

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 661 966**

51 Int. Cl.:

<b>C12N 15/09</b>	(2006.01) <b>D21C 5/00</b>	(2006.01)
<b>C12N 1/15</b>	(2006.01) <b>D21H 17/00</b>	(2006.01)
<b>C12N 1/19</b>	(2006.01) <b>D21H 21/10</b>	(2006.01)
<b>C12N 1/21</b>	(2006.01) <b>D06M 15/15</b>	(2006.01)
<b>C12N 5/10</b>	(2006.01) <b>A23K 10/32</b>	(2006.01)
<b>C12N 9/42</b>	(2006.01) <b>A23K 10/14</b>	(2006.01)
<b>C12P 19/14</b>	(2006.01)	
<b>C12P 21/02</b>	(2006.01)	
<b>D21C 5/02</b>	(2006.01)	
<b>D21H 17/22</b>	(2006.01)	

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **17.08.2010 PCT/JP2010/063844**
- 87 Fecha y número de publicación internacional: **24.02.2011 WO11021616**
- 96 Fecha de presentación y número de la solicitud europea: **17.08.2010 E 10809957 (3)**
- 97 Fecha y número de publicación de la concesión europea: **21.02.2018 EP 2468859**

54 Título: **Nueva proteína que tiene actividad  $\beta$ -glucosidasa, y usos de la misma**

30 Prioridad:

**20.08.2009 JP 2009190840**

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

**04.04.2018**

73 Titular/es:

**MEIJI SEIKA PHARMA CO., LTD. (100.0%)  
4-16, Kyobashi 2-chome  
Chuo-ku, Tokyo 104-8002, JP**

72 Inventor/es:

**YOKOYAMA, FUMIKAZU;  
YOKOYAMA, KENGO y  
MAZUKA, NOBUKO**

74 Agente/Representante:

**VALLEJO LÓPEZ, Juan Pedro**

ES 2 661 966 T3

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

## DESCRIPCIÓN

Nueva proteína que tiene actividad  $\beta$ -glucosidasa, y usos de la misma

5 **Campo de la técnica**

La presente invención se refiere a una nueva proteína que tiene actividad  $\beta$ -glucosidasa y usos de la misma. Específicamente, la presente invención se refiere a una nueva proteína que tiene una actividad  $\beta$ -glucosidasa derivada de *Acremonium cellulolyticus*, un análogo y una variante de la proteína, los polinucleótidos que codifican estas proteínas, y un método de producción y usos de estas proteínas.

**Técnica anterior**

La celulosa es un componente constitutivo esencial de células de plantas superiores, y existe ampliamente en la naturaleza. La celulosa es un polímero polisacárido de unidades de glucosa polimerizadas mediante un enlace  $\beta$ -1,4-glicosídico. En la naturaleza, la celulosa se encuentra en un estado cristalino o amorfo. Al enlazarse a otros componentes, tales como la lignina, hemicelulosas, y pectinas, de una manera complicada, la celulosa construye tejidos vegetales.

Celulasa es un término genérico de un grupo de enzimas que rompen la celulosa. En general, las celulasas producidas por microorganismos incluyen distintos tipos de componentes de celulasa. De acuerdo con la especificidad de sustrato, los componentes de celulasa se clasifican en tres tipos: celobiohidrolasa, endoglucanasa, y  $\beta$ -glucosidasa. Se cree que el *Aspergillus niger* que es un hongo filamentoso productor de celulasa, produce 4 tipos de celobiohidrolasas como máximo, 15 tipos de endoglucanasas, y 15 tipos de  $\beta$ -glucosidasas. Presumiblemente, múltiples enzimas de entre estas actúan con modos de reacción distintos compensándose entre ellas para presentar efectos sinérgicos, rompiendo de esta manera la celulosa que es un componente de las paredes de la célula vegetal. Se cree que la  $\beta$ -glucosidasa cataliza una reacción que libera glucosa a partir de los celooligosacáridos, celobiosa o glicósidos con aglicona unida a estos mediante un enlace  $\beta$ -D-glucopiranosil. La  $\beta$ -glucosidasa es una enzima importante en el último estadio de sacarificación de la celulosa y en la liberación de glucosa a partir de glicósidos.

La conversión en etanol de la biomasa tiene las ventajas de que: la materia prima está disponible más fácilmente, se puede evitar la combustión de la materia prima o perderse en el suelo, y el combustible de etanol es medioambientalmente limpio. La madera, los residuos agrícolas, los cultivos herbáceos los residuos sólidos municipales han atraído la atención como biomasa para la producción de etanol. Estos materiales están compuestos principalmente por celulosa, hemicelulosa y lignina. Una vez que la celulosa se convierte en glucosa, la glucosa fermenta fácilmente en etanol mediante levaduras. Por otra parte, la celobiosa no es fácilmente fermentable en etanol por las levaduras, y en consecuencia la celobiosa remanente produce una reducción del rendimiento de etanol. Lo que es más importante es que la celobiosa es un potente inhibidor de las endoglucanasas y celobiohidrolasas. Por esta razón, la acumulación de celobiosa durante la hidrólisis no es deseable para la producción de etanol. En general, los microorganismos productores de celulasa apenas pueden producir  $\beta$ -glucosidasas. Esto da lugar al principal problema de que la celobiosa producida por la hidrólisis enzimática se acumula.

Con el fin de promover la conversión de celobiosa en glucosa, el rendimiento de las  $\beta$ -glucosidasas aumenta por la sobre-expresión de  $\beta$ -glucosidasas en un huésped. Por lo tanto, es un medio eficaz para promover la sacarificación de biomasa en glucosa. En consecuencia, es deseable el aislamiento de un nuevo gen de  $\beta$ -glucosidasa que se introduzca y se exprese en microorganismos productores de celulasa.

Al mismo tiempo, se ha informado que el hongo filamentoso *Acremonium cellulolyticus* produce una celulasa que tiene una alta potencia de sacarificación (Bibliografía no Patente 1) y es muy útil para su uso en piensos y ensilaje (Bibliografía de Patentes 1 a 3). Además, también se ha examinado en detalle el componente de celulasa que contiene (Bibliografía de Patentes 4 a 10). Se ha revelado que se secretan distintos tipos de componentes de celulasa como en otros hongos filamentosos. Particularmente, se ha informado, en cuanto a la actividad  $\beta$ -glucosidasa de entre las celulasas del mismo, por ejemplo, que la actividad es significativamente mayor que la de las celulasas de *Trichoderma reesei* y similares (Bibliografía de Patente 11). Debido a dichas características, se ha enfocado la atención en el *Acremonium cellulolyticus* como diana para aislar a partir del mismo el gen de  $\beta$ -glucosidasa.

Sin embargo, hasta ahora se han aislado pocos genes de *Acremonium cellulolyticus* (Bibliografía de Patentes 9, 10). Además, los genes aislados no se han expresado todavía satisfactoriamente en hongos filamentosos distintos de *Acremonium cellulolyticus*.

El documento B6QHN4 (UniProt 2008) describe una proteína que tiene actividad  $\beta$ -glucosidasa. El documento US 2004/0091469 se refiere a una enzima  $\beta$ -glucosidasa derivada de *Acremonium cellulolyticus*. El documento US 2007/00911469 A1 describe una  $\beta$ -glucosidasa que tiene una mayor actividad que la del *Acremonium cellulolyticus*.

El documento EP 0 927 756 A1 que es idéntico al WO 97/33982 A1 describe una proteína que tiene actividad celulasa obtenida del género *Acremonium*, el ADN del mismo, y un vector de expresión para transformar un microorganismo. El Journal of Bioscience and Bioengineering, Vol. 107 No. 3, 256-261, 2009 describe una cepa mejorada de *Acremonium cellulolyticus* para la producción de celulasa mediante una mutación.

5

**Listado de citas bibliográficas**

[Bibliografía de Patentes]

- 10 [PTL 1] Publicación de Solicitud de Patente Japonesa sin examinar N° Hei 7-264994
- [PTL 2] Patente Japonesa N° 2531595
- [PTL 3] Publicación de Solicitud de Patente Japonesa sin examinar N° Hei 7-236431
- [PTL 4] Publicación de Solicitud de Patente Japonesa sin examinar N° 2001-17180
- 15 [PTL 5] Publicación Internacional N° WO 97/33982
- [PTL 6] Publicación Internacional N° WO 99/011767
- [PTL 7] Publicación de Solicitud de Patente Japonesa sin examinar N° 2000/69978
- [PTL 8] Publicación de Solicitud de Patente Japonesa sin examinar N° Hei 10-066569
- [PTL 9] Publicación de Solicitud de Patente Japonesa sin examinar N° 2002/101876
- 20 [PTL 10] Publicación Internacional N° WO 2002/026979
- [PTL 11] Publicación de Solicitud de Patente Japonesa examinada N° Sho 60-43954

[Bibliografía no Patente]

- 25 [NPL 1] "Agricultural and Biological Chemistry," (Japón), 1987, Vol. 51, p. 65

25

**Sumario de la invención**

**Problema técnico**

30 La presente invención se ha producido en vista de dichas circunstancias. Un objetivo de la presente invención es aislar un nuevo gen de β-glucosidasa de *Acremonium cellulolyticus*. Otro objetivo de la presente invención es aumentar el rendimiento de β-glucosidasa en un huésped que exprese el gen de β-glucosidasa aislado a un alto nivel.

35 **Solución al problema**

Para conseguir los objetivos descritos anteriormente, los presentes inventores han estudiado seriamente métodos de separación y purificación de una β-glucosidasa derivada de *Acremonium cellulolyticus*. Como resultado, finalmente los inventores han identificado satisfactoriamente una nueva β-glucosidasa diferente de la β-glucosidasa que se conocía hasta ahora a partir de *Acremonium cellulolyticus*. Además, el gen que codifica la β-glucosidasa identificada también se ha aislado satisfactoriamente. El gen de β-glucosidasa descubierto por los presentes inventores no se había aislado a pesar de intentos durante años para aislar el gen a partir de *Acremonium cellulolyticus*. La razón era presumiblemente que como la hidrofobia de la proteína codificada por el gen es alta, hacía que la separación y purificación de la misma fuera difícil.

45

Además, los presentes inventores han estudiado seriamente métodos de expresión del gen de β-glucosidasa aislado derivado de *Acremonium cellulolyticus* a un alto nivel en un huésped, produciendo de esta manera que el huésped produzca una β-glucosidasa que tiene una actividad excelente. Como resultado, introduciendo una modificación de múltiples bases en el gen de β-glucosidasa, se ha conseguido satisfactoriamente, por primera vez en el mundo, un alto nivel de expresión de un gen de β-glucosidasa en hongos filamentosos distintos de *Acremonium cellulolyticus*, y el producto de expresión presenta una alta actividad β-glucosidasa. Esto hace posible expresar una β-glucosidasa derivada de *Acremonium cellulolyticus* a alto nivel en un huésped, y aumentar la cantidad de β-glucosidasa producida. Los presentes inventores han descubierto que utilizando la β-glucosidasa o una preparación de celulasas obtenidas a partir del transformante preparado de esta manera, se puede llevar a cabo eficazmente la sacarificación de biomasa en glucosa y varios tratamientos y modificaciones en un sustrato basado en celulosa. Este descubrimiento ha permitido completar de la presente invención.

55

Específicamente, la presente invención se refiere a una nueva proteína que tiene actividad β-glucosidasa derivada de *Acremonium cellulolyticus*, un análogo o variante de la proteína, polinucleótidos que codifican esta proteína y un método de producción y los usos de esta proteína. Más específicamente, la presente invención proporciona lo siguiente.

60

- (1) Un polinucleótido de uno cualquiera de los siguientes puntos (i) a (v), que codifican una proteína que tiene actividad β-glucosidasa:

65

- (i) un polinucleótido que codifica una proteína que comprende una secuencia de aminoácidos de SEQ ID NO: 3;
- (ii) un polinucleótido que comprende la secuencia de base de las posiciones 1-2439 de SEQ ID NO: 1 o la secuencia de las posiciones de bases 218-2847 de SEQ ID NO: 2;
- (iii) un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 3 en la que 40 o menos aminoácidos se han sustituido, eliminado, insertado y/o añadido;
- (iv) un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos que tienen una identidad del 95 % o más con la secuencia de aminoácidos de SEQ ID NO: 3;
- (v) un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos de las posiciones 1-795 de SEQ ID NO: 3.
- (2) El polinucleótido de acuerdo con (1) se deriva de un hongo filamentoso.
- (3) El polinucleótido de acuerdo con (2), en el que el hongo filamentoso es *Acremonium cellulolyticus*.
- (4) Un polinucleótido de uno cualquiera de los siguientes (i) y (ii), que codifica una proteína que tiene una actividad β-glucosidasa:
- (i) un polinucleótido que codifica una proteína que codifica una secuencia de aminoácidos de SEQ ID NO: 5; y
- (ii) un polinucleótido que comprende una región codificante de una secuencia de bases de SEQ ID NO: 4.
- (5) Un polinucleótido que comprende la secuencia de bases de SEQ ID NO: 4 en la que 30 o menos bases se han sustituido, eliminado, insertado y/o añadido, codificando el polinucleótido una proteína que tiene actividad β-glucosidasa.
- (6) Un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos de las posiciones 1-795 de SEQ ID NO: 5.
- (7) Un vector de expresión que comprende el polinucleótido de acuerdo con uno cualquiera de (1) a (6).
- (8) Una célula huésped que comprende el vector de expresión de acuerdo con (7).
- (9) Una proteína que comprende una actividad β-glucosidasa seleccionada de entre (i) a (iv):
- (i) una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 3;
- (ii) una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 3 en la se han sustituido, eliminado, insertado y/o añadido 40 o menos aminoácidos;
- (iii) una proteína que comprende la secuencia de aminoácidos que tiene una identidad del 95 % o más con la secuencia de aminoácidos de SEQ ID NO: 3;
- (iv) una proteína de uno cualquiera de (i) a (iii) de la cual se ha retirado una secuencia de señal, en la que la secuencia de señal es una secuencia de aminoácidos de las posiciones -18 a -1 de la secuencia de aminoácidos de SEQ ID NO: 3.
- (10) Un método para la producción de la proteína de acuerdo con (9), comprendiendo el método las etapas de:
- cultivar la célula huésped de acuerdo con (8); y
- recolectar una proteína expresada por la célula huésped y/o un cultivo de la misma.
- (11) Una preparación de celulasas que comprende la proteína de acuerdo con (9).
- (12) Un método para la degradación o conversión de material celulósico, comprendiendo el método las etapas de:
- tratar el material celulósico con una cualquiera de entre la proteína de acuerdo con (9) y la preparación de celulasas de acuerdo con (11).
- (13) Un método para la producción de un material celulósico, comprendiendo el método las etapas de:
- tratar un material celulósico con una cualquiera de entre la proteína de acuerdo con (9) y la preparación de celulasas de acuerdo con (11); y
- recolectar un material celulósico degradado.
- (14) El método de acuerdo con (13), en el que el material celulósico degradado es un azúcar.
- (15) Una composición detergente que comprende una cualquiera de entre la proteína de acuerdo con (9) y la preparación de celulasas de acuerdo con (11).
- (16) Un método para el tratamiento de una fibra que contiene celulosa, comprendiendo el método las etapas de:
- poner en contacto una cualquiera de entre las proteínas de acuerdo con (9), la preparación de celulasas de acuerdo con (11), y la composición de detergente de acuerdo con (15) con la fibra que contiene celulosa.
- (17) Un método para destintar el papel desechado, caracterizándose el método porque una cualquiera de entre la proteína de acuerdo con (9) y la preparación de celulasas de acuerdo con (11) se utiliza en una etapa de destintado del tratamiento del papel desechado como un agente de decoloración.

(18) Un método para la producción de pasta de papel que tiene un drenaje mejorado, comprendiendo el método las etapas de:

5 tratar la pasta de papel con una cualquiera de entre la proteína de acuerdo con (9) y la preparación de la celulasa de acuerdo con (11).

(19) Un método para la producción de un pienso para animales que tiene una digestibilidad mejorada, comprendiendo el método la etapa de:

10 tratar un pienso para animales con una cualquiera de entre la proteína de acuerdo con (9) y la preparación de celulasas de acuerdo con (11).

### [Efectos ventajosos de la invención]

15 La presente invención proporciona un nuevo gen de  $\beta$ -glucosidasa derivado de *Acremonium cellulolyticus*, y un análogo y una variante del gen para expresar eficazmente una  $\beta$ -glucosidasa en un huésped. Adicionalmente, la presente invención proporciona un huésped que expresa la  $\beta$ -glucosidasa a un alto nivel, y que muestra una excelente actividad de la  $\beta$ -glucosidasa. Esto hace posible obtener una  $\beta$ -glucosidasa derivada de *Acremonium cellulolyticus* como una proteína purificada o una preparación de celulasas con un rendimiento alto.

20

### Breve descripción de los dibujos

[Fig. 1] La Fig. 1 es un esquema que muestra un mapa de las enzimas de restricción en un plásmido pBGLB.

### 25 Descripción de las realizaciones

#### Proteína que tiene actividad $\beta$ -glucosidasa y polinucleótido que codifica la proteína

30 La presente invención proporciona una nueva proteína que tiene una actividad  $\beta$ -glucosidasa y un polinucleótido que codifica la proteína. En la presente invención, la " $\beta$ -glucosidasa" significa una enzima que muestra una actividad  $\beta$ -glucosidasa, es decir la  $\beta$ -D-Glucósido glucohidrolasa EC3.2.1.21. La "actividad  $\beta$ -glucosidasa" significa una actividad de hidrólisis de celo-oligosacáridos, celobiosa, o glicósidos con aglicona unido a los mismos mediante una unión  $\beta$ -D-glucopiranosilo por un exo-mecanismo para producir glucosa.

35 El "polinucleótido" que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención incluye, por ejemplo, un ADN, un ARN, productos modificados o quimeras de los mismos, y es preferentemente un ADN. El ADN incluye un ADNc, un ADN genómico, y un ADN sintetizado químicamente. La secuencia de bases de un ADNc que han aislado los presentes inventores, y que codifica la nueva  $\beta$ -glucosidasa (a la que se hace referencia de aquí en adelante como "acBGLB") derivada de *Acremonium cellulolyticus*, se muestra en la SEQ ID NO: 1. La secuencia de bases del ADN genómico se muestra en la SEQ ID NO: 2. Además, la secuencia de aminoácidos de acBGLB codificada por estos ADN se muestran en la SEQ ID NO: 3.

40

45 Una realización preferida del polinucleótido de la presente invención es un polinucleótido que codifica la acBGLB que comprende la secuencia de aminoácidos de SEQ ID NO: 3. Un ejemplo de la misma incluye un polinucleótido que comprende una región codificante de la secuencia de bases de una cualquiera de las SEQ ID NO: 1 y 2.

Además, la presente invención comprende un polinucleótido que codifica una proteína funcionalmente equivalente a la acBGLB. Ejemplos de dicho polinucleótido incluye mutantes, derivados, alelos, variantes y homólogos de la acBGLB. En el presente documento, la frase "funcionalmente equivalente" significa que la proteína diana tiene una actividad  $\beta$ -glucosidasa. Preferentemente, cuando se comparan, la proteína diana tiene un 70 % o más, preferentemente un 80 % o más, más preferentemente un 90 % o más, y más preferentemente un 95 % o más de actividad  $\beta$ -glucosidasa de la acBGLB. La actividad  $\beta$ -glucosidasa de la proteína diana y la acBGLB se pueden evaluar como la actividad de producción de 1  $\mu$ mol de p-nitrofenol a partir de p-nitrofenil- $\beta$ -glucósido en 1 minuto cuando se mide por un método descrito en la bibliografía (Methods in ENZYMOLOGY, vol. 160, Biomass Part A Cellulose and Hemicellulose, Willis A. Wood ed. p 109-110).

55

Una realización del polinucleótido que codifica una proteína funcionalmente equivalente a la acBGLB es un polinucleótido que codifica una proteína que tiene actividad  $\beta$ -glucosidasa, comprendiendo la proteína una secuencia de aminoácidos de SEQ ID NO: 3 en el que uno o más aminoácidos se han sustituido, eliminado y/o añadido.

60

65 El número de restos de aminoácido modificados es preferentemente de 1 a 40, más preferentemente de 1 a 20, más preferentemente 1 a 8 y más preferentemente de 1 a 4. La modificación de aminoácidos es preferentemente una sustitución conservadora. Una "sustitución conservadora" significa que se sustituye al menos un resto de aminoácido con otro resto de aminoácido químicamente similar de tal manera que la actividad del polipéptido no cambia sustancialmente. Ejemplos de la misma incluye un caso en el que cierto resto de aminoácido hidrófobo se sustituye por otro resto de aminoácido hidrófobo, un caso en el que cierto resto de aminoácido polar se sustituye con otro

resto de aminoácido polar que tenga la misma carga, y otros casos. Ejemplos específicos de aminoácidos no polares (hidrófobos) incluyen alanina, valina, isoleucina, leucina, prolina, triptófano, fenilalanina, metionina, y similares. Ejemplos específicos de aminoácidos polares (neutros) incluyen glicina, serina, treonina, tirosina, glutamina, asparagina, cisteína, y similares. Ejemplos específicos de aminoácidos cargados positivamente (básicos) incluyen arginina, histidina, lisina, y similares. Ejemplos específicos de aminoácidos cargados negativamente (ácidos) incluyen ácido aspártico, ácido glutámico, y similares.

El polinucleótido que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención es preferentemente un polinucleótido que codifica una proteína que comprende una secuencia de aminoácidos de SEQ ID NO: 5 (por ejemplo, un polinucleótido que comprende una región codificante de una secuencia de bases de SEQ ID NO: 4), particularmente cuando se expresa en *Trichoderma viride*. En la secuencia de bases de SEQ ID NO: 4, se ha modificado un 13,2 % o más de bases en comparación con la secuencia de bases (SEQ ID NO: 1) del polinucleótido que codifica acBGLB. Además, con respecto a la secuencia de aminoácidos que se codifica, se modificaron 16 tipos de aminoácidos entre los 20 tipos de aminoácidos en la secuencia de bases. Se tuvo en consideración la frecuencia de distribución de codones que se utiliza en un huésped al determinar un codón correspondiente a cada aminoácido. Esto hace posible la expresión en *Trichoderma viride*, y el producto de expresión presenta satisfactoriamente una alta actividad  $\beta$ -glucosidasa. Una vez que se obtiene dicha secuencia preferida los expertos en la técnica podrían, basándose en esta secuencia, modificar adicionalmente la secuencia de bases para obtener un polinucleótido que se pueda expresar en *Trichoderma* de la misma manera que el polinucleótido que comprende la región codificante de la secuencia de bases de SEQ ID NO: 4. Por lo tanto, la presente invención proporciona un polinucleótido que comprende la secuencia de bases de SEQ ID NO: 4 en la que se han sustituido, eliminado, insertado y/o añadido una o más bases (preferentemente 30 bases o menos, más preferentemente 20 bases o menos, más preferentemente 10 bases o menos, y más preferentemente 5 bases o menos), codificando el polinucleótido una proteína que tiene una actividad  $\beta$ -glucosidasa y que se puede expresar en *Trichoderma viride*. Una realización preferida de dicho polinucleótido es un polinucleótido capaz de mejorar la actividad  $\beta$ -glucosidasa cuando se expresa en un transformante de *Trichoderma viride* 5 veces o más (preferentemente 7 veces o más) en comparación con una actividad  $\beta$ -glucosidasa en una cepa parental de *Trichoderma viride* (cepa de *Trichoderma viride* original no deficiente en un gen para biosíntesis de uracilo) (véase el Ejemplo 5).

En la presente invención, otra realización del polinucleótido que codifica una proteína funcionalmente equivalente a la acBGLB, es un polinucleótido que codifica una proteína que tiene una actividad  $\beta$ -glucosidasa, comprendiendo la proteína una secuencia de aminoácidos que tiene una identidad del 90 % o más con la secuencia de aminoácidos de SEQ ID NO: 3. En el presente documento, la "identidad" es un valor numérico que se calcula utilizando los parámetros por defecto (ajustes iniciales) en FASTA3 [Science, 227, 1435-1441 (1985), Proc. Natl. Acad. Sci. USA, 85, 2444-2448 (1988), <http://www.ddbj.nig.ac.jp/E-mail/homology-j.html>] que es un programa de búsqueda de homologías conocidos por los expertos en la técnica. La identidad puede ser preferentemente una identidad de un 95 % o más, más preferentemente una identidad de un 98 % o más, y particularmente preferentemente una identidad de un 99 % o más.

En la presente invención, otra realización del polinucleótido que codifica una proteína funcionalmente equivalente a la acBGLB es un polinucleótido que codifica una proteína que tiene una actividad  $\beta$ -glucosidasa, hibridándose el polinucleótido en condiciones rigurosas con un polinucleótido que comprende la secuencia de bases de una cualquiera de las SEQ ID NO: 1 y 2. En el presente documento "condiciones rigurosas" significa en las que se lleva a cabo un procedimiento de lavado en membrana después de la hibridación a alta temperatura en una solución que tiene una concentración baja en sales, y significa condiciones de lavado, por ejemplo a una concentración de 23 SSC (12 SSC: 15 mmol/l de citrato sódico y 150 mmol/l de cloruro sódico) y en una solución con un 0,5 % de SDS a 60 °C durante 20 minutos.

Además, la presente invención proporciona un polinucleótido que codifica la acBGLB o una proteína funcionalmente equivalente a la misma, en cuya secuencia de bases se ha eliminado una secuencia de señal. La secuencia de señal de la acBGLB es una secuencia de aminoácidos de las posiciones -18 a -1 de la secuencia de aminoácidos de SEQ ID NO: 3.

Se puede añadir cualquier polipéptido al extremo N y/o el extremo C de cada secuencia de aminoácidos correspondiente a una parte de proteína madura de la proteína de la presente invención de manera que no tenga influencia en la actividad  $\beta$ -glucosidasa. Ejemplos de dicha secuencia de polipéptido incluye una secuencia de señal, un marcador de detección (por ejemplo, un marcador FLAG), y un polipéptido de purificación [por ejemplo, glutatión S-transferasa (GST)].

El polinucleótido que codifica la proteína que tiene actividad  $\beta$ -glucosidasa de la presente invención se puede preparar utilizando medios convencionales para los expertos en la técnica. En la preparación de un ADN genómico que codifique la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención, por ejemplo, primero, se extrae un ADN genómico de un microorganismo diana tal como *Acremonium cellulolyticus*, por un método utilizado normalmente. El ADN genómico se digiere con una enzima de restricción apropiada y entonces se liga a un vector apropiado. De esta manera, se prepara una biblioteca de ADN de *Acremonium cellulolyticus*. Como vector, se

pueden utilizar distintos vectores como, por ejemplo, un vector plasmídico, un vector fago, un vector cósmido y un vector BAC. A continuación, se prepara una sonda apropiada basándose en la secuencia de bases (por ejemplo, la SEQ ID NO: 2) del polinucleótido que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención, y se puede aislar un ADN genómico deseado a partir de la biblioteca de ADN genómico mediante hibridación. De manera alternativa, se prepara un cebador basándose en la secuencia de bases (por ejemplo, la SEQ ID NO: 2) del polinucleótido que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención, y se lleva a cabo una PCR utilizando el ADN genómico de *Acremonium cellulolyticus* como matriz. Un fragmento de ADN amplificado de esta manera se liga en un vector apropiado. En consecuencia, se puede aislar un ADN genómico. Entre tanto, se prepara un ADNc que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención, por ejemplo, primero, se sintetiza un ADNc basándose en un ARNm extraído de un microorganismo diana tal como *Acremonium cellulolyticus*. El ADNc se digiere con una enzima de restricción y se liga en un vector apropiado. De esta manera, se prepara una biblioteca de ADNc de *Acremonium cellulolyticus*. A continuación, se prepara una sonda basándose en la secuencia de bases (por ejemplo, la SEQ ID NO: 1) del polinucleótido que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención, y se puede aislar un ADNc deseado de la biblioteca de ADNc mediante hibridación. De manera alternativa, se prepara un cebador basándose en la secuencia de bases (por ejemplo, la SEQ ID NO: 1) del polinucleótido que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención, y se lleva a cabo una PCR utilizando el ADNc de *Acremonium cellulolyticus* como matriz. Un fragmento de ADN amplificado de esta manera se liga en un vector apropiado. En consecuencia, se puede aislar un ADNc deseado. Además, el polinucleótido que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención se puede obtener artificialmente por síntesis química.

La presente invención proporciona un vector de expresión que comprende: el polinucleótido que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención, siendo el polinucleótido replicable en un microorganismo huésped; y la proteína codificada a partir de la secuencia de polinucleótido está en un estado expresable. El vector de expresión de la presente invención se puede construir basándose en un vector de auto-replicación, es decir, por ejemplo, un plásmido que existe como un elemento extra-cromosómico, y que se replica independientemente de la replicación del cromosoma del microorganismo huésped. De manera alternativa, el vector de expresión de la presente invención se puede replicar junto con el cromosoma del microorganismo huésped, después de introducirlo en el microorganismo huésped y se incorpora en el genoma del mismo. Como procedimiento y método para construir el vector de expresión de la presente invención, se puede utilizar cualquier procedimiento y cualquier método que se utiliza comúnmente en el campo de la modificación genética.

Para expresar la proteína que tiene una actividad  $\beta$ -glucosidasa después de que el vector de expresión de acuerdo con la presente invención se haya introducido en un microorganismo huésped, el vector de expresión de acuerdo con la presente invención comprende deseablemente, además del polinucleótido que codifica la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención, una secuencia de polinucleótido para regular la expresión, un marcador genético para seleccionar un microorganismo, y similar. Ejemplos de la secuencia de polinucleótido para regular la expresión incluyen secuencias de polinucleótido que codifican un promotor, un terminador y un péptido de señal. El promotor no está limitado particularmente, a condición de que presente actividad transcripcional en el microorganismo huésped. El promotor puede derivarse de un microorganismo homólogo o heterólogo con respecto al microorganismo huésped. Además, el péptido de señal no está limitado particularmente, a condición de que el péptido de señal contribuya a la secreción de la proteína en el microorganismo huésped. El péptido de señal se puede derivar de un microorganismo homólogo o heterólogo del microorganismo huésped. Además, el marcador genético se puede seleccionar según esté de acuerdo al método de selección del transformante. Por ejemplo, se puede utilizar un gen que codifique una resistencia a fármacos o un gen que complemente la auxotrofia.

Además, la presente invención proporciona un microorganismo transformado con el vector de expresión. El microorganismo huésped que se utiliza en la presente invención no está particularmente limitado. Ejemplos de los mismos incluyen hongos filamentosos, levaduras, *Escherichia coli*, actinomicetos, y similares. Ejemplos de las células de levadura incluyen las que pertenecen a los géneros *Saccharomyces*, *Hansenula*, y *Pichia*. Un ejemplo preferido de célula de levadura es *Saccharomyces cerevisiae*. Adicionalmente, ejemplos de hongos filamentosos incluyen los que pertenecen a los géneros *Humicola*, *Aspergillus*, *Trichoderma*, *Fusarium*, y *Acremonium*. Ejemplos preferidos de los hongos filamentosos incluyen *Humicola insolens*, *Aspergillus niger*, *Aspergillus oryzae*, *Trichoderma viride*, *Fusarium oxysporum*, y *Acremonium cellulolyticus*. La transformación de estos microorganismos con el vector de expresión de la presente invención se puede llevar a cabo de acuerdo con cualquier método utilizado en este campo.

La proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención (o una preparación de celulasas de la presente invención que se va a describir posteriormente) se puede recolectar en un cultivo (por ejemplo, células cultivadas, sobrenadante del cultivo) que se obtiene cultivando los transformantes preparados de esta manera en un medio apropiado. El cultivo del transformante y las condiciones del mismo pueden ser sustancialmente las mismas que las del organismo que se va a utilizar. Además, como método para la recolección de la proteína diana después de cultivar el transformante, se puede utilizar cualquier método que se utilice comúnmente en este campo. Por ejemplo, después del terminar el cultivo del transformante, se puede utilizar un fluido sobrenadante obtenido por retirada desde el cultivo por centrifugación o similar, como una enzima en bruto. Además, este fluido sobrenadante

se concentra por un método de ultrafiltración o similar, y se añade al mismo un antiséptico o similar. El resultado se puede utilizar como una enzima concentrada. Además, después de la concentración se puede preparar una enzima en polvo mediante un método de secado por pulverización o similar. La proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención (o la preparación de celulasas de la presente invención) se puede obtener, si fuera necesario, por purificación parcial o alta purificación de la enzima concentrada o la enzima en polvo. Como método de purificación se pueden utilizar métodos convencionales, por ejemplo un método de precipitación proteica por adición de sal con sulfato amónico o similar, un método de precipitación por disolvente orgánico con un alcohol o similar, un método de separación en membrana, y un método de separación cromatográfica utilizando un intercambiador iónico, un vehículo de cromatografía hidrófoba, un vehículo de filtración en gel, o similar, solos o en combinación según sea apropiado. La presente invención proporciona dicho método para la producción de la proteína que tiene actividad  $\beta$ -glucosidasa de la presente invención (o la preparación de celulasas de la presente invención).

#### Preparación de celulasas

La presente invención proporciona una preparación de celulasas que comprende la proteína descrita anteriormente que tiene una actividad  $\beta$ -glucosidasa de la presente invención. La preparación de celulasas de la presente invención puede comprender una proteína diferente a la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención. Como proteínas diferentes, la preparación de celulasas de la presente invención puede comprender, por ejemplo, una  $\beta$ -glucosidasa distinta de la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención, una hemicelulasa, endoglucanasa, celobiohidrolasa, aminopeptidasa, amilasa, carbohidrasa, carboxipeptidasa, catalasa, quitinasa, cutinasa, ciclodextrina glicosiltransferasa, desoxirribonucleasa, esterasa,  $\alpha$ -galactosidasa,  $\beta$ -galactosidasa, glucoamilasa,  $\alpha$ -glucosidasa, haloperoxidasa, invertasa, lacasa, lipasa, manosidasa, oxidasa, pectinasa, péptido glutaminasa, peroxidasa, fitasa, polifenol oxidasa, proteasa, ribonucleasa, transglutaminasa, orxilanasa. La proteína diferente de la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención comprendida en la preparación de celulasas de la presente invención se puede derivar del transformante que expresa la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención o puede ser una que se añade independientemente.

La preparación de celulasas de la presente invención se puede producir mientras que se mezcla con un vehículo o medio contenido generalmente, por ejemplo, un excipiente (por ejemplo, lactosa, cloruro sódico, sorbitol, o similares), un tensioactivo, un antiséptico, o similar. Además la preparación de celulasas de la presente invención puede producirse de una forma apropiada, por ejemplo, en polvo o líquido.

#### Usos de la proteína que tiene actividad $\beta$ -glucosidasa o preparación de celulasas

La presente invención proporciona un método para degradar o convertir un material celulósico, comprendiendo el método: tratar un material celulósico con una cualquiera de entre la proteína que tiene actividad  $\beta$ -glucosidasa de la presente invención y la preparación de celulasas de la presente invención. Además, la presente invención proporciona un método para producir un material celulósico degradado o convertido, comprendiendo el método: tratar un material celulósico; y después recolectar un material celulósico degradado. El material celulósico es normalmente una biomasa. Ejemplos del mismo incluye, pero no se limita a, paja de arroz, bagazo, rastrojo de maíz, hollejos de frutas tales como el coco, desechos de madera, y materiales obtenidos sometiendo estos a un pre-tratamiento apropiado. La proteína que tiene una actividad  $\beta$ -glucosidasa o la preparación de celulasas que se utiliza para el tratamiento del material celulósico puede estar en forma de un caldo de fermentación en bruto del cual se eliminan o no se eliminan las células, o puede estar en forma de una preparación semi-purificada o purificada. En el proceso de fermentación utilizando la biomasa, el transformante de la presente invención se puede utilizar como fuente de producción de la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención. Se pueden introducir en el transformante, distintos genes de celulosa o un gen que codifique otra enzima eficaz en el procesamiento de la biomasa. Los métodos de la presente invención se pueden utilizar para producir un azúcar (por ejemplo, monosacáridos, disacáridos, polisacáridos) como materia prima química o de fermentación a partir de la biomasa, por ejemplo. El azúcar obtenido de esta manera sirve como materia prima para producir, por ejemplo, etanol, plásticos, u otros productos o intermediarios.

Además, la presente invención proporciona una composición detergente que comprende una cualquiera de entre la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención y la preparación de celulasas de la presente invención. La composición detergente de la presente invención también puede comprender un tensioactivo (aniónico, no iónico, catiónico, anfotérico o tensioactivo zwitteriónico, o puede ser una mezcla de estos). Además, la composición detergente también puede comprender otros componentes detergentes conocidos en el campo, por ejemplo, un adyuvante de detergencia, un blanqueante, un activador del blanqueante, un inhibidor de corrosión, un agente secuestrante, un polímero de liberación en el suelo, un saborizante, otras enzimas (proteasas, lipasas, amilasas, y similares), un estabilizante enzimático, un auxiliar de formulación, un abrillantador óptico, y/o un agente espumante, y similares.

Además, la presente invención proporciona un método para tratar la fibra que contiene celulosa, comprendiendo el método las etapas de poner en contacto una cualquiera de entre la proteína que tiene una actividad  $\beta$ -glucosidasa

de la presente invención, la preparación de celulasas de la presente invención, y la composición detergente, con la fibra que contiene celulosa. Ejemplos de características de la fibra que contiene celulosa que se van a mejorar con el método de tratamiento de la presente invención incluye: (1) tacto y apariencia de la fibra mejorados por la reducción de peso, (2) variaciones locales del color proporcionado por la fibra que contiene fibra coloreada, es decir, apariencia y textura de lavado a la piedra proporcionadas a la fibra que contiene celulosa coloreada, (3) claridad del color de la fibra que contiene celulosa coloreada, (4) suavidad (tiempo más lento de comienzo de rigidez del material, reducción de la rigidez), y (5) eliminación de pelusa (tiempo más lento del comienzo de la formación de pelusa, reducción de pelusa).

Además, la presente invención proporciona un método para destintar papel desechado, caracterizado el método porque una cualquiera de la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención y la preparación de celulasas de la presente invención se utiliza en una etapa de destintado del tratamiento del papel desechado con un agente destintante.

Además, la presente invención proporciona un método para producir pasta de papel que tiene un drenaje mejorado, comprendiendo el método la etapa de: tratar la pasta de papel con una cualquiera de entre la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención y la preparación de celulasas de la presente invención. De acuerdo con la presente invención, el drenaje de la pasta de papel se puede mejorar sin una reducción significativa de la fuerza. Ejemplos de pasta de papel que es diana del tratamiento incluyen, pero no se limitan a, pasta de papel desechado, pasta de cartón reciclado, pasta de papel de envolver, pasta con sulfito, tratada termo-mecánicamente y otras pastas de alto rendimiento.

Además, la presente invención proporciona un método para producir un pienso para animales que tiene una digestibilidad mejorada, comprendiendo el método la etapa: de tratar un pienso para animales con una cualquiera de entre la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención y la preparación de celulasas de la presente invención. De acuerdo con el método de la presente invención, se puede mejorar la digestibilidad de glucano en el pienso en el cuerpo de un animal, por ejemplo.

#### Hongos filamentosos que expresan cantidades reducidas de la proteína que tiene actividad $\beta$ -glucosidasa

La presente invención proporciona un hongo filamentosos que expresa una cantidad reducida de la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención. El hongo filamentosos es preferentemente un hongo filamentosos que pertenece al género *Acremonium*, y más preferentemente *Acremonium cellulolyticus*. La expresión de la proteína que tiene una actividad  $\beta$ -glucosidasa de la presente invención (proteína endógena) en el hongo filamentosos se puede reducir utilizando técnicas generales, tales como, por ejemplo, un método de ARN de interferencia, método de ARN-ADN antisentido, y recombinación homóloga. Los métodos para preparar las moléculas de polinucleótido (por ejemplo, un ARNip, un ARN antisentido, un ADN antisentido, un polinucleótido que comprende una secuencia homóloga a un ADN diana para la recombinación, y similares) utilizados en estas técnicas, para preparar vectores que comprenden estos polinucleótidos, e introducir los vectores en huéspedes se conocen por los expertos en la técnica. Cuando la celulosa ampliamente distribuida entre las plantas y demás se degrada utilizando el hongo filamentosos producido de esta manera, no se produce glucosa que es un producto de degradación final en el procedimiento de degradación, sino que se produce selectivamente la celobiosa en la que dos moléculas de glucosa se unen por un enlace  $\beta$ -1,4. La celobiosa tiene un sabor dulce pero no se degrada en el cuerpo humano. En consecuencia, la celobiosa es útil como edulcorante para una alimentación sana y alimentos para pacientes diabéticos, una materia prima cosmética, o una materia prima para medicamentos. Utilizando el hongo filamentosos de la presente invención, se pueden proporcionar las materias primas de estos productos con un coste reducido.

#### **Ejemplos**

La presente invención se describirá más específicamente por medio de Ejemplos. Sin embargo, la presente invención no se limita a los Ejemplos posteriores mientras esté aún en la esencia de la presente invención.

#### **Ejemplo 1. Purificación de $\beta$ -glucosidasa de *Acremonium cellulolyticus***

Se preparó una enzima de celulasa en polvo secada por pulverización a partir de *Acremonium cellulolyticus*, y se disolvió en un tampón de Tris-HCl (0,05 M, pH 7,0) que contenía 0,5 M de  $(\text{NH}_4)_2\text{SO}_4$ . Se retiraron las impurezas por centrifugación refrigerante de altas prestaciones. Un sobrenadante obtenido de esta manera se purificó como material de partida para la purificación enzimática de acuerdo con un método que se presenta posteriormente.

#### (a) Cromatografía hidrófoba (Parte 1)

En un tampón Tris-HCl (0,05 M, pH 7,0) que contenía 0,5 M de  $(\text{NH}_4)_2\text{SO}_4$ , se adsorbió una proteína contenida en el sobrenadante a una HiTrap Butyl FF (fabricada por GE Healthcare). Después, en un tampón Tris-HCl (0,05 M, pH 7,0) que contenía 0,5 M de  $(\text{NH}_4)_2\text{SO}_4$ , la proteína adsorbida se sometió a una elución en gradiente lineal para fraccionar una fracción que indicaba la actividad  $\beta$ -glucosidasa.

## (b) Cromatografía hidrófoba (Parte 2)

Una proteína de la fracción obtenida en (a) se adsorbió de nuevo en una HiTrap Butyl FF (fabricada por GE Healthcare). Después la proteína adsorbida se sometió a una elución en un gradiente lineal por el mismo método que en (a) para fraccionar una fracción que indique la actividad  $\beta$ -glucosidasa.

## (c) Cromatografía de intercambio aniónico fuertemente básica

En un tampón Tris-HCl (0,05 M, pH 7,0) se adsorbió una proteína de la fracción obtenida en (b) a una MonoQ (fabricada por GE Healthcare). La proteína adsorbida se sometió a una elución en gradiente lineal en un tampón Tris-HCl (0,05 M, pH 7,0) que contenía 0 M a 1 M de NaCl para fraccionar una fracción que indique actividad  $\beta$ -glucosidasa.

**Ejemplo 2.** Determinación parcial de secuencias de aminoácidos de la  $\beta$ -glucosidasa purificada

La fracción que tiene una actividad  $\beta$ -glucosidasa fraccionada por la cromatografía de intercambio aniónico fuertemente básica del Ejemplo 1 se separó por electroforesis utilizando un 12 % de gel SDS-PAGE mini (fabricado por TEFCO) y se identificó la  $\beta$ -glucosidasa B (acBGLB) de *Acremonium cellulolyticus*. Una banda de acBGLB se cortó y entonces se carboximetiló de manera reductora, seguido por el tratamiento con lisil endopeptidasa. Este producto degradado se separó por electroforesis utilizando un 12 % de Gel SDS-PAGE mini (fabricado por TEFCO), y se transfirió a una membrana PVDF (fabricada por Millipore Corporation). Una banda del fragmento peptídico obtenido de esta manera se cortó. La secuencia de aminoácidos del extremo N del fragmento peptídico se determinó utilizando un secuenciador proteico Modelo 492 (fabricado por Applied Biosystems Inc.). Las secuencias de aminoácidos parciales ("BGLB-LE-1" y "BGLB-LE-2") de la acBGLB determinadas de esta manera se presentaron en las SEQ ID NO: 6 y 7, respectivamente.

**Ejemplo 3.** Clonación de un gen de acBGLB

## (1) Aislamiento del ADN genómico

Se cultivó la cepa ACCP-5-1 de *Acremonium cellulolyticus* en un medio (s) (un 2 % de caldo, un 0,5 % de extracto de levadura y un 2 % de glucosa) a 32 °C durante 2 días, y las células fúngicas se recolectaron por centrifugación. Se aisló un ADN genómico de las mismas a partir de las células fúngicas obtenidas, de acuerdo con el método de Horiuchi et al. (H. Horiuchi et. al., J. Bacteriol., 170, 272-278, (1988)).

## (2) Adquisición del fragmento genético de acBGLB

Basándose en las secuencias de aminoácidos parciales de la acBGLB, se prepararon los siguientes cebadores.

BGLB-F: CCNTTYGTNGGNAAYACNGCNGCNCC (SEQ ID NO: 8)  
BGLB-R: CATDATRTANCCNGGRAANCC (SEQ ID NO: 9)

Utilizando BGLB-F y BGLB-R como cebadores y el ADN genómico como matriz, se llevó a cabo una PCR. La PCR se llevó a cabo utilizando la polimerasa taq LA (fabricada por Takara Bio Inc.). La PCR se llevó a cabo en 35 ciclos de cada una de "94 °C durante 30 segundos, 53 °C durante 30 segundos, y 72 °C durante 2 minutos". El fragmento de ADN de 650 pb amplificado de esta manera se insertó en un vector plasmídico pCD2.1-TOPO utilizando el kit de clonación TOPO TA (fabricado por Invitrogen Corporation) de acuerdo con el protocolo adjunto al mismo. De esta manera se obtuvo un plásmido "TOPO-pBGLB-partial".

La secuencia del fragmento de ADN insertado clonado en el plásmido "TOPO-pBGLB-partial" se determinó utilizando el kit de secuenciación de ciclo BigDye Terminator v3.1 (fabricado por Applied Biosystems Inc.) y el Analizador Genético ABI PRISM (fabricado por Applied Biosystems Inc.) de acuerdo con los protocolos adjuntos a los mismos. La secuencia de bases obtenida de esta manera se tradujo en una secuencia de aminoácidos. Se llevó a cabo la búsqueda de homología en la secuencia de aminoácidos. Como resultado, la secuencia de aminoácidos presentaba una homología del 72 % con la  $\beta$ -glucosidasa (XP\_001216552) derivada de *Aspergillus terreus* y una homología del 88 % con la  $\beta$ -glucosidasa (XP\_002149046.1) derivada de *Penicillium marneffeii*. De esta manera, se determinó que el fragmento de ADN era parte del gen de la  $\beta$ -glucosidasa (familia de la glucósido hidrolasa 3).

## (3) Adquisición del gen de acBGLB de longitud completa por el método de PCR inversa

Se llevó a cabo el método de la PCR inversa de acuerdo con el método de Triglia et al. (T. Triglia et. al., Nucleic Acids Research, 16, 8186, (1988)). El ADN genómico de *Acremonium cellulolyticus* se digirió con Scal durante una noche y se preparó un ADN circular a partir del fragmento digerido utilizando Mighty Mix (fabricado por Takara Bio Inc.). La PCR se llevó a cabo utilizando el ADN circular como matriz y se prepararon los siguientes cebadores basándose en la información de secuencia de bases sobre el fragmento genético de acBGLB. Se obtuvieron una región 5' corriente arriba y una región 3' corriente abajo del gen acBGLB.

BGLB-inv-F: TAGGCGTTCGTTATGCGAAC (SEQ ID NO: 10)  
 BGLB-inv-R: AACGAGATTCAGATGGCG (SEQ ID NO: 11)

5 La región corriente arriba 5' y la región 3' corriente abajo se analizó por el método descrito en el Ejemplos 3- (2). Se determinó la secuencia del gen BGLB de longitud completa.

Basándose en la secuencia de bases obtenida por el método de la PCR inversa, se prepararon los siguientes cebadores. Se llevó a cabo la PCR utilizando el ADN genómico como matriz, para amplificar el gen BGLB.

10 pBGLB-F: CTGGACCTATATCCCGAT (SEQ ID NO: 12)  
 pBGLB-R: TGGTTTGTCCATACTGCGTC (SEQ ID NO: 13)

15 El ADN amplificado de esta manera se insertó en un vector plasmídico pCR2.1-TOPO con un kit de clonación TOPO TA (fabricado por Invitrogen Corporation) para obtener un plásmido "pBGLB". Una cepa TOP10 de *Escherichia coli* (fabricada por Invitrogen Corporation) se transformó con el plásmido "pBGLB" obtenido. De esta manera se obtuvo la cepa de "*Escherichia coli* TOP10/pBGLB".

(4) Preparación de un ADNc acBGLB y análisis de intrón de ADN genómico acBGLB

20 Se cultivó una cepa ACCP-5-1 de *Acremonium cellulolyticus* en un medio de inducción de celulasa a 32 °C durante 2 días, y se recolectaron las células fúngicas por centrifugación. Después de congelarlas con nitrógeno líquido, las células fúngicas resultantes se molieron utilizando un mortero de mano. El ARN total se aisló a partir de las células fúngicas con ISOGEN (Nippon Gene Co., Ltd.) de acuerdo con el protocolo adjunto al mismo. Además, se purificó un ARNm a partir del ARN total con el kit de purificación de ARNm (Pharmacia Corporation) de acuerdo con el protocolo adjunto al mismo.

25 Se sintetizó un ADNc a partir del ARNm obtenido de esta manera con el kit de síntesis de ADNc TimeSaver (Pharmacia Corporation) de acuerdo con el protocolo adjunto al mismo. Se prepararon los siguientes cebadores que contienen el codón de inicio y el codón de parada a partir de la secuencia genética acBGLB, y se llevó a cabo la PCR utilizando el ADNc como matriz.

30 BGLB-N: ATGTATTCCGCATTTCTTTTGCTGC (SEQ ID NO: 14)  
 BGLB-C: CTATTGTAGGCATTGAGAATACCAT (SEQ ID NO: 15)

35 La secuencia de bases (SEQ ID NO: 1) del ADNc amplificado de esta manera se analizó por el método descrito en el Ejemplo 3-(2), y se comparó con la secuencia de bases del ADN genómico de pBGLB. De esta manera, se determinaron las posiciones de los intrones en el ADN genómico.

(5) Estimación de la secuencia de aminoácidos de acBGLB

40 Los exones e intrones del ADN genómico de acBGLB aislados por el método descrito anteriormente se componían de 2630 pb que mostraban las posiciones 218 a 2847 de la secuencia de bases de SEQ ID NO: 2. Además, el ADN genómico de acBGLB contenía tres intrones que se muestran desde la 734<sup>a</sup> a la 792<sup>a</sup>, 1665<sup>a</sup> a 1717<sup>a</sup>, y 2523<sup>a</sup> a 2601<sup>a</sup> de secuencia de bases de SEQ ID NO: 2. La secuencia de aminoácidos de la acBGLB prevista a partir de la fase de lectura abierta (ORF) era como se muestra en la SEQ ID NO: 3. Una parte de la secuencia de aminoácidos prevista a partir de la ORF correspondía con la secuencia interna de la acBGLB determinada en el Ejemplo 2. Este hecho revelaba que el ADN genómico aislado codificaba la acBGLB. Nótese que, utilizando el software de predicción de secuencia de señal SignalP 3.0, se estimó que los restos de aminoácidos de -18 a -1 de la acBGLB eran una secuencia de señal.

50 **Ejemplo 4.** Expresión del gen acBGLB en *Trichoderma viride*

(1) Modificación de un codón del gen de acBGLB para la expresión adecuada en *Trichoderma viride*

55 Para expresar el gen acBGLB a alto nivel como una proteína activa en *Trichoderma viride*, se modificó el gen acBGLB. Como resultado de ensayo y error, se descubrió un ADN que comprendía la secuencia de bases de SEQ ID NO: 4 que se modificó a partir del gen acBGLB un 13,2 % o más de las bases. En el diseño de este gen de acBGLB modificado, se modificaron 16 tipos de aminoácidos entre los 20 tipos de aminoácidos en la secuencia de bases codificante; además, se tuvo en cuenta la frecuencia de distribución de codones utilizada en *Trichoderma viride*. Este gen acBGLB modificado se sintetizó artificialmente por Gene Design Inc. En la síntesis artificial, se hizo el diseño de manera que XbaI y SnaBI estaban contenidas en una secuencia corriente arriba del codón de inicio, y que Sall y XbaI estaban contenidas corriente abajo del codón de parada. De esta manera, se obtuvo un plásmido "pBHLBkai" en el que el gen acBGLB con el codón modificado se insertó en XbaI de pUC19.

65

## (2) Construcción del plásmido de expresión de BGLB con el codón modificado BGLBkai-pCB1

El plásmido "BGLBkai" se escindió con *Sna*BI y *Sall*, y se obtuvo un fragmento genético "BGLBkai-N" de aproximadamente 2,7 kpb. Entre tanto, para retirar un casete de resistencia a la higromicina B del pCB1-Eg3X (Publicación Internacional N° WO 98/11239), el pCB1-Eg3X se escindió con una enzima de restricción *Xba*I, y entonces se circularizó de nuevo utilizando el kit de ligadura de ADN TaKaRa Mighty Mix (fabricado por TAKARA SHUZO CO., LTD.). De esta manera se obtuvo un plásmido "pCB1-Eg3X-hphless". El "pCB1-Eg3X-hphless" se escindió con *Stu*I y *Xho*I, y se recolectó un fragmento de aproximadamente 7 kpb. Se ligó a este, aproximadamente 2,7 kpb del fragmento genético "BGLBkai-N" utilizando el kit de ligadura de ADN TaKaRa Mighty Mix (fabricado por TAKARA SHUZO CO., LTD.). De esta manera, se preparó un plásmido "BGLBkai-pCB1". Las condiciones de reacción tales como la enzima seguían las condiciones del protocolo adjunto al kit. El plásmido "BGLBkai-pCB1" se construyó de tal manera que expresara la acBGLB modificada utilizando su propio codón de inicio en el huésped *Trichoderma viride*.

(3) Preparación de un *Trichoderma viride* transformante con el plásmido "BGLBkai-pCB1"

Se transformó el *Trichoderma viride* con el plásmido "BGLBkai-pCB1" obtenido en el Ejemplo 4-(2) de acuerdo con el método descrito en la Publicación Internacional N° WO 2005/056787. Se llevó a cabo la transformación por un método de co-transformación utilizando la cepa 2 de *Trichoderma viride* deficiente en un gen para la biosíntesis de uracilo (*pyr4*) como huésped y un gen *pyr4* de *Neurospora crassa* como marcador genético. La cepa 2 de *Acremonium cellulosyticus* se cultivó en 50 ml de un medio formador de célula fúngica (un 1 % de extracto de levadura, un 1 % de extracto de malta, un 2 % de polipeptona, un 2,5 % de glucosa, un 0,1 % de fosfato hidrógeno dipotásico, un 0,05 % de sulfato magnésico heptahidrato, un 0,001 % de uridina (pH 7,0)) a 28 °C durante 24 horas, y se centrifugaron a 3000 rpm durante 10 minutos, y se recolectaron las células fúngicas. Las células fúngicas obtenidas se lavaron con 0,5 mol/l de sacarosa, y se suspendió en una solución enzimática formadora de protoplastos (1 mg/ml de  $\beta$ -glucuronidasa, 0,3 mg/ml de quitinasa, 0,3 mg/ml de zimoliasa, 0,5 mol/l de sacarosa) que se había filtrado en algodón. La mezcla se agitó a 30 °C durante 60 minutos, de manera que las hifas se formaron en protoplastos. Esta suspensión se filtró, y después de centrifugó a 2500 rpm durante 10 minutos para recolectar los protoplastos, que se lavaron con tampón SUTC (0,5 mol/l de sacarosa, 10 mmol/l de cloruro cálcico, 10 mmol/l de Tris-HCl (pH 7,5)).

Los protoplastos se suspendieron en 100  $\mu$ l de un tampón SUTC, al que se añadieron 10  $\mu$ g de una solución de ADN 10  $\mu$ l que contenía el plásmido "BGLBkai-pCB1" y 10  $\mu$ l de una solución de ADN que contiene el gen *pyr4*. El resultado se dejó en reposo en hielo durante 5 minutos. Después, se añadieron 400  $\mu$ l de una solución de PEG (un 60 % de PEG4000, 10 mmol/l de cloruro cálcico, 10 mmol/l de Tris-HCl (pH 7,5)), que se dejó en reposo en hielo durante 20 minutos. Después, se añadieron 10 ml de un tampón SUTC, que se centrifugó a 2500 rpm durante 10 minutos. Los protoplastos recolectados de esta manera se suspendieron en 1 ml de un tampón SUTC, y cada una de las soluciones de 200  $\mu$ l de la suspensión de protoplastos se cubrió con agar blando en un medio mínimo que contenía 0,5 mol/l de sacarosa, seguido por el cultivo a 28 °C durante 5 días. Posteriormente, las colonias en crecimiento se transfirieron de nuevo en un medio mínimo. Las colonias formadas se utilizaron como transformantes.

## (4) Cultivo e identificación de transformantes "BGLBkai-pCB1"

El plásmido "BGLBkai-pCB1" se introdujo en cada uno de los medios mínimos. Se seleccionó una línea que crecía en el medio, y se cultivó de acuerdo con el documento WO 98/11239 A. El sobrenadante del cultivo resultante se separó por electroforesis utilizando un 12 % de Gel SDS-PAGE mini (fabricado por TEFCO). Se seleccionó un sobrenadante de cultivo que tiene una banda detectable favorablemente que migraba la misma distancia que la de la acBGLB identificada en el Ejemplo 2.

## (5) Identificación de la secuencia de aminoácidos parcial de la acBGLB modificada recombinante

Para confirmar que la proteína expresada en una gran cantidad en el Ejemplo (4)-4 era la acBGLB modificada, se determinó la secuencia de aminoácidos parcial. Primero, la proteína del sobrenadante del cultivo se separó por electroforesis utilizando un 12 % de Gel SDS-PAGE mini (fabricado por TEFCO). Se trató una banda correspondiente con la acBGLB separada de acuerdo con el método del Ejemplo 2 con lisil endopeptidasa. El producto degradado se separó por electroforesis utilizando un 12 % de Gel SDS-PAGE mini (fabricado por TEFCO), y se transfirió en una membrana PVDF (fabricada por Millipore Corporation). Se cortó una banda del fragmento peptídico obtenido de esta manera. El extremo N de la secuencia de aminoácidos del fragmento peptídico se determinó utilizando un secuenciador proteico Modelo 492 (fabricado por Applied Biosystems Inc.). Como resultado, la secuencia de aminoácidos del extremo N se correspondía con la secuencia de aminoácidos parcial (SEQ ID NO: 6) de la acBGLB.

**Ejemplo 5.** Medición de la actividad enzimática del transformante de *Trichoderma viride*

La actividad  $\beta$ -glucosidasa se midió utilizando el sobrenadante del cultivo del transformante "BGLBkai-pCB1" obtenido en el Ejemplo 4-(4). El método de medición seguía el método descrito en la bibliografía (Methods in

5 ENZYMOLOGY, vol. 160, Biomass Part A Cellulose and Hemicellulose, Willis A. Wood ed. p 109-110). Nótese que la actividad  $\beta$ -glucosidasa se define como una actividad de producción de 1  $\mu$ mol de p-nitrofenol a partir de p-nitrofenil- $\beta$ -glucósido en 1 minuto, y se representó como actividad (U/ml) por ml de sobrenadante del cultivo. El resultado es como se muestra en la Tabla 1. Como es evidente en la Tabla 1, el transformante mostraba una actividad aproximadamente 7,5 veces más alta que una cepa parental (la cepa de *Trichoderma viride* original no deficiente en un gen de biosíntesis de uracilo).

[Tabla 1]

	Actividad $\beta$ -glucosidasa (U/ml)
Cepa parental	211
Transformante	1592

10 Esto revelaba que un microorganismo productor de celulasa que tiene una baja actividad  $\beta$ -glucosidasa sobre-expresaba  $\beta$ -glucosidasa derivada de *Acremonium cellulolyticus*, y que era posible de esta manera aumentar la actividad  $\beta$ -glucosidasa del microorganismo.

### 15 Aplicabilidad industrial

Como se ha descrito anteriormente, la presente invención hace posible obtener una  $\beta$ -glucosidasa derivada de *Acremonium cellulolyticus* como una proteína purificada o una preparación de celulasas con alto rendimiento. Utilizando la  $\beta$ -glucosidasa o preparación de celulasas obtenida de esta manera, se puede promover la sacarificación a partir de biomasa en glucosa, y se pueden llevar a cabo tratamientos y modificaciones en un sustrato basado en celulosa eficazmente. Además, se puede utilizar esta  $\beta$ -glucosidasa y preparación de celulasas con un coste reducido. Además, utilizando un hongo filamentoso que expresa una cantidad reducida de  $\beta$ -glucosidasa de la presente invención, se pueden producir eficazmente la celobiosa útil como edulcorante, una materia prima en cosmética, o una materia prima farmacológica.

### 25 LISTADO DE SECUENCIAS

<110> Meiji Seika Kaisha, Ltd.

<120> Proteína que tiene actividad beta-glucosidasa y su uso

<130> M0895

<150> JP 2009-190840

<151> 20-8-2009

<160> 15

<170> PatentIn versión 3.1

<210> 1

<211> 2442

<212> ADNc

<213> *Acremonium cellulolyticus*

<220>

<221> sig\_peptide

<222> (1)..(54)

<220>

<221> mat\_peptide

<222> (55)..(2439)

<220>

<221> CDS

<222> (1)..(2439)

<400> 1

ES 2 661 966 T3

atg	tat	tcc	gca	ttt	ctt	ttg	ctg	ctg	gct	tcg	gcc	acg	cct	ata	gtc	48
Met	Tyr	Ser	Ala	Phe	Leu	Leu	Leu	Leu	Ala	Ser	Ala	Thr	Pro	Ile	Val	
			-15					-10					-5			
agc	gcc	cag	tca	gct	tct	tgg	tcc	gca	gcc	tac	agt	aaa	gcc	acg	gct	96
Ser	Ala	Gln	Ser	Ala	Ser	Trp	Ser	Ala	Ala	Tyr	Ser	Lys	Ala	Thr	Ala	
	-1	1				5					10					
gct	ttg	agc	aaa	ctc	tct	caa	aat	gac	aaa	att	ggt	atg	gtg	aca	ggc	144
Ala	Leu	Ser	Lys	Leu	Ser	Gln	Asn	Asp	Lys	Ile	Gly	Met	Val	Thr	Gly	
15				20					25						30	
gtg	gga	tgg	ggg	aaa	ggt	cca	tgt	ggt	gga	aac	act	gcc	gcg	cca	tct	192
Val	Gly	Trp	Gly	Lys	Gly	Pro	Cys	Val	Gly	Asn	Thr	Ala	Ala	Pro	Ser	
				35				40						45		
gga	atc	tcg	ttt	cca	tca	ctc	tgt	att	caa	gat	agt	ccc	cta	ggc	ggt	240
Gly	Ile	Ser	Phe	Pro	Ser	Leu	Cys	Ile	Gln	Asp	Ser	Pro	Leu	Gly	Val	
			50					55					60			
cgt	tat	gcg	aac	ccc	gtc	aca	gcg	ttt	ccg	gca	ggc	acg	aat	gct	gga	288
Arg	Tyr	Ala	Asn	Pro	Val	Thr	Ala	Phe	Pro	Ala	Gly	Thr	Asn	Ala	Gly	
			65				70					75				
atg	acc	tgg	gat	cgg	acg	ttg	atg	aac	cag	aga	ggt	gcc	gct	ctt	ggt	336
Met	Thr	Trp	Asp	Arg	Thr	Leu	Met	Asn	Gln	Arg	Gly	Ala	Ala	Leu	Gly	
	80					85					90					
gca	gaa	tcc	aag	ggg	cta	ggt	gtc	cat	ggt	cag	tta	ggg	cct	gtg	gca	384
Ala	Glu	Ser	Lys	Gly	Leu	Gly	Val	His	Val	Gln	Leu	Gly	Pro	Val	Ala	
95				100					105						110	
ggt	ccc	cta	gga	aag	atc	gcg	cag	ggt	ggt	cgt	ggt	tgg	gaa	gga	ttt	432
Gly	Pro	Leu	Gly	Lys	Ile	Ala	Gln	Gly	Gly	Arg	Gly	Trp	Glu	Gly	Phe	
				115				120						125		
gga	acg	gat	cca	tac	ctc	agt	ggt	ggt	gct	atg	att	gag	act	att	tca	480
Gly	Thr	Asp	Pro	Tyr	Leu	Ser	Gly	Val	Ala	Met	Ile	Glu	Thr	Ile	Ser	
			130				135						140			
ggt	atg	cag	agt	tcg	ggt	act	cag	gca	tgc	gcg	aag	cac	tat	att	ggc	528
Gly	Met	Gln	Ser	Ser	Gly	Thr	Gln	Ala	Cys	Ala	Lys	His	Tyr	Ile	Gly	
		145				150						155				
aac	gag	caa	gag	cta	aac	agg	gaa	tcg	atg	agt	tct	aat	att	gat	gat	576
Asn	Glu	Gln	Glu	Leu	Asn	Arg	Glu	Ser	Met	Ser	Ser	Asn	Ile	Asp	Asp	
		160				165					170					
cgt	act	ttg	cac	gag	ctt	tac	ctg	tgg	ccc	ttt	gcc	gat	gcc	gtc	cgt	624
Arg	Thr	Leu	His	Glu	Leu	Tyr	Leu	Trp	Pro	Phe	Ala	Asp	Ala	Val	Arg	
175				180					185						190	
gcc	aat	ggt	gcc	agt	gtg	atg	tgc	tcc	tac	aac	caa	atc	aat	gga	aca	672
Ala	Asn	Val	Ala	Ser	Val	Met	Cys	Ser	Tyr	Asn	Gln	Ile	Asn	Gly	Thr	
				195					200					205		

ES 2 661 966 T3

ttt tcc tgt gag aat gaa gaa tcg atg aca ggt att ctg aag aca gag	720
Phe Ser Cys Glu Asn Glu Glu Ser Met Thr Gly Ile Leu Lys Thr Glu	
210 215 220	
ctc ggc ttt cca gga tac ata atg tct gac tgg gat gca cag cac acc	768
Leu Gly Phe Pro Gly Tyr Ile Met Ser Asp Trp Asp Ala Gln His Thr	
225 230 235	
aca gtt act agt gct aac tct gga ctt gat atg acg atg cca ggt agt	816
Thr Val Thr Ser Ala Asn Ser Gly Leu Asp Met Thr Met Pro Gly Ser	
240 245 250	
gat tat agt gat acg ccg agt agt gtc ctt tgg ggt caa aat ctg gcc	864
Asp Tyr Ser Asp Thr Pro Ser Ser Val Leu Trp Gly Gln Asn Leu Ala	
255 260 265 270	
aat gcc atc tca agt ggc caa gtt gcc cag tcg cgt ctc gac gat atg	912
Asn Ala Ile Ser Ser Gly Gln Val Ala Gln Ser Arg Leu Asp Asp Met	
275 280 285	
gtg act cga att ttg gct gct tgg tat ttg gtt ggg cag gat caa ggc	960
Val Thr Arg Ile Leu Ala Ala Trp Tyr Leu Val Gly Gln Asp Gln Gly	
290 295 300	
ttc cct gcg gtg gcc ttt aac tct tgg acc ggt ggg caa gca agt gtt	1008
Phe Pro Ala Val Ala Phe Asn Ser Trp Thr Gly Gly Gln Ala Ser Val	
305 310 315	
aat gtc aca tca aac cac aac caa gtt gcc cgt gca gtc gct cgc gat	1056
Asn Val Thr Ser Asn His Asn Gln Val Ala Arg Ala Val Ala Arg Asp	
320 325 330	
tct atc gtt ttg ctt aaa aat acc aat agc acg ctt ccg ttg aac aaa	1104
Ser Ile Val Leu Leu Lys Asn Thr Asn Ser Thr Leu Pro Leu Asn Lys	
335 340 345 350	
cca tcg agc att gct att att ggc act gac gcc cag aca aac cct tcc	1152
Pro Ser Ser Ile Ala Ile Ile Gly Thr Asp Ala Gln Thr Asn Pro Ser	
355 360 365	
ggg cca aac gct tgt act gat cgt ggt tgt gat act gga act ttg gct	1200
Gly Pro Asn Ala Cys Thr Asp Arg Gly Cys Asp Thr Gly Thr Leu Ala	
370 375 380	
atg ggt tgg ggt agt gga act tgc caa ttt cca tat ctc aca gat cct	1248
Met Gly Trp Gly Ser Gly Thr Cys Gln Phe Pro Tyr Leu Thr Asp Pro	
385 390 395	
cta aca gct att aaa act cga gct gcc agc gac ggg act acg atc acg	1296
Leu Thr Ala Ile Lys Thr Arg Ala Ala Ser Asp Gly Thr Thr Ile Thr	
400 405 410	
acg agc att agt gac aat ggc agt gcg gga gcc tca gtt gct caa agc	1344
Thr Ser Ile Ser Asp Asn Gly Ser Ala Gly Ala Ser Val Ala Gln Ser	
415 420 425 430	
gcc gag tat gca atc gtt ttc atc aat tca gac tct ggc gaa ggg tac	1392
Ala Glu Tyr Ala Ile Val Phe Ile Asn Ser Asp Ser Gly Glu Gly Tyr	
435 440 445	
ata aca gtc gaa ggc gtc gct ggt gac cgc aac aat ctc gac cca tgg	1440
Ile Thr Val Glu Gly Val Ala Gly Asp Arg Asn Asn Leu Asp Pro Trp	
450 455 460	
cac agt ggc aat gca tta gtg caa tcc gtc gcc gca gtc aac aag aag	1488
His Ser Gly Asn Ala Leu Val Gln Ser Val Ala Ala Val Asn Lys Lys	
465 470 475	
acg att gtc gtc att cat agc gtc ggg ccg gtc att ctt gaa acc ata	1536
Thr Ile Val Val Ile His Ser Val Gly Pro Val Ile Leu Glu Thr Ile	
480 485 490	
ttg gcg caa cct aac gtt gtg gcc gta gta tgg gct ggc ata cca gga	1584
Leu Ala Gln Pro Asn Val Val Ala Val Val Trp Ala Gly Ile Pro Gly	
495 500 505 510	
caa gag agc ggc tca gcc ctc acc gat att ctc tat ggg agt aca gct	1632
Gln Glu Ser Gly Ser Ala Leu Thr Asp Ile Leu Tyr Gly Ser Thr Ala	
515 520 525	
ccc agt gga aag cta acg tac acg att gcc aaa cag gct tcc gat tac	1680
Pro Ser Gly Lys Leu Thr Tyr Thr Ile Ala Lys Gln Ala Ser Asp Tyr	
530 535 540	
ggc act gca gtc gtc agt ggt agc gac aat tat cca gag gga ctt ttc	1728
Gly Thr Ala Val Val Ser Gly Ser Asp Asn Tyr Pro Glu Gly Leu Phe	



# ES 2 661 966 T3

<222> (218)..(733)

<220>  
<221> Intrón  
5 <222> (734)..(792)

<220>  
<221> Exón  
10 <222> (793)..(1664)

<220>  
<221> CDS  
<222> (793)..(1664)

15 <220>  
<221> Intrón  
<222> (1665)..(1717)

<220>  
20 <221> Exón  
<222> (1718)..(2522)

<220>  
25 <221> CDS  
<222> (1718)..(2522)

<220>  
<221> Intrón  
30 <222> (2523)..(2601)

<220>  
<221> Exón  
<222> (2602)..(2847)

35 <220>  
<221> CDS  
<222> (2602)..(2847)

40 <400> 2

ES 2 661 966 T3

ctggacctat attccccgat gaattatgga aagtggggtta catttcaagt tgaatatagct	60
tcggtaaatg tttccattaa aaattataaa aagagacttt cttcagaaaa gtagattttg	120
gttgtgcaat tgagaaaggt tcatcgcaat cagtttgcaa gctgaagcca gaaccttaac	180
gatcacaggc attccctggt tttctggtct gacagga atg tat tcc gca ttt ctt	235
	Met Tyr Ser Ala Phe Leu
	-15
ttg ctg ctg gct tcg gcc acg cct ata gtc agc gcc cag tca gct tct	283
Leu Leu Leu Ala Ser Ala Thr Pro Ile Val Ser Ala Gln Ser Ala Ser	
	-10 -5 -1 1
tgg tcc gca gcc tac agt aaa gcc acg gct gct ttg agc aaa ctc tct	331
Trp Ser Ala Ala Tyr Ser Lys Ala Thr Ala Ala Leu Ser Lys Leu Ser	
5 10 15 20	
caa aat gac aaa att ggt atg gtg aca ggc gtg gga tgg ggg aaa ggt	379
Gln Asn Asp Lys Ile Gly Met Val Thr Gly Val Gly Trp Gly Lys Gly	
	25 30 35
cca tgt gtt gga aac act gcc gcg cca tct gga atc tcg ttt cca tca	427
Pro Cys Val Gly Asn Thr Ala Ala Pro Ser Gly Ile Ser Phe Pro Ser	
	40 45 50
ctc tgt att caa gat agt ccc cta ggc gtt cgt tat gcg aac ccc gtc	475
Leu Cys Ile Gln Asp Ser Pro Leu Gly Val Arg Tyr Ala Asn Pro Val	
	55 60 65
aca gcg ttt ccg gca ggc acg aat gct gga atg acc tgg gat cgg acg	523
Thr Ala Phe Pro Ala Gly Thr Asn Ala Gly Met Thr Trp Asp Arg Thr	
	70 75 80
ttg atg aac cag aga ggt gcc gct ctt ggt gca gaa tcc aag ggg cta	571
Leu Met Asn Gln Arg Gly Ala Ala Leu Gly Ala Glu Ser Lys Gly Leu	
	85 90 95 100
ggt gtc cat gtt cag tta ggg cct gtg gca ggt ccc cta gga aag atc	619
Gly Val His Val Gln Leu Gly Pro Val Ala Gly Pro Leu Gly Lys Ile	
	105 110 115
gcg cag ggt ggt cgt ggt tgg gaa gga ttt gga acg gat cca tac ctc	667
Ala Gln Gly Gly Arg Gly Trp Glu Gly Phe Gly Thr Asp Pro Tyr Leu	
	120 125 130
agt ggt gtt gct atg att gag act att tca ggt atg cag agt tcg ggt	715
Ser Gly Val Ala Met Ile Glu Thr Ile Ser Gly Met Gln Ser Ser Gly	
	135 140 145
act cag gca tgc gcg aag gtgtgtatat ctccgcgaag gaaaccogta	763
Thr Gln Ala Cys Ala Lys	
	150
aataagaatg atctaataaa ccctgtcag cac tat att ggc aac gag caa gag	816
	His Tyr Ile Gly Asn Glu Gln Glu





<210> 3  
<211> 813  
<212> PRT  
<213> *Acremonium cellulolyticus*

5

<400> 3

Met Tyr Ser Ala Phe Leu Leu Leu Leu Ala Ser Ala Thr Pro Ile Val  
 -15 -10 -5  
 Ser Ala Gln Ser Ala Ser Trp Ser Ala Ala Tyr Ser Lys Ala Thr Ala  
 -1 1 5 10  
 Ala Leu Ser Lys Leu Ser Gln Asn Asp Lys Ile Gly Met Val Thr Gly  
 15 20 25 30  
 Val Gly Trp Gly Lys Gly Pro Cys Val Gly Asn Thr Ala Ala Pro Ser  
 35 40 45  
 Gly Ile Ser Phe Pro Ser Leu Cys Ile Gln Asp Ser Pro Leu Gly Val  
 50 55 60  
 Arg Tyr Ala Asn Pro Val Thr Ala Phe Pro Ala Gly Thr Asn Ala Gly  
 65 70 75  
 Met Thr Trp Asp Arg Thr Leu Met Asn Gln Arg Gly Ala Ala Leu Gly  
 80 85 90  
 Ala Glu Ser Lys Gly Leu Gly Val His Val Gln Leu Gly Pro Val Ala  
 95 100 105 110  
 Gly Pro Leu Gly Lys Ile Ala Gln Gly Gly Arg Gly Trp Glu Gly Phe  
 115 120 125  
 Gly Thr Asp Pro Tyr Leu Ser Gly Val Ala Met Ile Glu Thr Ile Ser  
 130 135 140  
 Gly Met Gln Ser Ser Gly Thr Gln Ala Cys Ala Lys His Tyr Ile Gly  
 145 150 155  
 Asn Glu Gln Glu Leu Asn Arg Glu Ser Met Ser Ser Asn Ile Asp Asp  
 160 165 170  
 Arg Thr Leu His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg  
 175 180 185 190  
 Ala Asn Val Ala Ser Val Met Cys Ser Tyr Asn Gln Ile Asn Gly Thr  
 195 200 205  
 Phe Ser Cys Glu Asn Glu Glu Ser Met Thr Gly Ile Leu Lys Thr Glu  
 210 215 220  
 Leu Gly Phe Pro Gly Tyr Ile Met Ser Asp Trp Asp Ala Gln His Thr  
 225 230 235  
 Thr Val Thr Ser Ala Asn Ser Gly Leu Asp Met Thr Met Pro Gly Ser  
 240 245 250  
 Asp Tyr Ser Asp Thr Pro Ser Ser Val Leu Trp Gly Gln Asn Leu Ala  
 255 260 265 270  
 Asn Ala Ile Ser Ser Gly Gln Val Ala Gln Ser Arg Leu Asp Asp Met  
 275 280 285  
 Val Thr Arg Ile Leu Ala Ala Trp Tyr Leu Val Gly Gln Asp Gln Gly  
 290 295 300  
 Phe Pro Ala Val Ala Phe Asn Ser Trp Thr Gly Gly Gln Ala Ser Val  
 305 310 315  
 Asn Val Thr Ser Asn His Asn Gln Val Ala Arg Ala Val Ala Arg Asp  
 320 325 330  
 Ser Ile Val Leu Leu Lys Asn Thr Asn Ser Thr Leu Pro Leu Asn Lys  
 335 340 345 350  
 Pro Ser Ser Ile Ala Ile Ile Gly Thr Asp Ala Gln Thr Asn Pro Ser  
 355 360 365  
 Gly Pro Asn Ala Cys Thr Asp Arg Gly Cys Asp Thr Gly Thr Leu Ala  
 370 375 380  
 Met Gly Trp Gly Ser Gly Thr Cys Gln Phe Pro Tyr Leu Thr Asp Pro  
 385 390 395  
 Leu Thr Ala Ile Lys Thr Arg Ala Ala Ser Asp Gly Thr Thr Ile Thr  
 400 405 410  
 Thr Ser Ile Ser Asp Asn Gly Ser Ala Gly Ala Ser Val Ala Gln Ser  
 415 420 425 430  
 Ala Glu Tyr Ala Ile Val Phe Ile Asn Ser Asp Ser Gly Glu Gly Tyr  
 435 440 445

Ile Thr Val Glu Gly Val Ala Gly Asp Arg Asn Asn Leu Asp Pro Trp  
 450 455 460  
 His Ser Gly Asn Ala Leu Val Gln Ser Val Ala Ala Val Asn Lys Lys  
 465 470 475  
 Thr Ile Val Val Ile His Ser Val Gly Pro Val Ile Leu Glu Thr Ile  
 480 485 490  
 Leu Ala Gln Pro Asn Val Val Ala Val Val Trp Ala Gly Ile Pro Gly  
 495 500 505 510  
 Gln Glu Ser Gly Ser Ala Leu Thr Asp Ile Leu Tyr Gly Ser Thr Ala  
 515 520 525  
 Pro Ser Gly Lys Leu Thr Tyr Thr Ile Ala Lys Gln Ala Ser Asp Tyr  
 530 535 540  
 Gly Thr Ala Val Val Ser Gly Ser Asp Asn Tyr Pro Glu Gly Leu Phe  
 545 550 555  
 Ile Asp Tyr Arg His Phe Asp Lys Ser Asn Ile Glu Pro Arg Tyr Glu  
 560 565 570  
 Phe Gly Tyr Gly Leu Ser Tyr Thr Thr Phe Gly Tyr Thr Asn Leu Ala  
 575 580 585 590  
 Ile Asp Ile Thr Val Ser Thr Gly Pro Thr Thr Gly Gln Ile Val Pro  
 595 600 605  
 Gly Gly Pro Ser Asp Leu Phe Glu Ser Val Gly Thr Val Thr Val Gln  
 610 615 620  
 Val Ala Asn Thr Gly Ser Val Ala Gly Ser Glu Val Ala Gln Leu Tyr  
 625 630 635  
 Ile Gly Leu Pro Ser Ser Ala Pro Ser Ser Pro Pro Lys Gln Leu Arg  
 640 645 650  
 Gly Phe Asp Lys Leu Ser Leu Ala Ala Gly Ala Ser Gly Thr Ala Thr  
 655 660 665 670  
 Phe Asp Leu Thr Arg Arg Asp Leu Ser Tyr Trp Asp Val Ser Lys Gln  
 675 680 685  
 Lys Trp Val Val Pro Ser Gly Ala Phe Thr Val Tyr Val Gly Ala Ser  
 690 695 700  
 Ser Arg Asp Ile Arg Leu Gln Gly Thr Phe Thr Pro Gly Gly Ser Ser  
 705 710 715  
 Thr Thr Ser Thr Ile Thr Ser Ser Lys Thr Ser Thr Thr Ile Ser Thr  
 720 725 730  
 Ser Val Thr Thr Ser Ser Thr Thr Ala Lys Thr Thr Thr Thr Ser  
 735 740 745 750  
 Ser Thr Thr Ser Ser Ala Gly Pro Thr Gln Thr Pro Tyr Gly Gln Cys  
 755 760 765  
 Gly Gly Gln Gly Trp Thr Gly Pro Thr Val Cys Ser Ser Gly Trp Thr  
 770 775 780  
 Cys Lys Val Thr Asn Gln Trp Tyr Ser Gln Cys Leu Gln  
 785 790 795

<210> 4  
 <211> 2442  
 <212> ADN  
 <213> *Acremonium cellulolyticus*

<220>  
 <221> sig\_peptide  
 <222> (1)..(54)

<220>  
 <221> mat\_peptide  
 <222> (55)..(2439)

<220>  
 <221> CDS  
 <222> (1)..(2439)

ES 2 661 966 T3

<400> 4

atg tac tcc gcc ttc ctt ttg ctg ctg gct tgg gcc acc cct atc gtc	48
Met Tyr Ser Ala Phe Leu Leu Leu Leu Ala Ser Ala Thr Pro Ile Val	
-15 -10 -5	
agc gcc cag tcc gct tct tgg tcc gcc gcc tac tcc aag gcc acc gct	96
Ser Ala Gln Ser Ala Ser Trp Ser Ala Ala Tyr Ser Lys Ala Thr Ala	

ES 2 661 966 T3

	-1	1				5					10								
	gct	ttg	agc	aag	ctc	tct	cag	aac	gac	aag	atc	ggt	atg	gtg	acc	ggc			144
	Ala	Leu	Ser	Lys	Leu	Ser	Gln	Asn	Asp	Lys	Ile	Gly	Met	Val	Thr	Gly			
	15					20					25				30				
	gtg	gga	tgg	ggt	aag	ggt	ccc	tgc	ggt	gga	aac	act	gcc	gcg	ccc	tct			192
	Val	Gly	Trp	Gly	Lys	Gly	Pro	Cys	Val	Gly	Asn	Thr	Ala	Ala	Pro	Ser			
					35					40					45				
	gga	atc	tcg	ttc	ccc	tcc	ctc	tgc	atc	cag	gat	tcc	ccc	ctc	ggc	ggt			240
	Gly	Ile	Ser	Phe	Pro	Ser	Leu	Cys	Ile	Gln	Asp	Ser	Pro	Leu	Gly	Val			
				50					55					60					
	cgt	tac	gcg	aac	ccc	gtc	acc	gcg	ttc	ccc	gcc	ggc	acc	aac	gct	gga			288
	Arg	Tyr	Ala	Asn	Pro	Val	Thr	Ala	Phe	Pro	Ala	Gly	Thr	Asn	Ala	Gly			
				65				70						75					
	atg	acc	tgg	gat	cgc	acc	ttg	atg	aac	cag	cgc	ggt	gcc	gct	ctt	ggt			336
	Met	Thr	Trp	Asp	Arg	Thr	Leu	Met	Asn	Gln	Arg	Gly	Ala	Ala	Leu	Gly			
							85					90							
	gcc	gaa	tcc	aag	ggt	ctc	ggt	gtc	cac	ggt	cag	ctc	ggt	cct	gtg	gcc			384
	Ala	Glu	Ser	Lys	Gly	Leu	Gly	Val	His	Val	Gln	Leu	Gly	Pro	Val	Ala			
	95					100					105				110				
	ggt	ccc	ctc	gga	aag	atc	gcg	cag	ggt	ggt	cggt	ggt	tgg	gaa	gga	ttc			432
	Gly	Pro	Leu	Gly	Lys	Ile	Ala	Gln	Gly	Gly	Arg	Gly	Trp	Glu	Gly	Phe			
					115					120					125				
	gga	acc	gat	ccc	tac	ctc	tcc	ggt	ggt	gct	atg	att	gag	act	att	tcc			480
	Gly	Thr	Asp	Pro	Tyr	Leu	Ser	Gly	Val	Ala	Met	Ile	Glu	Thr	Ile	Ser			
				130					135					140					
	ggt	atg	cag	tcc	tcg	ggt	act	cag	gcc	tgc	gcg	aag	cac	tac	att	ggc			528
	Gly	Met	Gln	Ser	Ser	Gly	Thr	Gln	Ala	Cys	Ala	Lys	His	Tyr	Ile	Gly			
				145				150						155					
	aac	gag	cag	gag	ctc	aac	cgc	gaa	tcg	atg	tcc	tct	aac	att	gat	gat			576
	Asn	Glu	Gln	Glu	Leu	Asn	Arg	Glu	Ser	Met	Ser	Ser	Asn	Ile	Asp	Asp			
	160					165						170							
	cggt	act	ttg	cac	gag	ctt	tac	ctg	tgg	ccc	ttc	gcc	gat	gcc	gtc	cggt			624
	Arg	Thr	Leu	His	Glu	Leu	Tyr	Leu	Trp	Pro	Phe	Ala	Asp	Ala	Val	Arg			
	175					180						185			190				
	gcc	aac	ggt	gcc	tcc	gtg	atg	tgc	tcc	tac	aac	cag	atc	aac	gga	acc			672
	Ala	Asn	Val	Ala	Ser	Val	Met	Cys	Ser	Tyr	Asn	Gln	Ile	Asn	Gly	Thr			
					195					200					205				
	ttc	tcc	tgc	gag	aac	gaa	gaa	tcg	atg	acc	ggt	att	ctg	aag	acc	gag			720
	Phe	Ser	Cys	Glu	Asn	Glu	Glu	Ser	Met	Thr	Gly	Ile	Leu	Lys	Thr	Glu			
				210					215					220					
	ctc	ggc	ttc	ccc	gga	tac	atc	atg	tct	gac	tgg	gat	gcc	cag	cac	acc			768
	Leu	Gly	Phe	Pro	Gly	Tyr	Ile	Met	Ser	Asp	Trp	Asp	Ala	Gln	His	Thr			
				225				230						235					
	acc	ggt	act	tcc	gct	aac	tct	gga	ctt	gat	atg	acc	atg	ccc	ggt	tcc			816
	Thr	Val	Thr	Ser	Ala	Asn	Ser	Gly	Leu	Asp	Met	Thr	Met	Pro	Gly	Ser			
				240			245						250						
	gat	tac	tcc	gat	acc	ccc	tcc	tcc	gtc	ctt	tgg	ggt	cag	aac	ctg	gcc			864
	Asp	Tyr	Ser	Asp	Thr	Pro	Ser	Ser	Val	Leu	Trp	Gly	Gln	Asn	Leu	Ala			
	255					260							265		270				
	aac	gcc	atc	tcc	tcc	ggc	cag	ggt	gcc	cag	tcg	cggt	ctc	gac	gat	atg			912
	Asn	Ala	Ile	Ser	Ser	Gly	Gln	Val	Ala	Gln	Ser	Arg	Leu	Asp	Asp	Met			
				275						280				285					
	gtg	act	cgc	att	ttg	gct	gct	tgg	tac	ttg	ggt	ggt	cag	gat	cag	ggc			960
	Val	Thr	Arg	Ile	Leu	Ala	Ala	Trp	Tyr	Leu	Val	Gly	Gln	Asp	Gln	Gly			
				290					295					300					
	ttc	cct	gcg	gtg	gcc	ttc	aac	tct	tgg	acc	ggt	ggt	cag	gcc	tcc	ggt			1008
	Phe	Pro	Ala	Val	Ala	Phe	Asn	Ser	Trp	Thr	Gly	Gly	Gln	Ala	Ser	Val			
				305				310						315					
	aac	gtc	acc	tcc	aac	cac	aac	cag	ggt	gcc	cggt	gcc	gtc	gct	cgc	gat			1056
	Asn	Val	Thr	Ser	Asn	His	Asn	Gln	Val	Ala	Arg	Ala	Val	Ala	Arg	Asp			
				320			325						330						
	tct	atc	ggt	ttg	ctt	aag	aac	acc	aac	agc	acc	ctt	ccc	ttg	aac	aag			1104
	Ser	Ile	Val	Leu	Leu	Lys	Asn	Thr	Asn	Ser	Thr	Leu	Pro	Leu	Asn	Lys			
	335					340						345			350				
	ccc	tcg	agc	att	gct	att	att	ggc	act	gac	gcc	cag	acc	aac	cct	tcc			1152

ES 2 661 966 T3

Pro	Ser	Ser	Ile	Ala	Ile	Ile	Gly	Thr	Asp	Ala	Gln	Thr	Asn	Pro	Ser		
				355					360				365				
ggt	ccc	aac	gct	tgc	act	gat	cgt	ggt	tgc	gat	act	gga	act	ttg	gct	1200	
Gly	Pro	Asn	Ala	Cys	Thr	Asp	Arg	Gly	Cys	Asp	Thr	Gly	Thr	Leu	Ala		
				370				375				380					
atg	ggt	tgg	ggt	tcc	gga	act	tgc	cag	ttc	ccc	tac	ctc	acc	gat	cct	1248	
Met	Gly	Trp	Gly	Ser	Gly	Thr	Cys	Gln	Phe	Pro	Tyr	Leu	Thr	Asp	Pro		
				385			390				395						
ctc	acc	gct	att	aag	act	cgc	gct	gcc	agc	gac	ggt	act	acc	atc	acc	1296	
Leu	Thr	Ala	Ile	Lys	Thr	Arg	Ala	Ala	Ser	Asp	Gly	Thr	Thr	Ile	Thr		
	400					405					410						
acc	agc	att	tcc	gac	aac	ggc	tcc	gcg	gga	gcc	tcc	ggt	gct	cag	agc	1344	
Thr	Ser	Ile	Ser	Asp	Asn	Gly	Ser	Ala	Gly	Ala	Ser	Val	Ala	Gln	Ser		
	415				420					425				430			
gcc	gag	tac	gcc	atc	ggt	ttc	atc	aac	tcc	gac	tct	ggc	gaa	ggt	tac	1392	
Ala	Glu	Tyr	Ala	Ile	Val	Phe	Ile	Asn	Ser	Asp	Ser	Gly	Glu	Gly	Tyr		
				435				440						445			
atc	acc	gtc	gaa	ggc	gtc	gct	ggt	gac	cgc	aac	aac	ctc	gac	ccc	ttg	1440	
Ile	Thr	Val	Glu	Gly	Val	Ala	Gly	Asp	Arg	Asn	Asn	Leu	Asp	Pro	Trp		
			450				455					460					
cac	tcc	ggc	aac	gcc	ctc	gtg	cag	tcc	gtc	gcc	gcc	gtc	aac	aag	aag	1488	
His	Ser	Gly	Asn	Ala	Leu	Val	Gln	Ser	Val	Ala	Ala	Val	Asn	Lys	Lys		
	465					470					475						
acc	att	gtc	gtc	att	cac	agc	gtc	ggt	ccc	gtc	att	ctt	gaa	acc	atc	1536	
Thr	Ile	Val	Val	Ile	His	Ser	Val	Gly	Pro	Val	Ile	Leu	Glu	Thr	Ile		
	480				485						490						
ttg	gcg	cag	cct	aac	ggt	gtg	gcc	gtc	gtc	tgg	gct	ggc	atc	ccc	gga	1584	
Leu	Ala	Gln	Pro	Asn	Val	Val	Ala	Val	Val	Trp	Ala	Gly	Ile	Pro	Gly		
	495			500						505				510			
cag	gag	agc	ggc	tcc	gcc	ctc	acc	gat	atc	ctc	tac	ggt	tcc	acc	gct	1632	
Gln	Glu	Ser	Gly	Ser	Ala	Leu	Thr	Asp	Ile	Leu	Tyr	Gly	Ser	Thr	Ala		
				515						520				525			
ccc	tcc	gga	aag	ctc	acc	tac	acc	att	gcc	aag	cag	gct	tcc	gat	tac	1680	
Pro	Ser	Gly	Lys	Leu	Thr	Tyr	Thr	Ile	Ala	Lys	Gln	Ala	Ser	Asp	Tyr		
			530					535				540					
ggc	act	gcc	gtc	gtc	tcc	ggt	agc	gac	aac	tac	ccc	gag	gga	ctt	ttc	1728	
Gly	Thr	Ala	Val	Val	Ser	Gly	Ser	Asp	Asn	Tyr	Pro	Glu	Gly	Leu	Phe		
	545					550						555					
att	gat	tac	cgc	cac	ttc	gac	aag	agc	aac	att	gaa	cct	cgc	tac	gaa	1776	
Ile	Asp	Tyr	Arg	His	Phe	Asp	Lys	Ser	Asn	Ile	Glu	Pro	Arg	Tyr	Glu		
	560				565						570						
ttc	ggc	tac	gga	ctg	tcc	tac	acc	acc	ttc	ggt	tac	acc	aac	ttg	gcc	1824	
Phe	Gly	Tyr	Gly	Leu	Ser	Tyr	Thr	Thr	Phe	Gly	Tyr	Thr	Asn	Leu	Ala		
	575			580						585				590			
att	gat	att	acc	ggt	tcg	acc	ggc	ccc	act	ggt	cag	atc	ggt	cct		1872	
Ile	Asp	Ile	Thr	Val	Ser	Thr	Gly	Pro	Thr	Thr	Gly	Gln	Ile	Val	Pro		
				595				600						605			
ggt	gga	cct	tct	gat	ctt	ttc	gag	tct	ggt	gga	acc	ggt	acc	gtc	cag	1920	
Gly	Gly	Pro	Ser	Asp	Leu	Phe	Glu	Ser	Val	Gly	Thr	Val	Thr	Val	Gln		
				610				615						620			
gtc	gcc	aac	acc	ggc	agc	ggt	gcc	ggc	tcc	gaa	ggt	gcc	cag	ctc	tac	1968	
Val	Ala	Asn	Thr	Gly	Ser	Val	Ala	Gly	Ser	Glu	Val	Ala	Gln	Leu	Tyr		
	625					630						635					
att	ggt	ctg	ccc	tcg	tcc	gcc	ccc	tcc	tcc	ccc	ccc	aag	cag	ttg	ogt	2016	
Ile	Gly	Leu	Pro	Ser	Ser	Ala	Pro	Ser	Ser	Pro	Pro	Lys	Gln	Leu	Arg		
	640					645						650					
ggt	ttc	gat	aag	ctt	tct	ctc	gct	gct	ggc	gct	agc	ggt	acc	gcc	acc	2064	
Gly	Phe	Asp	Lys	Leu	Ser	Leu	Ala	Ala	Gly	Ala	Ser	Gly	Thr	Ala	Thr		
	655				660					665				670			
ttc	gat	ttg	acc	cgc	cgc	gat	ttg	tcc	tac	tgg	gat	gtc	tcc	aag	cag	2112	
Phe	Asp	Leu	Thr	Arg	Arg	Asp	Leu	Ser	Tyr	Trp	Asp	Val	Ser	Lys	Gln		
				675						680				685			
aag	tgg	gtg	ggt	ccc	agc	gga	gcc	ttc	acc	gtc	tac	ggt	ggt	gcc	tcg	2160	
Lys	Trp	Val	Val	Pro	Ser	Gly	Ala	Phe	Thr	Val	Tyr	Val	Gly	Ala	Ser		
				690					695					700			

ES 2 661 966 T3

tcc	cgc	gat	att	cgc	ttg	cag	ggt	acc	ttc	acc	ccc	gga	ggt	agc	tcg	2208
Ser	Arg	Asp	Ile	Arg	Leu	Gln	Gly	Thr	Phe	Thr	Pro	Gly	Gly	Ser	Ser	
		705					710					715				
acc	act	tcg	act	atc	act	tcc	tct	aag	act	tct	act	act	atc	agc	act	2256
Thr	Thr	Ser	Thr	Ile	Thr	Ser	Ser	Lys	Thr	Ser	Thr	Thr	Ile	Ser	Thr	
		720					725					730				
tct	ggt	acc	acc	tcc	agc	tcc	acc	acc	gct	aag	acc	acc	acc	act	agc	2304
Ser	Val	Thr	Thr	Ser	Ser	Ser	Thr	Thr	Ala	Lys	Thr	Thr	Thr	Thr	Ser	
		735			740					745					750	
tcg	acc	acc	tcc	tct	gcc	ggt	ccc	acc	cag	acc	ccc	tac	gga	cag	tgc	2352
Ser	Thr	Thr	Ser	Ser	Ala	Gly	Pro	Thr	Gln	Thr	Pro	Tyr	Gly	Gln	Cys	
				755					760					765		
ggt	gga	cag	ggt	tgg	acc	ggc	cct	acc	gtg	tgc	tcc	tct	ggc	tgg	act	2400
Gly	Gly	Gln	Gly	Trp	Thr	Gly	Pro	Thr	Val	Cys	Ser	Ser	Gly	Trp	Thr	
			770				775						780			
tgc	aag	gtc	acc	aac	cag	tgg	tac	tct	cag	tgc	ctc	cag	tag			2442
Cys	Lys	Val	Thr	Asn	Gln	Trp	Tyr	Ser	Gln	Cys	Leu	Gln				
		785					790					795				

<210> 5  
 <211> 813  
 <212> PRT  
 <213> *Acremonium cellulolyticus*  
  
 <400> 5

5

Met Tyr Ser Ala Phe Leu Leu Leu Leu Ala Ser Ala Thr Pro Ile Val  
 -15 -10 -5  
 Ser Ala Gln Ser Ala Ser Trp Ser Ala Ala Tyr Ser Lys Ala Thr Ala  
 -1 1 5 10  
 Ala Leu Ser Lys Leu Ser Gln Asn Asp Lys Ile Gly Met Val Thr Gly  
 15 20 25 30  
 Val Gly Trp Gly Lys Gly Pro Cys Val Gly Asn Thr Ala Ala Pro Ser  
 35 40 45  
 Gly Ile Ser Phe Pro Ser Leu Cys Ile Gln Asp Ser Pro Leu Gly Val  
 50 55 60  
 Arg Tyr Ala Asn Pro Val Thr Ala Phe Pro Ala Gly Thr Asn Ala Gly  
 65 70 75  
 Met Thr Trp Asp Arg Thr Leu Met Asn Gln Arg Gly Ala Ala Leu Gly  
 80 85 90  
 Ala Glu Ser Lys Gly Leu Gly Val His Val Gln Leu Gly Pro Val Ala  
 95 100 105 110  
 Gly Pro Leu Gly Lys Ile Ala Gln Gly Gly Arg Gly Trp Glu Gly Phe  
 115 120 125  
 Gly Thr Asp Pro Tyr Leu Ser Gly Val Ala Met Ile Glu Thr Ile Ser  
 130 135 140  
 Gly Met Gln Ser Ser Gly Thr Gln Ala Cys Ala Lys His Tyr Ile Gly  
 145 150 155  
 Asn Glu Gln Glu Leu Asn Arg Glu Ser Met Ser Ser Asn Ile Asp Asp  
 160 165 170  
 Arg Thr Leu His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg  
 175 180 185 190  
 Ala Asn Val Ala Ser Val Met Cys Ser Tyr Asn Gln Ile Asn Gly Thr  
 195 200 205  
 Phe Ser Cys Glu Asn Glu Glu Ser Met Thr Gly Ile Leu Lys Thr Glu  
 210 215 220  
 Leu Gly Phe Pro Gly Tyr Ile Met Ser Asp Trp Asp Ala Gln His Thr  
 225 230 235  
 Thr Val Thr Ser Ala Asn Ser Gly Leu Asp Met Thr Met Pro Gly Ser  
 240 245 250  
 Asp Tyr Ser Asp Thr Pro Ser Ser Val Leu Trp Gly Gln Asn Leu Ala  
 255 260 265 270  
 Asn Ala Ile Ser Ser Gly Gln Val Ala Gln Ser Arg Leu Asp Asp Met  
 275 280 285  
 Val Thr Arg Ile Leu Ala Ala Trp Tyr Leu Val Gly Gln Asp Gln Gly  
 290 295 300

Phe Pro Ala Val Ala Phe Asn Ser Trp Thr Gly Gly Gln Ala Ser Val  
 305 310 315  
 Asn Val Thr Ser Asn His Asn Gln Val Ala Arg Ala Val Ala Arg Asp  
 320 325 330  
 Ser Ile Val Leu Leu Lys Asn Thr Asn Ser Thr Leu Pro Leu Asn Lys  
 335 340 345 350  
 Pro Ser Ser Ile Ala Ile Ile Gly Thr Asp Ala Gln Thr Asn Pro Ser  
 355 360 365  
 Gly Pro Asn Ala Cys Thr Asp Arg Gly Cys Asp Thr Gly Thr Leu Ala  
 370 375 380  
 Met Gly Trp Gly Ser Gly Thr Cys Gln Phe Pro Tyr Leu Thr Asp Pro  
 385 390 395  
 Leu Thr Ala Ile Lys Thr Arg Ala Ala Ser Asp Gly Thr Thr Ile Thr  
 400 405 410  
 Thr Ser Ile Ser Asp Asn Gly Ser Ala Gly Ala Ser Val Ala Gln Ser  
 415 420 425 430  
 Ala Glu Tyr Ala Ile Val Phe Ile Asn Ser Asp Ser Gly Glu Gly Tyr  
 435 440 445  
 Ile Thr Val Glu Gly Val Ala Gly Asp Arg Asn Asn Leu Asp Pro Trp  
 450 455 460  
 His Ser Gly Asn Ala Leu Val Gln Ser Val Ala Ala Val Asn Lys Lys  
 465 470 475  
 Thr Ile Val Val Ile His Ser Val Gly Pro Val Ile Leu Glu Thr Ile  
 480 485 490  
 Leu Ala Gln Pro Asn Val Val Ala Val Val Trp Ala Gly Ile Pro Gly  
 495 500 505 510  
 Gln Glu Ser Gly Ser Ala Leu Thr Asp Ile Leu Tyr Gly Ser Thr Ala  
 515 520 525  
 Pro Ser Gly Lys Leu Thr Tyr Thr Ile Ala Lys Gln Ala Ser Asp Tyr  
 530 535 540  
 Gly Thr Ala Val Val Ser Gly Ser Asp Asn Tyr Pro Glu Gly Leu Phe  
 545 550 555  
 Ile Asp Tyr Arg His Phe Asp Lys Ser Asn Ile Glu Pro Arg Tyr Glu  
 560 565 570  
 Phe Gly Tyr Gly Leu Ser Tyr Thr Thr Phe Gly Tyr Thr Asn Leu Ala  
 575 580 585 590  
 Ile Asp Ile Thr Val Ser Thr Gly Pro Thr Thr Gly Gln Ile Val Pro  
 595 600 605  
 Gly Gly Pro Ser Asp Leu Phe Glu Ser Val Gly Thr Val Thr Val Gln  
 610 615 620  
 Val Ala Asn Thr Gly Ser Val Ala Gly Ser Glu Val Ala Gln Leu Tyr  
 625 630 635  
 Ile Gly Leu Pro Ser Ser Ala Pro Ser Ser Pro Pro Lys Gln Leu Arg  
 640 645 650  
 Gly Phe Asp Lys Leu Ser Leu Ala Ala Gly Ala Ser Gly Thr Ala Thr  
 655 660 665 670  
 Phe Asp Leu Thr Arg Arg Asp Leu Ser Tyr Trp Asp Val Ser Lys Gln  
 675 680 685  
 Lys Trp Val Val Pro Ser Gly Ala Phe Thr Val Tyr Val Gly Ala Ser  
 690 695 700  
 Ser Arg Asp Ile Arg Leu Gln Gly Thr Phe Thr Pro Gly Gly Ser Ser  
 705 710 715  
 Thr Thr Ser Thr Ile Thr Ser Ser Lys Thr Ser Thr Thr Ile Ser Thr  
 720 725 730  
 Ser Val Thr Thr Ser Ser Ser Thr Thr Ala Lys Thr Thr Thr Ser  
 735 740 745 750  
 Ser Thr Thr Ser Ser Ala Gly Pro Thr Gln Thr Pro Tyr Gly Gln Cys  
 755 760 765  
 Gly Gly Gln Gly Trp Thr Gly Pro Thr Val Cys Ser Ser Gly Trp Thr  
 770 775 780  
 Cys Lys Val Thr Asn Gln Trp Tyr Ser Gln Cys Leu Gln  
 785 790 795



<213> Secuencia artificial  
 <220>  
 <223> Cebador:BGLB-R  
 5  
 <220>  
 <221> misc\_feature  
 <222> (10)..(10)  
 <223> "n" representa cualquier base  
 10  
 <220>  
 <221> misc\_feature  
 <222> (13)..(13)  
 <223> "n" representa una inosina  
 15  
 <220>  
 <221> misc\_feature  
 <222> (19)..(19)  
 <223> "n" representa una inosina  
 20  
 <400> 9  
 catdatrtan ccnggraanc c 21  
 25  
 <210> 10  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 30  
 <220>  
 <223> Cebador:BGLB-inv-F  
 <400> 10  
 taggcgttcg ttatgcgaac 20  
 35  
 <210> 11  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 40  
 <220>  
 <223> Cebador:BGLB-inv-R  
 <400> 11  
 aaacgagatt ccagatggcg 20  
 45  
 <210> 12  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 50  
 <220>  
 <223> Cebador:pBGLB-F  
 <400> 12  
 ctggacctat attccccgat 20  
 55  
 <210> 13  
 <211> 20  
 <212> ADN  
 <213> Secuencia artificial  
 60  
 <220>  
 <223> Cebador:pBGLB-R  
 65  
 <400> 13

# ES 2 661 966 T3

tggtttgtcc atactggtc 20

5 <210> 14  
<211> 25  
<212> ADN  
<213> Secuencia artificial

10 <220>  
<223> Cebador:BGLB-N

<400> 14  
atgtattcgg catttcttt gctgc 25

15 <210> 15  
<211> 25  
<212> ADN  
<213> Secuencia artificial

20 <220>  
<223> Cebador:BGLB-C

<400> 15  
ctattgtagg cattgagaat accat 25

25

REIVINDICACIONES

1. Un polinucleótido de uno cualquiera de los siguientes (i) a (v), que codifica una proteína que tiene una actividad  $\beta$ -glucosidasa:
- 5 (i) un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 3;  
(ii) un polinucleótido que comprende la secuencia de bases de las posiciones 1-2439 de SEQ ID NO: 1 o la secuencia de bases de las posiciones 218- 2847 de SEQ ID NO: 2;  
10 (iii) un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 3 en la que 40 o menos aminoácidos se han sustituido, eliminado, insertado y/o añadido;  
(iv) un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos que tiene una identidad del 95 % o más con la secuencia de aminoácidos de SEQ ID NO: 3;  
15 (v) un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos de las posiciones 1-795 de SEQ ID NO: 3.
2. El polinucleótido de acuerdo con la reivindicación 1, que se deriva de un hongo filamentos.
3. El polinucleótido de acuerdo con la reivindicación 2, en la que el hongo filamentos es *Acremonium cellulolyticus*.
- 20 4. Un polinucleótido de cualquiera de los siguientes (i) y (ii), que codifica una proteína que tiene actividad  $\beta$ -glucosidasa:
- (i) un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 5; y  
25 (ii) un polinucleótido que comprende una región codificante de la secuencia de bases de SEQ ID NO: 4.
5. Un polinucleótido que comprende la secuencia de bases de SEQ ID NO: 4 en la que 30 o menos bases se han sustituido, eliminado, insertado y/o añadido, codificando el polinucleótido una proteína que tiene una actividad  $\beta$ -glucosidasa.
- 30 6. Un polinucleótido que codifica una proteína que comprende la secuencia de aminoácidos de las posiciones 1-795 de SEQ ID NO: 5.
7. Un vector de expresión que comprende el polinucleótido de acuerdo con una cualquiera de las reivindicaciones 1 a 6.
- 35 8. Una célula huésped que comprende el vector de expresión de acuerdo con la reivindicación 7.
9. Una proteína que tiene actividad  $\beta$ -glucosidasa seleccionada de entre (i) a (iv):
- 40 (i) una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 3;  
(ii) una proteína que comprende la secuencia de aminoácidos de SEQ ID NO: 3 en la que 40 o menos aminoácidos se han sustituido, eliminado, insertado y/o añadido;  
45 (iii) una proteína que comprende la secuencia de aminoácidos que tiene un 95 % de identidad o más con la secuencia de aminoácidos de SEQ ID NO: 3;  
(iv) una proteína de cualquiera de (i) a (iii) de la que se ha retirado una secuencia de señal, en la que la secuencia de señal es una secuencia de aminoácidos de entre las posiciones -18 a -1 de la secuencia de aminoácidos de SEQ ID NO: 3.
- 50 10. Un método para la producción de la proteína de acuerdo con la reivindicación 9, comprendiendo el método las etapas de:
- cultivar la célula huésped de acuerdo con la reivindicación 8; y  
recolectar una proteína expresada por la célula huésped y/o un cultivo de la misma.
- 55 11. Una preparación de celulasa que comprende la proteína de acuerdo con la reivindicación 9.
12. Un método para la degradación o la conversión de un material celulósico, comprendiendo el método la etapa de tratar el material celulósico con una cualquiera de entre la proteína de acuerdo con la reivindicación 9 y la preparación de celulasa de acuerdo con la reivindicación 11.
- 60 13. Un método para la producción de un material celulósico degradado o convertido, comprendiendo el método las etapas de:
- 65 tratar un material celulósico con una cualquiera de la proteína de acuerdo con la reivindicación 9 y la preparación de celulasa de acuerdo con la reivindicación 11; y  
recolectar un material celulósico degradado.

14. El método de acuerdo con la reivindicación 13, en el que el material celulósico degradado es un azúcar.
- 5 15. Una composición detergente que comprende una cualquiera de la proteína de acuerdo con la reivindicación 9 y la preparación de celulasa de acuerdo con la reivindicación 11.
- 10 16. Un método para el tratamiento de una fibra que contiene celulosa, comprendiendo el método la etapa de poner una cualquiera de la proteína de acuerdo con la reivindicación 9, la preparación de celulasa de acuerdo con la reivindicación 11 y la composición detergente de acuerdo con la reivindicación 15 en contacto con la fibra que contiene celulosa.
- 15 17. Un método para el destintado de papel desechado, **caracterizado** el método **por que** una cualquiera de entre la proteína de la reivindicación 9 y la preparación de celulasa de acuerdo con la reivindicación 11 se utiliza en una etapa de destintado del tratamiento del papel desechado con un agente de destintado.
- 20 18. Un método para la producción de pasta de papel que tenga un drenaje mejorado, comprendiendo el método la etapa de tratar la pasta de papel con una cualquiera de entre la proteína de la reivindicación 9 y la preparación de celulasa de acuerdo con la reivindicación 11.
19. Un método para la producción de un pienso para animales que tenga una digestibilidad mejorada, comprendiendo el método la etapa de tratar un pienso para animales con una cualquiera de entre la proteína de acuerdo con la reivindicación 9 y la preparación de celulasa de acuerdo con la reivindicación 11.

Fig. 1

