia de ácido nucleico que funciona para dirigir la transcripción de un gen hacia 3' ("**a downstream gene**"). Generalmente el promotor será apropiado a la célula hospedera en la cual se está expresando el gen diana. El promotor, junto con otras secuencias de ácido nucleico reguladoras transcripcionales y traduccionales (denominadas asimismo "secuencias de control"), son necesarios para expresar un gen dado. En general, las secuencias reguladoras

transcripcionales y traduccionales incluyen, pero no se limitan a, secuencias promotoras, sitios de unión ribosómica, secuencias transcripcionales de iniciación y terminación, secuencias traduccionales de iniciación y terminación, y secuencias potenciadoras o activadoras.

[43] "Gen quimérico" o "constructo de ácido nucleico heterólogo", como se define en esta especificación, se refiere a un gen no nativo (es decir, uno que se ha introducido en un hospedero), que puede estar compuesto de partes de diferentes genes, incluyendo elementos reguladores. Un constructo de gen quimérico para transformación de una célula hospedera se compone normalmente de una región reguladora transcripcional (promotor) enlazada operacionalmente a una secuencia codificante de una proteína heteróloga, o, en un gen quimérico marcador seleccionable, a un gen marcador seleccionable que codifica una proteína que confiere, por ejemplo, resistencia antibiótica a células transformadas. Un gen quimérico típico de la presente divulgación, para transformación en una célula hospedera, incluye una región reguladora transcripcional que es constitutiva o inducible, una secuencia codificante de una proteína, y una secuencia terminadora. Un constructo de un gen quimérico puede incluir además una segunda secuencia de DNA que codifica un péptido señal si se desea la secreción de la proteína diana.

[44] Un ácido nucleico está "operacionalmente enlazado" cuando se pone en relación funcional con otra secuencia de ácido nucleico. Por ejemplo, DNA que codifica un líder secretor está operacionalmente enlazado a DNA para un polipéptido si se expresa como una pre-proteína que participa en la secreción del polipéptido; un promotor o potenciador está operacionalmente enlazado a una secuencia codificante si afecta la transcripción de la secuencia; o un sitio de unión a ribosoma está operacionalmente enlazado a una secuencia codificante si está posicionado de tal modo que facilite la traducción. Generalmente, "operacionalmente enlazado" significa que las secuencias de DNA que se están enlazando son contiguas, y, en el caso de un líder secretor, contiguas y en el marco de lectura. No obstante, los potenciadores no tienen que estar contiguos. El enlazamiento se lleva a cabo por ligación en sitios de restricción convenientes. Si tales sitios no existen, se usan de acuerdo con la práctica convencional los adaptadores oligonucleotídicos sintéticos, conectores o cebadores para PCR.

[45] Tal como se emplea en esta especificación, el término "gen" significa el segmento de DNA involucrado en producir una cadena polipeptídica, que puede o no puede incluir regiones que precedan y sigan a la región codificante, v.g.

secuencias 5' no traducidas (5' UTR) o secuencias "líder" y secuencias 3' UTR o secuencias "remolque", así como secuencias intervinientes (intrones) entre segmentos codificantes individuales (exones).

[46] En general, moléculas de ácido nucleico, que codifican la celulasa variante tal como CBH2, se hibridarán, bajo condiciones de astringencia moderada a alta, a la secuencia natural tal como la provista en esta especificación como SEQ ID NO:1. No obstante, en algunos casos, se emplea una secuencia nucleotídica codificante de CBH2 que posea un empleo de codón sustancialmente diferente, mientras que la proteína codificada por la secuencia nucleotídica codificante de CBH2 tiene la misma o sustancialmente la misma secuencia de aminoácidos que la proteína nativa. Por ejemplo, la secuencia codificante puede modificarse para facilitar una expresión más rápida de CBH2 en un sistema particular de expresión procariota o eucariota, de acuerdo con la frecuencia con la cual un codón particular se utiliza por el hospedero (Te'o y col., FEMS Microbiology Letters, 190: 13-19, 2000, por ejemplo, describe la optimización de genes para expresión en hongos filamentosos).

[47] Se considera que una secuencia de ácido nucleico es "hibrizable selectivamente" a una secuencia de ácido nucleico de referencia si las dos secuencias se hibridan específicamente una a otra bajo condiciones de astringencia moderada a alta y condiciones de lavado. Las condiciones de hibridación se basan en la temperatura de fusión (T_m) del complejo de unión a ácido nucleico o sonda. Por ejemplo, la "astringencia máxima" ocurre normalmente a alrededor de T_m 5°C (5°C por debajo de la T_m de la sonda); "alta astringencia" a alrededor de 5-10°C por debajo de la T_m ; "astringencia moderada" o "intermedia" a alrededor de 10-20°C por debajo de la T_m de la sonda; y "baja astringencia" a alrededor de 20-25°C por debajo de la T_m . Funcionalmente, pueden usarse condiciones de máxima astringencia para identificar secuencias que tienen una identidad estricta o una identidad casi estricta con la sonda de hibridización; mientras que las condiciones de alta astringencia se usan para identificar secuencias que tienen alrededor de 80% o más de identidad de secuencia con la sonda.

[48] Condiciones de hibridación de astringencia moderada a alta son bien conocidas en el arte (véase, por ejemplo, Sambrook, y col., 1989, Capítulos 9 y 11, y en Ausubel, F. M., y col., 1993, expresamente incorporados por referencia en esta especificación). Un ejemplo de condiciones de alta astringencia incluye hibridación a alrededor de 42°C en 50% formamida, 5xSSC, 5x solución de

Denhardt, 0.5% SDS y 100 ug/ml de DNA transportador desnaturalizado seguido por lavado dos veces en 2xSSC y 0.5% SDS a temperatura ambiente y dos veces más en 0.1x SSC y 0.5% SDS a 42°C.

[49] El término "recombinante" cuando se emplea con referencia, v.g., a 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 heterólogo o proteína o la alteración de un ácido nucleico nativo o proteína, o que la célula se deriva de una célula así modificada. Por consiguiente, por ejemplo, células recombinantes expresan genes que no se encuentran en la forma nativa (no recombinante) de la célula o expresan genes nativos que, de otro modo, se expresan anormalmente, se subexpresan o no se expresan en absoluto.

[50] Tal como se emplea en esta especificación, los términos "transformado", "establemente transformado " o "transgénico" con referencia a una célula significan que la célula tiene una secuencia de ácido nucleico no nativo (heterólogo) integrada en su genoma o como un plásmido episómico que se mantiene a través de múltiples generaciones.

[51] Tal como se emplea en esta especificación, 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 la transcripción como la traducción.

[52] El término "introducido" en el contexto de insertar 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, donde la secuencia de ácido nucleico puede incorporarse en el genoma de la célula (por ejemplo, DNA cromosómico, plasmídico, plástido, o mitocondrial), convertirse en un replicón autónomo, o expresarse transitoriamente (por ejemplo, mRNA transfectado).

[53] Se sigue que el término "expresión de CBH2" se refiere a la transcripción y traducción del gen *cbh2* o variantes del mismo, cuyos productos incluyen RNA precursor, mRNA, polipéptido, polipéptidos procesados pos-traduccionalmente, y derivados de los mismos, incluyendo 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 análisis para la expresión de CBH2 incluyen

Western blot para la proteína CBH2, análisis Northern blot y análisis de reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR) para *cbh2* mRNA, y análisis de Celulosa Esponjada con Ácido Fosfórico (PASC - *Phosphoric Acid Swollen Cellulose*) y análisis de hidrazida de ácido p-hidrobenczoico (PAHBAH - *p-hydroxybenzoic acid hydrazide*) como se describe en los siguientes: (a) PASC: (Karlsson, J. y col., (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.), páginas 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).

[54] El término "empalme alternativo" se refiere al proceso mediante el cual múltiples isoformas polipeptídicas se generan de un único gen, e implica el corte y empalme conjunto de exones no consecutivos durante el procesamiento de algunos, mas no todos, los transcritos del gen. Por consiguiente un exón particular puede conectarse a uno cualquiera de varios exones alternativos para formar RNAs mensajeros. Los mRNAs alternativamente empalmados producen polipéptidos ("variantes de empalme") en las cuales algunas partes son comunes mientras que otras partes son diferentes.

[55] El término "secuencia señal" se refiere a una secuencia de aminoácidos en la porción 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.

[56] Por el término "célula hospedera" se da a entender una célula que contiene un vector y apoya la replicación, y/o transcripción o la transcripción y traducción (expresión) del constructo de expresión. Células hospederas para uso en la presente divulgación pueden ser células procariotas, tales como células de *E. coli*, o células eucariotas tales como células de levadura, células vegetales, células de insectos, células anfibias, o células mamíferas. En general, las células hospederas son hongos filamentosos.

[57] El término "hongos filamentosos" significa cualquiera y todos los hongos filamentosos reconocidos por aquellas personas de habilidad en el arte. Un hongo preferido se selecciona del grupo conformado por *Aspergillus*, *Trichoderma*, *Fusarium*, *Chrysosporium*, *Penicillium*, *Humicola*, *Neurospora*, o formas sexuales alternativas de los mismos tales como *Emericella*, *Hypocrea*. Actualmente se ha

demostrado que el hongo industrial asexual *Trichoderma reesei* es un derivado clonal del ascomicete *Hypocrea jecorina* (véase, Kuhls y col., PNAS, 93:7755-7760, 1996).

[58] El término "celooligosacárido" se refiere a grupos oligosacáridos que contienen desde 2-8 unidades de glucosa y tienen enlaces beta-1,4, v.g., celobiosa.

[59] Los términos "celulasa", "enzimas celulolíticas" o "enzimas de celulasa" se refieren a una categoría de enzimas capaces de hidrolizar polímeros de celulosa en oligómeros celooligosacáridos más cortos, celobiosa y/o glucosa. Numerosos ejemplos de celulasas, tales como exoglucanasas, exocelobiohidrolasas, endoglucanasas, y glucosidasas se han obtenido de organismos celulolíticos, incluyendo particularmente hongos, plantas y bacterias. Las enzimas elaboradas 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 diferentes tipos de enzimas de celulasa actúan sinérgicamente para convertir celulosa y sus derivados en glucosa.

[60] Muchos microbios elaboran enzimas que hidrolizan celulosa, incluyendo el hongo de pudrición de la madera *Trichoderma*, las bacterias del compost *Thermomonospora*, *Bacillus*, y *Cellulomonas*; *Streptomyces*; y los hongos *Humicola*, *Aspergillus* y *Fusarium*.

[61] *Hypocrea jecorina* CBH2 es un miembro de la Familia 6 Glicosil Hidrolasa (de ahí Cel6) y, específicamente, fue el primer miembro de esa familia identificado en *Hypocrea jecorina* (de ahí Cel6A). La Familia 6 Glicosil Hidrolasa contiene tanto endoglucanasas y celobiohidrolasas/exoglucanasas, y esa CBH2 es la última. Por consiguiente, las expresiones CBH2, proteína tipo CBH2 y Cel6 celobiohidrolasas pueden usarse indistintamente en esta especificación.

[62] El término "dominio de unión a celulosa", tal como se emplea en esta especificación, se refiere a una porción de la secuencia de aminoácidos de una celulasa o a una región de la enzima que está involucrada en la actividad de unión a celulosa de una celulasa o derivado de la misma. Dominios de unión a celulosa generalmente funcionan uniendo no covalentemente la celulasa a la celulosa, un derivado de celulosa u otro polisacárido equivalente de la misma. Dominios de unión a celulosa permiten o facilitan la hidrólisis de las fibras de celulosa por la

región central catalítica estructuralmente distinta, y normalmente funcionan independientes del centro catalítico. Por consiguiente, un dominio de unión a celulosa no poseerá la actividad hidrolítica significativa atribuible a una parte central catalítica. En otras palabras, un dominio de unión a celulosa es un elemento estructural de la estructura terciaria proteica de la enzima de celulasa, que es distinto del elemento estructural que posee actividad catalítica. Dominio de unión a celulosa y módulo de unión a celulosa pueden emplearse indistintamente en esta especificación.

[63] Tal como se emplea en esta especificación, el término "tensioactivo" se refiere a cualquier compuesto generalmente reconocido en el arte como poseedor de cualidades tensioactivas. Así, por ejemplo, los tensioactivos comprenden tensioactivos aniónicos, catiónicos y no iónicos tales como los hallados comúnmente en detergentes. Los tensioactivos aniónicos incluyen alquilbencenosulfonatos lineales o ramificados; alquil o alquenil éter sulfatos con grupos alquilo lineales o ramificados o grupos alquenilo; alquil o alquenil sulfatos; olefinsulfonatos; y alcanosulfonatos. Tensioactivos anfólicos incluyen sulfonatos de sales de amonio cuaternario, y tensioactivos anfólicos tipo betaína. Tales tensioactivos anfólicos tienen tanto los grupos cargados positivos como los negativos en la misma molécula. Los tensioactivos no iónicos pueden incluir éteres de polioxialquileo, así como alcanolamidas de ácido graso más altas o su aducto de óxido de alquileo, monoésteres de glicerina de ácido graso, y similares.

[64] Tal como se emplea en esta especificación, el término "tela con contenido de celulosa" se refiere a cualesquiera telas cosidas o no cosidas, hilos o fibras hechos de celulosa con contenido de algodón o sin contenido de algodón o mezclas de celulosa con contenido de algodón o sin contenido de algodón incluyendo celulósicos naturales y celulósicos sintéticos (tales como yute, lino, ramio, rayón, y lyocell).

[65] Tal como se emplea en esta especificación, el término "tela con contenido de algodón" se refiere a telas cosidas o no cosidas, hilos o fibras hechos de solo algodón o mezclas de algodón incluyendo telas tejidas de algodón, tejidos de punto de algodón, telas de algodón para jeans, hilos de algodón, algodón en bruto y similares.

[66] Tal como se emplea en esta especificación, el término "composición de lavado a la piedra" se refiere a una formulación para uso en lavar a la piedra telas con contenido de celulosa. Las composiciones de lavado a la piedra se utilizan

para modificar telas con contenido de celulosa antes de la venta, es decir, durante el proceso de fabricación. En contraste, las composiciones detergentes están previstas para la limpieza de prendas sucias y no se emplean durante el proceso de fabricación.

[67] Tal como se emplea en esta especificación, el término "composición detergente" se refiere a una mezcla prevista para uso en un medio de lavado para lavar y blanquear telas sucias que contienen celulosa. En el contexto de la presente divulgación, tales composiciones pueden incluir, además de celulasas y tensioactivos, enzimas hidrolíticas adicionales, mejoradores de detergente, agentes blanqueadores, activadores de blanqueado, agentes azulantes y colorantes fluorescentes, inhibidores de formación de grumos, agentes enmascarantes, activadores de celulasa, antioxidantes, y solubilizantes.

[68] Tal como se emplea en esta especificación, el término "disminución o eliminación en la expresión del gen *cbh2*" significa que o bien que el gen *cbh2* se ha eliminado del genoma y, por consiguiente, no puede expresarse por el microorganismo hospedero recombinante; o bien que el gen *cbh2* o transcrito se ha modificado de tal modo que una enzima CBH2 funcional no se produce por el microorganismo hospedero o en niveles que son significativamente menores que el gen *cbh2* no modificado o transcrito.

[69] El término "gen *cbh2* variante" significa que la secuencia de ácido nucleico del gen *cbh2* de *H. jecorina* ha sido alterado retirando, agregando, y/o manipulando la secuencia codificante.

[70] Tal como se emplea en esta especificación, el término "purificar" se refiere generalmente a someter células, que contienen ácido nucleico transgénico o proteína, a la purificación bioquímica y/o cromatografía en columna.

[71] Tal como se emplea en esta especificación, los términos "activo" y "biológicamente activo" se refieren a una actividad biológica asociada con una proteína particular y se emplean indistintamente en esta especificación. Por ejemplo, la actividad enzimática asociada con una proteasa es proteólisis y, en consecuencia, una proteasa activa tiene actividad proteolítica. Se sigue que la actividad biológica de una proteína dada se refiere a cualquier actividad biológica atribuida normalmente a esa proteína por los expertos en el arte.

[72] Tal como se emplea en esta especificación, el término "enriquecido" significa

que la celulasa tal como CBH2 se encuentra en una concentración que es mayor en relación con la concentración de CBH2 hallada en una composición de celulasa fúngica, natural o de origen natural. Los términos enriquecido, elevado y mejorado pueden emplearse indistintamente en esta especificación.

[73] Una composición de celulasa fúngica natural es una producida por una fuente fúngica de origen natural y que consta de uno o más componentes BGL, CBH y EG en donde cada uno de estos componentes se halla en la relación producida por la fuente fúngica. En consecuencia, una composición CBH enriquecida tendría CBH en una relación alterada, en donde la relación de CBH a otros componentes de celulasa (es decir, EGs, beta-glucosidasas y otras endoglucanasas) es elevada. Esta relación puede aumentarse incrementando la CBH o disminuyendo (o eliminando) por lo menos otro componente por cualquier medio conocido en el arte.

[74] El término "aislado" o "purificado", tal como se emplea en esta especificación, se refiere a un ácido nucleico o aminoácido que se retira de por lo menos un componente con el cual está asociado naturalmente.

[75] Así, para ilustrar, un sistema de celulasas de origen natural puede purificarse en componentes sustancialmente puros por técnicas reconocidas de separación publicadas adecuadamente en la literatura, incluyendo cromatografía de intercambio iónico con un pH apropiado, cromatografía de afinidad, por exclusión de tamaño y similares. Por ejemplo, en la cromatografía de intercambio iónico (usualmente cromatografía de intercambio aniónico), es posible separar los componentes de celulasa eluyendo con un gradiente de pH, o un gradiente salino, o ambos, un gradiente de pH y un gradiente salino. La CBH purificada puede agregarse entonces a la solución enzimática resultando en una solución de CBH enriquecida. Es igualmente posible elevar la cantidad de CBH producida por un microbio usando métodos de genética molecular para sobreexpresar el gen que codifica CBH, posiblemente en conjunción con la eliminación de uno o más genes que codifican otras celulasas.

[76] Celulasas fúngicas pueden contener más de un componente CBH. Los diferentes componentes tienen generalmente diferentes puntos isoeléctricos los cuales permiten su separación mediante cromatografía de intercambio iónico, y similares. Un componente CBH único o una combinación de componentes CBH pueden emplearse en una solución enzimática.

[77] Cuando se emplea en soluciones enzimáticas, el componente CBH2 variante u homólogo se agrega generalmente en una cantidad suficiente para permitir la tasa más alta de liberación de azúcares solubles de la biomasa. La cantidad de componente CBH2 variante u homólogo agregado depende del tipo de biomasa que se va a sacarificar, la cual puede determinarse fácilmente por el artesano experto. Cuando se emplea, el porcentaje en peso del componente CBH2 variante u homólogo presente en la composición de celulasa es desde preferiblemente entre 1 y 100, siendo ejemplos ilustrativos alrededor de 1, preferiblemente alrededor de 5, preferiblemente alrededor de 10, preferiblemente alrededor de 15, o preferiblemente alrededor de 20 por ciento en peso a preferiblemente alrededor de 25, preferiblemente alrededor de 30, preferiblemente alrededor de 35, preferiblemente alrededor de 40, preferiblemente alrededor de 45 o preferiblemente alrededor de 50 por ciento en peso. Más aún, intervalos preferidos pueden ser alrededor de 0.5 a alrededor de 15 por ciento en peso, alrededor de 0.5 a alrededor de 20 por ciento en peso, desde alrededor de 1 a alrededor de 10 por ciento en peso, desde alrededor de 1 a alrededor de 15 por ciento en peso, desde alrededor de 1 a alrededor de 20 por ciento en peso, desde alrededor de 1 a alrededor de 25 por ciento en peso, desde alrededor de 5 a alrededor de 20 por ciento en peso, desde alrededor de 5 a alrededor de 25 por ciento en peso, desde alrededor de 5 a alrededor de 30 por ciento en peso, desde alrededor de 5 a alrededor de 35 por ciento en peso, desde alrededor de 5 a alrededor de 40 por ciento en peso, desde alrededor de 5 a alrededor de 45 por ciento en peso, desde alrededor de 5 a alrededor de 50 por ciento en peso, desde alrededor de 10 a alrededor de 20 por ciento en peso, desde alrededor de 10 a alrededor de 25 por ciento en peso, desde alrededor de 10 a alrededor de 30 por ciento en peso, desde alrededor de 10 a alrededor de 35 por ciento en peso, desde alrededor de 10 a alrededor de 40 por ciento en peso, desde alrededor de 10 a alrededor de 45 por ciento en peso, desde alrededor de 10 a alrededor de 50 por ciento en peso, desde alrededor de 15 a alrededor de 60 por ciento en peso, desde alrededor de 15 a alrededor de 65 por ciento en peso, desde alrededor de 15 a alrededor de 70 por ciento en peso, desde alrededor de 15 a alrededor de 75 por ciento en peso, desde alrededor de 15 a alrededor de 80 por ciento en peso, desde alrededor de 15 a alrededor de 85 por ciento en peso, desde alrededor de 15 a alrededor de 95 por ciento en peso. No obstante, cuando se emplea, el porcentaje en peso del componente de CBH2 variante u homólogo, relativo a cualesquiera componentes tipo EG presentes en la composición de celulasa es desde preferiblemente alrededor de 1, preferiblemente alrededor de 5, preferiblemente alrededor de 10, preferiblemente alrededor de 15, o preferiblemente alrededor de 20 por ciento en peso a preferiblemente alrededor de 25, preferiblemente alrededor de 30,

preferiblemente alrededor de 35, preferiblemente alrededor de 40, preferiblemente alrededor de 45 o preferiblemente alrededor de 50 por ciento en peso. Más aún, intervalos preferidos pueden ser alrededor de 0.5 a alrededor de 15 por ciento en peso, alrededor de 0.5 a alrededor de 20 por ciento en peso, desde alrededor de 1 a alrededor de 10 por ciento en peso, desde alrededor de 1 a alrededor de 15 por ciento en peso, desde alrededor de 1 a alrededor de 20 por ciento en peso, desde alrededor de 1 a alrededor de 25 por ciento en peso, desde alrededor de 5 a alrededor de 20 por ciento en peso, desde alrededor de 5 a alrededor de 25 por ciento en peso, desde alrededor de 5 a alrededor de 30 por ciento en peso, desde alrededor de 5 a alrededor de 35 por ciento en peso, desde alrededor de 5 a alrededor de 40 por ciento en peso, desde alrededor de 5 a alrededor de 45 por ciento en peso, desde alrededor de 5 a alrededor de 50 por ciento en peso, desde alrededor de 10 a alrededor de 20 por ciento en peso, desde alrededor de 10 a alrededor de 25 por ciento en peso, desde alrededor de 10 a alrededor de 30 por ciento en peso, desde alrededor de 10 a alrededor de 35 por ciento en peso, desde alrededor de 10 a alrededor de 40 por ciento en peso, desde alrededor de 10 a alrededor de 45 por ciento en peso, desde alrededor de 10 a alrededor de 50 por ciento en peso, desde alrededor de 15 a alrededor de 20 por ciento en peso, desde alrededor de 15 a alrededor de 25 por ciento en peso, desde alrededor de 15 a alrededor de 30 por ciento en peso, desde alrededor de 15 a alrededor de 35 por ciento en peso, desde alrededor de 15 a alrededor de 40 por ciento en peso, desde alrededor de 15 a alrededor de 45 por ciento en peso, desde alrededor de 15 a alrededor de 50 por ciento en peso.

II. Celulasas

[78] Las celulasas son conocidas en el arte como enzimas que hidrolizan celulosa (enlaces beta-1,4-glucano o beta D-glucosídicos) lo cual resulta en la formación de glucosa, celobiosa, celooligosacáridos, y similares. Como ya se expuso, las celulasas se han dividido tradicionalmente en tres clases principales: endoglucanasas (EC 3.2.1.4) ("EG"), exoglucanasas o celobiohidrolasas (EC 3.2.1.91) ("CBH") y beta-glucosidasas (EC 3.2.1.21) ("BG").

[79] Ciertos hongos producen sistemas completos de celulasas, los cuales incluyen exo-celobiohidrolasas o celulasas tipo CBH, endoglucanasas o celulasas tipo EG y beta-glucosidasas o celulasas tipo BG. Sin embargo, a veces estos sistemas carecen de celulasas tipo CBH, y las celulasas bacterianas incluyen además normalmente pocas o ningunas celulasas tipo CBH. Adicionalmente, se ha demostrado que los componentes EG y los componentes CBH interactúan

sinérgicamente para degradar la celulosa de manera más eficiente. Los diferentes componentes, es decir, las varias endoglucanasas y exocelobiohidrolasas en un sistema de celulasas multi-componente o completo, generalmente tienen diferentes propiedades, tales como punto isoelectrico, peso molecular, grado de glicosilación, especificidad del sustrato y patrones de acción enzimática.

[80] Se cree que las celulasas tipo endoglucanasa hidrolizan enlaces beta-1,4-glucosídicos internos en regiones de baja cristalinidad de la celulosa y que las celulasas tipo exo-celobiohidrolasa hidrolizan celobiosa del extremo reductor o no reductor de la celulosa. Se sigue que la acción de los componentes endoglucanasa puede facilitar en gran medida la acción de las exocelobiohidrolasas creando nuevos extremos de cadena, los cuales son reconocidos por componentes exo-celobiohidrolasa. Además, se ha demostrado que las celulasas 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 carbohidratos, tales como celobiosa. Esto produce glucosa como el solo producto para el microorganismo y reduce o elimina la celobiosa, la cual inhibe las celobiohidrolasas y las endoglucanasas.

[81] Las celulasas encuentran asimismo un sinnúmero de usos en composiciones detergentes incluyendo potenciar la capacidad limpiadora, como agente suavizante y para mejorar el tacto de las telas de algodón (Hemmpel, ITB Dyeing/Printing/Finishing 3:5-14, 1991; Tyndall, Textile Chemist and Colorist 24:23-26, 1992; y Kumar y col., Textile Chemist and Colorist, 29:37-42, 1997). Aunque el mecanismo no es parte de la divulgación, las propiedades de suavizado y restauración del color de la celulosa se ha atribuido a los componentes alcalinos de la endoglucanasa en composiciones de celulosa, como se ejemplifica por las Patentes de los Estados Unidos Nos. 5,648,263, 5,691,178, y 5,776,757, las cuales divulgan que composiciones detergentes que contienen una composición de celulosa enriquecida con un componente alcalino de endoglucanasa especificado imparten restauración del color y una suavidad mejorada a las prendas tratadas, en comparación con composiciones de celulosa no enriquecidas con tal componente. Además, se ha demostrado que el uso de tales componentes alcalinos de endoglucanasa en composiciones detergentes complementan los requerimientos de pH de la composición detergente (v.g., exhibiendo actividad máxima con un pH alcalino de 7.5 a 10, como se describe en las Patentes de los Estados Unidos Nos. 5,648,263, 5,691,178, y 5,776,757).

[82] Se ha demostrado igualmente que las composiciones de celulasa degradan telas que contienen algodón, con el resultado de una pérdida reducida de resistencia en la tela (Patente de los Estados Unidos No. 4,822,516), lo cual contribuye a la renuencia a usar composiciones de celulasa en aplicaciones detergentes comerciales. Se ha sugerido que las composiciones de celulasa que comprenden componentes de endoglucanasa exhiben una pérdida reducida de resistencia para telas con contenido de algodón, comparadas con composiciones que comprenden un sistema completo de celulasas.

[83] De igual modo, se ha demostrado que las celulasas son útiles en la degradación de biomasa de celulasa a etanol (en donde la celulasa degrada celulosa a glucosa, y levadura u otros microbios fermentan posteriormente la glucosa en etanol), en el tratamiento de pulpa mecánica (Pere y col., In Proc. Tappi Pulping Conf., Nashville, Tenn., 27-31, páginas 693-696, 1996), para uso como aditivo de forrajes (WO 91/04673) y en la molienda húmeda de granos.

[84] La mayoría de las CBHs y EGs tienen una estructura multi-dominio conformada por un dominio central, separado de un dominio de unión a celulosa (CBD – *cellulose binding domain*) por un péptido conector (Suurnakki y col., 2000). El dominio central contiene el sitio activo mientras que el CBD interactúa con celulosa uniéndola a él (van Tilbeurgh y col., FEBS Lett. 204:223-227, 1986; Tomme y col., Eur. J. Biochem. 170:575-581, 1988). Los CBDs son particularmente importantes en la hidrólisis de celulosa cristalina. Se ha demostrado que la capacidad de las celobiohidrolasas de degradar celulosa cristalina disminuye claramente cuando el CBD está ausente (Linder y Teeri, J. Biotechnol. 57:15-28, 1997). No obstante, el papel exacto y el mecanismo de acción de los CBDs es aún materia de especulación. Se ha sugerido que el CBD potencia la actividad enzimática aumentando simplemente la concentración enzimática efectiva en la superficie de la celulosa (Stahlberg y col., Bio/Technol. 9:286-290, 1991), y/o aflojando cadenas simples de celulosa de la superficie de la celulosa (Torino y col., EMBO J. vol. 15, no. 21, páginas 5739-5751, 1996). La mayoría de los estudios concernientes a los efectos de los dominios de celulasa sobre diferentes sustratos se han llevado a cabo con proteínas centrales ("**core proteins**") de celobiohidrolasas, ya que sus proteínas centrales pueden producirse fácilmente por proteólisis limitada con papaína (Tomme y col., 1988). Numerosas celulasas se han descrito en la literatura científica, ejemplos de las cuales incluyen: de *Trichoderma reesei*: Shoemaker, S. y col., Bio/Technology, 1:691-696, 1983, que divulga CBH1; Teeri, T. y col., Gene, 51:43-52, 1987, que divulga CBH2. Asimismo se han descrito celulasas de especies distintas a

Trichoderma v.g., Ooi y col., Nucleic Acids Research, vol. 18, No. 19, 1990, que divulga la secuencia de cDNA que codifica la endoglucanasa F1-CMC producida por *Aspergillus aculeatus*; Kawaguchi T y col., Gene 173(2):287-8, 1996, que divulga la clonación y secuenciación del cDNA que codifica beta-glucosidasa 1 de *Aspergillus aculeatus*; Sakamoto y col., Curr. Genet. 27:435-439, 1995, que divulga la secuencia de cDNA que codifica la endoglucanasa CMCase-1 de *Aspergillus kawachii* IFO 4308; Saarilahti y col., Gene 90:9-14, 1990, que divulga una endoglucanasa de *Erwinia carotovora*; Spilliaert R, y col., Eur J Biochem. 224(3):923-30, 1994, que divulga la clonación y secuenciación de bg1A, que codifica una beta-glucanasa termoestable de *Rhodothermus marinus*; y Halldorsdottir S y col., Appl Microbiol Biotechnol. 49(3):277-84, 1998, que divulga la clonación, secuenciación y sobreexpresión de un gen *Rhodothermus marinus* que codifica una celulasa termoestable de la familia 12 glicosil hidrolasa. Sin embargo, subsiste una necesidad de identificación y caracterización de celulasas novedosas, con propiedades mejoradas, tales como un mayor rendimiento bajo condiciones de estrés térmico o en la presencia de tensioactivos, una mayor actividad específica, un patrón alterado de escisión del sustrato, y/o alto nivel de expresión *in vitro*.

[85] El desarrollo de composiciones de celulasa novedosas y mejoradas que contienen cantidades variables de celulasas tipo CBH, tipo EG y tipo BG es de interés para uso: (1) en composiciones para degradar pulpa de madera u otra biomasa en azúcares (v.g., para la producción de bioquímicos tales como biocombustibles); (2) en composiciones detergentes que exhiben una capacidad limpiadora mejorada (3) funcionan como agente suavizante y/o mejoran el tacto de telas de algodón (v.g., "lavado a la piedra" o "biodesfibrilación") ("**stone washing**" or "**biopolishing**"); y/o (3) en composiciones para forrajes, por ejemplo.

[86] También se proveen en esta especificación preparaciones de celulasa entera que comprenden variantes de celulasa. Tal como se emplea en esta especificación, la expresión "preparación de celulasa entera" se refiere a composiciones que contienen celulasa de origen natural y 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 tipo celobiohidrolasa, uno o más componentes tipo endoglucanasa, y uno o más componentes tipo beta-glucosidasa, en donde cada uno de estos componentes se encuentra en la relación producida por la fuente. Una composición de origen natural es una que es producida por un organismo no modificado con respecto a las enzimas celulolíticas, de modo que la relación de las enzimas componentes permanece

inalterada de las producidas por el organismo natural. Una composición "de origen no natural" abarca aquellas composiciones producidas: (1) combinando enzimas celulolíticas componentes en una relación de origen natural o en una relación de origen no natural, es decir, alterada; o (2) modificando un organismo para sobreexpresar o subexpresar una o más enzimas celulolíticas; o (3) modificando un organismo de modo que por lo menos se elimine una enzima celulolítica. Por consiguiente, en algunas incorporaciones, la preparación de celulasa entera puede tener eliminadas una o más de las diferentes EGs y/o CBHs, y/o beta-glucosidasa. Por ejemplo, EG1 puede eliminarse sola o en combinación con otras EGs y/o CBHs.

[87] En general, la preparación de celulasa entera contiene enzimas incluyendo, pero no limitadas a: (i) endoglucanasas (EG) o 1,4- β -d-glucan-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 -d-glucano celobiohidrolasas (exo-celobiohidrolasas, CBH) (EC 3.2.1.91), y (iii) β -glucosidasa (BG) o β -glucósido glucohidrolasas (EC 3.2.1.21).

[88] En la presente divulgación, la preparación de celulasa entera puede ser de cualquier microorganismo que sea útil para la hidrólisis de un material celulósico. En algunas incorporaciones, la preparación de celulasa entera es una celulasa entera de hongos filamentosos. "Hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión *Eumycota* y *Oomycota*. En algunas incorporaciones, la preparación de celulasa entera es una celulasa entera de la especie *Acremonium*, *Aspergillus*, *Emericella*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Scytalidium*, *Thielavia*, *Tolypocladium*, o *Trichoderma*. En algunas incorporaciones, 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 de

Trichoderma harzianum, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* v.g., RL-P37 (Sheir-Neiss y col., Appl. Microbiol. Biotechnology, 20 (1984) páginas 46-53; Montenecourt B.S., Can., 1-20, 1987), QM9414 (ATCC No. 26921), NRRL 15709, ATCC 13631, 56764, 56466, 56767, o *Trichoderma viride* v.g., ATCC 32098 y 32086. En algunas incorporaciones, la preparación de celulasa entera es una celulasa entera *Trichoderma reesei* RutC30, la cual se encuentra disponible de la American Type Culture Collection como *Trichoderma reesei* ATCC 56765.

[89] Ejemplos de preparaciones comerciales de celulasa apropiadas para uso en la presente divulgación incluyen, por ejemplo, CELLUCLAST™ (disponible de Novozymes A/S) y la enzima LAMINEX™, IndiAge™ y Primafast™ LAMINEX BG, ACCELLERASE™ 100 y ACCELLERASE™ 1500 (disponibles de Genencor Division, Danisco US. Inc.)

[90] En la presente divulgación, la preparación de celulasa entera puede provenir de cualquier método de cultivo de microorganismos conocido en el arte, resultante en la expresión de enzimas capaces de hidrolizar un material celulósico. La fermentación puede incluir cultivo en frascos de agitación, fermentación a pequeña o gran escala, tal como fermentaciones continuas, en lote, en lote alimentado o fermentaciones en estado sólido en fermentadores de laboratorio o fermentadores industriales, realizadas en un medio apropiado y bajo condiciones que permitan expresar o aislar la celulasa.

[91] Generalmente, el microorganismo se cultiva en un medio de cultivo celular apropiado para la producción de enzimas capaces de hidrolizar un material celulósico. El cultivo tiene lugar en un nutriente apropiado constituido por fuentes de carbono y nitrógeno y sales inorgánicas, usando procedimientos conocidos en el arte. Medios de cultivo apropiados, intervalos de temperatura y otras condiciones apropiadas para crecimiento y producción de celulasa son conocidos en el arte. Como ejemplo no limitante, el intervalo de temperatura normal para la producción de celulasas por *Trichoderma reesei* es 24°C a 28°C.

[92] Generalmente, la preparación de celulasa entera se usa como se produce por fermentación sin ninguna o mínima recuperación y/o purificación. Por ejemplo, una vez que las celulasas son secretadas por una célula en el medio de cultivo celular, puede usarse el medio de cultivo celular que contiene las celulasas. En algunas incorporaciones la preparación de celulasa entera comprende el contenido no fraccionado de material de fermentación, incluyendo medio de cultivo celular,

enzimas extracelulares y células. Alternativamente, la preparación de celulasa entera puede procesarse por cualquier método conveniente, v.g., por precipitación, centrifugación, afinidad, filtración o cualquier otro método conocido en el arte. En algunas incorporaciones, la preparación de celulasa entera puede concentrarse, por ejemplo, y luego utilizarse sin purificación posterior. En algunas incorporaciones la preparación de celulasa entera contiene agentes químicos que disminuyen la viabilidad celular o eliminan las células. En algunas incorporaciones, las células se lisan o permeabilizan usando métodos conocidos en el arte.

III. Biología Molecular

[93] En una incorporación esta divulgación posibilita la expresión de genes *cbh2* variantes bajo el control de un promotor funcional en un hongo filamentoso. Por consiguiente, esta divulgación se basa en técnicas rutinarias en el campo de la genética recombinante (véase, v.g., Sambrook y col., *Molecular Cloning, A Laboratory Manual*, 2nd ed., 1989; Kriegler, *Gene Transfer and Expression: A Laboratory Manual*, 1990; y Ausubel y col., eds., *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, Greene Publishing and Wiley-Interscience, New York, 1994).

Métodos o de Mutar Secuencias de Ácido Nucleico *cbh2*

[94] Cualquier método conocido en el arte que pueda introducir mutaciones está contemplado por la presente divulgación.

[95] La presente divulgación se refiere a la expresión, purificación y/o aislamiento y uso de la variante CBH2. Estas enzimas se preparan preferiblemente por métodos recombinantes utilizando el gen *cbh2* de *H. jecorina*. El caldo de fermentación puede usarse con o sin purificación.

[96] Luego del aislamiento y clonación del gen *cbh2* de *H. jecorina*, se emplean otros métodos conocidos en el arte, tales como mutagénesis dirigida al sitio, para hacer las sustituciones, adiciones o eliminaciones que correspondan a los aminoácidos sustituidos en la variante CBH2 expresada. Nuevamente, mutagénesis dirigida al sitio y otros métodos de incorporar cambios de aminoácidos en proteínas expresadas en el nivel DNA son conocidos en el arte (Sambrook y col., *supra*; y Ausubel y col., *supra*).

[97] DNA que codifica una variante de secuencia de aminoácidos de la *H. jecorina*

CBH2 se prepara por una variedad de métodos conocidos en el arte. Estos métodos incluyen, pero no se limitan a, la preparación por mutagénesis dirigida al sitio (o mediada por oligonucleótidos), mutagénesis por PCR, y mutagénesis de casete de un DNA preparado anteriormente que codifica la *H. jecorina* CBH2.

[98] La mutagénesis dirigida al sitio es un método preferido para preparar variantes de sustitución. Esta técnica es bien conocida en el arte (véase, v.g., Carter y col., *Nucleic Acids Res.* 13:4431-4443 (1985) y Kunkel y col., *Proc. Natl. Acad. Sci. USA* 82:488 (1987)). En síntesis, cuando se lleva a cabo mutagénesis de DNA dirigida al sitio, el DNA inicial se altera hibridando primeramente un oligonucleótido, que codifica la mutación deseada, a una sola hebra de tal DNA inicial. Luego de la hibridación, se emplea una DNA polimerasa para sintetizar una segunda hebra completa, usando el oligonucleótido hibridado como cebador, y usando la hebra única del DNA inicial como molde. Por consiguiente, el oligonucleótido que codifica la mutación deseada se incorpora en el DNA de doble hebra resultante.

[99] Las mutagénesis por PCR es igualmente apropiada para elaborar variantes de secuencias de aminoácidos del polipéptido inicial, es decir, *H. jecorina* CBH2. Véase Higuchi, en *PCR Protocols*, páginas 177-183 (Academic Press, 1990); y Vallette y col., *Nuc. Acids Res.* 17:723-733 (1989). Véase, asimismo, por ejemplo Cadwell y col., *PCR Methods and Applications*, Vol 2, 28-33 (1992). En síntesis, cuando pequeñas cantidades de DNA molde se utilizan como material de partida en una PCR, pueden emplearse cebadores que difieren ligeramente en secuencia de la región correspondiente en un DNA molde para generar cantidades relativamente grandes de un fragmento de DNA específico, que difiere de la secuencia molde únicamente en las posiciones en las que los cebadores difieren del molde.

[100] Otro método para preparar variantes, la mutagénesis de casete, se basa en la técnica descrita por Wells y col., *Gene* 34:315-323 (1985). El material de partida es el plásmido (u otro vector) que contiene el DNA polipeptídico inicial que se va a mutar. Se identifica(n) el codón (los codones) en el DNA de partida que se va a mutar. Tiene que haber un sitio único de endonucleasa de restricción en cada lado del sitio(s) de mutación identificado(s). Si no existen tales sitios de restricción, éstos pueden generarse usando el método ya descrito de mutagénesis mediada por oligonucleótido para introducirlos en localizaciones apropiadas en el DNA polipeptídico inicial. El DNA plasmídico se corta en estos sitios para linearizarlo. Un oligonucleótido de doble hebra, que codifica la secuencia del DNA entre los

sitios de restricción pero contiene la(s) mutación (mutaciones) deseada(s), se sintetiza usando procedimientos estándar, en donde las dos hebras del oligonucleótido se sintetizan separadamente y luego se hibridan entre sí empleando técnicas estándar. Este oligonucleótido de doble hebra se denomina el casete. Este casete se diseña para tener extremos 5' y 3', que son compatibles con los extremos del plásmido linearizado, de modo que pueda ligarse directamente al plásmido. Este plásmido contiene ahora la secuencia de DNA mutado.

[101] Alternativa, o adicionalmente, puede determinarse la secuencia de aminoácidos deseada que codifica una variante CBH2, y una secuencia de ácido nucleico que codifica tal secuencia de aminoácidos variante puede generarse sintéticamente.

[102] La (las) variante(s) CBH2(s) así preparada(s) puede(n) someterse a modificaciones adicionales, dependiendo a menudo del uso previsto de la celulasas. Tales modificaciones pueden involucrar alteración adicional de la secuencia de aminoácidos, fusión a un polipéptido(s) heterólogo(s) y/o modificaciones covalentes.

IV. Ácidos Nucleicos cbh2 y Polipéptidos CBH2

A. Ácidos Nucleicos de Tipo cbh2 Variante

[103] La secuencia de ácido nucleico para el cbh2 natural se muestra en SEQ ID NO:1. La divulgación abarca una molécula de ácido nucleico que codifica las celulasas variantes descritas en esta especificación. El ácido nucleico puede ser una molécula de DNA.

[104] Una vez que las secuencias de DNA, que codifican las variantes CBH2, se han clonado en constructos de DNA, el DNA se usa para transformar microorganismos. El microorganismo que se va a transformar con el propósito de expresar una variante CBH2 de acuerdo con la presente divulgación puede incluir ventajosamente una cepa derivada de *Trichoderma sp.* Por consiguiente, un modo preferido para preparar celulasas variantes CBH2 de acuerdo con la presente divulgación incluye transformar una célula hospedera *Trichoderma sp.* con un constructo de DNA, que contiene por lo menos un fragmento de DNA que codifica una porción o la totalidad de la variante CBH2. Por lo general, el constructo de DNA estará fijado funcionalmente a un promotor. La célula hospedera

transformada se cultiva entonces bajo condiciones tales que expresen la proteína deseada. Posteriormente, el producto proteína deseada puede purificarse a una homogeneidad sustancial.

[105] No obstante, en realidad, puede ser que el mejor vehículo de expresión para un DNA dado, que codifica una variante CBH2, difiera de *H. jecorina*. Por lo tanto, es posible que sea muchísimo más ventajoso expresar una proteína en un hospedero de transformación que tenga similitud filogenética con el organismo fuente para la variante CBH2. En una incorporación alternativa, puede usarse *Aspergillus niger* como un vehículo de expresión. Para una descripción de las técnicas de transformación con *A. niger*, véase WO 98/31821, cuya divulgación se incorpora por referencia en su totalidad.

[106] Por consiguiente, se provee la presente descripción de un sistema de expresión de *Aspergillus* spp. con fines de ilustración únicamente y como una opción para expresar la variante CBH2 de la divulgación. Una persona experta en el arte, sin embargo, puede estar inclinada a expresar el DNA que codifica la variante CBH2 en una célula hospedera diferente si es apropiado, y debe entenderse que la fuente de la variante CBH2 debe considerarse al determinar el hospedero de expresión óptimo. Adicionalmente, el trabajador experto en el campo estará en capacidad de seleccionar el mejor sistema de expresión para un gen particular mediante técnicas rutinarias, utilizando las herramientas disponibles en el arte.

B. Polipéptidos de CBH2 Variante

[107] Las CBH2 variantes de esta divulgación tienen secuencias de aminoácidos que se derivan de la secuencia de aminoácidos de una CBH2 precursora. La secuencia de aminoácidos de la variante CBH2 difiere de la secuencia de aminoácidos de la CBH2 precursora por la sustitución, eliminación o inserción de uno o más aminoácidos de la secuencia de aminoácidos precursora. En una incorporación preferida, la CBH2 precursora es *Hypocrea jecorina* CBH2. La secuencia de aminoácidos madura de *H. jecorina* CBH2 se muestra en SEQ ID NO:3. Por lo tanto, esta divulgación se dirige a CBH2 variantes que contienen residuos aminoácidos en posiciones que son equivalentes al residuo particular identificado en *H. jecorina* CBH2. Un residuo (aminoácido) de un homólogo CBH2 es equivalente a un residuo de *Hypocrea jecorina* CBH2 si es homólogo (es decir, correspondiente en posición en cualquiera estructura primaria o terciaria) o es funcionalmente análogo a un residuo específico o porción de ese residuo en

Hypocrea jecorina CBH2 (es decir, que tiene la misma capacidad funcional o similar de combinar, reaccionar, o interactuar química o estructuralmente). Tal como se emplea en esta especificación, se pretende que la numeración corresponda a la de la secuencia de aminoácidos de CBH2 madura (SEQ ID NO:3).

[108] El alineamiento de secuencias de aminoácidos para determinar la homología se determina preferiblemente usando un "algoritmo de comparación de secuencias." El alineamiento óptimo de secuencias para comparación puede conducirse, v.g., por el algoritmo de homología local de Smith & Waterman, Adv. Appl. Math. 2:482 (1981), por el algoritmo de alineamiento de homología de Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), por el método de búsqueda de similitud de Pearson & Lipman, Proc. Nat'l Acad. Sci. USA 85:2444 (1988), por implementaciones computarizadas de estos algoritmos (GAP, BESTFIT, FASTA, y TFASTA en el 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 por Chemical Computing Group, Montreal, Canadá.

[109] Un ejemplo de un algoritmo que es apropiado para determinar la similitud de secuencias es el algoritmo BLAST, el cual se describe en Altschul, y col., J. Mol. Biol. 215:403-410 (1990). El software para efectuar análisis BLAST se encuentra públicamente disponible a través del National Center for Biotechnology Information (www.ncbi.nlm.nih.gov). Este algoritmo involucra primeramente identificar pares de secuencias de alta puntuación (HSPs - *high scoring sequence pairs*) identificando palabras cortas de longitud W en la secuencia problema que coincidan o satisfagan alguna puntuación T umbral de valor positivo cuando se alineen con una palabra de la misma longitud en una secuencia de base de datos. Estos aciertos de palabras de vecindad inicial actúan como puntos de partida para encontrar HSPs más largos que los contengan. Los aciertos de palabras se extienden en ambas direcciones a lo largo de cada una de las dos secuencias que se están comparando, en la medida en que la puntuación de alineamiento acumulativo pueda aumentarse. La extensión de los aciertos de palabras se suspende cuando: el puntaje de alineamiento acumulativo cae fuera de la cantidad X de un valor máximo alcanzado; el puntaje acumulativo tiende a cero o por debajo; o se alcanza el extremo de cualquiera de las dos secuencias. Los parámetros del algoritmo BLAST W, T y X determinan la sensibilidad y velocidad del alineamiento. El programa BLAST utiliza como valores predeterminados una longitud de palabra (W) de 11, la matriz de puntuación BLOSUM62 (véase

Henikoff & Henikoff, Proc. Natl. Acad. Sci. USA 89:10915 (1989)) las alineaciones (B) de 50, expectativa (E) de 10, M'5, N'-4, y una comparación de ambas hebras.

[110] El algoritmo BLAST realiza entonces un análisis estadístico de la similitud entre dos secuencias (véase, v.g., Karlin & Altschul, Proc. Nat'l. Acad. Sci. USA 90:5873-5787 (1993)). Una medida de similitud provista por el algoritmo BLAST es la probabilidad de suma mínima (P(N)), que ofrece una indicación de la probabilidad mediante la cual ocurrirá casualmente una coincidencia entre dos secuencias de nucleótidos o aminoácidos. Por ejemplo, una secuencia de aminoácidos se considera similar a una proteasa si la probabilidad de suma mínima en una comparación de la secuencia de aminoácidos de prueba a una secuencia de aminoácidos de la proteasa es menor de alrededor de 0.1, más preferiblemente menor de alrededor de 0.01 y, muy preferiblemente, menor de alrededor de 0.001.

[111] Para efectos de la presente divulgación, el grado de identidad puede determinarse apropiadamente por medio de programas de computador conocidos en el arte, tales como GAP provisto en paquete de programas 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), usando GAP con las siguientes configuraciones para comparación de secuencias polinucleotídicas: penalización de creación GAP de 5.0 y penalización de extensión GAP de 0.3.

[112] Un alineamiento estructural entre una *T. reesei* CBH2 y otras celulasas puede usarse para identificar posiciones equivalentes/correspondientes en otras celulasas que tienen un grado de homología moderado a alto, v.g., alrededor de 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 *T. reesei* CBH2 (SEQ ID NO: 3). Un método de obtener el citado alineamiento estructural es usar el programa Pile Up del paquete GCG utilizando valores predeterminados de penalizaciones de espacios, es decir, una penalización por creación de espacios de 3.0 y una penalización por extensión de espacios de 0.1. Otros métodos de alineamiento estructural incluyen el análisis de clúster hidrofóbico (Gaboriaud y col., FEBS Letters, 224:149-155, 1987) y concatenado inverso ("**reverse threading**") (Huber y Torda, Protein Science, 7:142-149, 1998).

[113] Un alineamiento ejemplar de la forma madura de varias celulasas de referencia se provee como 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 empleando los algoritmos de alineamiento de secuencias múltiples ClustalW y MUSCLE. Una matriz que muestra la identidad porcentual de celulasas del alineamiento de secuencias de Figura 3 se provee en la Tabla 1.

Tabla 1. Matriz de Identidad Porcentual de Celulasa*

ID porcentual	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 en la fila superior y la columna izquierda corresponden a las SEQ ID NOS de las secuencias alineadas de Figura 3.

[114] Normalmente las búsquedas de secuencias se llevan a cabo empleando el programa BLASTN cuando se evalúa una secuencia de ácido nucleico dada, relativa a secuencias de ácido nucleico en el GenBank DNA Sequences 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 contra secuencias de aminoácidos en el GenBank Protein Sequences y otras bases de datos públicas. Tanto BLASTN como BLASTX se ejecutan usando parámetros predeterminados de una penalización de espacio abierta de 11.0, y una penalización de espacio extendida de 1.0, y utilizan la matriz BLOSUM-62. (Véase, v.g., Altschul, y col., 1997.)

V. Expresión de Variantes CBH2 Recombinantes

[115] Los métodos de la divulgación se basan en el uso de células para expresar

la variante CBH2, no requiriéndose un método particular de expresión de CBH2. La variante CBH2 se secreta preferiblemente de las células. La divulgación provee células hospederas que se han transducido, transformado o transfectado con un factor de expresión que comprende por una secuencia de ácido nucleico que codifica la variante CBH2. Las condiciones de cultivo, tales como temperatura, pH y similares, son las previamente utilizadas para la célula hospedera progenitora antes de la transducción, transformación o transfección y serán evidentes para los expertos en el arte.

[116] En un método, una célula fúngica filamentosa o una célula de levadura se transfecta con un vector de expresión que tiene un promotor o un fragmento de promotor biológicamente activo de uno o más (v.g., una serie) de potenciadores que funciona en la línea de células hospederas, enlazados operacionalmente a un segmento de DNA que codifica la variante CBH2, de modo que la variante CBH2 se exprese en la línea de células.

A. Constructos de Ácido Nucleico/Vectores de Expresión

[117] Fragmentos polinucleotídicos naturales o sintéticos que codifican una variante CBH2 ("secuencias de ácido nucleico que codifican CBH2") pueden incorporarse en constructos de ácido nucleico heterólogo o vectores, capaces de introducción y replicación en una célula fúngica filamentosa o una célula de levadura. Los vectores y métodos divulgados en esta especificación son apropiados para uso en células hospederas para la expresión de una variante CBH2. Puede emplearse cualquier vector siempre y cuando sea replicable y viable en células en las cuales se introduce. Grandes números de vectores y promotores apropiados son conocidos por los expertos en el arte, y se encuentran disponibles en el comercio. La clonación y los vectores de expresión se describen igualmente en Sambrook y col., 1989, Ausubel F M y col., 1989, y Strathern y col., The Molecular Biology of the Yeast *Saccharomyces*, 1981, cada uno de los cuales se incorpora expresamente por referencia en esta especificación. Vectores de expresión apropiados para hongos se describen en van den Hondel, C. A. M. J. J. y col., (1991) In: Bennett, J. W. y Lasure, L. L. (eds.) *More Gene Manipulations in Fungi*. Academic Press, páginas 396-428. La secuencia de DNA apropiada puede insertarse en un plásmido o vector (denominados colectivamente en esta especificación "vectores") por una variedad de procedimientos. En general, la secuencia de DNA se inserta en un sitio(s) apropiado(s) de endonucleasa de restricción por procedimientos estándar. Los citados procedimientos y procedimientos relacionados de subclonación se consideran dentro del campo de

acción del conocimiento de los expertos en el arte.

[118] Hongos filamentosos recombinantes que incluyen la secuencia codificante para la variante CBH2 pueden producirse introduciendo un constructo de ácido nucleico heterólogo que comprende la secuencia que codifica la variante CBH2 en células de una cepa seleccionada de los hongos filamentosos.

[119] Una vez obtenida la forma deseada de una secuencia de ácido nucleico del *cbh2* variante, ésta puede modificarse en una variedad de modos. Donde la secuencia involucra regiones flanqueantes no codificantes, las regiones flanqueantes pueden someterse a resección, mutagénesis, etc. Por lo tanto, pueden realizarse transiciones, transversiones, eliminaciones e inserciones en la secuencia de origen natural.

[120] Una secuencia seleccionada que codifique un *cbh2* variante puede insertarse en un vector apropiado de acuerdo con técnicas recombinantes bien conocidas y usarse para transformar hongos filamentosos capaces de expresión de CBH2. Debido a la degeneración inherente del código genético, otras secuencias de ácido nucleico, las cuales codifican sustancialmente la misma secuencia de aminoácidos o una secuencia funcionalmente equivalente, pueden usarse para clonar y expresar la variante CBH2. Por consiguiente, se observa que tales sustituciones en la región codificante caen dentro de las variantes de secuencias cubiertas por la presente divulgación. Cualquiera y todas estas variantes de secuencias pueden utilizarse en la misma manera como se describe en esta especificación para una secuencia de ácido nucleico que codifica una CBH2 progenitora.

[121] La presente divulgación incluye asimismo constructos de ácido nucleico recombinante constituidos por una o más de las secuencias de ácido nucleico que codifican la CBH2 variante, como ya se describió. Los constructos contienen un vector, tal como un plásmido o vector vírico, en el cual se ha insertado una secuencia de la divulgación, en una orientación directa o inversa.

[122] Los constructos de ácido nucleico heterólogo pueden incluir la secuencia codificante para el *cbh2* variante (i) en aislamiento; (ii) en combinación con secuencias codificantes adicionales; tales como secuencias codificantes de proteína de fusión o péptido señal, donde la secuencia que codifica el *cbh2* es la secuencia codificante dominante; (iii) en combinación con secuencias no codificantes, tales como intrones y elementos de control, tales como elementos

promotores y terminadores o regiones no traducidas 5' y/o 3', efectivos para expresión de la secuencia codificante en un hospedero apropiado; y/o (iv) en un entorno vector u hospedero en el cual la secuencia que codifica el cbh2 es un gen heterólogo.

[123] En un aspecto de la presente divulgación, se emplea un constructo de ácido nucleico heterólogo para transferir una secuencia de ácido nucleico, que codifica la CBH2 variante, a una célula *in vitro*, prefiriéndose las líneas fúngicas filamentosas y de levaduras. Para la producción a largo plazo de la CBH2 variante, se prefiere la expresión estable. Se sigue que, para poner en práctica la divulgación, puede usarse cualquier método efectivo para generar transformantes estables.

[124] Vectores apropiados se equipan normalmente con una secuencia de ácido nucleico que codifica un marcador seleccionable, sitios de inserción, y elementos de control apropiados, tales como secuencias promotoras y terminadoras. El vector puede comprender secuencias reguladoras, incluyendo, por ejemplo, secuencias no codificantes, tales como intrones y elementos de control, es decir, elementos promotores y terminadores o regiones no traducidas 5' y/o 3', efectivos para la expresión de la secuencia codificante en células hospederas (y/o en un entorno vector o célula hospedera en el cual una secuencia codificante de antígeno de proteína soluble modificada no se expresa normalmente), enlazados operacionalmente con la secuencia codificante. Grandes números de vectores y promotores apropiados son conocidos por los expertos en el arte, muchos de los cuales se encuentran disponibles en el comercio y/o se describe en Sambrook, y col., (supra).

[125] Promotores ejemplares incluyen promotores constitutivos y promotores inducibles, ejemplos de los cuales incluyen un promotor CMV, un promotor temprano del virus SV40, un promotor RSV, un promotor EF-1 α , un promotor que contiene el elemento sensible a la tetraciclina (TRE - *tet responsive element*) en el sistema de activación o desactivación de tetraciclina ("**tet-on or tet-off system**"), como se describe (ClonTech y BASF), el promotor de beta actina y el promotor de metalotionina que pueden regularse por aumento mediante la adición de ciertas sales de metales. Una secuencia promotora es una secuencia de DNA que es reconocida por el hongo filamentoso particular para propósitos de expresión. Está operacionalmente conectada a la secuencia de DNA que codifica un polipéptido CBH2 variante. Tal conexión incluye posicionamiento del promotor con respecto al codón de iniciación de la secuencia de DNA que codifica el polipéptido de CBH2

variante en los vectores de expresión divulgados. La secuencia promotora contiene secuencias de control de transcripción y traducción que median la expresión del polipéptido de CBH2 variante. Ejemplos incluyen los promotores de los genes que codifican la glucoamilasa, α -amilasa, o α -glucosidasa de *Aspergillus niger*, *A. awamori* o *A. oryzae*; los genes de *A. nidulans* *gpdA* o *trpC*; los genes de *Neurospora crassa* *cbh1* o *trp1*; los genes que codifican la proteinasa aspártica de *A. niger* o *Rhizomucor miehei*; los genes *cbh1*, *cbh2*, *egl1*, *egl2* de *H. jecorina* (*T. reesei*), u otros genes que codifican celulasa.

[126] La selección del marcador seleccionable apropiado dependerá de la célula hospedera, y marcadores apropiados para diferentes hospederos son bien conocidos en el arte. 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*. Marcadores seleccionables ejemplares adicionales incluyen, pero no se limitan a *trpc*, *trp1*, *oliC31*, *niaD* o *leu2*, los cuales se incluyen en constructos de ácido nucleico heterólogo, usados para transformar una cepa mutante tal como *trp*-, *pyr*-, *leu*- y similares.

[127] Los citados marcadores seleccionables confieren a los transformantes la capacidad de utilizar un metabolito que usualmente no es metabolizado por los hongos filamentosos. Por ejemplo, el gen *amdS* de *H. jecorina*, que codifica la enzima acetamidasa, que permite a células transformantes crecer en acetamida como fuente de nitrógeno. El marcador seleccionable (v.g. *pyrG*) puede restaurar la capacidad de una cepa mutante auxotrófica de crecer en un medio mínimo selectivo o el marcador seleccionable (v.g. *olic31*) puede conferir a los transformantes la capacidad de crecer en la presencia de un fármaco inhibitorio o antibiótico.

[128] La secuencia que codifica el marcador seleccionable se clona en cualquier plásmido apropiado usando métodos generalmente empleados en el arte. Plásmidos ejemplares incluyen *pUC18*, *pBR322*, *pRAX* y *pUC100*. El plásmido *pRAX* contienen las secuencias AMAL de *A. nidulans*, que hacen posible replicarlo en *A. niger*.

[129] La práctica de la presente divulgación empleará, a menos que se indique lo contrario, técnicas convencionales de biología molecular, microbiología, DNA recombinante, e inmunología, las cuales se encuentran al alcance de la habilidad en el arte. Tales técnicas se explican completamente en la literatura. Véase, por ejemplo, Sambrook y col., 1989; Freshney, *Animal Cell Culture*, 1987; Ausubel, y

col., 1993; y Coligan y col., Current Protocols in Immunology, 1991.

B. Células Hospederas Y Condiciones de Cultivo para la Producción de CBH2

(i) Hongos Filamentosos

[130] Por lo tanto, la presente divulgación provee hongos filamentosos que comprenden células que se han modificado, seleccionado y cultivado en una manera efectiva para resultar en la producción de CBH2 variante o expresión relativa a los hongos progenitores no transformados correspondientes.

[131] Ejemplos de especies de hongos filamentosos progenitores que pueden tratarse y/o modificarse para expresión de CBH2 variante incluyen, pero no se limitan a *Trichoderma*, v.g., *Trichoderma reesei*, *Trichoderma longibrachiatum*, *Trichoderma viride*, *Trichoderma koningii*; *Penicillium* sp., *Humicola* sp., incluyendo *Humicola insolens*; *Aspergillus* sp., *Chrysosporium* sp., *Fusarium* sp., *Hypocrea* sp., y *Emericella* sp.

[132] Células que expresan CBH2 se cultivan bajo condiciones empleadas normalmente para cultivar la línea fúngica progenitora. Generalmente, las células se cultivan en medio estándar que contienen sales fisiológicas y nutrientes, tal como se describe en Pourquie, J. y col., Biochemistry and Genetics of Cellulose Degradation, eds. Aubert, J. P. y col., Academic Press, páginas 71-86, 1988 e Ilmen, M. y col., Appl. Environ. Microbiol. 63:1298-1306, 1997. Las condiciones de cultivo son también estándar, v.g., los cultivos se incuban a 28° C en cultivos en frascos de agitación o fermentadores, hasta que se consiguen niveles deseados de expresión de CBH2.

[133] Condiciones de cultivo preferidas para un hongo filamentoso dado pueden encontrarse en la literatura científica y/o de la fuente de los hongos tales como la American Type Culture Collection (ATCC; www.atcc.org/). Una vez se ha establecido el crecimiento fúngico, las células se exponen a condiciones efectivas para provocar o permitir la expresión de CBH2 variante.

[134] En casos en los que una secuencia codificante de CBH2 está bajo control de un promotor inducible, el agente inductor, v.g., un azúcar, una sal de metal o antibiótico, se agrega al medio en una concentración efectiva para inducir la expresión de CBH2.

[135] En una incorporación, la cepa comprende *Aspergillus niger*, la cual es una cepa útil para obtener una proteína sobreexpresada. Por ejemplo dgr246 de *A. niger var awamori* es conocida por secretar cantidades elevadas de celulasas secretadas (Goedegebuur y col., Curr. Genet (2002) 41: 89-98). Son conocidas otras cepas de *Aspergillus niger var awamori* tales como GCDAP3, GCDAP4 y GAP3-4 (Ward y col., Appl. Microbiol. Biotechnol. 39:738-743, 1993).

[136] En otra incorporación, la cepa consta de *Trichoderma reesei*, la cual es una cepa útil para obtener proteína sobreexpresada. Por ejemplo, RL-P37, descrita por Sheir-Neiss, y col., Appl. Microbiol. Biotechnol. 20:46-53 (1984), es conocida por secretar cantidades elevadas de enzimas de celulasa. Equivalentes funcionales de RL-P37 incluyen la cepa RUT-C30 de *Trichoderma reesei* (ATCC No. 56765) y la cepa QM9414 (ATCC No. 26921). Se contempla que estas cepas serán igualmente útiles en sobreexpresar la CBH2 variante.

[137] Donde se desea obtener la CBH2 variante en la ausencia de actividad celulolítica natural potencialmente prejudicial, es útil obtener una cepa de células hospederas de *Trichoderma* que haya tenido eliminados uno o más genes de celulasa antes de la introducción del constructo o plásmido de DNA que contiene el fragmento de DNA que codifica la CBH2 variante. Tales cepas pueden prepararse por el método divulgado en la Patente de los Estados Unidos No. 5,246,853 y WO 92/06209, divulgaciones que se incorporan por este medio por referencia en su totalidad. Expresando una celulasa CBH2 variante en un microorganismo hospedero que carece de uno o más genes de celulasa, se simplifican los procedimientos de identificación y purificación subsiguiente. Puede eliminarse cualquier gen de *Trichoderma sp.* que se haya clonado, por ejemplo, los genes *cbh1*, *cbh2*, *egl1*, y *egl2*, así como aquellos que codifican la proteína EG III y/o EGV (véase v.g., la Patente de los Estados Unidos No. 5,475,101 y WO 94/28117, respectivamente).

[138] La eliminación de los genes puede llevarse a cabo insertando una forma del gen deseado, que se va a eliminar o alterar, en un plásmido por métodos conocidos en el arte. El plásmido de eliminación se corta entonces en un sitio(s) apropiado(s) de enzimas de restricción, interno(s) a la región codificante del gen deseado, y la secuencia que codifica el gen o parte del mismo se reemplaza con un marcador seleccionable. Secuencias de DNA flanqueantes del locus del gen que se va a eliminar o alterar, preferiblemente entre alrededor de 0.5 a 2.0 kb, permanecen a cualquiera de los lados del gen marcador seleccionable. Un

plásmido de eliminación apropiado tendrá generalmente sitios únicos de enzimas de restricción presentes en él para permitir que el fragmento que contiene el gen eliminado, incluyendo secuencias de DNA flanqueantes, y el gen marcador seleccionable se extraigan como un elemento lineal único.

[139] Un marcador seleccionable tiene que elegirse de modo que permita la detección del microorganismo transformado. Será apropiado cualquier gen marcador seleccionable que se exprese en el microorganismo seleccionado. Por ejemplo, con *Aspergillus sp.*, el marcador seleccionable se elige de modo que la presencia del marcador seleccionable en los transformantes no afecte significativamente las propiedades de los mismos. Tal marcador seleccionable puede ser un gen que codifica un producto analizable. Por ejemplo, puede usarse una copia funcional de un gen de *Aspergillus sp.* que, si falta en la cepa hospedera, resulta en la cepa hospedera que despliega un fenotipo auxotrófico. Análogamente, existen marcadores seleccionables para *Trichoderma sp.*

[140] En una incorporación, una cepa derivada pyrG de *Aspergillus sp.* se transforma con un gen pyrG funcional, el cual provee así un marcador seleccionable para transformación. Una cepa derivada pyrG puede obtenerse por selección de cepas de *Aspergillus sp.*, que son resistentes a ácido fluoroorótico (FOA - *fluoroorotic acid*). El gen pyrG codifica orotidina-5'-monofosfato decarboxilasa, una enzima requerida para la biosíntesis de uridina. Cepas con un gen pyrG intacto crecen en un medio que carece de uridina, pero son sensibles a ácido fluoroorótico. Es posible seleccionar cepas derivadas pyrG que carecen de una enzima orotidina monofosfato decarboxilasa funcional y requieren uridina para crecimiento seleccionando con respecto a resistencia FOA. Usando la técnica de selección FOA es también posible obtener cepas que requieren uridina, las cuales carecen de una orotato pirofosforribosil transferasa funcional. Es posible igualmente transformar estas células con una copia funcional del gen que codifica esta enzima (Berges & Barreau, Curr. Genet. 19:359-365 (1991), y van Hartingsveldt y col., (1986) Development of a homologous transformation system for *Aspergillus niger* based on the pyrG gene. Mol. Gen. Genet. 206:71-75). La selección de cepas derivadas se lleva a cabo fácilmente usando la técnica de resistencia FOA ya mencionada y, por consiguiente, el gen pyrG se emplea preferiblemente como un marcador seleccionable.

[141] En una segunda incorporación, una cepa derivada pyr4 de *Hyprocrea sp.* (*Hyprocrea sp.* (*Trichoderma sp.*)) se transforma con un gen pyr4 funcional, el cual provee así un marcador seleccionable para transformación. Una cepa derivada

pyr4 puede obtenerse por selección de cepas de *Hyprocrea sp.* (*Trichoderma sp.*), que son resistentes a ácido fluoroorótico (FOA). El gen pyr4 codifica orotidina-5'-monofosfato decarboxilasa, una enzima requerida para la biosíntesis de uridina. Cepas con un gen pyr4 intacto crecen en un medio que carece de uridina pero son sensibles a ácido fluoroorótico. Es posible seleccionar cepas derivadas pyr4, que carecen de una enzima orotidina monofosfato decarboxilasa funcional y requieren uridina para crecimiento, seleccionando con respecto a resistencia FOA. Usando la técnica de selección FOA es igualmente posible obtener cepas que requieren uridina, las cuales 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 usando la técnica de resistencia FOA ya mencionada y, por consiguiente, el gen pyr4 se emplea preferiblemente como un marcador seleccionable.

[142] Para transformar pyrG - *Aspergillus sp.* o pyr4 - *Hyprocrea sp.* (*Trichoderma sp.*) de modo que carezcan de la capacidad de expresar uno o más genes de celulasa, un fragmento único de DNA que comprende un gen de celulasa alterado o eliminado se aísla entonces del plásmido de eliminación y se usa para transformar el hospedero apropiado *pyr-Aspergillus* o *pyr-Trichoderma*. Los transformantes se identifican entonces y se seleccionan basándose en su capacidad de expresar el producto génico pyrG o pyr4, respectivamente y, por consiguiente, satisfacer la auxotrofia de uridina de la cepa hospedera. Análisis Southern blot se lleva entonces a cabo en los transformantes resultantes para identificar y confirmar un caso de integración de cruce doble, que reemplaza parte o la totalidad de la región codificante de la copia genómica del gen que se va a eliminar con los marcadores seleccionables pyr apropiados.

[143] Aunque los vectores plasmídicos específicos ya descritos se refieren a la preparación de transformantes pyr, la presente divulgación no se limita a estos vectores. Varios genes pueden eliminarse y reemplazarse en la cepa *Aspergillus sp.* o *Hyprocrea sp.* (*Trichoderma sp.*) utilizando las técnicas anteriores. Adicionalmente, pueden usarse cualesquiera marcadores seleccionables, como ya se discutió. De hecho, cualquier hospedero, v.g., el gen *Aspergillus sp.* o *Hyprocrea sp.*, que se ha clonado e identificado así, puede eliminarse del genoma empleando la estrategia ya descrita.

[144] Como se especifica anteriormente, las cepas hospederas utilizadas pueden ser los derivados de *Hyprocrea sp.* (*Trichoderma sp.*), que carecen o tienen un

gen o genes no funcionales, correspondientes al marcador seleccionable elegido. Por ejemplo, si el marcador seleccionable de pyrG se escoge para *Aspergillus sp.*, entonces se usa una cepa específica derivada pyrG como receptor en el procedimiento de transformación. Asimismo, por ejemplo, si el marcador seleccionable de pyr4 se escoge para una *Hyprocrea sp.*, entonces la cepa específica derivada pyr4 se usa en un receptor en el procedimiento de transformación. Análogamente, pueden emplearse 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 correspondiente cepa receptora tiene que ser una cepa derivada tal como argB-, trpC-, niaD-, respectivamente.

[145] DNA que codifica la CBH2 variante se prepara entonces para inserción en un microorganismo apropiado. De acuerdo con la presente divulgación, DNA que codifica una CBH2 variante incluye el DNA necesario para codificar una proteína que tiene actividad celulolítica funcional. El fragmento de DNA que codifica la CBH2 variante puede fijarse 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*.

[146] Se contempla asimismo que más de una copia del DNA que codifica una CBH2 variante pueda recombinarse en la cepa para facilitar la sobreexpresión. El DNA que codifica la CBH2 variante puede prepararse por la construcción de un vector de expresión que transporta el DNA que codifica la variante. El vector de expresión que transporta el fragmento de DNA insertado que codifica la CBH2 variante puede ser cualquier vector que es capaz de replicarse autónomamente en un organismo hospedero dado o de integrarse en el DNA del hospedero, normalmente un plásmido. En incorporaciones preferidas se contemplan dos tipos de vectores de expresión para obtener la expresión de genes. El primero contiene secuencias de DNA en las cuales el promotor, la región codificante del gen, y la secuencia terminadora se originan todos del gen que se va a expresar. El truncado de genes puede obtenerse donde se desee, eliminando secuencias indeseadas de DNA (v.g., que codifican dominios no deseados) para dejar el dominio que se va a expresar bajo control de sus propias secuencias reguladoras transcripcionales y traduccionales. Un marcador seleccionable puede estar contenido igualmente en el vector, permitiendo la selección para integración en el hospedero de múltiples copias de las novedosas secuencias génicas.

[147] El segundo tipo de vector de expresión se pre-ensambla y contiene

secuencias requeridas para transcripción de alto nivel y un marcador seleccionable. Se contempla que la región codificante de un gen o parte del mismo pueda insertarse en este vector de expresión de propósito general, de modo que esté bajo el control transcripcional de las secuencias promotoras y terminadoras de casetes de expresión.

[148] Por ejemplo, en *Aspergillus*, pRAX es tal vector de expresión de propósito general. Genes o parte de los mismos pueden insertarse hacia 3' del promotor glaa fuerte.

[149] Por ejemplo, en *Hypocrea*, pTEX es tal vector de expresión de propósito general. Genes o parte de los mismos pueden insertarse hacia 3' del promotor cbh1 fuerte.

[150] En el vector, la secuencia de DNA que codifica la CBH2 variante de la presente divulgación debe conectarse operacionalmente a secuencias transcripcionales y traduccionales, es decir, una secuencia promotora apropiada y una secuencia señal en el marco de lectura al gen estructural. El promotor puede ser cualquier secuencia de DNA que muestre actividad transcripcional en la célula hospedera y pueda derivarse de genes que codifican proteínas homólogas o heterólogas a la célula hospedera. Un péptido señal opcional posibilita la producción extracelular de la CBH2 variante. El DNA que codifica la secuencia señal es preferiblemente aquel que se asocia naturalmente con el gen que se va a expresar; no obstante, se contempla en la presente divulgación la secuencia señal de cualquier fuente conveniente, por ejemplo una exo-celobiohidrolasa o endoglucanasa de *Trichoderma*.

[151] Los procedimientos usados para ligar las secuencias de DNA que codifican la CBH2 variante de la presente divulgación con el promotor, y la inserción en vectores apropiados son bien conocidos en el arte.

[152] El vector o constructo de DNA ya descritos pueden introducirse en la célula hospedera de acuerdo con técnicas conocidas, tales como transformación, transfección, microinyección, microporación, bombardeo biolístico y similares.

[153] En la técnica de transformación preferida, hay que tener en cuenta que la permeabilidad de la pared celular a DNA en *Hypocrea sp.* (*Trichoderma sp.*) es muy baja. Por consiguiente, la absorción de la secuencia deseada de DNA, gen o fragmento génico es a lo mejor mínima. Existen muchos métodos de 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 usado) antes del proceso de transformación.

[154] El método preferido en la presente divulgación para preparar *Aspergillus sp.* o *Hyprocrea sp.* (*Trichoderma sp.*) para transformación involucra la preparación de protoplastos del micelio fúngico. Véase Campbell y col., Improved transformation efficiency of *A. niger* using homologous *niaD* gene for nitrate reductase. Curr. Genet. 16:53-56; 1989. El micelio puede obtenerse de esporas vegetativas germinadas. El micelio se trata con una enzima que digiere la pared celular resultando en protoplastos. A continuación los protoplastos se protegen por la presencia de un estabilizante osmótico en el medio de suspensión. Estos estabilizantes incluyen sorbitol, manitol, cloruro de potasio, sulfato de magnesio y similares. Usualmente la concentración de estos estabilizantes varía entre 0.8 M y 1.2 M. Es preferible emplear alrededor de una solución 1.2 M de sorbitol en el medio de suspensión.

[155] La absorción del DNA en la cepa hospedera, (*Aspergillus sp.* o *Hyprocrea sp.* (*Trichoderma sp.*), depende de la concentración de los iones de calcio. Generalmente se usan entre alrededor de 10 mM CaCl_2 y 50 mM CaCl_2 en una solución de absorción. Además de la necesidad del ión de calcio en la solución de absorción, otros ítems generalmente incluidos son un sistema amortiguador tal como amortiguador TE (10 Mm Tris, pH 7.4; 1 mM EDTA) o amortiguador 10 mM MOPS, pH 6.0 (ácido morfolinopropanosulfónico) y polietilén glicol (PEG). Se cree que el polietilén glicol actúa para fusionar las membranas celulares permitiendo así que el contenido del medio se transfiera al citoplasma de la célula hospedera, a modo de ejemplo la cepa *Aspergillus sp.* o *Hyprocrea sp.*, y DNA plasmídico se transfiera al núcleo. Frecuentemente, esta fusión deja múltiples copias del DNA del plásmido integradas en el cromosoma hospedero.

[156] Usualmente una suspensión que contiene los protoplastos de *Aspergillus sp.* o células que se han sometido a un tratamiento de permeabilidad con una densidad de 10^5 a $10^6/\text{mL}$, preferiblemente $2 \times 10^5/\text{mL}$ se emplea en la transformación. Análogamente, una suspensión que contiene los protoplastos de *Hyprocrea sp.* (*Trichoderma sp.*) o células que se han sometido a un tratamiento de permeabilidad con una densidad de 10^8 a $10^9/\text{mL}$, preferiblemente $2 \times 10^8/\text{mL}$ se usa en la transformación. Un volumen de 100 μL de estos protoplastos o células en una solución apropiada (v.g., 1.2 M sorbitol; 50 mM CaCl_2) se mezclan con el DNA deseado. Generalmente se agrega una alta concentración de PEG a la

solución de absorción. De 0.1 a 1 volumen de 25% PEG 4000 puede agregarse a la suspensión de protoplastos. Sin embargo, es preferible agregar alrededor de 0.25 volúmenes a la suspensión de protoplastos. Aditivos tales como sulfóxido de dimetilo, heparina, espermidina, cloruro de potasio y similares pueden agregarse asimismo a la solución de absorción y ayudar en la transformación.

[157] Generalmente, la mezcla se incuba luego a aproximadamente 0°C durante un período de entre 10 a 30 minutos. A continuación se agrega PEG adicional a la mezcla para potenciar aún más la absorción del gen deseado o secuencia de DNA. El 25% PEG 4000 se agrega generalmente en volúmenes de 5 a 15 veces el volumen de la mezcla de transformación; no obstante, pueden ser apropiados volúmenes mayores y menores. El 25% PEG 4000 es preferiblemente alrededor de 10 veces el volumen de la mezcla de transformación. Una vez agregado el PEG, la mezcla de transformación se incuba entonces a temperatura ambiente o sobre hielo antes de la adición de una solución de sorbitol y CaCl₂. La suspensión de protoplastos se agrega entonces a alícuotas líquidas de un medio de crecimiento. Este medio permite el crecimiento de transformantes únicamente. Puede usarse cualquier medio de crecimiento en la presente divulgación que sea apropiado para desarrollar los transformantes deseados. No obstante, si se están seleccionando transformantes Pyr⁺, es preferible usar un medio de crecimiento que no contiene uridina. Las colonias subsiguientes se transfieren y purifican en un medio de crecimiento agotado de uridina.

[158] En esta fase, transformantes estables pueden distinguirse de transformantes inestables por su tasa de crecimiento más rápida y, en *Trichoderma*, por ejemplo, la formación de colonias circulares con un perfil suavizado, más que irregular, en medio de cultivo sólido que carece de uridina. Adicionalmente, en algunos casos puede hacerse una prueba adicional de estabilidad desarrollando los transformantes en medio sólido no selectivo (es decir, que contiene uridina), recolectando las esporas de este medio de cultivo y determinando el porcentaje de estas esporas que germinarán posteriormente y crecerán en medio selectivo que carece de uridina.

[159] En una incorporación particular del método ya mencionado, la(s) CBH2(s) variante(s) se recuperan en forma activa de la célula hospedera luego del crecimiento en medio líquido como resultado del procesamiento postraducciona apropiado de la CBH2 variante.

(ii) Levadura

[160] La presente divulgación contempla igualmente el uso de levadura como célula hospedera para la producción de CBH2. Varios otros genes que codifican enzimas hidrolíticas se han expresado en varias cepas de la levadura *S. cerevisiae*. Éstos incluyen secuencias que codifican dos endoglucanasas (Penttila y col., Yeast vol. 3, páginas 175-185, 1987), dos celobiohidrolasas (Penttila y col., 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, no. 1, páginas 209-213, 1996), una alpha-amilasa del trigo (Rothstein y col., Gene 55:353-356, 1987), etc. Además, un casete del gen de celulasa que codifica la *Butyrivibrio fibrisolvens* endo-[beta]-1,4-glucanasa (END1), la *Phanerochaete chrysosporium* celobiohidrolasa (CBH1), la *Ruminococcus flavefaciens* celodextrinasa (CEL1) y la *Endomyces fibuliger* celobiasa (BgII) se expresó exitosamente en una cepa de laboratorio de *S. cerevisiae* (Van Rensburg y col., Yeast, vol. 14, páginas 67-76, 1998).

C. Introducción de una Secuencia de Ácido Nucleico que Codifica una CBH2 en Células Hospederas

[161] La divulgación provee además células y composiciones de células que se han modificado genéticamente para incluir una secuencia de ácido nucleico que codifica una CBH2 variante provista exógenamente. Una célula progenitora o línea celular puede modificarse genéticamente (es decir, transducirse, transformarse o transfectarse) con un vector de clonación o un vector de expresión. El vector puede estar, por ejemplo, en la forma de un plásmido, una partícula vírica, un bacteriófago, etc, como se describe con más detalle anteriormente.

[162] Los métodos de transformación de la presente divulgación pueden resultar en la integración estable de la totalidad o parte del vector de transformación en el genoma del hongo filamentoso. Sin embargo, se contempla también la transformación resultante en el mantenimiento de un vector de transformación extra-cromosómico auto-replicante.

[163] Muchos métodos de transfección estándar pueden usarse para producir líneas celulares de *Trichoderma reesei* que expresan grandes cantidades de la proteína heteróloga. Algunos de los métodos publicados para la introducción de constructos de DNA en cepas productoras de 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 y col., 1985, The John Innes Foundation, Norwich, UK y para *Bacillus Brigidi*, DeRossi, Bertarini, Riccardi y Matteuzzi, 1990, FEMS Microbiol. Lett. 55: 135-138).

[164] Otros métodos para introducir un constructo de ácido nucleico heterólogo (vector de expresión) en hongos filamentosos (v.g., *H. jecorina*) incluyen, pero no se limitan a, el uso de una pistola de partículas o genes, permeabilización de las paredes celulares de hongos filamentosos antes del proceso de transformación (v.g., mediante el empleo de elevadas concentraciones de álcali, v.g., 0.05 M a 0.4 M CaCl_2 o acetato de litio), fusión de protoplastos o transformación mediada por *Agrobacterium*. Un método ejemplar para transformación de hongos filamentosos por tratamiento de protoplastos o esferoplastos con polietilén glicol y CaCl_2 se describe en Campbell, E. I. y col., Curr. Genet. 16:53-56, 1989 y Penttila, M. y col., Gene, 63:11-22, 1988.

[165] Puede emplearse cualquiera de los procedimientos bien conocidos para introducir secuencias nucleotídicas extrañas en células hospederas. Éstos incluyen el uso de transfección con fosfato de calcio, polibreno, fusión de protoplastos, electroporación, biolística, liposomas, microinyección, vectores de plasma, vectores víricos y cualquiera de los demás métodos bien conocidos para introducir DNA genómico clonado, cDNA, DNA sintético u otro material genético extraño en una célula hospedera (véase, v.g., Sambrook y col., supra). También útil es el método de transfección mediado por *Agrobacterium*, descrito en la Patente de los Estados Unidos No. 6,255,115. Sólo es necesario que el procedimiento de ingeniería genética particular utilizado pueda introducir exitosamente por lo menos un gen en la célula hospedera, capaz de expresar el gen heterólogo.

[166] Además, constructos de ácido nucleico heterólogo que contienen una secuencia de ácido nucleico que codifica una CBH2 variante pueden transcribirse *in vitro*, y el RNA resultante introducirse en la célula hospedera por métodos bien conocidos, v.g., mediante inyección.

[167] La divulgación incluye asimismo transformantes novedosos y útiles de hongos filamentosos tales como *H. jecorina* y *A. niger* para uso en producir

composiciones de celulasas fúngicas. La divulgación incluye transformantes de hongos filamentosos, especialmente hongos que comprenden la secuencia que codifica la CBH2 variante, o eliminación de la secuencia que codifica el cbh endógeno.

[168] Luego de la introducción de un constructo de ácido nucleico heterólogo que comprende la secuencia codificante para un cbh2 variante, las células genéticamente modificadas pueden cultivarse en medio nutriente convencional modificado según sea apropiado para activar promotores, seleccionar transformantes o amplificar la expresión de una secuencia de ácido nucleico que codifica una CBH2 variante. Las condiciones de cultivo, tales como temperatura, pH y similares, son aquellas previamente usadas para la célula hospedera seleccionada para expresión, y serán evidentes a los expertos en el arte.

[169] Generalmente se considera que la progenie de células, en las cuales se han introducido tales constructos de ácido nucleico heterólogo, comprende la secuencia de ácido nucleico que codifica la CBH2 variante hallada en el constructo de ácido nucleico heterólogo.

[170] La divulgación incluye además transformantes útiles y novedosos de hongos filamentosos tales como *H. jecorina* para uso en producir composiciones de celulasas fúngicas. De igual forma *Aspergillus niger* puede usarse en producir la CBH2 variante. La divulgación incluye transformantes de hongos filamentosos, especialmente hongos que comprenden la secuencia que codifica el cbh2 variante, o eliminación de la secuencia que codifica el cbh2 endógeno.

[171] Transformantes estables de hongos filamentosos pueden distinguirse generalmente de transformantes inestables por su tasa de crecimiento más rápida y, en *Trichoderma*, por ejemplo, la formación de colonias circulares con un perfil suave, más bien que irregular, en medio de cultivo sólido. Adicionalmente, en algunos casos, puede realizarse una prueba de estabilidad desarrollando los transformantes en medio no selectivo sólido, recolectando las esporas de este medio de cultivo y determinando el porcentaje de estas esporas que germinarán posteriormente y crecerán sobre el medio selectivo.

VI. Aislamiento y Purificación de Proteína CBH2 Recombinante

[172] En general, una proteína CBH2 variante producida en cultivo celular se secreta en el medio y puede purificarse o aislarse, v.g., eliminando componentes

indeseados del medio de cultivo celular. Sin embargo, en algunos casos, una proteína CBH2 variante puede producirse en una forma celular que requiera recuperación de un lisado de células. En tales casos la proteína CBH2 variante se purifica de las células en las cuales se produjo usando técnicas empleadas rutinariamente por los expertos en el arte. Los ejemplos incluyen, pero no se limitan a, cromatografía de afinidad (Tilbeurgh y col., FEBS Lett. 16:215, 1984), métodos cromatográficos de intercambio iónico (Goyal y col., Bioresource Technol. 36:37-50, 1991; Fliess y col., Eur. J. Appl. Microbiol. Biotechnol. 17:314-318, 1983; Bhikhabhai y col., J. Appl. Biochem. 6:336-345, 1984; Ellouz y col., J. Chromatography 396:307-317, 1987), incluyendo usar materiales de intercambio iónico con alto poder de resolución (Medve y col., J. Chromatography A 808:153-165, 1998), cromatografía de interacción hidrofóbica (Tomaz y Queiroz, J. Chromatography A 865:123-128, 1999), y partición de dos fases (Brumbauer, y col., Bioseparation 7:287-295, 1999).

[173] Normalmente, la proteína CBH2 variante se fracciona para segregar proteínas que tengan propiedades seleccionadas, tales como afinidad de unión a agentes de unión particulares, v.g., anticuerpos o receptores; o que tengan un intervalo de peso molecular seleccionado, o intervalo de puntos isoeléctricos.

[174] Una vez se ha logrado la expresión de una proteína CBH2 variante dada, la proteína CBH2 así producida se purifica de las células o del cultivo celular. Procedimientos ejemplares apropiados para la citada purificación incluyen los siguientes: cromatografía en columna de afinidad de anticuerpo, cromatografía de intercambio iónico; precipitación en etanol; HPLC de fase inversa; cromatografía en sílice o en una resina de intercambio catiónico tal como DEAE; cromatoenfocado; SDS-PAGE; precipitación en sulfato de amonio; y filtración en gel usando, v.g., Sephadex G-75. Pueden emplearse varios métodos de purificación de proteínas, y tales métodos son conocidos en el arte y se describen v.g. en Deutscher, Methods in Enzymology, vol. 182, No. 57, páginas. 779, 1990; Scopes, Methods Enzymol. 90: 479-91, 1982. Los pasos de purificación seleccionados dependerán, v.g., de la naturaleza del proceso de producción utilizado y la proteína particular producida.

VII. Utilidad de cbh2 y CBH2

[175] Puede observarse que los ácidos nucleicos de cbh variante, la proteína CBH2 variante y composiciones que incluyen actividad de proteína CBH2 variante tienen utilidad en una amplia variedad de aplicaciones, algunas de las cuales se

describen a continuación.

[176] Composiciones de celulasa novedosas y mejoradas que incluyen cantidades variables de celulasas tipo BG, tipo EG y tipo CBH variante tienen utilidad en composiciones detergentes que exhiben una capacidad limpiadora mejorada, funcionan como agente suavizante y/o mejoran el tacto de telas de algodón (v.g., "lavado a la piedra" o "biodesfibrilación"), en composiciones para degradar pulpa de madera en azúcares (v.g., para producción de bio-etanol), y/o en composiciones de forrajes. El aislamiento y caracterización de celulasa de cada tipo ofrece la posibilidad de controlar los aspectos de tales composiciones.

[177] En un enfoque, la celulasa de la divulgación tiene utilidad en composiciones detergentes o en el tratamiento de telas para mejorar el tacto y la apariencia.

[178] Puesto que la tasa de hidrólisis de productos celulósicos puede aumentarse usando un transformante, que tiene por lo menos una copia adicional del gen cbh insertado en el genoma, productos que contienen celulosa o heteroglucanos pueden degradarse a una tasa más rápida y en una mayor medida. Productos hechos de celulosa tales como papel, algodón, pañales celulósicos y similares pueden degradarse más eficientemente en un relleno sanitario. Por consiguiente, el producto de fermentación obtenible de los transformantes o los transformantes solos pueden usarse en composiciones para ayudar a degradar por licuefacción una variedad de productos de celulosa que se agregan a los rellenos sanitarios atestados.

[179] La sacarificación y fermentación separada es un proceso mediante el cual celulosa presente en biomasa, v.g., rastrojo de maíz, se convierte en glucosa y, posteriormente, cepas de levadura convierten glucosa en etanol. La sacarificación y fermentación simultánea es un proceso por el cual celulosa presente en biomasa, v.g., rastrojo de maíz, se convierte en glucosa y, al mismo tiempo y en el mismo reactor, cepas de levadura convierten glucosa en etanol. Por consiguiente, en otro enfoque, la celulasa tipo CBH2 variante de la divulgación tiene utilidad en la degradación de biomasa en etanol. La producción de etanol de fuentes fácilmente disponibles de celulosa ofrece una fuente estable de combustibles renovables.

[180] Materias primas a base de celulosa están constituidas por desechos agrícolas, pastos y maderas y demás biomasa de bajo valor, tales como desechos municipales (v.g., papel reciclado, residuos vegetales, etc.). Etanol puede

producirse de la fermentación de cualquiera de esas materias primas celulósicas. No obstante, la celulosa tiene que convertirse primero en azúcares antes de que pueda haber conversión en etanol.

[181] Una gran variedad de materias primas puede usarse con la CBH variante inventiva, y la seleccionada para uso puede depender de la región en donde se está haciendo la conversión. Por ejemplo, en el Medio Oeste de los Estados Unidos pueden predominar desechos 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 cualquier biomasa celulósica disponible puede usarse en cualquier región.

[182] Los métodos de la presente divulgación pueden emplearse en la producción de monosacáridos, disacáridos, y polisacáridos como materias primas químicas o de fermentación con microorganismos para la producción de productos orgánicos, químicos y combustibles, plásticos, y otros productos o intermedios. En particular, el valor de residuos de procesamiento (grano seco de destilerías, granos agotados de cervecerías, bagazo de caña de azúcar, etc.) puede aumentarse por solubilización parcial o completa de celulosa o hemicelulosa. Aparte de etanol, algunos químicos que pueden producirse de celulosa y hemicelulosa incluyen acetona, acetato, glicina, lisina, ácidos orgánicos (v.g., ácido láctico), 1,3-propanodiol, butanodiol, glicerol, etilen glicol, furfural, polihidroxialcanoatos, cis, ácido cis-mucónico, forrajes para animales y xilosa.

[183] Una composición de celulasa, que contiene una cantidad mejorada de celobiohidrolasa, tiene utilidad en la producción de etanol. El etanol de este proceso puede usarse además como un potenciador de octanaje o directamente como combustible en lugar de gasolina, lo cual es ventajoso porque el etanol como una fuente de combustible es más amigable con el medio ambiente que los productos derivados del petróleo. Se sabe que el uso de etanol mejorará la calidad del aire y reducirá posiblemente los niveles locales de ozono y smog. Más aún, la utilización de etanol en lugar de gasolina puede ser de importancia estratégica para amortiguar el impacto de los cambios súbitos en la energía no renovable y los suministros petroquímicos.

[184] Etanol puede producirse mediante procesos de sacarificación y fermentación de biomasa celulósica como árboles, plantas herbáceas, desechos sólidos municipales y residuos agrícolas y forestales. Sin embargo, la relación de enzimas de celulasa individuales con una mezcla de celulasas de origen natural,

producidas por un microbio, puede no ser la más eficiente para la rápida conversión en glucosa de la celulosa contenida en la biomasa. Se sabe que las endoglucanasas actúan para producir nuevos extremos de cadena de celulosa, los cuales son ellos mismos sustratos para la acción de celobiohidrolasas y mejoran así la eficiencia de hidrólisis del sistema completo de celulasas. Por consiguiente, el uso de una actividad de celobiohidrolasa incrementada u optimizada puede mejorar en gran medida la producción de etanol.

[185] Por lo tanto, la celobiohidrolasa inventiva encuentra uso en la hidrólisis de celulosa en sus componentes de azúcar. En una incorporación, una celobiohidrolasa variante se agrega a la biomasa antes de la adición de un organismo fermentador. En una segunda incorporación, una celobiohidrolasa variante se agrega a la biomasa al mismo tiempo que un organismo fermentador. Opcionalmente, pueden estar presentes otros componentes de celulosa en cualquiera de las dos incorporaciones.

[186] En otra incorporación, la materia prima celulósica puede pretratarse. El pretratamiento puede ser por temperatura elevada y la adición de cualquiera de ácido diluido, ácido concentrado o solución de álcali diluido. La solución de pretratamiento se agrega durante un tiempo suficiente para hidrolizar por lo menos parcialmente los componentes de hemicelulosa y luego se neutralizan.

[187] El principal producto de la acción de CBH2 sobre la celulosa es celobiosa la cual está disponible para conversión en glucosa mediante actividad de BG (por ejemplo en un producto de celulosa fúngica). Ya sea por el pretratamiento de la biomasa celulósica o por la acción enzimática sobre la biomasa, pueden obtenerse de la biomasa otros azúcares, aparte de la glucosa y la celobiosa. El contenido de hemicelulosa de la biomasa puede convertirse (por hemicelulasas) en azúcares tales como xilosa, galactosa, mannososa y arabinosa. Por consiguiente, en un proceso de conversión de biomasa, la sacarificación enzimática puede producir azúcares que se hacen disponibles para conversiones biológicas o químicas en otros intermedios o productos finales. En consecuencia, los azúcares generados de biomasa tienen uso en una variedad de procesos, aparte de la generación de etanol. Ejemplos de tales conversiones son fermentación de glucosa en etanol (como se reseña por M. E. Himmel y col., páginas 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-diaceto-D-gluconato (Patente de los Estados Unidos No. 6,599,722), ácido láctico (R. Datta y S-P. Tsai páginas 224-236, *ibid*), succinato (R. R. Gokarn, M. A. Eiteman y J. Sridhar páginas 237-

263, *ibid*), 1,3-propanodiol (A-P. Zheng, H. Biebl y W-D. Deckwer páginas 264-279, *ibid*), 2,3-butanodiol (C. S. Gong, N. Cao y G. T. Tsao páginas 280-293, *ibid*), y las conversiones biológicas y químicas de xilosa en xilitol (B. C. Saha y R. J. Bothast páginas 307-319, *ibid*). Véase asimismo, por ejemplo, WO 98/21339.

[188] Las composiciones detergentes de esta divulgación pueden emplear, además de la composición de celulasa (independientemente del contenido de celobiohidrolasa, es decir, libre de celobiohidrolasa, sustancialmente libre de celobiohidrolasa, o mejorada con celobiohidrolasa), un tensioactivo, incluyendo tensioactivos aniónicos, no iónicos y anfotéricos, una hidrolasa, agentes mejoradores de detergente, agentes blanqueadores, agentes azulantes y colorantes fluorescentes, inhibidores de formación de grumos, solubilizantes, tensioactivos catiónicos y similares. Todos estos componentes son conocidos en el arte de los detergentes. Como ya se describió, la composición de celulasa puede agregarse a la composición detergente en un diluyente líquido, en gránulos, en emulsiones, en geles, en pastas, y similares. Las citadas formas son bien conocidas por el artesano experto. Cuando se emplea una composición detergente sólida, la composición de celulasa se formula preferiblemente en forma de gránulos. Preferiblemente, los gránulos pueden formularse con el fin de contener un agente protector de celulasa. Para una discusión más exhaustiva, véase la Patente de los Estados Unidos No. 6,162,782 titulada "Detergent compositions containing cellulase compositions deficient in CBH2 type components," la cual se incorpora aquí por referencia.

[189] Preferiblemente las composiciones de celulasa se emplean desde alrededor de 0.00005 por ciento en peso a alrededor de 5 por ciento en peso relativo a la composición detergente total. Más preferiblemente, las composiciones de celulasa se emplean desde alrededor de 0.0002 por ciento en peso a alrededor de 2 por ciento en peso relativo a la composición detergente total.

[190] Adicionalmente la secuencia de ácido nucleico de CBH2 variante tiene utilidad en la identificación y caracterización de secuencias de ácido nucleico relacionadas. Un número de técnicas útiles para determinar (predecir o confirmar) la función de genes relacionados o productos génicos incluyen, pero no se limitan a, (A) análisis DNA/RNA, tales como (1) sobreexpresión, expresión ectópica, y expresión en otras especies; (2) bloqueo de genes (genética inversa, bloqueo direccionado, silenciamiento génico inducido por virus (VIGS - *viral induced gene silencing*, véase Baulcombe, 100 Years of Virology, Calisher y Horzinek eds., Springer-Verlag, New York, N.Y. 15:189-201, 1999); (3) análisis del estado de

metilación del gen, especialmente regiones reguladoras flanqueantes; y (4) hibridación *in situ*; (B) análisis de productos génicos tales como (1) expresión de proteínas recombinantes; (2) producción de antisueros, (3) inmunolocalización; (4) análisis bioquímicos para actividad catalítica u otra actividad; (5) estado de fosforilación; y (6) interacción con otras proteínas mediante análisis de dos híbridos de levadura; (C) análisis de vías, tal como poner un gen o un producto génico en una vía particular bioquímica o vía de señalización, basándose en su fenotipo de sobreexpresión, o por homología de secuencias con los genes relacionados; y (D) otros análisis que pueden realizarse igualmente para determinar o confirmar la participación del gen aislado y su producto en una vía metabólica particular o de señalización, y ayudar a determinar la función del gen.

DIVULGACIÓN EXPERIMENTAL

[191] La presente divulgación se describe con mayor detalle en los siguientes Ejemplos, los cuales en ninguna forma pretenden limitar el ámbito de la divulgación, tal como se reivindica. La finalidad de las Figuras es ser parte integrante de la especificación y describir la divulgación. Los siguientes Ejemplos se dan con el fin de ilustrar, no para limitar, la divulgación reivindicada.

[192] En la siguiente divulgación experimental se aplican las siguientes abreviaturas: M (molar); mM (milimolar); μ M (micromolar); nM (nanomolar); mol (moles); mmol (milimoles); μ mol (micromoles); nmol (nanomoles); gm (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); MW (peso molecular); sec (segundos); min(s) (minuto/minutos); h(s) y hr(s) (hora/horas); °C (grados Centígrados); QS (cantidad suficiente); ND (no hecho); NA (no aplicable); rpm (revoluciones por minuto); H₂O (agua); dH₂O (agua desionizada); HCl (ácido clorhídrico); aa (aminoácido); bp (par de bases); kb (par de kilobases); kD (kilodaltons); cDNA (DNA copia o complementario); DNA (ácido desoxirribonucleico); ssDNA (DNA de una sola hebra); dsDNA (DNA de doble hebra); dNTP (desoxirribonucleótido trifosfato); RNA (ácido ribonucleico); MgCl₂ (cloruro de magnesio); NaCl (cloruro de sodio); w/v (peso a volumen); v/v (volumen a volumen); g (gravedad); OD (densidad óptica); CNPG (cloro-nitro-fenil-beta-D-glucósido); CNP (2-cloro-4-nitrofenol); APB (bagazo pretratado con ácido - *acid-pretreated bagasse*); PASC (celulosa esponjada con ácido fosfórico - *phosphoric acid swollen cellulose*); PCS (rastajo de maíz pre-tratado con ácido - *acid-pretreated corn stover*); Pi o PI (índice de rendimiento - *performance index*); PAGE (electroforesis en gel de

poliacrilamida - *polyacrylamide gel electrophoresis*); PCR (reacción en cadena de la polimerasa - *polymerase chain reaction*); RT-PCR (PCR de transcripción inversa); y HPLC (cromatografía líquida de alta presión - *high pressure liquid chromatography*).

EJEMPLO 1

Modificación Química de una Preparación de Celulasa *Trichoderma sp.* y Análisis para Examinar CBH2 Variantes

[193] Este Ejemplo describe el tratamiento de una preparación comercial de celulasa *Trichoderma sp.* (complejo enzimático BG LAMINEX (Genencor Division, Danisco US, Inc.) con anhídrido succínico para acetilar los residuos lisina. La acetilación de residuos lisina del complejo enzimático BG LAMINEX altera la carga neta de las proteínas (v.g., mayor carga negativa). Otras modificaciones químicas similares pueden usarse igualmente para convertir la carga positiva de la lisina en un grupo de carga negativa (por ejemplo con tratamiento con anhídrido acetoxisuccínico, anhídrido maleico, anhídrido tartárico y anhídrido ftálico) o incluso a dos cargas negativas (por ejemplo con tratamiento con anhídrido trimelítico, anhídrido cis-aconítico y anhídrido 4-nitroftálico). Otras modificaciones químicas pueden emplearse para eliminar las cargas positivas de los residuos lisina, resultantes en un residuo no cargado (por ejemplo, con tratamiento con anhídrido acético, anhídrido butírico, anhídrido isobutírico, anhídrido hexanoico, anhídrido valérico, anhídrido isovalérico y anhídrido piválico).

[194] Los residuos lisina en una preparación de celulasa se modificaron usando anhídrido succínico, empleando una variedad 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 del complejo enzimático BG LAMINEX en 1 mL de amortiguador 500 mM HEPES pH 8. Una solución de anhídrido succínico (Aldrich) se preparó disolviendo el polvo en DMSO a una concentración final de 500mg/mL antes de la adición del complejo enzimático. Se agregó una alícuota de anhídrido succínico, de modo que se lograra una relación de > 1:100 lisina a ácido succínico en el tubo de reacción. Otro tubo de reacción se preparó con DMSO y enzima únicamente, usando volúmenes similares, para servir como control de proteína no modificada. Los tubos se sometieron a rotación en espiral ("**were vortexed**") y se dejaron a temperatura ambiente de un día para otro. Al día siguiente, un volumen 1:10 de 1M de glicina, con pH 3, se agregó a cada tubo para extinguir la reacción de

anhídrido succínico.

[195] Se confirmó la modificación química comparando proteínas modificadas y no modificadas en geles nativos. Se analizaron alícuotas de cada reacción (modificada y no modificada químicamente) en un gradiente 8-25 %, geles nativos corridos con pH 8.8 a 100 voltios (Phast System gels, GE Healthcare). Las proteínas se visualizaron luego de la tinción del gel con azul de Coomassie, para confirmar que la modificación había sido exitosa. La tinción reveló cambios en la migración de bandas de proteína, confirmando los cambios en carga de los diferentes componentes de proteína de la preparación de celulasa. Las muestras modificadas de la preparación de celulasa de *Trichoderma sp.* fueron más negativamente cargadas que las muestras no modificadas.

[196] Para aislar las proteínas modificadas y no modificadas (control), se desalinizaron alícuotas de 80µl de cada muestra empleando columnas de desalinización giratorias ("**spin desalt columns**") (Pierce). Se midió la absorbencia a 280nm de las muestras desalinizadas (incluyendo el control sin modificación) usando un espectrofotómetro NanoDrop™ (Thermo), en duplicados después de una dilución 1:10 de la muestra para determinar la concentración total de proteína de las muestras.

Determinaciones del Potencial Zeta

[197] Este Ejemplo describe cómo determinar el potencial zeta de una enzima y un sustrato. La presencia de una carga en la superficie de una partícula influye en la distribución de iones en la región interfacial circundante. El resultado es una concentración creciente de iones contrarios de carga opuesta a la de la partícula cerca de la superficie de la partícula. A medida que se aleja de la superficie de la partícula, la distribución heterogénea de iones se hará eventualmente homogénea. La distancia a la cual se obtiene una distribución homogénea se denomina la longitud de Debye ($1/\kappa$) o distancia de apantallamiento ("**screening distance**"), y depende de la resistencia iónica, como se muestra en la siguiente expresión, donde ϵ_0 es la permitividad de espacio libre ($8.854 \times 10^{-12} \text{ F m}^{-1}$), ϵ_r es la permitividad del líquido, k es la constante de Boltzmann ($1.38 \times 10^{-23} \text{ J K}^{-1}$), T es la temperatura en Kelvin, e es la carga electrónica ($1.6022 \times 10^{-19} \text{ C}$), I es la fuerza iónica molar, y NA es la constante de Avogadro ($6.022 \times 10^{23} \text{ mol}^{-1}$).

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

La fuerza iónica molar puede calcularse a partir de la siguiente ecuación, en donde C es la concentración de especie iónica y Z es la valencia.

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

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

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

La capa líquida que circunda la partícula existe como dos partes; una región interna (capa de Stern) en la que los iones están unidos fuertemente y una región externa (difusa) en la que están asociados menos firmemente. En la capa difusa existe una frontera dentro de la cual los iones y partículas forman una entidad estable. Cuando una partícula se mueve, los iones en esta frontera se mueven con ella. Aquellos iones más allá de la frontera no se desplazan con la partícula. El potencial eléctrico en esta frontera, también llamado la superficie de corte hidrodinámico, se define como el potencial zeta.

[198] En la dispersión electroforética de la luz, el potencial zeta ζ se calcula a partir de la movilidad electroforética medida u usando la ecuación de Henry mostrada a continuación, en la que ϵ es la constante dieléctrica, η es la viscosidad de la solución, κ es la longitud inversa de Debye, a es el radio de la partícula, y $f(ka)$ es la función de Henry.

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

[199] Las unidades de κ son longitud recíproca, siendo $1/\kappa$ el "espesor" de la doble capa eléctrica (longitud de Debye). El parámetro a se refiere al radio de la partícula y, por consiguiente, κa es la relación del radio de la partícula al espesor 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 desde $f(0) = 1$ (límite de Hückel) a $f(\infty) = 1.5$ (límite de Smoluchowski). Para partículas pequeñas tales como proteínas en un medio dieléctrico bajo (o fuerza iónica baja), el límite de Hückel de $f(ka) = 1$ es el modelo más apropiado.

[200] Se midieron potenciales zeta de proteínas con el Zetasizer NS (Malvern Instruments, UK) de acuerdo con el principio descrito más arriba. Potenciales zeta de telas manchadas con sangre/leche/tinta (BMI - *blood/milk/ink*) se midieron con el SurPass (Anton-Paar, Austria) usando la implementación de potencial de flujo del principio anterior. De la definición de carga superficial, usualmente expresada en Culombios:

$$q_s = 4pe_o e_r a (1 + ka) z \quad (0.5)$$

Esto puede expresarse igualmente como la carga neta z multiplicada por la carga elemental e $1.6 \cdot 10^{-19} \text{C}$:

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

Por consiguiente, el cambio esperado en el potencial zeta debido a un incremento de carga neta viene dado por:

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

Es también posible medir potenciales zeta usando una técnica de geles nativos (Sparks y col., Journal of Lipid Research, 33:123-130, 1992), como se describe en el Ejemplo 1. La movilidad electroforética medida con geles nativos es usualmente menor que en solución debido al retardo causado por la matriz del gel. Potenciales zeta calculados de esta manera son usualmente menores, comparados con métodos a base de solución. Por consiguiente, nos referimos a ellos como potenciales zeta aparentes, cuando se obtienen mediante la técnica de geles nativos.

[201] La carga efectiva en una formulación dada se expresa como su potencial zeta. El uso del potencial zeta como una escala común de carga permite la comparación de variantes de enzima con diferentes pliegues (v.g., serina proteasas, metaloproteasas, etc.), así como interacciones con diferentes sustratos (v.g., BMI microswatch - micromuestra de tela manchada con sangre/leche/ tinta) bajo las condiciones de interés (v.g., detergente HDL AATCC). Aunque se prefiere el potencial zeta para comparar diferentes pliegues de proteína, las movilidades electroforéticas o cargas medidas proporcionan asimismo una escala absoluta y son adecuadas para comparaciones. El rendimiento BMI, como una función del potencial zeta de la enzima, está perfectamente descrito como una distribución normal estándar, indicada por la línea sólida, con una media μ igual a -9.68 mV , desviación estándar σ de 11.39 mV y un valor pico de $0.4056 [\text{A600-background}]$. Esta distribución es indicada en coordenadas reducidas estándar por la actividad BMI dividida 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 usualmente como $(X - \mu)/\sigma$ en donde X en este caso es el potencial zeta.

Tabla 1-1. Actividad BMI Microswatch de Proteasas		
Nivel de Rendimiento	Puntuación Z	Ventana de Potencial Zeta ζ * para Actividad BMI Microswatch
90%	± 0.46	$-14.92 < \zeta < -4.44$
80%	± 0.65	$-17.08 < \zeta < -2.28$
70%	± 0.84	$-19.25 < \zeta < -0.11$
60%	± 1.00	$-21.07 < \zeta < +1.71$
50%	± 1.18	$-23.12 < \zeta < +3.76$

*Media μ = -9.68 mV, desviación estándar σ = 11.39 mV, potencial zeta ζ = $Z \cdot \sigma + \mu$

Amortiguador de referencia: 5mM HEPES pH 8.0, 2.5mM NaCl

[202] La distribución normal es única para cada mancha de sustrato bajo condiciones de reacción dadas (pH, conductividad, tipo de sal, quelantes de detergente, etc.). Diferentes beneficios o resultados favorables siguen una distribución normal con una propiedad física que se mantiene a través de enzimas de varios pliegues, como es por ejemplo el caso de niveles de expresión y potenciales zeta para variantes de escalera de cargas ASP y NprE. En una distribución normal, el valor pico ocurre en la media. La comparación de cargas de enzima y sustrato en una escala común de potencial zeta revela que el rendimiento BMI óptimo ocurre cuando el potencial zeta medio de la enzima, en este caso -9.68 mV, coincide esencialmente con el potencial zeta de la mancha sustrato, en este caso -8.97 mV, medido bajo las mismas condiciones.

[203] Niveles de rendimiento de distribuciones normales estándar se describen convenientemente en términos de sus puntuaciones z como se indica en la Tabla 1-1 (Véase, Abramowitz y Stegun, Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables, Dover, New York, 9th Ed., 1964). La conversión en potenciales zeta es un conocimiento claro dado de la media y la desviación estándar, que define la distribución para una aplicación dada. En este ejemplo el rendimiento de limpieza medido para un pliegue de proteína se limita a valores de potencial zeta entre -40 mV y + 20 mV. Variantes con un rendimiento de limpieza por encima de 80% de su pliegue óptimo (*es decir*, $z = + 0.65$), se limitan a valores de potencial zeta entre -17.08 mV y -2.28 mV. Variantes con un rendimiento de limpieza por encima de 90% de su pliegue óptimo (*es decir*, $z = + 0.46$), se limitan a valores de potencial zeta entre -14.92 mV y -4.44 mV.

[204] Diferentes manchas de sustrato (v.g., pasto, fluidos corporales, tomate) tienen diferentes potenciales zeta bajo la misma formulación, y la misma mancha de sustrato tiene diferentes potenciales zeta bajo diferentes formulaciones (v.g., North American HDL, European powder dishwashing detergent). No obstante, aunque la carga de la mancha del sustrato varía, se espera que la desviación estándar de la distribución normal permanezca constante. El conocimiento de los potenciales zeta de enzima y sustrato en una formulación detergente dada permite la rápida identificación del nivel de rendimiento esperado para esa variante, así como la dirección y magnitud del cambio de carga necesario, con el fin de alcanzar niveles óptimos de rendimiento. La medición del potencial zeta del sustrato en el medio de reacción deseado permite la optimización de la reacción enzimática sobre el sustrato particulado en ese medio. Cualquier reacción enzimática en cualquier medio puede optimizarse usando un proceso similar.

Optimización de Reacciones en Sustratos que Exhiben Carga Variable

[205] La conversión de celulosa se evaluó por técnicas conocidas en el arte (véase, v.g., Baker y col., Appl Biochem Biotechnol, 70-72:395-403, 1998). En los experimentos se empleó un análisis de conversión celulósica estándar. En este análisis, enzima y sustrato amortiguado se pusieron en recipientes y se incubaron a una temperatura durante un tiempo. La reacción se extinguió con 100 mM de glicina suficiente, pH 11, para llevar el pH de la mezcla de reacción a por lo menos pH 10. Una vez se extinguió la reacción, una alícuota de la mezcla de reacción se filtró a través de una membrana de 0.2 micrones para retirar los sólidos. La solución filtrada se analizó entonces para detectar la presencia de azúcares solubles mediante HPLC, de acuerdo con los métodos descritos en Baker y col., más arriba.

Determinación del Potencial Zeta de Rastrojo de Maíz Pretratado (PCS)

[206] Rastrojo de maíz se pretrató con 2% w/w H₂SO₄, como se describe (Schell y col., J Appl Biochem Biotechnol, 105:69 - 86 [2003]) y se siguió por múltiples lavados con agua desionizada para obtener un pH de 4.5. Para obtener una concentración final de 50mM se agregó acetato sódico, y la solución se valoró volumétricamente a pH 5.0. La concentración de celulosa en la mezcla de reacción fue aproximadamente 7%.

[207] Alícuotas de PCS antes y después de la sacarificación por una mezcla de celulasas comercial, Spezyme CP e Indige 44L, se dosificaron en tubos de

centrifugación Eppendorf de 1.5 mL y ocuparon alrededor de una tercera parte del volumen. Las muestras se centrifugaron a 6.000 rpm durante 5 min, el sobrenadante se cambió por agua Milli-Q™, y el proceso se repitió 5 veces. Se preparó una solución patrón concentrada de 100 mg/mL en agua Milli-Q™ del rastrojo de maíz lavado ("*the rinsed corn stover*"). Esta solución patrón se diluyó a 1 mg/mL en un amortiguador 50 mM de acetato sódico, pH 5.0, para mediciones de potencial zeta. Una alícuota de 1 mL de cada muestra de sustrato se transfirió a una cubeta limpia, desechable, Zetasizer NS™ de Malvern Instruments (RU).

[208] Tabla 1-2 indica que durante el curso de la reacción de sacarificación la carga del sustrato PCS, expresada como potencial zeta, se hizo casi dos veces negativa. Sin entrar en limitaciones teóricas, existen muchas explicaciones para un aumento de la carga negativa neta incluyendo, pero lo limitado a, enriquecimiento en lignina, la porción no reactiva de este sustrato, así como unión no productora o acumulación indeseada de depósitos ("*fouling*") de celulasas enteras y otras proteínas. Existe un potencial zeta óptimo de la enzima para el rendimiento (v.g., grado de reacción y tasa de reacción), que coincide con el potencial zeta del sustrato bajo condiciones del medio de reacción. Diferentes pretratamientos de biomasa pueden influir dramáticamente la carga inicial del sustrato. Si la enzima o el sustrato resultan incompatibles en cuanto al potencial zeta durante el curso de la reacción, la interacción enzima-sustrato ya no será óptima. Este efecto será dramático para cambios de aproximadamente 10 mV, los cuales son el caso para conversión de biomasa.

[209] Estrategias para remediar esta situación incluyen, pero no se limitan a, suministrar una mezcla de enzimas que abarque varias cargas; un método de proceso con lote alimentado, en el que enzimas que poseen diferentes cargas con el nuevo óptimo se suministran en diferentes tiempos de reacción y/o grados de conversión; control de la carga superficial del sustrato mediante la adición de agentes de formulación, particularmente tensioactivos (iónicos y no iónicos) u otras proteínas; control de la carga superficial del sustrato mediante ajustes de pH; ajustes de fuerza iónica durante la reacción con el fin de cambiar los óptimos de carga de la enzima; filtración de membrana, particularmente ósmosis inversa y nanofiltración, para controlar la fuerza iónica de principio a fin de la reacción; adición de quelantes para controlar la fuerza iónica mediante la eliminación de sales; y control de la carga del sustrato biomasa mediante procesos de pretratamiento.

Tabla 1-2. Potencial Zeta de Rastrojo de Maíz Pretratado con Ácido	
Condición PCS	Potencial Zeta
Inicial (antes de la sacarificación)	-12.0 ± 7.00 mV
Después de la sacarificación por Spezyme CP	-22.2 ± 8.67 mV
Después de la sacarificación por Indiage 44L	-22.7 ± 6.84 mV

[210] En los ejemplos descritos a continuación se usaron los siguientes análisis. Todas las desviaciones de los protocolos provistos más adelante se indican en los ejemplos. En estos experimentos, se utilizó un espectrofotómetro para medir la absorbencia de los productos formados luego de concluidas de las reacciones.

Análisis de Hexocinasa para la Medición de Glucosa Residual

[211] La glucosa residual de sobrenadantes de cultivo de *H. jecorina* que expresan CBH2 variantes se midió empleando un análisis de hexocinasa. Un volumen de 5 µl de sobrenadante se agregó a un análisis de hexocinasa de 195 µl glucosa (Instrumentation Laboratory, Breda, Países Bajos) en una placa de microtitulación de 96 pozos (Costar Flat Bottom PS). Las placas se incubaron a temperatura ambiente durante 15 min. Luego de la incubación, se midió la absorbencia a 340 nm OD. Los sobrenadantes de cultivos que expresan glucosa residual, se excluyeron de la combinación ("**pooling**") para estudios posteriores.

Análisis HPLC para Determinación del Contenido de Proteína

[212] La concentración de proteínas CBH2 variantes de sobrenadantes de cultivo combinados se determinó usando una HPLC Agilent 1100 (Hewlet Packard) equipada con una columna Proswift RP 2H (Dionex). 10 microlitros de muestra, mezclados con 50 µl de 10% acetonitrilo en agua desmineralizada filtrada, se inyectaron luego de equilibrar la columna HPLC con 10% acetonitrilo, que contenía 0.01% ácido trifluoroacético. Los compuestos se eluyeron usando un gradiente de 10% a 30% acetonitrilo de 0.3 min a 1 min, seguido por un gradiente de 30% a 65% de 1 min a 4 mins. Se determinaron las concentraciones de proteína de CBH2 variantes de una curva de calibración generada, usando CBH2 natural purificada (6.25, 12.5, 25, 50 µg/ml). Para calcular el índice de rendimiento (P_i o PI), se promediaron la relación de la proteína total (promedio), producida por una variante, y la proteína total (promedio), producida por la natural con la misma dosis.

Determinación de la Actividad Específica por Análisis de Hidrólisis con Celulosa Esponjada con Ácido Fosfórico (PASC)

[213] Hidrólisis de Celulosa: Se preparó celulosa esponjada con ácido fosfórico (PASC) de Avicel, de acuerdo con un método publicado (Walseth, Tappi 35:228, 1971; y Wood, Biochem J, 121:353-362, 1971). Este material se diluyó con amortiguador y agua para obtener una mezcla 1% w/v, de modo que la concentración final de acetato sódico fue 50mM, pH 5.0. 100 microlitros de una suspensión al 1% de PASC en 50 mM amortiguador de acetato sódico (pH 5.0) se dispensaron en una placa de microtitulación de 96 pozos (Costar Flat Bottom PS). 10 microlitros de un sobrenadante de cultivo 5 mg/ml de una cepa eliminada de CBH2 se agregó a la PASC, y 5, 10, 15, o 20 µl de sobrenadantes de cultivo combinados de células de *H. jecorina*, que expresan CBH2 naturales o CBH2 variantes, se agregaron a ella. La eliminación del gen CBH2 de *Hypocrea jecorina* (también conocida como *Trichoderma reesei*) se describe en las Patentes de los Estados Unidos Nos. 5,861,271 y 5,650,322. Se agregaron volúmenes compensadores de amortiguador acetato para compensar diferencias en el volumen total. La placa de microtitulación se selló y se incubó en un incubador con termoestato a 50°C bajo agitación continua a 900 rpm. Después de dos horas, se detuvo la reacción de hidrólisis por la adición de 100 µl de amortiguador glicina, pH 10, a cada pozo. Los productos de la reacción de hidrólisis se examinaron con el análisis PAHBAH.

[214] Análisis PAHBAH: Alícuotas de 150 µl de reactivo de azúcar reductor PAHBAH (5% w/v hidrazida de ácido p-hidroxibenzoico (PAHBAH, Sigma # H9882, disuelto en 0.5 N HC1), (Lever, Anal Biochem, 47:273-279, 1972) se agregaron a todos los pozos de una placa de microtitulación vacía. 10 microlitros de los sobrenadantes de la reacción de hidrólisis se agregaron a la placa de reacción PABA. Todas las placas se sellaron e incubaron a 69°C bajo agitación continua de 900 rpm. Después de una hora, las placas se pusieron sobre hielo durante cinco minutos y se centrifugaron a 720 x g a temperatura ambiente durante cinco minutos. Muestras de 80 µL de las mezclas de reacción PAHBAH desarrolladas se transfirieron a una placa fresca (leída), y se midió la absorbencia a 410 nm en un espectrofotómetro. Un estándar de celobiosa se incluyó como control. Se generó una curva de respuesta a la dosis para la proteína CBH2 natural. Para calcular el índice de rendimiento (P_i o PI), se promediaron la relación del azúcar total (promedio) producido por una variante y el azúcar total (promedio) producido por la natural con la misma dosis.

Determinación de la Actividad Específica por Hidrólisis de Rastrojo de Maíz Pretratado (PCS) con Ácido Diluido

[215] *Rastrojo de Maíz Pretratado (PCS - Pretreated Corn Stover)*: se pretrató rastrojo de maíz con 2% w/w H_2SO_4 como se describe (Schell y col., J Appl Biochem Biotechnol, 105:69-86, 2003) y se siguió por múltiples lavados con agua desionizada para obtener una pasta que tenía un pH de 4.5. A continuación se agregó amortiguador acetato sódico (pH 5.0) a una concentración final de 50mM acetato sódico y, si era necesario, esta mezcla se valoró entonces volumétricamente a un pH 5.0 usando 1N NaOH. La concentración de celulosa en la mezcla de reacción fue aproximadamente 7%. Se agregaron 65 microlitros de esta suspensión de celulosa por pozo a una placa de microtitulación de 96 pozos (Nunc Flat Bottom PS). Se agregaron al PCS 10 microlitros de un sobrenadante de cultivo 5 mg/ml de una cepa eliminada de CBH2, y 5, 10, 15, o 20 μl de sobrenadantes de cultivo agrupados de células de *H. jecorina* que expresan CBH2 naturales o CBH2 variantes se agregaron a ella. Para compensar diferencias en el volumen total, se agregaron volúmenes compensadores de amortiguador acetato. Después de sellarlas, las placas se pusieron en un incubador con termoeestado a 50°C bajo agitación continua de 1.300 rpm durante 5 minutos. Las placas se incubaron luego a 50°C agitando al mismo tiempo a 220 rpm bajo 80% de humedad, para evitar la desecación. Después de 7 días las placas se pusieron sobre hielo durante 5 min y se detuvo la reacción de hidrólisis por la adición de 100 μl de amortiguador glicina, pH 10, a cada pozo. Los productos de reacción de hidrólisis se examinaron con el análisis PAHBAH.

[216] *Análisis PAHBAH*: alícuotas de 150 μl de reactivo de azúcar reductor PAHBAH (5% w/v hidrazida de ácido p-hidroxibenzoico (PAHBAH, Sigma # H9882, disuelta en 0.5 N HC1), (Lever, Anal Biochem, 47:273-279, 1972), se agregaron a todos los pozos de una placa de microtitulación vacía. Se agregaron 10 microlitros de los sobrenadantes de la reacción de hidrólisis a la placa de reacción PABA. Todas las placas se sellaron e incubaron a 69°C bajo agitación continua de 900 rpm. Después de una hora las placas se pusieron sobre hielo durante cinco minutos y se centrifugaron a 720 x g a temperatura ambiente durante cinco minutos. Muestras de 80 μL de las mezclas de reacción PAHBAH desarrolladas se transfirieron a una placa fresca (leída), y se midió la absorbencia a 410 nm en un espectrofotómetro. Un estándar de celobiosa se incluyó como control. Se generó una curva de respuesta a la dosis para la proteína CBH2 natural. Para calcular el índice de rendimiento (P_i o PI), se promediaron la relación del azúcar total (promedio) producido por una variante y el azúcar total (promedio)

producido por la natural con la misma dosis.

Estabilidad de CBH2 Variantes en Presencia de Etanol

[217] Se examinó la estabilidad de CBH2 naturales y CBH2 variantes en la presencia de 4.5% etanol (EtOH) a 49°C. Sobrenadantes de cultivo combinados (80 µL) de células *H. jecorina*, que expresan CBH2 variantes, se agregaron a una placa de 96 pozos (Greiner V-bottom PS), que contenía 10 µl de 40.5% EtOH por pozo. Las placas se sellaron e incubaron en un incubador con termostato a 49°C durante 16 horas con agitación a 900 rpm. Luego de la incubación, las placas se pusieron sobre hielo durante 5 minutos. Se determinó la actividad de CBH2 residual usando el análisis de hidrólisis en celulosa esponjada con ácido fosfórico (PASC), como ya se describió.

[218] Para calcular la actividad residual, el valor del producto formado por la adición de 5, 10, 15 y 20 µl de CBH2 incubada en EtOH al análisis PASC de actividad residual se dividió por el valor del producto formado por la adición de 5, 10, 15 y 20 µl CBH2 libre de EtOH al análisis PASC. Los valores individuales de estas cuatro relaciones se promediaron entonces para producir la actividad residual promedio. Para determinar el valor PI para la variante, el valor de la actividad residual promedio para las variantes se dividió entonces por el promedio de los valores de actividad residual de los controles CBH2 natural.

Termoestabilidad de CBH2 Variantes

[219] Se analizó la termoestabilidad de CBH2 naturales y CBH2 variantes a 53°C. Sobrenadante de cultivo combinado (80 µL) de células *H. jecorina*, que expresan CBH2 variantes, se agregó a una placa de 96 pozos (Greiner V-bottom PS). Las placas se sellaron e incubaron en un incubador con termostato a 53°C durante 16 horas con agitación a 900 rpm. Luego de la incubación, las placas se pusieron sobre hielo durante 5 minutos. Se determinó la actividad de CBH2 residual usando el análisis de hidrólisis con celulosa esponjada con ácido fosfórico (PASC), como se describe más arriba.

[220] Para calcular la actividad residual, el valor del producto formado por la adición de 5, 10, 15 y 20 µl de CBH2 tratada térmicamente al análisis PASC de actividad residual se dividió por el valor del producto formado por la adición de 5, 10, 15 y 20 µl CBH2 no calentada al análisis PASC. Los valores individuales de estas cuatro relaciones se promediaron entonces para producir la actividad

residual promedio. Para determinar el valor PI para la variante, el valor de la actividad residual promedio para las variantes se dividió entonces por el promedio de los valores de actividad residual de los controles CBH2 naturales.

EJEMPLO 2

Evaluación de la Unión de Lignina

[221] Lignina, un biopolímero complejo de fenilpropanoide, es el principal constituyente no carbohidrato de la madera, el cual se une a fibras de celulosa para endurecer y fortalecer las paredes celulares de las plantas. Debido a que se enlaza cruzadamente con otros componentes de la pared celular, la lignina minimiza la accesibilidad de la celulosa y la hemicelulosa a enzimas que degradan la celulosa. Por consiguiente, la lignina se asocia generalmente con una digestibilidad reducida de toda la biomasa vegetal. En particular, la unión de celulosa a lignina reduce la degradación de celulosa por las celulasas. La lignina es hidrófoba y, aparentemente, de carga negativa. Por lo tanto, se contempla la adición de cargas negativas a las celulasas para reducir su unión a lignina.

[222] Como se describe en esta especificación, se preparó una reacción para cuantificar el efecto de la modificación química sobre la capacidad de una preparación de celulasa *Trichoderma sp.* de unir un componente de polímeros vegetales, a saber, lignina. Se recuperó lignina mediante una digestión extensiva de bagazo de caña de azúcar, pretratado con ácido, por celulasas (100 mg Laminex BG/ g de celulosa), seguida por hidrólisis de las celulasas por serina proteasa no específica, exactamente como se describe en Berlin y col., (Applied Biochemistry and Biotechnology, 121:163-170, 2005), excepto que no se realizaron sonicación, secamiento, molienda y cribado, y se agregó al procedimiento un lavado con ácido (0.1 N HCl) para extraer la proteasa, seguido por lavados repetidos con amortiguador acetato (50 mM acetato sódico, pH 5) para devolver la muestra a un pH 5. En síntesis, 50 µL de 1.16% lignina (recuperada de la sacarificación completa del bagazo), preparados en 50mM amortiguador acetato sódico a pH 5 se combinaron con 4 µl de una preparación de celulasa *Trichoderma sp.* modificada o no modificada, desalinizada. Tubos de microcentrifuga, que contenían la muestra de reacción, se incubaron a temperatura ambiente durante 1 hora, y luego se centrifugaron a alta velocidad para separar los materiales solubles de los insolubles. Luego se recogieron 10 µl del sobrenadante de cada tubo. Los tubos de reacción se re-mezclaron e incubaron durante 2 horas más, luego de las cuales se recogieron segundas alícuotas de 10 µl del sobrenadante de cada tubo. Las muestras de sobrenadante

se analizaron por electroforesis en gel de poliacrilamida SDS (SDS-PAGE). La reducción de la intensidad de la banda en preparaciones de celulasa *Trichoderma sp.* fue indicativa de una reducción en la unión de lignina.

EJEMPLO 3

Evaluación de la Unión de Bagazo

[223] El bagazo es la biomasa que queda luego de que la caña de azúcar se ha triturado para extraer su jugo. Una solución que contenía 2% de celulosa de bagazo (tratado con ácido, 28% de sólidos, 57% glucano) se preparó en 50 mM acetato sódico a pH 5. Muestras de preparaciones de celulasa de *Trichoderma sp.* modificada químicamente o no modificada se diluyeron 10 veces en el mismo amortiguador acetato sódico. Alícuotas de las enzimas diluídas se mezclaron con solución de bagazo o amortiguador solo y se incubaron durante 1 hr a temperatura ambiente. Se recogió el sobrenadante y se analizó para detectar actividad de un componente de celulasa, es decir, beta-glucosidasa.

[224] Se midió la actividad de beta-glucosidasa usando el análisis cloro-nitro-fenil-beta-D-glucósido (CNPG – *chloro-nitro-phenyl-beta-D-glucoside*). El análisis CNPG es un análisis cinético en el cual β -glucosidasa convierte CNPG en el producto coloreado 2-cloro-4-nitrofenol (CNP – *2-chloro-4-nitrophenol*). La OD se mide a 405 nm durante un período de 10 minutos a 37°C. Las tasas se obtienen como V_{max} usando el software SpectraMax y luego se convierten a actividad específica (μ M CNP/sec/mg Proteína). En síntesis, 200 μ l de 50 mM amortiguador acetato sódico, pH 5.0, se agregan a cada pozo de una placa de microtitulación de 96 pozos. La placa se cubrió y se puso en un Thermomixer Eppendorf a 37°C durante 15 minutos para permitir equilibrar a temperatura. Luego de equilibrar, se agregaron 5 μ l de las muestras de enzima, diluída serialmente en 50 mM amortiguador acetato sódico, pH 5.0. Una solución patrón concentrada 10 mM CNPG se diluyó 1:5 usando 50 mM amortiguador acetato sódico, pH 5.0, luego 20 μ l de la solución CNPG diluída (2 mM) se agregaron a cada pozo que contenía muestras de enzima. La placa de microtitulación se transfirió a un espectrofotómetro (SpectraMAX, tipo 340; Molecular Devices) ajustado a 37°C, y se leyó la OD a 405 nm durante 0-15 min, leyendo en intervalos ≤ 9 sec.

[225] La cantidad de actividad de beta-glucosidasa de las muestras de enzima de celulasa, que quedó sin unir al sustrato bagazo, fue considerablemente mayor en el caso de la preparación de celulasa *Trichoderma sp.* químicamente modificada. En particular, como se determinó mediante el análisis CNPG, menos de 50% de la

beta-glucosidasa no modificada quedó sin unir (v.g., 50% unida) al sustrato bagazo, mientras que aproximadamente 80% de la bglu quedó sin unir (v.g., 20% unida) al sustrato bagazo. Tomados conjuntamente, los datos de unión de celulasa modificada indican que disminuir las cargas positivas en la celulasa lleva a una unión reducida a un sustrato polimérico de carga más negativa. En este caso, el sustrato polimérico vegetal fue lignina, que quedó en la biomasa tratada con ácido. Se demostró que biomasa tratada con ácido de rastrojo de maíz, un biopolímero vegetal de composición química similar, adopta una carga crecientemente más negativa durante el curso de la sacarificación, como se determinó por la medición del potencial zeta (véase, Tabla 1-2).

EJEMPLO 4

Sacarificación de Bagazo Pretratado con Ácido

[226] Se evaluó la sacarificación de la celulosa presente en bagazo pretratado con ácido (APB - *acid-pretreated bagasse*) que contenía diferentes cantidades de lignina adicional usando preparaciones de celulasa *Trichoderma sp.* químicamente modificada y no modificada, y se analizó mediante HPLC para monitorear la liberación de azúcares, DP1 a DP7. Los resultados se muestran en FIG. 1 como conversión porcentual del sustrato polimérico. En una placa de microtitulación, 200 μ L de APB (3.5% glucano) se prepararon en 50 mM amortiguador acetato sódico a pH 5, y se ajustaron a cantidades variables de lignina. Se agregaron a los pozos 20 microlitros de solución de enzimas de celulasa (LAMINEX BG modificada y no modificada). Las placas se cubrieron con selladores de aluminio y se pusieron en incubadores a 50°C, con agitación, durante 24 hrs o 48 hrs. La reacción se terminó agregando 100 μ l 100 mM glicina, pH 10, a cada pozo. Luego de mezclar concienzudamente, los contenidos de los pozos de la placa de microtitulación se filtraron a través de una placa filtrante Millipore de 96 pozos (0.45 μ m, PES). El filtrado se diluyó en una placa que contenía 100 μ l 10 mM glicina, pH 10, y se midió por HPLC la cantidad de azúcares solubles (DP1 a DP7) producidos. La HPLC Agilent serie 1100 se equipó con una columna de desionización/protección (Biorad Catálogo No. 125-0118) y una columna para carbohidratos con resina iónica a base de plomo Aminex (Aminex Catálogo No. HPX-87P). La fase móvil usada fue agua, con un caudal de 0.6 ml/min. Estándares de azúcares solubles (DP1-DP7) obtenidos de Sigma se diluyeron todos en agua Milli-Q a 100 mg/mL y se usaron para convertir el área pico para los azúcares individuales en concentraciones efectivas de azúcar. El porcentaje de conversión se calculó dividiendo los azúcares medidos de la HPLC por conversión 100% de celulosa en glucosa.

[227] La unión de celulasa a lignina disminuirá su eficiencia de degradar celulosa. Esto se demuestra como una reducción en la conversión de celulosa en la presencia de cantidades crecientes de lignina presente en las reacciones de sacarificación. Esta tendencia persiste en las preparaciones de celulasa no modificada. Sin embargo, hay un incremento del 10% en la conversión de celulosa en las muestras de celulasa modificadas, comparadas con las muestras de celulasa no modificada. Este resultado indica que aumentar la carga negativa de la celulasa reduce la unión no productiva de celulasa a lignina.

EJEMPLO 5

CBH2 Modificada Químicamente Aumentó la Sacarificación de APB

[228] *Trichoderma* CBH1 purificada, CBH2 variante, EG1, EG2 y beta-glucosidasa se modificaron químicamente, como se describe en el Ejemplo 1. La CBH2 variante usada en este experimento tiene múltiples sustituciones P98L/M134V/T154A/I2112V/S316P/S413Y, con números que corresponden a la celulasa CBH2 madura natural), como se describe en la Publicación de los Estados Unidos No. 2006/0205042. La secuencia de aminoácidos de la CBH2 variante madura es como sigue:

QACSSVWGQCGGQNWSGPTCCASGSTCVYSNDYYSQCLPGAASSSSSTRAA
STTSRVSPITTSRSSSATPPPGSTTTTRVPPVSGGTATYSGNPFVGVTLWANAYYASEVS
SLAIPSLTGAMATAAAAVAKVPSFVWLDTLDKTPLMEQTLADIRAANKNGGNYAG
QFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDVRTLLVIEPDS
LANLVTNLGTPKCANAQSAYLECINYAVTQLNLPNVAMYLDAGHAGWLGW PANQ
DPAAQLFANVYKNASSPRALRGLATNVANYNGWNITSPPPYTQGNVYNEKLYIHA
IGPLLANHGWSNAFFITDQGRSGKQPTGQQQWGDWCNVIGTGFGIRPSANTGDSLLD
SFVWVKPGGECDGTSDSSAPRFDYHICALPDALQPAPQAGAWFQAYFVQLLTNANPS
FL (SEQ ID NO:14).

[229] Modificaciones químicas de CBH1, CBH2 variante, EG1, EG2 y beta-glucosidasa se verificaron por su movilidad cambiada en el gel nativo, comparadas con las proteínas no modificadas. CBH1, CBH2 variante, EG1, EG2 y beta-glucosidasa modificadas tienen cargas más negativas. Todas las concentraciones de proteína se midieron usando un espectrofotómetro NanoDrop™ (Thermo). Se estableció una reacción de sacarificación en una placa de microtitulación, en cada pozo de una placa de microtitulación, 150 µL de APB (7% glucano preparada como se describe más arriba para PCS) se prepararon en 50 mM amortiguador acetato sódico a pH 5, 20 µl de la mezcla de enzimas de 21 µg de proteína total se

agregaron de modo que la relación de proteína final a celulosa en cada pozo es 20mg/g. Se elaboraron 6 mezclas de enzimas agregando Bglu modificada o no modificada purificada, CBH2 variante, EG1, o EG2 a un fondo de *T. reesei* en el cual los genes que codifican celobiohidrolasa I (CBHI, Cel7a), celobiohidrolasa II (CBHII, Cel6a), endoglucanasa I (EGI, Cel7b), y endoglucanasa II (EGII, Cel5a) se habían inactivado (véase, US 2007/0128690). En cada mezcla, se agregaron 72.5% de fondo *T. reesei*, 2.5% Bglu, 15% CBH2 variante, 5% EGI, 5% EG2, las cuatro primeras mezclas tienen una proteína que no es modificada, la quinta mezcla tiene toda la proteína no modificada, y la sexta mezcla tiene toda la proteína modificada. La placa se incubó a 50°C durante 72 horas. Se terminó la reacción agregando 100 µl 100 mM glicina, pH 10, a cada pozo. Luego de mezclar concienzudamente, los contenidos de los pozos de la placa de microtitulación se filtraron a través de una placa filtrante Millipore de 96 pozos (0.45 µm, PES). El filtrado se diluyó en una placa que contenía 100 µl 10 mM glicina, pH 10, y se midió por HPLC la cantidad de azúcares solubles (DP1 a DP7) producidos. La HPLC Agilent serie 1100 se equipó con una columna de desionización/protección (Biorad Catálogo No. 125-0118) y una columna para carbohidratos con resina iónica a base de plomo Aminex (Aminex Catálogo No. HPX-87P). La fase móvil usada fue agua con un caudal de 0.6 ml/min. Estándares de azúcares solubles (DP1-DP7) obtenidos de Sigma se diluyeron todos en agua Milli-Q a 100 mg/mL y se usaron para convertir el área pico para los azúcares individuales en concentraciones efectivas de azúcar. El porcentaje de conversión se calculó dividiendo los azúcares medidos de la HPLC por conversión 100% de celulosa en glucosa.

[230] FIG. 2 A muestra que para la sexta mezcla de enzimas (EG2, EG1, CBH2 variante y beta-glucosidasa modificadas) con toda la proteína modificada tiene la máxima conversión de celulosa, y la quinta mezcla de enzimas con toda la proteína no modificada tiene la conversión mínima. Comparando las primeras cuatro mezclas de enzimas, la segunda mezcla de enzimas con la CBH2 no modificada arrojó la concentración mínima. FIG. 2B muestra las ventajas de proteínas modificadas sobre proteínas no modificadas en la conversión de celulosa.

EJEMPLO 6

Preparación de Variantes de Escalera de Cargas CBH2 *T. reesei*

[231] Como se determinó durante el desarrollo de la presente divulgación, la succinilación de residuos de lisina superficial de CBH2 mejoró el rendimiento

sobre APB, y sobre rastrojo de maíz pretratado. La carga de CBH2 modificada fue alrededor de -17 comparada con la CBH2 variante no modificada.

Con esto en mente, se diseñó una escalera de cargas de CBH2 para la determinación de la carga superficial óptima en aplicaciones de rendimiento de celulasa.

[232] SEQ ID NO: 1 divulga la secuencia de DNA que codifica la *Hypocrea jecorina* CBH2 de referencia:

```
atgattgtcggcattctcaccacgctggctacgctggccacactgcagctagtgtgcctctagaggagcggcaagcttgctcaagcgt
ctggggccaatgtggtggccagaattggtcgggtcgcgacttgctgtgcttcggaagcacatgcgtctactccaacgactattactccc
agtgtcttcccggcgtgcaagctcaagctcgtccacgcgcgcgcgctgcagcacttctcgagtatccccacacatcccggtcga
gctccgcgacgcctccacctggttctactactaccagagtacctccagtcggatcgggaaccgctacgtattcaggcaaccctttgttg
gggtcactccttgggccaatgcatattacgcctctgaagttagcagcctcgtattcctagcttgactggagccatggccactgctgcag
cagctgtcgcaaaggttccctctttatgtggctagatactcttgacaagaccctctcatggagcaaaccctggccgacatccgcaccg
ccaacaagaatggcggtaactatgccggacagtttggtgtatgacttgcggatcgcgattgcgctgcccttgccctgaatggcgaa
tactctattgccgatggtggcgctgcgcaaatataagaactatatcgacaccattcgtcaaattgctgtggaatattccgatatccggaccct
cctggttattgagcctgactctcttccaacctggtgaccaacctcggtactccaaagtgtgccaatgctcagtcagcctaccttgagtgc
atcaactacgccgtcacacagctgaaccttccaaatgttgcatgtatttgacgctggccatgcaggatggcttggtggcgccggcaaa
ccaagaccggcgctcagctatttgcaaatgtttacaagaatgcatcgtctccgagagctcttcgaggattggcaaccaatgtcgcca
actacaacgggtggaacattaccagcccccatcgtaacgcaaggcaacgctgtctacaacgagaagctgtacatccacgctattg
gacctcttctgccaatcacggctggtccaacgccttctcatcactgatcaagtcgatcgggaaagcagcctaccggacagcaaca
gtggggagactggtgcaatgtgatcggcaccggatttggtattcgccatccgcaaacactggggactcgttgctggattcgtttgtctg
ggtcaagccaggcggcgagtgtgacggcaccagcgacagcagtcgccacgatttgactcccactgtgcgctcccagatgccttgc
aaccggcgctcaagctggtgcttgggtccaagcctactttgtgcagcttctcacaacgcaaaccatcgttctctgtaa*
```

[233] SEQ ID NO:2 divulga la secuencia de la proteína de longitud total *Hypocrea jecorina* CBH2:

```
MIVGILTTLATLATLAASVPLEERQACSSVWGQCGGQNWSGPTCCASGSTCVYSNDYYSQCL
PGAASSSSSTRAASTTSRVSPSTTSRSSATPPPGSTTTTRVPPVGSATYSGNPFVGVTPWA
NAYYASEVSSLAIPSLTGAMATAAAVAKVPSFMWLDTLDKTPLMEQTLADIRTANKNGGNY
AGQFVVYDLDPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTLLVIEPDSL
NLVTNLGTPKCANAQSAYLECINYAVTQLNLPNVAMYLDAGHAGWLGWPANQDPAAQLFANV
YKNASSPRALRGLATNVANYNGWNITSPPSYTQGNNAVYNEKLYIHAI GPLL ANHGSNAFFI
TDQGRSGKQPTGQQQWGDWCNVIGTGFGIRPSANTGDSLLDSFVWVKPGGECGTSDSSAPR
FDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL*
```

[234] SEQ ID NO:3 divulga la secuencia de la proteína madura *Hypocrea jecorina* CBH2:

QACSSVWGQCGGQNWSGPTCCASGSTCVYSNDYYSQCLPGAASSSSSTRAASTTSRVSPPTS
RSSSATPPPGSTTTTRVPPVGSGTATYSGNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAA
AAVAKVPSFMWLDTLDKTPLMEQTLADIRTANKNGGNYAGQFVVYDLPDRDCAALASNGEYS
IADGGVAKYKNYIDTIRQIVVEYSDIRTLLVIEPDSLNLVTNLGTPKCANAQSAYLECINY
AVTQLNLPNVAMYLDAGHAGWLGW PANQDPAAQLFANVYKNASSPRALRGLATNVANYNGWN
ITSPPSYTOGNAVYNEKLYIHAIGPLL ANHGWSNAFFITDQGRSGKOPTGQQQWGDWCNVIG
TGF GIRPSANTGDSL LDSFVWVKPGGEC DGTSDSSAPRFDSHCALPDALQPAPQAGAWFQAY
FVQLLTNANPSFL*

[235] Residuos seleccionados para inducir mutagénesis incluían residuos lisina, arginina, asparagina, y glutamina expuestos, no conservados, los cuales se seleccionaron para sustitución para introducir cargas negativas. Lisinas succiniladas en CBH2 modificada se identificaron por espectrometría de masas y se seleccionaron para mutagénesis a glutamato, resultando en una diferencia de carga de -2 por cada sustitución. Otros residuos se seleccionaron para sustitución mediante análisis de la estructura tridimensional de CBH2 combinada con alineamiento de aminoácidos de secuencias CBH2 homólogas (véase, v.g., Figura 3 de la Publicación de los Estados Unidos No. US 2006/0205042, que se incorpora en esta especificación por referencia). Residuos superficiales, que fueron altamente variables en el alineamiento de secuencias de aminoácidos CBH2, fueron candidatos para mutagénesis. No obstante, se evitó la acumulación de sustituciones en estrecha proximidad. La arginina se reemplazó con glutamina (carga -1), y glutamina y asparagina se sustituyeron con las respectivas variantes carboxilo (carga -1). Adicionalmente, residuos aspartato y glutamato se seleccionaron para sustitución a los respectivos residuos amina para completar la escalera de cargas (carga +1). Sustituciones específicas CBH2 se relacionan en Tabla 6-1, con todas las posiciones mostradas, con la excepción de R63 y R77, localizadas en el dominio catalítico CBH2. Una carga positiva neta puede crearse por eliminación de un residuo de carga negativa, o por la introducción de un residuo de carga positiva. Análogamente puede crearse una carga negativa neta por eliminación de un residuo de carga positiva o por la introducción de un residuo de carga negativa.

Tabla 6-1. Sustituciones 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			

[236] Para la preparación de una escalera de cargas CBH2, se diseñaron 10 variantes de carga CBH2 (C-1 a C-10), que abarcan un intervalo de +8 a -32, comparadas con CBH2 natural en pasos de 4, como se muestra en Tabla 6-2.

Tabla 6-2. Escalera de Cargas 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
D189N	D189N	WT	K157E	K157E	K157E	K157E	K157E	K157E	K157E	K157E
E208Q	E208Q		K129E	K129E	K129E	K129E	K129E	K129E	K129E	K129E
D211N	D211N			K288E	K288E	K288E	K288E	K288E	K288E	K288E
D405N	D405N			K194E	K194E	K194E	K194E	K194E	K194E	K194E
E244Q					K356E	K356E	K356E	K356E	K356E	K356E
D277N					K327E	K327E	K327E	K327E	K327E	K327E
D151N						R153Q	R153Q	R153Q	R153Q	R153Q
E146Q						R294Q	R294Q	R294Q	R294Q	R294Q
						R203Q	R203Q	R203Q	R203Q	R203Q
						R378Q	R378Q	R378Q	R378Q	R378Q
							N382D	N382D	N382D	N382D
							N344D	N344D	N344D	N344D
							N237D	N237D	N237D	N237D
							N339D	N339D	N339D	N339D
								N289D	N289D	N289D
								N161D	N161D	N161D
								Q204E	Q204E	Q204E

Tabla 6-2. Escalera de Cargas CBH2										
								Q147E	Q147E	Q147E
									N285D	N285D
									N197D	N197D
									N254D	N254D
									N247D	N247D
										Q239E
										Q281E
										R63Q
										R77Q

[237] Las secuencias de aminoácidos de las variantes se retradujeron a DNA y codón optimizado para expresión en *Trichoderma reesei* usando software GeneDesigner (DNA2.0). Se sintetizaron los genes variantes cbh2 optimizados en el codón, y se amplificó el DNA de las variantes de carga superficial CBH2 (SCVs - *surface charge variants*) de los constructos DNA2.0 por PCR usando cebadores: GGHTK22 directo 5'-CACCATGATCGTGGGAATTCTTACTACTC-3' (SEQ ID NO: 15); y GGTHK23 inverso 5'-CTACAAAAACGAAGGGTTTCGCATT-3' (SEQ ID NO: 16). En un experimento, se utilizó mutagénesis dirigida al sitio para introducir mutaciones K129E y K157E (cbh2 variante de carga C-3) en el DNA genómico de CBH2 natural. CBH2 variante de carga C-3 se clonó en pTrex3GM y se expresó como se describe a continuación.

[238] Los productos PCR se purificaron y clonaron en pENTR/TOPO para transformación de células TOP10 de *E. coli*. DNA plasmídico se aisló de colonias individuales, y se verificó la secuencia correcta. CBH2 SCVs se clonaron en pTrex3GM, pTTTpyr(pcbh1), y pTTTpyr(pstpl) como se muestra en la Tabla 6-3.

Tabla 6-3. Clones de Expresión de Variantes de Carga Superficial CBH2			
CBH2 Variante	Vectores de Destino		
	pTrex3gM	pTTTpyr (P _{cbh1})	pTTTpyr(P _{stpl})
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
----------------	----	----	----

[239] Variantes de carga superficial CBH2 (SCV - *surface charge variants*) en *T. reesei*. Se realizó transformación biolística de *T. reesei* con el vector de expresión pTrex3gM, que contenía el marco de lectura abierto de la variante de carga cbh2 C3 (con mutaciones K129E y K157E), usando el siguiente protocolo. Se empleó *T. reesei* en el cual se habían inactivado los genes que codifican la celobiohidrolasa I (CBHI, Cel7a), la celobiohidrolasa II (CBHII, Cel6a), la endoglucanasa I (EGI, Cel7b) y la endoglucanasa II (EGII, Cel5a)). La transformación de la cepa *Trichoderma reesei* mediante el método de transformación biolística se llevó a cabo usando un Sistema de Transferencia de Partículas Biolistic® PDS-1000/he de Bio-Rad (Hercules, CA), siguiendo las instrucciones del fabricante (véase WO 05/001036 y US 2006/0003408). Los transformantes se transfirieron a placas nuevas de selección de acetamida. Se inocularon transformantes estables en placas de microtitulación filtrantes (Millipore), que contenían 200ul/pozo de medio mínimo de glicina (6.0 g/L glicina; 4.7 g/L (NH₄)₂SO₄; 5.0 g/L KH₂PO₄; 1.0 g/L MgSO₄·7H₂O; 33.0 g/L PIPPS; p.H. 5.5) con adición pos-estéril de ~2% de una mezcla de glucosa/soforosa como fuente de carbono, 10 ml/L de 100g/L de CaCl₂, 2.5 ml/L de elementos trazadores de *T. reesei* (400X): 175g/L 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 desarrollaron en cultivo líquido durante 5 días en una cámara rica en O₂, alojada en un incubador a 28° C. Las muestras de sobrenadante de la placa de microtitulación filtrante se obtuvieron usando un múltiple de vacío. Las muestras se corrieron en geles 4-12% NuPAGE (Invitrogen), de acuerdo con las instrucciones del fabricante. Se tiñó el gel con tinte Simply Blue (Invitrogen). La expresión de variantes de carga CHB2 adicionales puede realizarse empleando este método.

EJEMPLO 7

Generación de Variantes de Carga CBH2 en *T. reesei*

[240] El plásmido pTTTpyr-cbh2 que contenía la secuencia codificante de la proteína *Hypocrea jecorina* CBH2 (SEQ ID NO:1) se envió a BASECclear (Leiden, Países Bajos), Gene Art AG (Regensburg, Alemania), y Sloning BioTechnology GmbH (Puchheim, Alemania) para la generación de Bibliotecas de Evaluación de Sitios (SELS - *Site Evaluation Libraries*). El mapa del plásmido de pTTTpyr-cbh2 se muestra en Figura 5. Se hizo una solicitud a los proveedores para la generación de bibliotecas posicionales en cada uno de los sitios en la proteína madura *Hypocrea jecorina* CBH2 (SEQ IDNO:3). La secuencia de aminoácidos de

la proteína de longitud total CBH2 se muestra en SEQ ID NO:2.

[241] SEQ ID NO: 1 divulga la secuencia de DNA que codifica *Hypocrea jecorina* CBH2:

atgattgtcggcattctcaccacgctgggtacgctggccacactcgcagctagtgtgcctctagaggagcggcaagcttgctcaagcgt
ctggggccaatgtggtggccagaattgtcgggtccgacttgctgtgcttccggaagcacatgcgtctactccaacgactattactccc
agtgtcttccggcgctgcaagctcaagctcgtccacgcgcgcgcgctgcagcacttctcgagtatccccacaacatcccggcgga
gctccgcgacgcctccacctggttctactactaccagagtacctccagtcggatcgggaaccgctacgtattcaggcaaccctttgttg
gggtcactccttggccaatgcatattacgcctctgaagttagcagcctcgctattcctagcttgactggagccatggccactgctgcag
cagctgtcgcgaaaggttccctcttttatgtggctagatactcttgacaagacccctctcatggagcaaaccttggccgacatccgcaccg
ccaacaagaatggcggtaactatgccggacagtttggtgtatgacttgccggatcgcgattgcgctgcccttgccctgaatggcgaa
tactctattgccgatggtggcgctgcgcaaatataagaactatatcgacaccattcgtcaaattgctgtggaatattccgatatccggaccct
cctggttattgagcctgactctcttgcgaacctggtgaccaacctcggtactccaaagtgtgccaatgctcagtcagcctaccttgagtgc
atcaactacgcgcgtcacacagctgaaccttccaaatgttgcgatgtatttgacgctggccatgcaggatggcttggctggccggcaaa
ccaagacccggcgcgtcagctatttgcaaatgtttacaagaatgcatcgtctccgagagctcttcgcggattggcaaccaatgtcgcca
actacaacgggtggaacattaccagccccccatcgtaacgcaaggcaacgctgtctacaacgagaagctgtacatccacgctattg
gacctcttcttgcgaatcacggctggtccaacgccttcttcatcactgatcaaggtcgatcgggaaagcagcctaccggacagcaaca
gtggggagactggtgcaatgtgatcggcaccggatttggtattcgccatccgcaaacactggggactcgttgctggattcgttgtctg
ggtcaagccagggcgagtggtgacggcaccagcgacagcagtcgccacgatttgactcccactgtgcgctcccagatgccttgc
aaccggcgcccaagctggtgcttgggtccaagcctactttgtgcagcttctcacaacgcaaacccatcgttctgttaa*

[242] SEQ ID NO:2 divulga la secuencia de proteína de longitud total *Hypocrea jecorina* CBH2

MIVGILTTLATLATLAASVPLEERQACSSVWGQCGQNWSGPTCCASGSTCVYSNDYYSQCL
PGAASSSSSTRAASTTSRVSPPTSRSSSATPPPGSTTTTRVPPVSGTATYSGNPFVGVTPWA
NAYYASEVSSLAIPSLTGAMATAAAVAKVPSFMWLDTLDKTPLMEQTLADIRTANKNGGNY
AGQFVVYDLPDRDCAALASNGEYSIADGGVAKYKNYIDTIRQIVVEYSDIRTLLVIEPDSL
NLVTNLGTPKCANASAYLECINYAVTQLNLPNVAMYLDAGHAGWLGW PANQDPAAQLFANV
YKNASSPRALRGLATNVANYNGWNITSPPSYTQGNVYNEKLYIHAIGPLL ANHGWSNAFFI
TDQGRSGKQPTGQQQWGDWCNVIGTGFGIRPSANTGDSLDSFVWVKPGGEC DGTSDSSAPR
FDSHCALPDALQPAPQAGAWFQAYFVQLLTNANPSFL*

[243] SEQ ID NO:3 divulga la secuencia de proteína madura *Hypocrea jecorina* CBH2:

QACSSVWGQCGGQNWSGPTCCASGSTCVYSNDYYSQCLPGAASSSSSTRAASTTSRVSPPTS
 RSSSATPPPGSTTTTRVPPVGS GTATYSGNPFVGVTPWANAYYASEVSSLAIPSLTGAMATAA
 AAVAKVPSFMWLDTLDKTPLMEQTLADIRTANKNGGNYAGQFVVYDLPDRDCAALASNGEYS
 IADGGVAKYKNYIDTIRQIVVEYSDIRTLLVIEPDSLNLVTNLGTPKCANAQSAYLECINY
 AVTQLNLPNVAMYLDAGHAGWLGW PANQDPAAQLFANVYKNASSPRALRGLATNVANYNGWN
 ITSPPSYTOGNAVYNEKLYIHAIGPLL ANHGWSNAFFITDQGRSGKQPTGQQQWGDWCNVIG
 TGF GIRPSANTGDSL LDSFVWVKPGGECDGTSDSSAPRFDSHCALPDALQPAPQAGAWFQAY
 FVQLLTNANPSFL*

[244] Los plásmidos pTTTpyr-cbh2 purificados (p_{cbh1} , Amp^R, Acetamide^R) que contenían los marcos de lectura abiertos, que codifican las secuencias de CBH2 variante, se obtuvieron de los proveedores especificados anteriormente. Protoplastos de la cepa *H. jecorina* ($\Delta eg1$, $\Delta eg2$, $\Delta cbh1$ $\Delta cbh2$) se transformaron con los constructos pTTTpyr y se desarrollaron en agar selectivo, que contenía acetamida a 28°C durante 7 días. En síntesis, se llevó a cabo transformación biolística de *H. jecorina* empleando el siguiente protocolo y una cepa en la cual se habían inactivado los genes que codifican la celobiohidrolasa I (CBHI, Cel7a), la celobiohidrolasa II (CBHII, Cel6a), la endoglucanasa I (EGI, Cel7b) y la endoglucanasa II (EGII, Cel5a). La transformación de la *H. jecorina* por el método de transformación biolística se realizó usando un Sistema de Transferencia de Partículas Biolistic® PDS-1000/he de Bio-Rad (Hercules, CA), siguiendo las instrucciones del fabricante (véase WO 05/001036 y US 2006/0003408). Se cosecharon las esporas, se esparcieron en agar de acetamida, y se incubaron a 28°C durante 7 días. Se cosecharon las esporas en 15% glicerol y se almacenaron a -20°C para uso posterior. Para la producción de la proteína CBH2 variante, se agregó un volumen de 10 µl de suspensión de esporas a 200 µl de medio mínimo glicina suplementado con una mezcla al 2% de glucosa/soforosa en una placa de filtro PVDF: 6.0 g/L 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 pos-estéril de una mezcla de ~2% glucosa/soforosa como fuente de carbono, 10 ml/L de 100g/L de CaCl₂, 2.5 ml/L de elementos trazadores de *T. reesei* (400X): 175g/L 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 CBH2 variante se desarrolló en cuadruplicado. Luego de sellar la placa con una membrana permeable al oxígeno, las placas se incubaron a 28°C durante 6 días, agitando simultáneamente a 220 rpm. Se recogió el sobrenadante transfiriendo el medio de cultivo a una placa de microtitulación bajo presión reducida y se examinó para detectar la presencia de glucosa residual usando el análisis de hexocinasa, como se describe en el Ejemplo 1.

EJEMPLO 8

Expresión, Actividad y Rendimiento de CBH2 Variantes

[245] Se examinaron variantes de carga *H. jecorina* CBH2 con respecto a varias propiedades de interés. En particular, se analizaron las variantes de celulasa con respecto a la expresión de proteína utilizando el análisis HPLC (HPLC), actividad específica usando el análisis de hidrólisis PASC (Act. PASC) y el análisis de hidrólisis PCS (Act. PCS), estabilidad en la presencia de etanol (relación de EtOH) y termoestabilidad (relación térmica), como se describe en el Ejemplo 1. Los datos de rendimiento para las variantes de carga CBH2 se muestran en la Tabla 8-1. El índice de desempeño (PI - *Performance Index*) es la relación de rendimiento de la celulasa variante a la celulasa progenitora o de referencia. Varios términos expuestos más adelante se emplean para describir la mutación: mutaciones "up" tienen un PI > 1; mutaciones neutras tienen un PI ≥ 0.5, mutaciones no nocivas tienen un PI > 0.05; mutaciones nocivas tienen un PI = 0.05; mutaciones combinables son aquellas mutaciones para las cuales la variante tiene valores de índice de Rendimiento ≥ 0.5 para por lo menos una propiedad. Las mutaciones combinables son mutaciones que pueden combinarse para proporcionar proteínas con índices de desempeño apropiados para una o más propiedades deseadas. Las posiciones en las cuales ocurren mutaciones se clasifican en la siguiente forma: Posiciones no restrictivas tienen ≥ 20% de mutaciones neutras para por lo menos una propiedad; y las posiciones restrictivas tienen < 20% de mutaciones neutras para actividad y estabilidad. Las posiciones completamente restrictivas no tienen mutaciones neutras para actividad o estabilidad.

[246] Estos datos pueden usarse para genomanipular cualquier CBH2. Incluso si la CBH2 que se va a genomanipular tiene un aminoácido diferente del de *Hypocrea jecorina* CBH2 en una posición particular, estos datos pueden emplearse para encontrar una sustitución que alterará las propiedades deseadas, identificando la sustitución de mejor elección, incluyendo una sustitución al amino de *Hypocrea jecorina* CBH2 natural.

[247] Tabla 8-11 muestra valores de Índice de Rendimiento (Pi o PI) para variantes de *Hypocrea jecorina* CBH2. Los Índices de Rendimiento menores de o iguales a 0.05 se fijaron en 0.05 y se indican en cursiva negrita en la Tabla.

Tabla 8-1: Índices de Rendimiento de Variantes de Carga CBH2					
Variante	HPLC	PASC sp.ac.	PCS sp.ac.	Res EtOH	Res Heat
R63A	1.21	1.03	0.94	0.97	1.00

Tabla 8-1: Índices de Rendimiento de Variantes de Carga CBH2					
Variante	HPLC	PASC sp.ac.	PCS sp.ac.	Res EtOH	Res Heat
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
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

Tabla 8-1: Índices de Rendimiento de Variantes de Carga CBH2					
Variante	HPLC	PASC sp.ac.	PCS sp.ac.	Res EtOH	Res Heat
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
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

Tabla 8-1: Índices de Rendimiento de Variantes de Carga CBH2					
Variante	HPLC	PASC sp.ac.	PCS sp.ac.	Res EtOH	Res Heat
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
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

Tabla 8-1: Índices de Rendimiento de Variantes de Carga CBH2					
Variante	HPLC	PASC sp.ac.	PCS sp.ac.	Res EtOH	Res Heat
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
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.D.	1.17	1.15
N382E	2.21	0.16	N.D.	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

Tabla 8-1: Índices de Rendimiento de Variantes de Carga CBH2					
Variante	HPLC	PASC sp.ac.	PCS sp.ac.	Res EtOH	Res Heat
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 del Cambio de Carga sobre la Actividad de Variantes CBH2

[248] En este ejemplo, se evaluó el efecto del cambio de carga sobre la actividad de CBH2 en un análisis, con rastroyo de maíz pretratado (PCS), de la actividad de la celulasa. En síntesis, se determinó el número de ganadores PCS en las CBH2 SELs como una propiedad del cambio de carga. En Tabla 9-1, se determinó la relación de los ganadores observados a esperados (*o/e* - *observed to expected*) en el análisis PCS. Los valores en cursiva negrita son significativamente diferentes del promedio de 10 distribuciones aleatorias más o menos el número de desviaciones estándar (*sd* - *standard deviations*) listadas en las columnas respectivas.

Tabla 9-1: Efecto de la Carga sobre la Actividad de Variantes CBH2 en un Análisis PCS				
PCS	o/e 1 sd	o/e 2 sd	o/e 3 sd	Resultados
Cambio de carga				
-2.00	1.20	1.20	1.20	>90% de confianza más de lo esperado
-1.00	1.40	1.40	1.40	>99% de confianza más de lo esperado
0.00	1.03	1.03	1.03	como se esperaba
1.00	0.46	0.46	0.46	>95% de confianza menos de lo esperado
2.00	0.00	0.00	0.00	como se esperaba

[249] Como se muestra en Tabla 9-1 y Figura 4, disminuir la carga (v.g., -1, -2) resulta en una frecuencia significativamente más alta de CBH2 ganadoras en el análisis PCS, mientras que aumentar la carga (v.g., +1) resulta en una frecuencia significativamente menor de CBH2 ganadoras en el análisis PCS. En conclusión, la actividad CBH2 sobre PCS se correlaciona con la reducción en carga.

Reivindicaciones

1. Una variante de celulasa aislada, en donde la citada variante es una forma madura que tiene actividad celulasa, y en donde la citada variante es una variante de una celulasa progenitora cuya secuencia de aminoácidos es por lo menos 97% idéntica a la secuencia de aminoácidos de SEQ ID NO: 3, variación que consiste de una sustitución en la posición 294 numerada por correspondencia con la secuencia de aminoácidos de SEQ ID NO:3, y en 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, en donde la sustitución en una o más posiciones comprende la eliminación de una o más cargas positivas.
3. La variante de celulasa aislada de la Reivindicación 1, en donde la sustitución es R294Q numerada por correspondencia con la secuencia de aminoácidos de SEQ ID NO:3.
4. La variante de celulasa aislada de la Reivindicación 1, en donde la variante es la variante de sustitución R294Q, R294A, R294C, R294D, R294E, R294F, R294G, R294I, R294L, R294M, R294N, R294P, R294S, R294T, R294V, R294W o R294Y de la secuencia de aminoácidos de SEQ ID NO:3.
5. La variante de celulasa aislada de la Reivindicación 1, en donde la carga neta más negativa es -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 contactar la citada biomasa con la variante de celulasa de la Reivindicación 1.
7. Un método para producir un combustible, que comprende:
contactar una composición de biomasa con una composición enzimática que contiene la variante de celulasa de la Reivindicación 1 para producir una solución de azúcar; y cultivar la solución de azúcar con un microorganismo fermentador bajo condiciones suficientes para producir un combustible.