

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 396 542**

51 Int. Cl.:

**A21D 8/04** (2006.01)

**C11D 3/386** (2006.01)

**C12N 9/42** (2006.01)

**C12N 15/113** (2010.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **30.07.2002 E 02752640 (9)**

97 Fecha y número de publicación de la concesión europea: **24.10.2012 EP 1436309**

54 Título: **Beta-glucosidasa BGL3 y ácidos nucleicos que la codifican**

30 Prioridad:

**21.09.2001 US 957880**

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

**22.02.2013**

73 Titular/es:

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

72 Inventor/es:

**DUNN-COLEMAN, NIGEL;  
GOEDEGEBUUR, FRITS;  
WARD, MICHAEL y  
YAO, JIAN**

74 Agente/Representante:

**PONTI SALES, Adelaida**

ES 2 396 542 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

Beta-glucosidasa BGL3 y ácidos nucleicos que la codifican

5 **Apoyo del gobierno**

[0001] Partes de este trabajo han sido financiadas por el subcontrato n ° ZCO-30017-01 con el Laboratorio Nacional de Energías Renovables bajo el contrato principal n ° DE-AC36-99GO10337 con el Departamento de Energía de EE.UU. Por consiguiente, el Gobierno de Estados Unidos tiene algunos derechos en esta invención.

10

**Campo de la invención**

[0002] La presente invención se refiere a secuencias de ácidos nucleicos de *bgl3* aisladas, que codifican polipéptidos que tienen actividad de beta-glucosidasa. La invención también se refiere a construcciones de ácidos nucleicos, vectores y células huésped que comprenden las secuencias de ácidos nucleicos, así como a procedimientos para producir polipéptidos BGL3 recombinantes.

15

**Referencias**20 **[0003]**

- Altschul, S. F., y col., *J. Mol. Biol.* 215: 403 – 410, 1990.  
 Altschul, S. F., y col., *Nucleic Acids Res.* 25: 3389 – 3402, 1997.  
 Aro, N., y col., *J. Biol. Chem.*, 10.1074 / M003624200, April 13, 2001.  
 25 Aubert, y col., Ed., pág. 11 y siguientes, Academic Press, 1988.  
 Ausubel G. M., y col. *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, John Wiley & Sons, New York, N.Y., 1993.  
 Baldwin, D., y col., *Curr. Opin. Plant Biol.* 2 (2): 96 – 103, 1999.  
 Baulcombe, D., *Arch. Virol. Suppl.* 15: 189 – 201, 1999.  
 30 Bhikhabhai, R. y col., *J. Appl. Biochem.* 6: 336, 1984.  
 Brumbauer, A. y col., *Bioseparation* 7: 287 – 295, 1999.  
 Carter y col., *Nucl. Acids Res.* 13: 4331, 1986.  
 Chen y col., *Biochem. Biophys. Acta.* 1121: 54 – 60, 1992.  
 Coligan, J. E. y col., eds., *CURRENT PROTOCOLS IN IMMUNOLOGY*, 1991.  
 35 Collen, A., y col., *Journal of Chromatography A* 910: 275 – 284, 2001.  
 Coughlan, y col., *BIOCHEMISTRY AND GENETICS OF CELLULOSE DEGRADATION*.  
 Cummings y Fowler, *Curr. Genet.* 29: 227 – 233, 1996.  
 Dayhoff y col. en *Atlas of Protein Sequence and Structure*, Volumen 5, Suplemento 3, Capítulo 22, pág. 345 – 352, 1978.  
 40 Deutscher, M.P., *Methods Enzymol.* 182: 779 – 80, 1990.  
 Doolittle, R. F., *OF URFs AND ORFs*, University Science Books, CA, 1986.  
 Ellouz, S. y col., *J. Chromatography* 396: 307, 1987.  
 Fields y Song, *Nature* 340: 245 – 246, 1989.  
 Filho, y col. *Can. J. Microbiol.* 42: 1 – 5, 1996.  
 45 Fliess, A., y col., *Eur. J. Appl. Microbiol. Biotechnol.* 17: 314, 1983.  
 Freer, y col. *J. Biol. Chem.* 268: 9337 – 9342, 1993.  
 Freshney, R. I., ed., *ANIMAL CELL CULTURE*, 1987.  
 Goyal, A. y col. *Bioresource Technol.* 36: 37, 1991.  
 Halldorsdottir, S y col., *Appl Microbiol Biotechnol.* 49 (3): 277 – 84, 1998.  
 50 Hu y col., *Mol. Cell Biol.* 11: 5792 – 9, 1991.  
 Hemmpel, W.H. *ITB Dyeing / Printing / Finishing* 3: 5 – 14, 1991.  
 Herr y col., *Appl. Microbiol. Biotechnol.* 5: 29 – 36, 1978.  
 Jakobovits, A, y col., *Ann. N. Y. Acad. Sci.* 764: 525 – 35, 1995.  
 Jakobovits, A, *Curr. Opin. Biotechnol.* 6 (5): 561 – 6, 1995.  
 55 Jones y col., *Nature* 321: 522 – 525, 1986.  
 Kawaguchi, T y col., *Gene* 173 (2): 287 – 8, 1996.  
 Knowles, J. y col., *TIBTECH* 5, 255 – 261, 1987.  
 Kohler y Milstein, *Nature* 256: 495, 1975.  
 Krishna, S. y col., *Bioresource Tech.* 77: 193 – 196, 2001.  
 60 Kumar, A., y col., *Textile Chemist and Colorist* 29: 37 – 42, 1997.  
 Lehtio, J. y col., *FEMS Microbiology Letters* 195: 197 – 204, 2001.  
 Li y Ljungdahl *Appl. Environ. Microbiol.* 62: 209 – 213, 1996.

- Linder, M. y Teeri, T.T., *Biotechnol.* 57: 15 – 28, 1997.  
 Medve, J. y col., *J. Chromatography A* 808: 153, 1998.  
 Ohmiya y col., *Biotechnol. Gen. Engineer. Rev.* 14: 365 – 414, 1997.  
 Ooi y col., *Nucleic Acids Res.* 18 (19): 5884, 1990.  
 5 Ortega y col., *International Biodeterioration and Biodegradation* 47: 7 – 14, 2001.  
 Penttila y col., *Yeast* 3: 175 – 185, 1987.  
 Penttila y col., *Gene* 63: 103 – 112, 1988.  
 Pere, J., y col., en *Proc. Tappi Pulping Conf., Nashville, TN*, 27 – 31, pág. 693 – 696, 1996.  
 Riechmann y col., *Nature* 332: 323 – 327, 1988.  
 10 Rothstein y col., *Gene* 55: 353 – 356, 1987.  
 Saarihahti y col., *Gene* 90: 9 – 14, 1990.  
 Sakamoto y col., *Curr. Genet.* 27: 435 – 439, 1995.  
 Saloheimo M, y col., *Gene* 63: 11 – 22, 1988.  
 Sambrook y col., *MOLECULAR CLONING: A LABORATORY MANUAL* (Second Edition), Cold Spring Harbor Press, Plainview, N.Y., 1989.  
 15 Schulein, *Methods Enzymol.*, 160, 25, páginas 234 y siguientes, 1988.  
 Scopes, *Methods Enzymol.* 90 Pt E: 479 – 90, 1982.  
 Spilliaert R, y col., *Eur. J. Biochem.* 224 (3): 923 – 30, 1994.  
 Stahlberg, J. y col., *Bio / Technol.* 9: 286 – 290, 1991.  
 20 Strathern y col., eds. (1981) *The Molecular Biology of the Yeast Saccharomyces*.  
 Suurnakki, A. y col., *Cellulose* 7: 189 – 209, 2000.  
 Te'o, J. y col., *FEMS Microbiology Letters* 190: 13 – 19, 2000.  
 Tilbeurgh, H. y col., *FEBS Lett.* 16: 215, 1984.  
 Timberlake y col., *Cell* 1: 29 – 37, 1981.  
 25 Tomaz, C. y Queiroz, J., *J. Chromatography A* 865: 123 – 128, 1999.  
 Tomme, P. y col., *Eur. J. Biochem.* 170: 575 – 581, 1988.  
 Tormo, J. y col., *EMBO J.* 15: 5739 – 5751, 1996.  
 Tyndall, R.M., *Textile Chemist and Colorist* 24: 23 – 26, 1992.  
 Van Rensburg y col., *Yeast* 14: 67 – 76, 1998.  
 30 Van Tilbeurgh, H. y col., *FEBS Lett.* 204: 223 – 227, 1986.  
 Verhoeyen y col., *Science* 239: 1534 – 1536, 1988.  
 Warrington, y col., *Genomics* 13: 803 – 808, 1992.  
 Wells y col., *Gene* 34: 315, 1985.  
 Wells y col., *Philos. Trans. R. Soc. London SerA* 317: 415, 1986.  
 35 Wood, *Biochem. Soc. Trans.*, 13, pág. 407 – 410, 1985.  
 Wood y col., *METHODS IN ENZYMOLOGY*, 160, 25, pág. 87 y siguientes, Academic Press, New York, 1988.  
 Zoller y col., *Nucl. Acids Res.* 10: 6487, 1987.

#### **Antecedentes de la invención**

- 40 **[0004]** La celulosa y hemicelulosa son los materiales vegetales más abundantes producidos por fotosíntesis. Los pueden degradar y usar como una fuente de energía numerosos microorganismos incluyendo bacterias, levaduras y hongos, que producen enzimas extracelulares capaces de hidrólisis de los sustratos polímeros a azúcares monómeros (Aro y col., 2001). Puesto que los límites de las energías no renovables están cerca, el  
 45 potencial de la celulosa para convertirse en una fuente principal de energía renovable es enorme (Krishna y col., 2001). La utilización eficaz de la celulosa mediante procesos biológicos es un procedimiento para superar la escasez de alimentos, piensos y combustibles (Ohmiya y col., 1997).
- [0005]** Las celulasas son enzimas que hidrolizan la celulosa (enlaces beta-1,4-glucano o beta-D-glucosídico) produciendo la formación de glucosa, celobiosa, celooligosacáridos y similares. Las celulasas tradicionalmente se han dividido en tres clases principales: endoglucanasas (EC 3.2.1.4) ("EG"), exoglucanasas o celobiohidrolasas (EC 3.2.1.91) ("CBH") y beta-glucosidasas ([beta]-D-glucósido glucohidrolasa; EC 3.2.1.21) ("BG"). (Knowles y col., 1987; Shulein, 1988). Las endoglucanasas actúan principalmente en las partes amorfas de la fibra de celulosa, mientras que las celobiohidrolasas también son capaces de degradar la celulosa cristalina (Nevalainen y Penttila, 1995). Por  
 55 lo tanto, es necesaria la presencia de una celobiohidrolasa en un sistema de celulasas para la solubilización eficaz de la celulosa cristalina (Suurnakki, y col. 2000). La beta-glucosidasa actúa para liberar unidades de D-glucosa de la celobiosa, celooligosacáridos, y otros glucósidos (Freer, 1993).
- [0006]** Se sabe que las celulasas son producidas por un gran número de bacterias, levaduras y hongos. Algunos hongos producen un sistema de celulasas completo capaz de degradar formas cristalinas de la celulosa, de modo que las celulasas son producidas fácilmente en grandes cantidades por fermentación. Los hongos filamentosos tienen una función especial puesto que muchas levaduras, tales como *Saccharomyces cerevisiae*,
- 60

carecen de la capacidad para hidrolizar la celulosa. Véase, p. ej., Aro y col., 2001; Aubert y col., 1988; Wood y col., 1988, y Coughlan, y col.

5 [0007] Las clasificaciones de celulasas fúngicas en CBH, EG y BG se pueden ampliar más para incluir múltiples componentes dentro de cada clasificación. Por ejemplo, se han aislado múltiples CBH, EG y BG de una variedad de fuentes fúngicas incluyendo *Trichoderma reesei* que contiene genes conocidos para 2 CBH, es decir, CBH I y CBH II, al menos 5 EG, es decir, EG I, EG II, EG III, EG IV y EG V, y al menos 2 BG, es decir, BG1 y BG2.

10 [0008] Con el fin de convertir eficazmente celulosa cristalina en glucosa, es necesario el sistema de celulasa completo que comprende componentes de cada una de las clasificaciones de CBH, EG y BG, siendo los componentes aislados menos eficaces en la hidrólisis de la celulosa cristalina (Filho y col., 1996). Se ha observado una relación sinérgica entre los componentes de la celulasa de diferentes clasificaciones. En particular, las celulasas de tipo EG y las celulasas de tipo CBH interactúan de forma sinérgica para degradar la celulosa más eficazmente. Véase, por ejemplo, Wood, 1985.

15 [0009] Es conocido en la materia que las celulasas son útiles en el tratamiento de productos textiles con el propósito de potenciar la capacidad de limpieza de composiciones de detergentes, para usar como un agente suavizante, para mejorar el tacto y el aspecto de las telas de algodón y similar (Kumar y col., 1997).

20 [0010] Se han descrito composiciones de detergentes que contienen celulasas con mejores características de limpieza (patente de EE.UU. n.º 4.435.307; solicitudes de GB n.º 2.095.275 y 2.094.826) y para usar en el tratamiento de la tela para mejorar el tacto y el aspecto del producto textil (patentes de EE.UU. n.º 5.648.263, 5.691.178. y 5.776.757; solicitud de GB n.º 1.358.599; The Shizuoka Prefectural Hamamatsu Textile Industrial Research Institute Report. Vol. 24. pp. 54 – 61. 1986).

25 [0011] Por lo tanto, las celulasas producidas en hongos y bacterias han recibido una atención significativa. En particular, se ha mostrado que la fermentación de *Trichoderma spp.* (p. ej., *Trichoderma longibrachiatum* o *Trichoderma reesei*) produce un sistema de celulasa completo capaz de degradar las formas cristalinas de la celulosa. La patente de EE.UU. n.º 5.475.101 describe la purificación y la clonación molecular de una enzima particularmente útil denominada EGIII que se obtiene de *Trichoderma longibrachiatum*.

30 [0012] TAKASHIMA S. Y COL. (*J. BIOCHEM.* vol. 124, n.º 4, enero 1999, páginas 728 – 736) describen la clonación y caracterización de *bgl4* y *bgl2* de *H. grisea* y *T. reesei* respectivamente.

35 [0013] IWASHITA K. Y COL. (*APPL. ENVIRON. MICROBIOL.* vol. 65, n.º 12, enero 1999, páginas 5546 – 5553) describen la clonación y caracterización de *bglA* de *A. kawachii*.

40 [0014] El documento US-A-6 022725 describe un procedimiento para expresar beta-glucosidasa extracelular en un hongo filamentoso mediante expresión de una secuencia de ADN fúngica que codifica la beta-glucosidasa potenciada, eliminada o alterada en un microorganismo huésped recombinante. También se describen composiciones de celulasa fúngica recombinante que contienen la expresión de la beta-glucosidasa potenciada, eliminada o alterada.

45 [0015] BRUMBAUER ANIKO Y COL. (*BIOSEPARATION*, vol. 7, n.º 6, 1999, páginas 287 – 295) describen el fraccionamiento de enzimas celulasa presentes en un medio de cultivo de *T. reesei*.

50 [0016] VINZANT T.B. Y COL. (*APPLIED BIOCHEMISTRY AND BIOTECHNOLOGY*, vol. 91 – 93, abril 2001 (2001 – 04), páginas 99 – 107) describen el uso de electroforesis 2D para analizar enzimas celulíticas de caldos de *T. reesei* comerciales.

55 [0017] FOREMAN P.K. Y COL. (*JOURNAL OF BIOLOGICAL CHEMISTRY, AMERICAN SOCIETY OF BIOLOGICAL CHEMISTS, BIRMINGHAM* US, vol. 278, n.º 34, 22 de agosto 2003 (2003 – 0822), páginas 31988 – 31997) describen la secuenciación parcial de 5100 clones de ADNc aleatorios de *T. reesei*. Se identificaron 12 genes que se consideró que podían funcionar en la degradación de biomasa.

60 [0018] Aunque se han descrito previamente composiciones de celulasa, siguen siendo necesarias composiciones de celulasa nuevas y mejoradas para usar en detergentes domésticos, composiciones de lavado a la piedra, o detergentes para la ropa, etc. Las celulasas que presentan resistencia a los tensioactivos (p. ej., sulfonatos de alquilo lineales, LAS), rendimiento mejorado en condiciones de estrés térmico, aumento o disminución de la capacidad celulítica, y/o alto nivel de expresión in vitro, tiene un interés particular.

**Resumen de la invención**

[0019] La invención proporciona una proteína celulasa aislada, identificada en el presente documento como BGL3, y ácidos nucleicos que codifican BGL3.

[0020] En un aspecto, los polipéptidos o proteínas BGL3 comprenden una secuencia que tiene una identidad de secuencia de al menos 98 % o más con la secuencia presentada en la SEQ ID NO: 2.

[0021] En el presente documento se describen (i) fragmentos de BGL3, preferiblemente de al menos aproximadamente 20 – 100 aminoácidos de longitud, más preferiblemente aproximadamente 100 – 200 aminoácidos de longitud, y (ii) una composición farmacéutica que comprende BGL3. En varias realizaciones, el fragmento corresponde al dominio N-terminal de BGL3 o al dominio C-terminal de BGL3.

[0022] También se describe en el presente documento un polinucleótido aislado que tiene una secuencia que codifica BGL3, una secuencia complementaria a la secuencia que codifica *bg/3*, y una composición que comprende el polinucleótido. El polinucleótido puede ser ARNm, ADN, ADNc, ADN genómico, o un análogo de sentido contrario de los mismos.

[0023] Un polinucleótido de *bg/3* puede comprender una molécula de ácido nucleico aislada que hibrida con el complemento del ácido nucleico presentado como SEQ ID NO: 1 en condiciones de restricción de moderada a alta, en el que la molécula de ácido nucleico codifica un polipéptido BGL3 que presenta actividad de beta-glucosidasa.

[0024] El polinucleótido puede codificar una proteína BGL3 que tiene una identidad de secuencia de al menos 98 % o más con la secuencia presentada como SEQ ID NO: 1. En una realización específica, el polinucleótido comprende una secuencia sustancialmente idéntica a la SEQ ID NO: 1. Los fragmentos del polinucleótido preferiblemente tienen al menos aproximadamente 15 – 30 nucleótidos de longitud.

[0025] En el presente documento se describen vectores de expresión recombinantes que contienen una secuencia de ácido nucleico que codifica BGL3 o un fragmento o una variante de empalme de la misma, operativamente unida a elementos reguladores eficaces para la expresión de la proteína en un huésped seleccionado. En un aspecto relacionado, la invención incluye una célula huésped que contiene el vector.

[0026] En el presente documento se describe un procedimiento para producir BGL3 por técnicas recombinantes, mediante cultivo de células huésped procariotas o eucariotas que comprenden la secuencia de ácido nucleico que codifica BGL3 en condiciones eficaces para promover la expresión de la proteína, y posteriormente recuperar la proteína de la célula huésped o del medio de cultivo celular.

[0027] También se describe un anticuerpo específicamente inmunorreactivo con BGL3.

[0028] También se describen procedimientos analíticos para detectar ácidos nucleicos de *bg/3* y proteínas BGL3.

**Breve descripción de las figuras**

[0029]

La figura 1 es una representación de una cadena de la secuencia de ácido nucleico (SEQ ID NO: 1) del ADNc de *bg/3* de *T. reesei*, en la que la secuencia no codificante está indicada en negrilla.

La figura 2 muestra la secuencia de aminoácidos predicha (SEQ ID NO: 2) y la secuencia señal (SEQ ID NO: 3) basada en la secuencia de nucleótidos proporcionada en la figura 1, en la que la secuencia señal está indicada en negrilla.

**Descripción detallada de la invención****I. Definiciones**

[0030] Salvo que se indique lo contrario, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que el que tendría para un experto en la materia de la presente invención. Se dirige en particular a los profesionales a Sambrook y col., 1989, y Ausubel FM y col., 1993, para las definiciones y términos de la materia. Debe entenderse que esta invención no está limitada a la metodología, protocolos y reaccionantes particulares descritos, ya que estos pueden variar.

5 [0031] El término “polipéptido” como se usa en el presente documento se refiere a un compuesto formado de una sola cadena de restos de aminoácidos unidos por enlaces peptídicos. El término “proteína” como se usa en el presente documento puede ser sinónimo del término “polipéptido” o se puede referir, además, a un complejo de dos o más polipéptidos.

10 [0032] La expresión “molécula de ácido nucleico” incluye moléculas ARN, ADN y ADNc. Se entenderá que, como resultado de la degeneración del código genético, se puede producir una multitud de secuencias de nucleótidos que codifican una proteína dada, tal como BGL3. La presente invención contempla todas las posibles secuencias de nucleótidos variantes que codifican BGL3, todas las cuales son posibles dada la degeneración del código genético.

15 [0033] Una construcción o secuencia de ácido nucleico “heteróloga” tiene una parte de la secuencia que no es natural de la célula en la que se expresa. Heterólogo con respecto a una secuencia de control se refiere a una secuencia de control (es decir, promotor o potenciador) que no funciona en la naturaleza para regular el mismo gen cuya expresión está regulando actualmente. En general, las secuencias de ácidos nucleicos heterólogas no son endógenas de la célula o parte del genoma en el que están presentes, y se han añadido a la célula por infección, transfección, transformación, microinyección, electroporación o similares. Una construcción de ácido nucleico “heteróloga” puede contener una combinación de secuencia de control / secuencia codificante de ADN que es la misma o es diferente de una combinación de secuencia de control / secuencia codificante de ADN encontrada en la célula natural.

20 [0034] Como se usa en el presente documento, el término “vector” se refiere a una construcción de ácido nucleico diseñada para la transferencia entre diferentes células huésped. Un “vector de expresión” se refiere a un vector que tiene la capacidad de incorporarse y expresar fragmentos de ADN heterólogos en una célula extraña. Muchos vectores de expresión procariontes y eucariotes están disponibles en el comercio. La selección de los vectores de expresión adecuados depende del conocimiento de los expertos en la materia.

25 [0035] Por consiguiente, un “casete de expresión” o “vector de expresión” es una construcción de ácido nucleico generada de forma recombinante o sintética, con una serie de elementos de ácido nucleico específicos que permiten la transcripción de un ácido nucleico particular en una célula diana. El casete de expresión recombinante se puede incorporar en un plásmido, cromosoma, ADN mitocondrial, ADN plásmido, virus o fragmento de ácido nucleico. Típicamente, la parte de casete de expresión recombinante de un vector de expresión incluye, entre otras secuencias, una secuencia de ácido nucleico para transcribir y un promotor.

30 [0036] Como se usa en el presente documento, el término “plásmido” se refiere a una construcción de ADN bicatenaria (bc) circular usada como un vector de clonación, y que forma un elemento genético de autorreplicación extracromosómico en muchas bacterias y algunos eucariotes.

35 [0037] Como se usa en el presente documento, la expresión “secuencia de nucleótidos que codifica un marcador seleccionable” se refiere a una secuencia de nucleótidos que es capaz de expresión en células y en la que la expresión de un marcador seleccionable confiere a las células que contienen el gen expresado, la capacidad de crecer en presencia del agente selectivo correspondiente, o en las correspondientes condiciones de crecimiento selectivas.

40 [0038] Como se usa en el presente documento, el término “promotor” se refiere a una secuencia de ácido nucleico que funciona para dirigir la transcripción de un gen en la dirección 3'. El promotor en general será adecuado para la célula huésped en la que el gen diana se va a expresar. El promotor junto con otras secuencias de ácido nucleico reguladoras de la transcripción y traducción (denominadas también “secuencias de control”) son necesarias para expresar un gen dado. En general las secuencias reguladoras de la transcripción y traducción incluyen, pero sin limitar, secuencias promotoras, sitios de unión al ribosoma, secuencias de inicio y fin de la transcripción, secuencias de inicio y fin de la traducción, y secuencias potenciadoras o activadoras.

45 [0039] “Gen quimérico” o “construcción de ácido nucleico heteróloga”, como se definen en el presente documento, se refieren a un gen no natural (es decir, uno que se ha introducido en un huésped) que puede estar compuesto por partes de genes diferentes, incluyendo elementos reguladores. Una construcción de gen quimérico para transformar una célula huésped típicamente está compuesta de una región reguladora de la transcripción (promotor) operativamente unida a una secuencia que codifica una proteína heteróloga, o en un gen quimérico con marcador seleccionable, a un gen marcador seleccionable que codifica una proteína que confiere resistencia a antibióticos a las células transformadas. Un gen quimérico típico de la presente invención, para la transformación en una célula huésped, incluye una región reguladora de la transcripción que es constitutiva o inducible, una secuencia que codifica proteína, y una secuencia de terminación. Una construcción de gen quimérico puede incluir también una  
55 60 segunda secuencia de ADN que codifica un péptido señal, si se desea la secreción de la proteína diana.

**[0040]** Un ácido nucleico está “operativamente unido” cuando está situado en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, el ADN que codifica una secuencia líder secretora está operativamente unido a ADN para un polipéptido, si es expresado como una preproteína que participa en la secreción del polipéptido; un promotor o potenciador está operativamente unido a una secuencia codificante si afecta a la transcripción de la secuencia; o un sitio de unión al ribosoma está operativamente unido a una secuencia codificante, si está situado de forma que facilita su traducción. En general, “operativamente unido” significa que las secuencias de ADN que se están uniendo son contiguas, y, en el caso de una secuencia líder secretora, contiguas y en el marco de lectura. Sin embargo, los potenciadores no tienen que estar contiguos. La unión se lleva a cabo mediante ligado a sitios de de restricción convenientes. Si dichos sitios no existen, se usan adaptadores oligonucleótidos sintéticos, conectores o cebadores para la PCR, de acuerdo con la práctica convencional.

**[0041]** Como se usa en el presente documento, el término “gen” significa el segmento de ADN implicado en la producción de una cadena de polipéptido, que puede incluir o no regiones precedentes y siguientes a la región codificante, p. ej., secuencias 5’ no traducidas (5’ UTR) o “líder” y secuencias 3’ UTR o “remolque”, así como secuencias intercaladas (intrones) entre segmentos codificantes individuales (exones).

**[0042]** En general, las moléculas de ácido nucleico que codifican BGL3 o un análogo u homólogo de la misma, hibridarán en condiciones de restricción de moderadas a altas, con la secuencia proporcionada en el presente documento como SEQ ID NO: 1. Sin embargo, en algunos casos se usa una secuencia de nucleótidos que codifica BGL3 que tiene un uso de codón sustancialmente diferente, mientras que la proteína codificada por la secuencia de nucleótidos que codifica BGL3 tiene la misma o sustancialmente la misma secuencia de aminoácidos que la proteína natural. Por ejemplo, la secuencia codificante puede estar modificada para facilitar la expresión más rápida de BGL3 en un sistema de expresión procariota o eucariota particular, de acuerdo con la frecuencia con la que el huésped usa un codón particular. Te'o, y col. (2000), por ejemplo, describen la optimización de genes para la expresión en hongos filamentosos.

**[0043]** Una secuencia de ácido nucleico se considera que es “selectivamente hibridable” con una secuencia de ácido nucleico de referencia, si las dos secuencias hibridan específicamente entre sí en condiciones de hibridación y lavado de restricción de moderada a alta. Las condiciones de hibridación se basan en la temperatura de fusión ( $T_m$ ) del complejo de unión al ácido nucleico o sonda. Por ejemplo, la “restricción máxima” típicamente se produce a aproximadamente  $T_m - 5^\circ\text{C}$  ( $5^\circ$  por debajo de la  $T_m$  de la sonda); “restricción alta” a aproximadamente  $5 - 10^\circ$  por debajo de la  $T_m$ ; “restricción intermedia” a aproximadamente  $10 - 20^\circ$  por debajo de la  $T_m$  de la sonda; y “restricción baja” a aproximadamente  $20 - 25^\circ$  por debajo de la  $T_m$ . Funcionalmente, se pueden usar condiciones de restricción máxima para identificar secuencias que tienen una identidad estricta o identidad cercana a la estricta con la sonda de hibridación; mientras que las condiciones de restricción alta se usan para identificar secuencias que tienen una identidad de secuencia de aproximadamente 80 % o más con la sonda.

**[0044]** Las condiciones de hibridación de restricción moderada a alta son bien conocidas en la materia (véase, por ejemplo, Sambrook, y col., 1989, capítulos 9 y 11, y en Ausubel, F.M., y col., 1993, incorporados expresamente por referencia en el presente documento). Un ejemplo de condiciones de restricción alta incluye la hibridación a aproximadamente  $42^\circ\text{C}$  en formamida al 50 %, 5X SSC, 5X disolución de Denhardt, SDS al 0,5 % y ADN vehículo desnaturalizado  $100\ \mu\text{g} / \text{ml}$  seguido de lavado 2 veces en 2X SSC y SDS al 0,5 % a temperatura ambiente y dos veces adicionales en 0,1 X SSC y SDS al 0,5 % a  $42^\circ\text{C}$ .

**[0045]** Como se usa en el presente documento, “recombinante” incluye la referencia a una célula o vector, que se ha modificado por la introducción de una secuencia de ácido nucleico heteróloga o que la célula deriva de una célula así modificada. Así, por ejemplo, las células recombinantes expresan genes que no se encuentran en forma idéntica en la forma natural (no recombinante) de la célula o expresan genes naturales que de otra manera son expresados de forma anormal, expresados en defecto o no son expresados en absoluto, como resultado de la intervención humana deliberada.

**[0046]** Como se usa en el presente documento, los términos “transformada”, “establemente transformada” o “transgénica” con referencia a una célula significa que la célula tiene una secuencia de ácido nucleico no natural (heteróloga) integrada en su genoma o como un plásmido episomal que es mantenido a lo largo de múltiples generaciones.

**[0047]** Como se usa en el presente documento, el término “expresión” se refiere al procedimiento por el que un polipéptido es producido basándose en la secuencia de ácido nucleico de un gen. El procedimiento incluye tanto transcripción como traducción.

**[0048]** El término “introducido” en el contexto de insertar una secuencia de ácido nucleico en una célula, significa “transfección” o “transformación” o “transducción” e incluye la referencia a la incorporación de una secuencia de

ácido nucleico en una célula eucariota o procariota en la que la secuencia de ácido nucleico se puede incorporar en el genoma de la célula (por ejemplo, cromosoma, plásmido, plástido o ADN mitocondrial), convertido en un replicón autónomo, o expresado de forma transitoria (por ejemplo, ARNm transfectado).

5 **[0049]** En consecuencia, la frase “expresión de BGL3” se refiere a la transcripción y traducción del gen *bg13*, cuyos productos incluyen ARN precursor, ARNm, polipéptido, polipéptidos procesados postraduccionalmente, y derivados de los mismos, incluyendo BGL3 de especies relacionadas tales como *Trichoderma longibrachiatum (reesei)*, *Trichoderma viride*, *Trichoderma koningii*, *Hypocrea jecorina* e *Hypocrea schweinitzii*. A modo de ejemplo, los ensayos para la expresión de BGL3 incluyen transferencia Western para la proteína BGL3, análisis de transferencia Northern y ensayos de reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR) para el ARNm de BGL3, y ensayos de actividad de glucosidasa, como describen Chen y col. (1992) y Herr y col. (1978).

15 **[0050]** La expresión “empalme alternativo” se refiere al procedimiento por el que se generan múltiples isoformas de polipéptidos a partir de un solo gen, e implica el empalme entre sí de exones no consecutivos durante el procesamiento de algunos, pero no todos, los transcritos del gen. Por lo tanto, un exón particular se puede conectar a cualquiera de varios exones alternativos para formar ARN mensajeros. Los ARNm empalmados de forma alternativa producen polipéptidos (“variantes de empalme”) en los que algunas partes son comunes mientras que otras partes son diferentes.

20 **[0051]** La expresión “secuencia señal” se refiere a una secuencia de aminoácidos en la parte N-terminal de una proteína que facilita la secreción de la forma madura de la proteína fuera de la célula. La forma madura de la proteína extracelular carece de secuencia señal que es escindida durante el proceso de secreción.

25 **[0052]** Por la expresión “célula huésped” se entiende una célula que contiene un vector que soporta la replicación, y/o la transcripción o transcripción y traducción (expresión) de la construcción de expresión. Las células huésped para usar en la presente invención pueden ser células procariotas, tales como *E. coli*, o células eucariotas como células de levaduras, plantas, insectos, anfibios o mamífero. En general, las células huésped son hongos filamentosos.

30 **[0053]** La expresión “hongos filamentosos” significa todos y cada uno de los hongos filamentosos reconocidos por los expertos en la materia. Un hongo filamentoso preferido se selecciona del grupo de *Aspergillus*, *Trichoderma*, *Fusarium*, *Chrysosporium*, *Penicillium*, *Humicola*, *Neurospora*, o formas sexuales alternativas de los mismos tales como *Emericella*, *Hypocrea*.

35 **[0054]** El término “celooligosacárido” se refiere a grupos de oligosacáridos que contienen 2 – 8 unidades de glucosa y que tienen uniones  $\beta$ -1,4, p. ej., celobiosa.

40 **[0055]** El término “celulasa” se refiere a una categoría de enzimas capaces de hidrolizar polímeros de celulosa a oligómeros de celooligosacáridos más cortos, celobiosa y/o celulosa. Se han obtenido numerosos ejemplos de celulasas, tales como exoglucanasas, exocelobiohidrolasas, endoglucanasas, y glucosidasas a partir de organismos celulóticos, en particular incluyendo hongos, plantas y bacterias.

45 **[0056]** La expresión “dominio de unión a la celulosa” como se usa en el presente documento, se refiere a la parte de la secuencia de aminoácidos de una celulasa o a una región de la enzima que está implicada en la actividad de unión a la celulosa, de una celulasa o derivado de la misma. Los dominios de unión a la celulosa en general funcionan por unión no covalente de la celulasa a la celulosa, un derivado de celulosa u otro polisacárido equivalente del mismo. Los dominios de unión a la celulosa permiten o facilitan la hidrólisis de fibras de celulosa por la región central catalítica estructuralmente distinta, y típicamente funcionan independiente del centro catalítico. Por lo tanto, un dominio de unión a la celulosa no tendrá la actividad hidrolítica significativa atribuible a un centro catalítico. En otras palabras, un dominio de unión a la celulosa es un elemento estructural de la estructura terciaria de la proteína enzimática celulasa que es distinto del elemento estructural que tiene actividad catalítica.

50 **[0057]** Como se usa en el presente documento, el término “tensioactivo” se refiere a cualquier compuesto reconocido en general en la materia como que tiene cualidades tensioactivas. Así, por ejemplo, los tensioactivos comprenden tensioactivos aniónicos, catiónicos y no iónicos tales como los que se encuentran habitualmente en los detergentes. Los tensioactivos aniónicos incluyen alquilbencenosulfonatos lineales o ramificados, alquil- o alquenil-éter-sulfatos que tienen grupos alquilo o grupos alquenilo lineales o ramificados; sulfatos de alquilo o alquenilo; oleinsulfonatos; y alcanosulfonatos. Los tensioactivos anfólicicos incluyen sales de sulfonatos de amonio cuaternario, y tensioactivos anfólicicos de tipo betaína. Dichos tensioactivos anfólicicos tienen grupos cargados tanto positivos como negativos en la misma molécula. Los tensioactivos no iónicos pueden comprender éteres polioxoalquilenos, así como alcanolamidas de ácido graso superior o aducto de óxido de alquilenos del mismo, monoésteres de glicerina y ácido graso, y similares.

**[0058]** Como se usa en el presente documento, la expresión “tela que contiene celulosa” se refiere a cualquier tela cosida o no cosida, hilos o fibras hechas de algodón o que no son de algodón que contienen celulosa o mezclas de algodón y no algodón que contienen celulosa, incluyendo productos celulósicos naturales y celulósicos hechos por el hombre (tales como yute, lino, ramio, rayón y liocel).

**[0059]** Como se usa en el presente documento, la expresión “tela que contiene algodón” se refiere a telas cosidas o no cosidas, hilos o fibras hechas de algodón puro o mezclas de algodón que incluyen telas tejidas de algodón, telas tricotadas de algodón, tejanos de algodón, hilos de algodón, algodón en bruto y similares.

**[0060]** Como se usa en el presente documento, la expresión “composición de lavado a la piedra” se refiere a una formulación para usar en el lavado a la piedra de telas que contienen celulosa. Las composiciones de lavado a la piedra se usan para modificar las telas que contienen celulosa antes de la venta, es decir, durante el procedimiento de fabricación. A diferencia de esto, las composiciones de detergentes están destinadas a la limpieza de prendas sucias y no se usan durante el procedimiento de fabricación.

**[0061]** Como se usa en el presente documento, la expresión “composición de detergente” se refiere a una mezcla que está destinada a usar en un medio de lavado para el lavado de telas sucias que contienen celulosa. En el contexto de la presente invención, dichas composiciones pueden incluir, además de celulasas y tensioactivos, enzimas hidrolíticas adicionales, mejoradores, agentes blanqueantes, activadores de blanqueadores, agentes añiles y colorantes fluorescentes, inhibidores del apelmazamiento, agentes enmascarantes, activadores de celulasa, antioxidantes y solubilizantes.

**[0062]** Como se usa en el presente documento, la expresión “disminución o eliminación de la expresión del gen *bgl3*” significa que el gen *bgl3* se ha eliminado del genoma y por lo tanto no puede ser expresado por el microorganismo huésped recombinante; o que el gen *bgl3* se ha modificado de modo que el microorganismo huésped recombinante no produce una enzima BGL3 funcional.

**[0063]** Las expresiones “*bgl3* alterado” o “gen *bgl3* alterado” significa que la secuencia de ácido nucleico del gen se ha alterado eliminando, añadiendo y/o manipulando la secuencia codificante o se ha modificado la secuencia de aminoácidos de la proteína expresada.

**[0064]** Como se usa en el presente documento, el término “purificar” se refiere en general a someter al ácido nucleico transgénico o a las células que contienen proteínas a purificación bioquímica y/o cromatografía en columna.

**[0065]** Como se usa en el presente documento, las expresiones “activo” y “biológicamente activo” se refieren a una actividad biológica asociada con una proteína particular, tal como la actividad enzimática asociada con una proteasa. En consecuencia, la actividad biológica de una proteína dada se refiere a cualquier actividad biológica típicamente atribuida a esa proteína por el experto en la materia.

## II. Organismos diana

### A. Hongos filamentosos

**[0066]** Los hongos filamentosos incluyen todas las formas filamentosas de la subdivisión Eumycota y Oomycota. Los hongos filamentosos se caracterizan por micelio vegetativo que tienen una pared celular compuesta de quitina, glucano, chitosán, manano y otros polisacáridos complejos, con crecimiento vegetativo por elongación de la hifa y catabolismo de carbono que es obligadamente aerobio.

**[0067]** En la presente invención, la célula fúngica filamentosa original puede ser una célula de una especie de, pero sin limitar, *Trichoderma*, p. ej., *Trichoderma longibrachiatum* (reesei), *Trichoderma viride*, *Trichoderma koningii*, *Trichoderma harzianum*; *Penicillium* sp.; *Humicola* sp., incluyendo *Humicola insolens*; *Chrysosporium* sp., incluyendo *C. lucknowense*; *Gliocladium* sp.; *Aspergillus* sp.; *Fusarium* sp., *Neurospora* sp., *Hypocrea* sp., y *Emericella* sp. Como se usa en el presente documento, el término “*Trichoderma*” o “*Trichoderma* sp.” se refiere a cualquier cepa de hongo que se haya clasificado previamente como *Trichoderma* o se clasifique actualmente como *Trichoderma*.

**[0068]** En una realización preferida, la célula fúngica filamentosa original es una célula de *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus aculeatus*, o *Aspergillus nidulans*.

**[0069]** En otra realización preferida, la célula fúngica filamentosa original es una célula de *Trichoderma reesei*.

### III. Celulasas

**[0070]** Las celulasas son conocidas en la materia como enzimas que hidrolizan celulosa (enlaces beta-1,4-glucano o beta-D-glucosídico) dando como resultado la formación de glucosa, celobiosa, celooligosacáridos, y similares. Como se ha expuesto antes, las celulasas se han dividido tradicionalmente en tres clases principales: endoglucanasas (EC 3.2.1.4) ("EG"), exoglucanasas o celobiohidrolasas (EC 3.2.1.91) ("CBH") y beta-glucosidasas (EC 3.2.1.21) ("BG"). (Knowles, y col., 1987; Schulein, 1988).

**[0071]** Algunos hongos producen sistemas de celulasa completos que incluyen exocelobiohidrolasas o celulasas de tipo CBH, endoglucanasas o celulasas de tipo EG y beta-glucosidasas de celulasas de tipo BG (Schulein, 1988). Sin embargo, a veces estos sistemas carecen de celulasas de tipo CBH y las celulasas bacterianas también incluyen típicamente poco o nada de celulasas de tipo CBH. Además, se sabe que los componentes de EG y los componentes de CBH interactúan de forma sinérgica para degradar la celulosa de forma más eficaz. Véase, p. ej., Wood, 1985. Los diferentes componentes, es decir, las diferentes endoglucanasas y exocelobiohidrolasas en un sistema de celulasa de multicomponentes o completo, en general tienen diferentes propiedades, tales como el punto isoelectrónico, peso molecular, grado de glicosilación, especificidad de sustrato y patrones de acción enzimáticos.

**[0072]** Se cree que las celulasas de tipo endoglucanasa hidrolizan enlaces internos beta-1,4-glucosídicos en regiones de baja cristalinidad de la celulosa y las celulasas de tipo exocelobiohidrolasa hidrolizan la celobiosa desde el extremo reductor o no reductor de la celulosa. Se deduce que la acción de los componentes de las endonucleasas puede facilitar mucho la acción de las exocelobiohidrolasas creando nuevos extremos de cadena que son reconocidos por componentes de las exocelobiohidrolasas. Además, se ha mostrado que las celulasas de tipo beta-glucosidasa catalizan la hidrólisis de alquil- y/o aril-beta-D-glucósidos tales como metil-beta-D-glucósido y p-nitrofenil-glucósido así como glucósidos que contienen solo restos carbohidrato, tales como celobiosa. Esto da glucosa como el único producto para el microorganismo y reduce o elimina la celobiosa que inhibe las celobiohidrolasas y las endoglucanasas.

**[0073]** Por consiguiente, se considera que las celulasas de tipo beta-glucosidasa son una parte integral del sistema de celulasa porque dirigen la reacción general a la glucosa. Se ha mostrado que la mayor expresión de BG en *T. reesei* mejora la degradación de la celulosa en glucosa. Véase el documento EP0562003. Además, las beta-glucosidasas pueden catalizar la hidrólisis de una serie de sustratos diferentes, y por lo tanto son útiles en una variedad de aplicaciones diferentes. Algunas beta-glucosidasas se pueden añadir a las uvas durante la preparación de vino para potenciar el aroma potencial del producto de vino acabado. Otra aplicación más puede ser usar la beta-glucosidasa en la fruta para potenciar el aroma de la misma. Alternativamente, la beta-glucosidasa se puede usar directamente en aditivos alimentarios o en el procesamiento del vino para potenciar el sabor y el aroma.

**[0074]** Las celulasas también tienen una serie de usos en composiciones de detergentes que incluyen potenciar la capacidad de limpieza, como agente suavizante y para mejorar el tacto de las telas de algodón (Hemmpel, 1991; Tyndall, 1992; Kumar y col., 1997). Aunque el mecanismo no es parte de la invención, las propiedades de restauración de la suavidad y el color de la celulosa se han atribuido a los componentes de la endoglucanasa alcalina en las composiciones de celulasa, como se ilustra en las patentes de EE.UU. n.º 5.648.263, 5.691.178 y 5.776.757, que describen que las composiciones de detergente que contienen una composición de celulasa enriquecida en un componente de endoglucanasa alcalina específico, imparten restauración del color y mejora de la suavidad a las prendas tratadas comparado con las composiciones de celulasa que no están enriquecidas en dicho componente. Además, el uso de dichos componentes de endoglucanasa alcalina en composiciones de detergentes se ha mostrado que complementa los requisitos de pH de la composición de detergente (p. ej., presentando la actividad máxima a un pH alcalino de 7,5 a 10, como se describe en las patentes de EE.UU. n.º 5.648.263, 5.691.178 y 5.776.757).

**[0075]** También se ha mostrado que las composiciones de celulasas degradan las telas que contienen algodón, dando como resultado una menor pérdida de resistencia en la tela (patente de EE.UU. n.º 4.822.516), contribuyendo a la renuncia al uso de composiciones de celulasas en aplicaciones de detergentes comerciales. Se ha sugerido que las composiciones de celulasas que comprenden componentes de endoglucanasa presentan menor pérdida de la resistencia para las telas que contienen algodón comparado con las composiciones que comprenden un sistema de celulasas completo.

**[0076]** También se ha mostrado que las celulasas son útiles en la degradación de biomasa por la celulasa a etanol (en la que la celulasa degrada la celulosa a glucosa y levaduras u otros microbios fermentan además la glucosa en etanol), en el tratamiento de la pasta mecánica (Pere y col., 1996) para usar como un aditivo de piensos (documento WO 91 / 04673) y en la molienda en húmedo del grano.

[0077] La mayoría de las CBH y EG tienen una estructura de multidominios que consiste en un dominio central separado de un dominio de unión a la celulosa (CBD) por un péptido conector (Suurnakki y col., 2000). El dominio central contiene el sitio activo mientras que el CBD interacciona con la celulosa uniendo la enzima al mismo (van Tilbeurgh y col., 1986; Tomme y col., 1988). Los CBD son particularmente importantes en la hidrólisis de celulosa cristalina. Se ha mostrado que la capacidad de las celobiohidrolasas para degradar la celulosa cristalina claramente disminuye cuando está ausente el CBD (Linder y Teeri, 1997). Sin embargo, la función exacta y el mecanismo de acción de los CBD es una cuestión sobre la que todavía se especula. Se ha sugerido que el CBD potencia la actividad enzimática simplemente aumentando la concentración enzimática eficaz en la superficie de la celulosa (Stahlberg y col. 1991), y/o liberando cadenas de celulosa sencillas de la superficie de celulosa (Tormo y col., 1996). La mayoría de los estudios relacionados con los efectos de los dominios de celulasa en los diferentes sustratos, se han llevado a cabo con las proteínas centrales de las celobiohidrolasas, ya que sus proteínas centrales se pueden producir fácilmente por proteólisis limitada con papaína (Tomme y col., 1988). (Nota: en relación con el párrafo anterior, no es relevante para las BG).

[0078] Se han descrito numerosas celulasas en la bibliografía científica, cuyos ejemplos incluyen: *Trichoderma reesei*: Shoemaker, S. y col., *Bio / Technology*, 1: 691 – 696, 1983, que describen *CBHI*; Teeri, T. y col., *Gene*, 51: 43 – 52, 1987, que describen *CBHII*; Penttilä, M. y col., *Gene*, 45: 253 – 263, 1986, que describen *EGI*; Saloheimo, M. y col., *Gene*, 63: 11 – 22, 1988, que describen *EGII*; Okada, M. y col., *Appl. Environ. Microbiol.*, 64: 555 – 563, 1988, que describen *EGIII*; Saloheimo, M. y col., *Eur. J. Biochem.*, 249: 584 – 591, 1997, que describen *EGIV*; Saloheimo, A. y col., *Molecular Microbiology*, 13: 219 – 228, 1994, que describen *EGV*; Barnett, C.C., y col., *Bio / Technology*, 9: 562 – 567, 1991, que describen *BGL1*, y Takashima, S. y col., *J. Biochem.*, 125: 728 – 736, 1999, que describen *BGL2*. También se han descrito celulasas de especies distintas de tricodermas, p. ej., Ooi y col., 1990, que describen la secuencia de ADNc que codifica la endoglucanasa F1-CMC producida por *Aspergillus aculeatus*; Kawaguchi T. y col., 1996, que describen la clonación y secuenciación del ADNc que codifica la beta-glucosidasa 1 de *Aspergillus aculeatus*; Sakamoto y col., 1995, que describen la secuencia de ADNc que codifica la endoglucanasa CMCasa-1 de *Aspergillus kawachii* IFO 4308; Saarilahti y col., 1990 que describen una endoglucanasa de *Erwinia carotovora*; Spilliaert R, y col., 1994, que describen la clonación y secuenciación de *bglA*, que codifica una beta-glucanasa termoestable de *Rhodothermus marinus*; y Halldorsdottir S y col., 1998, que describen la clonación, secuenciación y exceso de expresión de un gen de *Rhodothermus marinus* que codifica una celulasa termoestable de la familia 12 de glicosil hidrolasas. Sin embargo, sigue siendo necesaria la identificación y caracterización de nuevas celulasas, con propiedades mejoradas, tales como mejor comportamiento en condiciones de estrés térmico o en la presencia de tensioactivos, mayor actividad específica, patrón de escisión del sustrato alterado, y/o nivel de expresión alto in vitro.

[0079] El desarrollo de composiciones de celulasa nuevas y mejoradas que comprenden diferentes cantidades de celulasas de tipo CHB, tipo EG y tipo BG tiene interés para usar: (1) en composiciones de detergentes que presentan una mejor capacidad de limpieza, función como un agente suavizante y/o de mejora del tacto de las telas de algodón (p. ej., “lavado a la piedra” o “biopulido”); (2) en composiciones para degradar la pasta de madera u otras biomásas en azúcares (p. ej., producción de bioetanol); y/o (3) en composiciones de piensos.

#### IV. Procedimientos para identificar nuevas secuencias

[0080] Se analizan los marcos de lectura abiertos (ORF) después de la secuenciación parcial o total del genoma de *T. reesei* o de los clones de bibliotecas de ADNc de ARNm de *T. reesei* y se analizan además usando el software de análisis de secuencia, y determinando la homología con secuencias conocidas en bases de datos (públicas / privadas).

#### V. Ácidos nucleicos *bgl3* y polipéptidos BGL3

##### A. Ácidos nucleicos *bgl3*

[0081] Las moléculas de ácido nucleico descritas en el presente documento incluyen la secuencia codificante natural, la secuencia de ADNc para *bgl3* presentada en el presente documento como SEQ ID NO: 1, y homólogos de las mismas en otras especies, variantes alélicas y de empalme naturales, fragmentos de ácido nucleico, y derivados de los mismos biológicamente activos (funcionales), tales como variantes de la secuencia de aminoácidos de la molécula natural y secuencias que codifican proteínas de fusión. Las secuencias se denominan colectivamente en el presente documento “secuencias de ácido nucleico que codifican BGL3”.

[0082] Se llevó a cabo una búsqueda con BLASTN Basic (<http://www.ncbi.nlm.nih.gov/BLAST>) de la base de datos de secuencias de ácidos nucleicos no redundantes, el 17 de agosto de 2001, con la secuencia del gen *bgl3* presentada como SEQ ID NO: 1, que indicó cuatro secuencias que producían alineamiento significativos (es decir, con un valor E menor de  $10^{-5}$ ). Estas cuatro secuencias se anotaron como sigue. ADN para la beta-D-glucosidasa de

*Aspergillus kawachii* (número de acceso en GenBank AB003470); gen de beta-glucosidasa de *Pichia capsulata* (bglN) (número de acceso en GenBank U16259); ARNm para beta-glucosidasa de *Aspergillus aculeatus* (número de acceso en GenBank D64088); similar a la secuencia del gen de beta-glucosidasa de *Aspergillus niger* (número de acceso en GenBank AF121777). La secuencia que codifica BGL3 (la secuencia que no está en negrilla en la Fig. 1) se alineó con la secuencia codificante del gen beta-glucosidasa de *A. kawachii* (número de acceso en GenBank AB003470 con regiones no traducidas, incluyendo intrones no eliminados), usando el programa CLUSTAL-W que puso de manifiesto una identidad de nucleótidos de 65 % entre las dos secuencias.

**[0083]** La secuencia del ácido nucleico de *bg/3* puede ser una secuencia de ADN o ARN, derivada del ADN genómico, ADNc, ARNm, o se puede sintetizar entera o en parte. El ADN puede ser bicatenario o monocatenario, y si es monocatenario puede ser la cadena codificante o la cadena no codificante (sentido contrario, complementaria). La secuencia de ácido nucleico se puede clonar, por ejemplo, aislando el ADN genómico de una fuente adecuada y amplificando y clonando la secuencia de interés usando una reacción en cadena de la polimerasa (PCR). Alternativamente, la secuencia de ácido nucleico se puede sintetizar, completamente o en parte, en especial cuando se desea proporcionar secuencias preferidas del huésped para la expresión óptima. Por lo tanto, todo o una parte del gen estructural deseado (la parte del gen que codifica un polipéptido o proteína) se puede sintetizar usando codones preferidos por un huésped seleccionado.

**[0084]** Debido a la degeneración inherente del código genético se pueden usar secuencias de ácido nucleico distintas de la forma nativa que codifican sustancialmente la misma secuencia o una secuencia de aminoácidos funcionalmente equivalente para clonar y/o expresar las secuencias de ácido nucleico que codifican BGL3. Por lo tanto, para una secuencia dada de ácido nucleico que codifica BGL3, se apreciará que como resultado de la degeneración del código genético, se puede producir una serie de secuencias codificantes que codifican una proteína que tiene la misma secuencia de aminoácidos. Por ejemplo, el triplete CGT codifica el aminoácido arginina. La arginina es codificada alternativamente por CGA, CGC, CGG, AGA y AGG. Por lo tanto, se apreciará que dichas sustituciones en la región codificante están dentro de las variantes de la secuencia de ácido nucleico descritas en el presente documento. Se pueden usar todas y cada una de estas variantes de secuencias de la misma forma descrita en el presente documento para la forma natural de la secuencia de ácido nucleico que codifica BGL3.

**[0085]** Una "variante" de la secuencia de ácido nucleico que codifica BGL3 puede codificar una "variante" de la secuencia de aminoácidos de BGL3 que está alterada por uno o más aminoácidos respecto a la secuencia de polipéptido natural, o puede estar truncada por eliminación de uno o más aminoácidos de cualquiera de los extremos de la secuencia de polipéptido. Igualmente, la expresión "forma modificada de", con respecto a BGL3, significa un derivado o forma variante de la secuencia natural de ácido nucleico que codifica la proteína BGL3 o la secuencia natural de aminoácidos de BGL3.

**[0086]** Igualmente, los polinucleótidos descritos en el presente documento incluyen secuencias que codifican proteínas BGL3 naturales y variantes de empalme de las mismas, secuencias complementarias de la secuencia que codifica la proteína natural, y nuevos fragmentos de polinucleótidos que codifican BGL3. Una secuencia de ácido nucleico que codifica BGL3 puede contener una o más secuencias de intrones si es una secuencia de ADN genómico.

**[0087]** En una realización general, la secuencia de nucleótidos que codifica BGL3 tiene una identidad de secuencia de al menos 98 % o más con la secuencia que codifica *bg/3* presentada en el presente documento como SEQ ID NO: 1.

**[0088]** En otra realización, una secuencia de nucleótidos que codifica BGL3 hibridará en condiciones de restricción alta con la secuencia de nucleótidos presentada como SEQ ID NO: 4, como se define en las reivindicaciones.

**[0089]** Se apreciará que algunas variantes de secuencia de ácido nucleico que codifican BGL3 pueden hibridar o no de forma selectiva con la secuencia original. A modo de ejemplo, en situaciones en las que la secuencia codificante se ha optimizado basándose en la degeneración del código genético, se puede producir una secuencia codificante variante que codifica una proteína BGL3, pero no hibrida con una secuencia natural de ácido nucleico que codifica BGL3 en condiciones de restricción alta. Esto ocurriría, por ejemplo, cuando la variante de la secuencia incluye un codón diferente para cada uno de los aminoácidos codificados por el nucleótido original.

**[0090]** Como entenderán mejor los expertos en la materia, en algunos casos puede ser ventajoso producir secuencias de nucleótidos que tienen codones no naturales, p. ej., inosina u otro análogo de nucleótido no natural. Se pueden seleccionar los codones preferidos por un huésped eucariota particular, por ejemplo, para aumentar la tasa de expresión de la proteína BGL3 o para producir transcritos de ARN recombinante que tengan las propiedades deseadas, tales como una semivida más larga que los transcritos producidos a partir de la secuencia natural. Por lo tanto, una secuencia natural de nucleótidos que codifica BGL3 se puede transformar con el fin de alterar la

secuencia codificante por una variedad de razones, que incluyen, pero sin limitar, alteraciones que modifican la clonación, procesamiento y/o expresión de la proteína BGL3 por una célula.

5 [0091] Se prefieren en particular las sustituciones, adiciones y eliminaciones del ácido nucleico que son silenciosas, de modo que no alteran las propiedades o actividades del polinucleótido o polipéptido natural.

10 [0092] Las variaciones se pueden hacer usando procedimientos conocidos en la materia tales como mutagénesis mediada por oligonucleótido (dirigida al sitio), y mutagénesis por PCR. Se puede llevar a cabo mutagénesis dirigida al sitio (Carter y col., 1986; Zoller y col., 1987), mutagénesis en casete (Wells y col., 1985), mutagénesis por selección de restricción (Wells y col., 1986) u otras técnicas conocidas en el ADN clonado para producir el ADN variante que codifica el polipéptido BGL3.

15 [0093] Sin embargo, en algunos casos, puede ser ventajoso expresar las variantes de *bg/3* que carezcan de propiedades o actividades del polinucleótido *bg/3* natural o del polipéptido BGL3. En dichos casos, se pueden generar formas mutantes o modificadas de la secuencia natural de ácido nucleico que codifica BGL3, usando técnicas usadas rutinariamente por los expertos en la materia.

#### B. Polipéptidos BGL3

20 [0094] En el presente documento se describe un polipéptido BGL3, que tiene una secuencia del polipéptido BGL3 de longitud completa o madura natural, que comprende la secuencia presentada como SEQ ID NO: 2. Un polipéptido BGL3 puede ser el polipéptido BGL3 maduro, parte de una proteína de fusión o un fragmento o variante de la secuencia del polipéptido BGL3 presentada como SEQ ID NO: 2.

25 [0095] Normalmente, un polipéptido BGL3 comprende una región que tiene una identidad de secuencia de al menos 98 % o más con la secuencia del polipéptido BGL3 presentada como SEQ ID NO: 2, usando un programa de alineamiento de secuencias, como se detalla en el presente documento.

30 [0096] Típicamente, una “forma modificada de” una proteína BGL3 natural o una proteína BGL3 “variante” tiene una secuencia derivada que contiene al menos una sustitución, adición, eliminación o inserción de aminoácido, respectivamente.

35 [0097] Es bien conocido en la materia que se pueden hacer algunas sustituciones de aminoácidos en secuencias de proteínas sin afectar a la función de la proteína. En general, las sustituciones de aminoácidos conservativas o sustituciones de aminoácidos similares son toleradas sin afectar a la función de la proteína. Los aminoácidos similares pueden ser los que son similares en tamaños y/o propiedades de carga, por ejemplo, el aspartato y glutamato, y la isoleucina y valina, son los dos pares de aminoácidos similares. La similitud entre pares de aminoácidos se ha evaluado en la materia en una serie de formas. Por ejemplo, Dayhoff y col. (1978), proporcionan tablas de frecuencias de sustituciones de aminoácidos que se pueden usar como una medida de la similitud de aminoácidos. Las tablas de frecuencia de Dayhoff y col. se basan en comparaciones de secuencias de aminoácidos para proteínas que tienen la misma función a partir de una variedad de fuentes evolutivamente diferentes.

40 [0098] En el presente documento se describen fragmentos y variantes de la secuencia del polipéptido BGL3 presentada como SEQ ID NO: 2. Un fragmento es un polipéptido variante que tiene una secuencia de aminoácidos que como parte es totalmente igual, pero no es toda la secuencia de aminoácidos, de los polipéptidos descritos previamente. Los fragmentos pueden ser “independientes” o estar comprendidos en un polipéptido más largo del cual el fragmento forma una parte o una región, lo más preferiblemente como una sola región continua. Los fragmentos preferidos son fragmentos biológicamente activos que son los fragmentos que median actividades de los polipéptidos de la invención, incluyendo aquellos con actividad similar o actividad mejorada o con una menor actividad. También se describen los fragmentos que son antigénicos o inmunogénicos en un animal, en particular en un ser humano. En este aspecto, la invención incluye (i) fragmentos de BGL3, preferiblemente de al menos aproximadamente 20 – 100 aminoácidos de longitud, más preferiblemente aproximadamente 100 – 200 aminoácidos de longitud, y (ii) una composición farmacéutica que comprende BGL3. En diferentes realizaciones, el fragmento corresponde al dominio N-terminal de BGL3 o al dominio C-terminal de BGL-3.

55 [0099] Los polipéptidos BGL3 también incluyen polipéptidos que pueden variar de la secuencia del polipéptido BGL3 presentada como SEQ ID NO: 2. Estas variantes pueden ser variantes de sustitución, inserción o eliminación. Las variantes típicamente presentan la misma actividad biológica cualitativa que el análogo natural, aunque también se pueden seleccionar variantes que tienen características modificadas como se describe con más detalle a continuación.

60

**[0100]** Una “sustitución” resulta del reemplazo de uno o más nucleótidos o aminoácidos por diferentes nucleótidos o aminoácidos, respectivamente.

5 **[0101]** Una “inserción” o “adición” es el cambio en una secuencia de nucleótidos o aminoácidos que ha dado como resultado la adición de uno o más nucleótidos o restos de aminoácidos, respectivamente, comparado con la secuencia natural.

10 **[0102]** Una “eliminación” se define como un cambio en la secuencia de nucleótidos o aminoácidos en el que uno o más nucleótidos o restos de aminoácidos, respectivamente, están ausentes.

15 **[0103]** Las sustituciones de aminoácidos normalmente son de restos individuales; las inserciones normalmente serán del orden de aproximadamente 1 a 20 aminoácidos, aunque se pueden tolerar inserciones considerablemente mayores. Las eliminaciones están en el intervalo de aproximadamente 1 a aproximadamente 20 restos, aunque en algunos casos las eliminaciones pueden ser mucho mayores.

**[0104]** Las sustituciones, eliminaciones, inserciones o cualquier combinación de las mismas, se pueden usar para obtener un derivado final. En general estos cambios se hacen en unos pocos aminoácidos para minimizar la alteración de la molécula. Sin embargo, en algunas circunstancias se pueden tolerar cambios mayores.

20 **[0105]** Las sustituciones de aminoácidos pueden ser el resultado de sustituir un aminoácido por otro aminoácido que tiene propiedades estructurales y/o químicas similares, tales como la sustitución de una isoleucina por una valina, es decir, sustituciones de aminoácidos conservativas. Las inserciones o eliminaciones pueden estar opcionalmente en el intervalo de 1 a 5 aminoácidos.

25 **[0106]** Las sustituciones en general se hacen de acuerdo con “sustituciones conservativas” conocidas. Una “sustitución conservativa” se refiere a la sustitución de un aminoácido de una clase por un aminoácido de la misma clase, donde una clase se define por las propiedades fisicoquímicas comunes de la cadena lateral y frecuencias de sustitución altas en proteínas homólogas encontradas en la naturaleza (determinado, por ejemplo, por una matriz de intercambio de frecuencia de Dayhoff estándar o matriz BLOSUM). (Véase en general, Doolittle, R.F., 1986).

30 **[0107]** Una “sustitución no conservativa” se refiere a la sustitución de un aminoácido de una clase por un aminoácido de otra clase.

35 **[0108]** Las variantes del polipéptido BGL3 típicamente presentan la misma actividad biológica cualitativa que el análogo natural, aunque las variantes también se seleccionan para modificar las características del polipéptido BGL3, según sea necesario. Por ejemplo, los sitios de glicosilación, y más en particular se pueden alterar o eliminar uno o más sitios de glicosilación unidos a O o unidos a N. Los expertos en la materia apreciarán que los cambios de aminoácidos pueden alterar los procesos postraduccionales del polipéptido BGL3, tal como cambiar el número o posición de los sitios de glicosilación o alterar las características de anclaje a la membrana o las características de secreción, u otras características de localización celular.

40 **[0109]** También está incluido en la definición de los polipéptidos BGL3 otros polipéptidos BGL3 relacionados. Por lo tanto, se pueden usar secuencias de sondas o cebadores de la reacción en cadena de la polimerasa (PCR) degeneradas para encontrar otros polipéptidos relacionados. Se pueden diseñar secuencias de sondas o cebadores para: toda o parte de la secuencia del polipéptido BGL3, o secuencias fuera de la región codificante. Como se sabe en general en la materia, los cebadores de la PCR preferidos tienen de aproximadamente 15 a aproximadamente 35 nucleótidos de longitud, siendo preferidos de aproximadamente 20 a aproximadamente 30, y pueden contener inosina, según sea necesario. Las condiciones para la reacción PCR son conocidas en general en la materia.

45 **[0110]** Las modificaciones covalentes de los polipéptidos BGL3 también están incluidas dentro del alcance de la invención como se define en las reivindicaciones. Por ejemplo, la invención proporciona polipéptidos BGL3 que son una proteína madura y pueden comprender aminoácidos amino o carboxilo terminales adicionales, o aminoácidos dentro del polipéptido maduro (por ejemplo, cuando la forma madura de la proteína tiene más de una cadena polipeptídica). Dichas secuencias pueden, por ejemplo, tener una función en el procesamiento de la proteína desde un precursor a una forma madura, permitir el transporte de proteína, acortar o prolongar la semivida de la proteína, o facilitar la manipulación de la proteína en ensayos o en la producción.

50 **[0111]** También se contemplan las modificaciones dirigidas a alterar un sitio activo, alterar el pH óptimo, la temperatura óptima y/o la afinidad por el sustrato de la enzima BGL3.

60 **[0112]** La figura 2 muestra la secuencia de aminoácidos predicha (SEQ ID NO: 2) de un ejemplo de polipéptido BGL3 basado en la secuencia de nucleótidos proporcionada en la figura 1. El peso molecular predicho del

polipéptido BGL3 codificado es 93,3 kDa. Un péptido señal predicho de 16 aminoácidos precede al extremo amino maduro de BGL3 como se proporciona en la figura (Nielsen, H., Engelbrecht, J., Brunak, S., von Heijne, G., *Protein Engineering*, 10: 1 – 6, 1997).

5 **[0113]** Se llevó a cabo una búsqueda con BLASTP Basic (<http://www.ncbi.nlm.nih.gov/BLAST>) de la base de datos de secuencias de proteínas no redundantes, el 17 de agosto de 2001, con la secuencia de aminoácidos de BGL3 que indicaba una identidad de secuencia de 56 % con el número de acceso en GenBank AAA86880 (precursor del antígeno H de *Ajellomyces capsulatus*) y una identidad de secuencia de 54 % con el número de acceso en GenBank P48825 (precursor de beta-glucosidasa 1 de *Aspergillus aculeatus*) y el número de acceso en GenBank AAF21242 (precursor de beta-glucosidasa de *Coccidioides immitis*). Las diez secuencias que tenían la mayor identidad pero menor de 56 % de identidad con BGL3 se anotaron todas como beta-glucosidasas. Estas similitudes de secuencia indican que BGL3 es un miembro de la familia de glicosil hidrolasa 3 (Henrissat, B. y Bairoch, A. (1993) *Biochem. J.* 293: 781 – 788).

#### 15 C. Anticuerpos dirigidos contra BGL3.

**[0114]** En el presente documento se describen anticuerpos dirigidos contra BGL3. Los anticuerpos pueden ser anticuerpos policlonales, monoclonales, humanizados, biespecíficos o heteroconjugados.

20 **[0115]** Los procedimientos de preparación de anticuerpos policlonales son conocidos para el experto en la materia. El agente inmunizante puede ser un polipéptido BGL3 o una proteína de fusión del mismo. Puede ser útil conjugar el antígeno con una proteína que se sabe que es inmunógena en el mamífero que se va a inmunizar. El protocolo de inmunización lo puede determinar el experto en la materia basándose en protocolos estándar o la experimentación rutinaria.

25 **[0116]** Alternativamente, los anticuerpos dirigidos contra BGL3 pueden ser anticuerpos monoclonales. Los anticuerpos monoclonales se pueden producir en células inmunizadas en un animal o usando procedimientos de ADN recombinante (Véase, p. ej., Kohler y col., 1975; patente de EE.UU. n ° 4.816.567).

30 **[0117]** Un anticuerpo dirigido contra BGL3 puede comprender además un anticuerpo humanizado o humano. La expresión “anticuerpo humanizado” se refiere a formas humanizadas de anticuerpos no humanos (p. ej., murino) que son anticuerpos quiméricos, cadenas de inmunoglobulina o fragmentos de las mismas (tales como Fv, Fab, Fab', F (ab')<sub>2</sub> u otras secuencias parciales de anticuerpos de unión al antígeno) que contienen alguna parte de la secuencia derivada de anticuerpo no humano. Los procedimientos para humanizar anticuerpos no humanos son bien conocidos en la materia, como se detalla más en Jones y col., 1986; Riechmann y col., 1988; y Verhoeyen y col., 1988. Los procedimientos para producir anticuerpos humanos también son conocidos en la materia. Véase, p. ej., Jakobovits, A. y col., 1995 y Jakobovits, A., 1995.

#### 40 **VI. Expresión de BGL3 recombinante**

**[0118]** Los procedimientos descritos en el presente documento se basan en el uso de células para expresar BGL3, sin que la expresión BGL3 requiere un procedimiento particular.

45 **[0119]** En el presente documento se describen células huésped que se han transducido, transformado o transfectado con un vector de expresión que comprende una secuencia de ácido nucleico que codifica BGL3. Las condiciones de cultivo, tales como la temperatura, pH y similares, son las usadas previamente para la célula huésped original antes de la transducción, transformación o transfección, y serán evidentes para el experto en la materia.

50 **[0120]** En un procedimiento, una célula de hongo filamentoso o célula de levadura se transfecta con un vector de expresión que tiene un promotor o fragmento de promotor biológicamente activo o uno o más (p. ej., una serie) de potenciadores que funcionan en la línea celular huésped, operativamente unidos a un segmento de ADN que codifica BGL3, de modo que BGL3 se expresa en la línea celular.

#### 55 **A. Construcciones de ácidos nucleicos / vectores de expresión**

**[0121]** Se pueden incorporar fragmentos de polinucleótidos naturales o sintéticos que codifican BGL3 (“secuencias de ácido nucleico que codifican BGL3”) en construcciones o vectores de ácido nucleico heterólogos, capaces de introducción y replicación en una célula de hongo filamentoso o célula de levadura. Los vectores y procedimientos descritos en el presente documento son adecuados para usar en células huésped para la expresión de BGL3. Se puede usar cualquier vector siempre que se pueda replicar y sea viable en las células en las que se introduce. Los expertos en la materia conocen un gran número de dichos vectores y promotores adecuados, y están disponibles en

el comercio. También se describe la clonación y vectores de expresión en Sambrook y col., 1989, Ausubel FM y col., 1989, y Strathern y col., 1981. Se describen vectores de expresión adecuados para hongos en van den Hondel, C.A.M.J.J. y col. (1991) en: Bennett, J.W. y Lasure, L.L. (eds.) *More Gene Manipulations in Fungi*. Academic Press, pág. 396 – 428. La secuencia de ADN adecuada se puede insertar en un plásmido o vector (denominados de forma colectiva en el presente documento "vectores") por una variedad de procedimientos. En general, la secuencia de ADN se inserta en un sitio o sitios de endonucleasas de restricción adecuados, por procedimientos estándar. Dichos procedimientos y procedimientos de subclonación relacionados se considera que están dentro del alcance del conocimiento de los expertos en la materia.

**[0122]** Los hongos filamentosos recombinantes que comprenden la secuencia que codifica BGL3 se pueden producir introduciendo una construcción de ácido nucleico heteróloga que comprende la secuencia que codifica BGL3 en las células de una cepa seleccionada de los hongos filamentosos.

**[0123]** Una vez que se obtiene la forma deseada de una secuencia de ácido nucleico de *bg/3*, homólogo, variante o fragmento de la misma, se puede modificar en una variedad de formas. Cuando la secuencia implica regiones flanqueadoras no codificantes, las regiones flanqueadoras se pueden someter a resección, mutagénesis, etc. Por lo tanto, se pueden realizar transiciones, transversiones, eliminaciones e inserciones en la secuencia natural.

**[0124]** Una secuencia que codifica *bg/3* seleccionada se puede insertar en un vector adecuado de acuerdo con técnicas recombinantes bien conocidas y usar para transformar hongos filamentosos capaces de expresar BGL3. Debido a la degeneración inherente del código genético, se pueden usar otras secuencias de ácido nucleico que codifican sustancialmente la misma secuencia de aminoácidos o una funcionalmente equivalente, para clonar y expresar BGL3. Por lo tanto, se observará que dichas sustituciones en la región codificante están dentro de las variantes de secuencias descritas en el presente documento.

**[0125]** Todas y cada una de las variantes de secuencia se pueden usar de la misma forma descrita en el presente documento para una secuencia original de ácido nucleico que codifica BGL3. También se describen en el presente documento construcciones de ácido nucleico que comprenden una o más secuencias de ácido nucleico que codifican BGL3 como se ha descrito antes. Las construcciones comprenden un vector, tal como un plásmido o vector vírico, en el que se ha insertado una secuencia de la invención, en una orientación directa o inversa.

**[0126]** Las construcciones de ácido nucleico heterólogas pueden incluir la secuencia que codifica *bg/3*, o una variantes, fragmento o variante de empalme de la misma: (i) aisladas; (ii) en combinación con secuencias codificantes adicionales; tales como secuencias de proteína de fusión o codificantes de péptido señal, en las que la secuencia que codifica *bg/3* es la secuencia codificante dominante; (iii) en combinación con secuencias no codificantes, tales como intrones y elementos de control, tales como elementos promotores y terminadores o regiones 5' y/o 3' no traducidas, eficaces para la expresión de la secuencia codificante en un huésped adecuado; y/o (iv) en un vector o entorno de huésped en el que la secuencia que codifica *bg/3* es un gen heterólogo.

**[0127]** Se puede usar una construcción de ácido nucleico heteróloga para transferir una secuencia de ácido nucleico que codifica BGL3 a una célula in vitro, prefiriéndose las líneas de levaduras y hongos filamentosos establecidas. Para la producción de alto rendimiento, a largo plazo de BGL3, se prefiere la expresión estable. Se deduce que se puede usar cualquier procedimiento eficaz para generar transformantes estables.

**[0128]** Los vectores adecuados típicamente están equipados con una secuencia de ácido nucleico que codifica un marcador seleccionable, sitios de inserción y elementos de control adecuados, tales como secuencias de promotores y de terminación. El vector puede comprender secuencias reguladoras, incluyendo, por ejemplo, secuencias no codificantes, tales como intrones y elementos de control, es decir, elementos promotores y terminadores o regiones 5' y/o 3' no traducidas, eficaces para la expresión de la secuencia codificante en células huésped (y/o en un vector o entorno de célula huésped en el que una secuencia codificante de antígeno de proteína soluble modificada normalmente no se expresa) operativamente unida a la secuencia codificante. Los expertos en la materia conocen un gran número de vectores y promotores adecuados, muchos de los cuales están disponibles en el comercio y/o se describen en Sambrook y col. (véase antes).

**[0129]** Los promotores de ejemplo incluyen tanto promotores constitutivos como promotores inducibles, cuyos ejemplos incluyen un promotor de CMV, un promotor temprano de SV40, un promotor de RSV, un promotor de EF-1 $\alpha$ , un promotor que contiene el elemento de respuesta a tet (TRE) en el sistema tet-on o tet-off como se ha descrito (ClonTech y BASF), el promotor de beta-actina y el promotor de metalotionina que puede favorecer la expresión por adición de determinadas sales de metales. Una secuencia promotora es una secuencia de ADN que es reconocida por el hongo filamentosos particular para los propósitos de expresión. Está operativamente unido a la secuencia de ADN que codifica un polipéptido BGL3. Dicha unión comprende la colocación del promotor con respecto al codón de inicio de la secuencia de ADN que codifica el polipéptido BGL3 en los vectores de expresión descritos. La secuencia

promotora contiene secuencias de control de la transcripción y traducción que median la expresión del polipéptido BGL3. Los ejemplos incluyen los promotores de los genes que codifican glucoamilasa, alfa-amilasa, o alfa-glucosidasa de *Aspergillus niger*, *A. awamori* o *A. oryzae*; los genes *gpdA* o *trpC* de *A. nidulans*; los genes *cbh1* o *trp1* de *Neurospora crassa*; los genes que codifican la aspártico proteasa de *A. niger* o *Rhizomucor miehei*; los genes *cbh1*, *cbh2*, *egl1*, *egl2* de *T. reesei* u otros genes que codifican celulasas.

**[0130]** La elección del marcador seleccionable adecuado dependerá de la célula huésped, y los marcadores adecuados para los diferentes huéspedes son bien conocidos en la materia. Los genes de marcadores seleccionables típicos incluyen *argB* de *A. nidulans* o *T. reesei*, *amdS* de *A. nidulans*, *pyr4* de *Neurospora crassa* o *T. reesei*, *pyrG* de *Aspergillus niger* o *A. nidulans*. Los ejemplos adicionales de marcadores seleccionables incluyen, pero sin limitar *trpc*, *trp1*, *oliC31*, *niaD* o *leu2*, que se incluyen en construcciones de ácido nucleico heterólogas usadas para transformar una cepa mutante tal como *trp-*, *pyr-*, *leu-* y similares.

**[0131]** Dichos marcadores seleccionables confieren a los transformantes la capacidad de usar un metabolito que normalmente no es metabolizado por los hongos filamentosos. Por ejemplo, el gen *amdS* de *T. reesei* que codifica la enzima acetamidasa que permite que las células transformantes se desarrollen en acetamida como fuente de nitrógeno. El marcador seleccionable (p. ej., *pyrG*) puede restablecer la capacidad de una cepa mutante auxótrofa para crecer en un medio mínimo selectivo, o el marcador seleccionable (p. ej., *oliC31*) puede conferir a los transformantes la capacidad de crecer en presencia de un fármaco o antibiótico inhibidor.

**[0132]** La secuencia que codifica el marcador seleccionable se clona en cualquier plásmido adecuado usando procedimientos usados en general en la materia. Los plásmidos de ejemplo incluyen pUC18, pBR322, y pUC100.

**[0133]** La práctica de la presente invención usará, salvo que se indique lo contrario, técnicas convencionales de biología molecular, microbiología, ADN recombinante e inmunología, que se basa en la experiencia en la materia. Dichas técnicas se explican íntegramente en la bibliografía. Por ejemplo, véase, Sambrook y col., 1989; Freshney, 1987; Ausubel, y col., 1993; y Coligan y col., 1991.

## **B. Células huésped y condiciones de cultivo para la producción potenciada de BGL3**

### **(i) Hongos filamentosos**

**[0134]** Por lo tanto, la presente invención proporciona hongos filamentosos que comprenden células que se han modificado, seleccionado y cultivado de una forma eficaz para que den como resultado la producción o expresión potenciada de BGL3 con respecto a los correspondientes hongos originales no transformados.

**[0135]** Los ejemplos de especies de hongos filamentosos originales que se pueden tratar y/o modificar para potenciar la expresión de BGL3 incluyen, pero sin limitar, *Trichoderma*, p. ej., *Trichoderma reesei*, *Trichoderma longibrachiatum*, *Trichoderma viride*, *Trichoderma koningii*; *Penicillium sp.*, *Humicola sp.*, incluyendo *Humicola insolens*; *Aspergillus sp.*, *Chrysosporium sp.*, *Fusarium sp.*, *Hypocrea sp.*, y *Emericella sp.*

**[0136]** Las células que expresan BGL3 se cultivan en condiciones que se usan típicamente para cultivar la línea fúngica original. En general, las células se cultivan en un medio estándar que contiene sales fisiológicas y nutrientes, tal como se describe en Pourquie, J. y col., *Biochemistry and Genetics of Cellulose Degradation*, eds. Aubert, J. P. y col., Academic Press, pág. 71 – 86, 1988 y Ilmen, M. y col., *Appl. Environ. Microbiol.* 63: 1298 – 1306, 1997. Las condiciones de cultivo también son estándar, p. ej., los cultivos se incuban a 28 °C en cultivos con agitación o fermentadores hasta lograr los niveles deseados de expresión de BGL3.

**[0137]** Las condiciones de cultivo preferidas para un hongo filamentoso dado se pueden encontrar en la bibliografía científica y/o en la fuente de los hongos tal como la American Type Culture Collection (ATCC; "http://www.atcc.org/"). Después de establecerse el crecimiento fúngico, las células se exponen a condiciones eficaces para producir o permitir el exceso de expresión de BGL3.

**[0138]** En los casos en los que una secuencia que codifica BGL3 esté bajo el control de un promotor inducible, el agente inductor, p. ej., un azúcar, sal de metal o antibióticos, se añade al medio en una concentración eficaz para inducir un nivel alto de expresión de BGL3.

### **(ii) Levaduras**

**[0139]** La presente invención también contempla el uso de levaduras como una célula huésped para la producción de BGL3. Se han expresado algunos otros genes que codifican enzimas hidrolíticas en diferentes cepas de la levadura *S. cerevisiae*. Estas incluyen secuencias que codifican dos endoglucanasas (Penttila y col., 1987), dos

celobiohidrolasas (Penttila y col., 1988) y una beta-glucosidasa de *Trichoderma reesei* (Cummings y Fowler, 1996), una xilanasa de *Aureobasidium pullulans* (Li y Ljungdahl, 1996), una alfa-amilasa del trigo (Rothstein y col., 1987), etc. Además se expresó con éxito un casete de gen de celulasa que codificaba la endo-[beta]-1,4-glucanasa (END1) de *Butyrivibrio fibrisolvens*, celobiohidrolasa (CBH1) de *Phanerochaete chrysosporium*, la celodextrinasa (CEL1) de *Ruminococcus flavefaciens* y la celobiasa (Bgl1) de *Endomyces fibrilizer*, en una cepa de laboratorio de *S. cerevisiae* (Van Rensburg y col., 1998).

#### Introducción de una secuencia de ácido nucleico que codifica BGL3 en células huésped

10 **[0140]** En el presente documento se describen células y composiciones celulares que se han modificado genéticamente para que comprendan una secuencia de ácido nucleico que codifica BGL3 proporcionada de forma exógena. Se puede modificar una célula o línea celular original (es decir, transducir, transformar o transfectar) con un vector de clonación o un vector de expresión. El vector puede ser, por ejemplo, en forma de un plásmido, una partícula vírica, un fago, etc., como se ha descrito con más detalle antes.

15 **[0141]** Se pueden usar diferentes procedimientos para suministrar un vector de expresión en células in vitro. Después de construir un vector adecuado, se usa para transformar cepas de hongos o levaduras. Los procedimientos generales para introducir ácidos nucleicos en células para expresar secuencias de ácido nucleico heterólogas son conocidas por el experto en la materia. Dichos procedimientos incluyen, pero sin limitar, electroporación; microinyección nuclear o microinyección directa en células individuales; fusión de protoplastos bacterianos con células intactas; uso de policonaciones, p. ej., Polybrene o poliornitina; fusión de membrana con liposomas, transfección mediada por lipofectamina o lipofección; bombardeo a alta velocidad con microproyectiles recubiertos de ADN; incubación con precipitado de ADN-fosfato de calcio; transfección mediada por DEAE-dextrano; infección con ácidos nucleicos víricos modificados; y similares.

20 **[0142]** Los procedimientos preferidos para introducir una construcción de ácido nucleico heteróloga (vector de expresión) en hongos filamentosos (p. ej., *T. reesei*) incluyen, pero sin limitar, el uso de una pistola de partículas o genes, permeabilización de paredes celulares de hongos filamentosos antes del procedimiento de transformación (p. ej., usando concentraciones altas de álcali, p. ej., CaCl<sub>2</sub> o acetato de litio de 0,05 M a 0,4 M), fusión de protoplastos o transformación mediada por agrobacterias. Un procedimiento de ejemplo para transformar hongos filamentosos por tratamiento de protoplastos o esferoplastos con polietilenglicol y CaCl<sub>2</sub> se describe en Campbell, E.I. y col., *Curr. Genet.* 16: 53 – 56, 1989 y Penttila, M. y col., *Gene*, 63: 11 – 22, 1988.

25 **[0143]** Además, las construcciones de ácido nucleico heterólogas que comprenden una secuencia de ácido nucleico que codifica BGL3 se pueden transcribir in vitro, e introducir el ARN resultante en la célula huésped por procedimientos bien conocidos, p. ej., por inyección.

30 **[0144]** Después de introducir una construcción de ácido nucleico heteróloga que comprende la secuencia que codifica *bg13*, las células genéticamente modificadas se pueden cultivar en un medio nutriente convencional modificado según sea adecuado para activar los promotores, seleccionar transformantes o amplificar la expresión de una secuencia de ácido nucleico que codifica BGL3. Las condiciones de cultivo, tales como temperatura, pH y similares, son las usadas previamente para la célula huésped seleccionada para la expresión, y serán evidentes para el experto en la materia.

35 **[0145]** La progenie de las células en las que se han introducido construcciones de ácido nucleico heteróloga, en general se considera que comprende la secuencia de ácido nucleico que codifica BGL3 encontrada en la construcción de ácido nucleico heteróloga.

40 **[0146]** También se describen transformantes nuevos y útiles de hongos filamentosos tales como *Trichoderma reesei* para usar en la producción de composiciones de celulasa fúngicas. Por ejemplo, transformantes de hongos filamentosos, en especial hongos que comprenden la secuencia que codifica *bg13*, que comprenden una forma modificada de la secuencia que codifica *bg13* o eliminación de la secuencia que codifica *bg13*.

45 **[0147]** Los transformantes estables de los hongos filamentosos en general se pueden distinguir de los transformantes inestables por su tasa de crecimiento más rápida y la formación de colonias circulares con un contorno liso más que irregular en medio de cultivo sólido. Además, en algunos casos, se puede hacer un ensayo adicional de estabilidad mediante el cultivo de los transformantes en medio sólido no selectivo, recolección de las esporas de este medio de cultivo y determinación del porcentaje de estas esporas que posteriormente germinan y crecen en un medio selectivo.

60

VII. Análisis de las secuencias que codifican el ácido nucleico de BGL3 y/o la expresión de la proteína

**[0148]** Con el fin de evaluar la expresión de BGL3 por una línea celular que se ha transformado con una construcción de ácido nucleico que codifica BGL3, se pueden llevar a cabo ensayos a nivel de proteínas, a nivel del ARN o usando bioensayos funcionales en particular de la actividad y/o producción de glucosidasa.

**[0149]** En una aplicación de ejemplo, el ácido nucleico de *bgl3* y las secuencias de proteínas descritas en el presente documento, una cepa genéticamente modificada de hongos filamentosos, p. ej., *Trichoderma reesei*, se modifican para producir una mayor cantidad de BGL3. Dichos hongos filamentosos genéticamente modificados serían útiles para producir un producto de celulasa con un mayor aumento de la capacidad celulolítica. En un procedimiento, esto se logra introduciendo la secuencia codificante para *bgl3* en un huésped adecuado, p. ej., un hongo filamentoso tal como *Trichoderma reesei*.

**[0150]** Por consiguiente, en el presente documento se describen procedimientos para expresar BGL3 en un hongo filamentoso u otro huésped adecuado, introduciendo un vector de expresión que contiene la secuencia de ADN que codifica BGL3 en células del hongo filamentoso u otro huésped adecuado.

**[0151]** En otro aspecto, se describen procedimientos para modificar la expresión de BGL3 en un hongo filamentoso u otro huésped adecuado. Dicha modificación incluye la eliminación de la expresión, o la expresión de una forma alterada de BGL3. Una forma alterada de BGL3 puede tener una secuencia de aminoácidos alterada o una secuencia de ácido nucleico alterada.

**[0152]** En general, los ensayos usados para analizar la expresión de BGL3 incluyen, transferencia Northern, inmunotransferencia (análisis de ADN o ARN), RT-PCR (reacción en cadena de la polimerasa con transcriptasa inversa), o hibridación in situ, usando una sonda marcada de forma adecuada (basada en la secuencia que codifica el ácido nucleico) y transferencia Southern convencional y autorradiografía.

**[0153]** Además, la producción y/o expresión de BGL3 se puede medir en una muestra directamente, por ejemplo, mediante ensayos de la actividad, expresión y/o producción de la glucosidasa. Dichos ensayos se describen, por ejemplo, en Chen y col. (1992), Herr y col. (1978), y patente de EE.UU. n° 6.184.018 (Li y col.; 2001), cada uno de los cuales se incorpora expresamente por referencia en el presente documento. La capacidad de BGL3 para hidrolizar sustratos aislados solubles e insolubles, se puede medir usando ensayos descritos en Suurnakki y col. (2000) y Ortega y col. (2001). Los sustratos útiles para ensayar las actividades de celobiohidrolasa, endoglucanasa o  $\beta$ -glucosidasa, incluyen celulosa cristalina, papel de filtro, celulosa hinchada con ácido fosfórico, hidroxietilcelulosa, carboximetilcelulosa, celooligosacáridos, metilumbeliferil-lactósido, metilumbeliferil-celobiósido, ortonitrofenil-lactósido, paranitrofenil-lactósido, ortonitrofenil-celobiósido, paranitrofenil-celobiósido, ortonitrofenil-glucósido, paranitrofenil-glucósido, metilumbeliferil-glucósido. Los tres últimos son particularmente útiles para ensayar  $\beta$ -glucosidasas. Los ensayos de  $\beta$ -glucosidasas son bien conocidos en la materia. Véase Cummings y Fowler (1996).

**[0154]** Además, la expresión de proteínas se puede evaluar por procedimientos inmunológicos, tales como tinción inmunohistoquímica de células, secciones de tejidos o inmunoensayo de medio de cultivo tisular, p. ej., por transferencia Western o ELISA. Dichos inmunoensayos se pueden usar para evaluar cualitativa y cuantitativamente la expresión de BGL3. Los detalles de dichos procedimientos son conocidos para el experto en la materia y hay muchos reactivos disponibles en el comercio para poner en práctica dichos procedimientos.

**[0155]** Se puede usar una forma purificada de BGL3 para producir anticuerpos monoclonales o policlonales específicos para la proteína expresada para usar en diferentes inmunoensayos (Véase, por ejemplo, Hu y col., 1991). Los ensayos de ejemplo incluyen ELISA, inmunoensayos competitivos, radioinmunoensayos, transferencia Western, ensayos de inmunofluorescencia indirectos y similares. En general, se pueden usar anticuerpos y/o kits disponibles en el comercio para el inmunoensayo cuantitativo del nivel de expresión de proteínas glucosidasa.

VIII. Aislamiento y purificación de la proteína BGL3 recombinante

**[0156]** En general, una proteína BGL3 producida en cultivo celular es secretada al medio y se puede purificar o aislar, p. ej., eliminando los componentes no deseados del medio de cultivo celular. Sin embargo, en algunos casos, se puede producir una proteína BGL3 en una forma celular que necesita la recuperación de un lisato celular. En dichos casos, la proteína BGL3 se purifica a partir de células en las que se produjo usando técnicas usadas rutinariamente por los expertos en la materia. Los ejemplos incluyen, pero sin limitar, cromatografía de afinidad (Tilbeurgh y col., 1984), procedimientos cromatográficos de intercambio iónico (Goyal y col., 1991; Fliess y col., 1983; Bhikhabhai y col., 1984; Ellouz y col., 1987), incluyendo materiales que usan intercambio iónico con alto poder de resolución (Medve y col., 1998), cromatografía de interacción hidrófoba (Tomaz y Queiroz, 1999), y reparto en dos fases (Brumbauer, y col., 1999).

**[0157]** Típicamente, la proteína BGL3 se fracciona para segregar proteínas que tienen las propiedades seleccionadas, tales como la afinidad de unión a agentes de unión particulares, p. ej., anticuerpos o receptores; o que tienen un intervalo de peso molecular seleccionado, o un intervalo de puntos isoelectrónicos.

**[0158]** Una vez lograda la expresión de una proteína BGL3 dada, la proteína BGL3 así producida se purifica de las células o el cultivo celular. Los procedimientos de ejemplo adecuados para dicha purificación incluyen los siguientes: cromatografía en columna de afinidad con anticuerpos, cromatografía de intercambio iónico; precipitación con etanol; HPLC de fase inversa; cromatografía en sílice o en una resina de intercambio catiónico tal como DEAE; cromatoenfoco; SDS-PAGE; precipitación con sulfato amónico; y filtración en gel usando, p. ej., Sephadex G-75. Se pueden usar diferentes procedimientos de purificación de proteínas y dichos procedimientos son conocidos en la materia y se describen, p. ej. en Deutscher, 1990; Scopes, 1982. La etapa o etapas de purificación seleccionadas dependerán, p. ej., de la naturaleza del procedimiento de producción usado y de la proteína particular producida.

#### IX. Utilidad de *bgl3* y BGL3

**[0159]** Se puede observar que el nucleótido *bgl3*, la proteína BGL3 y las composiciones que comprenden actividad de la proteína BGL3 tienen utilidad en una amplia variedad de aplicaciones, algunas de las cuales se describen a continuación.

**[0160]** Las composiciones nuevas y mejoradas de celulasas que comprenden diferentes cantidades de celulasas de tipo CBH, tipo EG y tipo BG tienen utilidad en composiciones de detergentes que presentan una capacidad de limpieza potenciada, funcionan como agente suavizante y/o mejoran el tacto de las telas de algodón (p. ej., "lavado a la piedra" o "biopolido"), en composiciones para degradar la pasta de madera en azúcares (p. ej., para la producción de bioetanol), y/o en composiciones de piensos. El aislamiento y caracterización de las celulasas de cada tipo proporciona la capacidad de controlar los aspectos de dichas composiciones.

**[0161]** En un procedimiento preferido, la celulasa de la invención es útil en composiciones de detergentes o en el tratamiento de telas para mejorar el tacto y el aspecto.

**[0162]** Las  $\beta$ -glucosidasas de la invención se pueden usar en una variedad de aplicaciones diferentes. Por ejemplo, la  $\beta$ -glucosidasa se puede añadir a las uvas durante la elaboración del vino para potenciar el aroma potencial del producto de vino acabado. Otra aplicación más puede ser usar la  $\beta$ -glucosidasa en la fruta para potenciar el aroma de la misma. Alternativamente, el producto de fermentación recombinante aislado que contiene la  $\beta$ -glucosidasa potenciada, se puede usar directamente en aditivos de alimentos o procesamiento de vino para potenciar el sabor o el aroma.

**[0163]** Puesto que la velocidad de hidrólisis de los productos celulósicos puede aumentar usando un transformante que tenga al menos una copia adicional del gen *bgl3* insertado en el genoma, los productos que contienen celulosa o heteroglucanos pueden ser degradados a una velocidad más rápida y en una mayor extensión. Los productos hechos de celulosa tales como el papel, algodón, pañales de celulosa y similares se pueden degradar más eficazmente en un vertedero. Por lo tanto, el producto de fermentación que se puede obtener de los transformantes o los transformantes solos se pueden usar en composiciones para ayudar a degradar por licuefacción una variedad de productos de celulosa que aumentan los vertederos repletos.

**[0164]** La sacarificación y fermentación simultáneas es un procedimiento por el que la celulosa presente en la biomasa, p. ej., residuos del maíz, se convierte en glucosa, y al mismo tiempo y en el mismo reactor, las cepas de levadura convierten la glucosa en etanol. Por lo tanto, en otro procedimiento preferido, la celulasa de tipo glucosidasa de la invención es útil en la degradación de la biomasa en etanol. La producción de etanol a partir de fuentes fácilmente disponibles de celulosa proporciona una fuente de combustible renovable, estable. Una composición de celulasa que contiene una mayor cantidad de  $\beta$ -glucosidasa es útil en la producción de etanol. El etanol de este proceso después se puede usar como un potenciador de octanos o directamente como un combustible en lugar de gasolina, lo cual es ventajoso porque el etanol como fuente de combustible perjudica menos el medioambiente que los productos derivados del petróleo. Se sabe que el uso de etanol mejora la calidad del aire y posiblemente reduce los niveles de ozono local y la contaminación. Además, el uso de etanol en lugar de gasolina puede tener una importancia estratégica para amortiguar el impacto de los cambios repentinos en los suministros de energía no renovable y productos petroquímicos.

**[0165]** El etanol se puede producir por procesos de sacarificación y fermentación a partir de biomasa celulósica tal como árboles, plantas herbáceas, residuos sólidos municipales y residuos agrícolas y forestales. Sin embargo, un problema principal que se encuentra en este procedimiento es la falta de  $\beta$ -glucosidasa en el sistema para convertir la celobiosa en glucosa. Se sabe que la celobiosa actúa como un inhibidor de las celobiohidrolasas y

endoglucanasas y por lo tanto reduce la velocidad de hidrólisis para el sistema entero de celulasas. Por lo tanto, el uso de una mayor actividad de  $\beta$ -glucosidasa para convertir rápidamente la celobiosa en glucosa potenciaría mucho la producción de etanol.

5 **[0166]** En un procedimiento alternativo, se prefiere una composición de celulasas que es deficiente en o que carece de  $\beta$ -glucosidasa. La eliminación del gen de la  $\beta$ -glucosidasa de esta invención sería particularmente útil para preparar composiciones de celulasas para usar en detergentes. Además, dichas composiciones son útiles para producir celobiosa y otros celooligosacáridos. La eliminación del gen *bgl3* de cepas de *T. reesei* sería particularmente útil para preparar composiciones de celulasas para usar en detergentes y para aislar celobiosa. Las enzimas celulasas se han usado en una variedad de composiciones de detergentes para limpiar la ropa de forma enzimática. Sin embargo, se sabe en la materia que el uso de enzimas celulasas puede producir degradación de las fibras de celulosa en la ropa. Una posibilidad para disminuir el efecto de degradación es producir un detergente que no contenga  $\beta$ -glucosidasa. Por lo tanto, la eliminación de esta proteína haría que el sistema de celulasas inhibiera los otros componentes por acumulación de celobiosa. Los microorganismos modificados de esta invención son particularmente adecuados para preparar dichas composiciones porque el gen *bgl3* se puede eliminar dejando el resto de los componentes CBH y EG que producen una mejor limpieza y beneficios suavizantes en la composición sin los efectos de degradación.

20 **[0167]** Las composiciones de detergentes de esta invención pueden usar además de la composición de celulasas (independientemente del contenido de  $\beta$ -glucosidasa, es decir, exenta de  $\beta$ -glucosidasa, sustancialmente exenta de  $\beta$ -glucosidasa, o potenciada en  $\beta$ -glucosidasa), un tensioactivo, incluyendo tensioactivos aniónicos, no iónicos y anfóteros, una hidrolasa, agentes mejoradores, agentes blanqueantes, agentes añiles y colorantes fluorescentes, inhibidores del apelmazamiento, solubilizantes, tensioactivos catiónicos y similares. Todos estos componentes son conocidos en el campo de los detergentes. La composición de celulasas descrita antes se puede añadir a la composición de detergente en un diluyente líquido, en gránulos, en emulsiones, en geles, en pastas y similares. Dichas formas son bien conocidas para el experto en la materia. Cuando se usa una composición de detergente sólida, la composición de celulasas preferiblemente se formula en forma de gránulos. Preferiblemente, los gránulos se pueden formular para que contengan un agente protector de celulasas. Para una discusión más completa, véase la patente de EE.UU. n° 6.162.782 titulada "Detergent compositions containing cellulase compositions deficient in CBH I type components" (Composiciones de detergentes que contienen composiciones de celulasas deficientes en componentes CBH de tipo I).

35 **[0168]** En otra realización más, las composiciones de detergentes también pueden contener niveles potenciados de beta-glucosidasa o beta-glucosidasa alterada. En relación con esto, depende realmente del tipo de producto que uno desea usar en las composiciones de detergentes para dar los efectos adecuados.

40 **[0169]** Preferiblemente, las composiciones de celulasas se usan de aproximadamente 0,00005 por ciento en peso a aproximadamente 5 por ciento en peso, respecto a la composición de detergente total. Más preferiblemente, las composiciones de celulasas se usan de aproximadamente 0,0002 por ciento en peso a aproximadamente 2 por ciento en peso con respecto a la composición de detergente total.

45 **[0170]** La eliminación del gen *bgl3* también proporcionaría acumulación de celobiosa en el sistema de celulasas, del cual después se puede purificar. En relación con esto, se describe la posibilidad de aislar celobiosa de microorganismos de una forma fácil y eficaz.

**[0171]** Las partes de la secuencia del ácido nucleico de *bgl3* que son capaces de unirse a celulosa, se pueden usar para generar proteínas de superficie quiméricas bacterianas, permitiendo la inmovilización de la célula entera sobre filtros de celulosa y otros soportes sólidos fibrosos como describen Lehtio y col., 2001.

50 **[0172]** Además, la secuencia del ácido nucleico de *bgl3* es útil en la identificación y caracterización de secuencias de ácido nucleico relacionadas. Una serie de técnicas útiles para determinar (predecir o confirmar) la función de genes relacionadas o productos génicos incluye, pero sin limitar, (A) análisis de ADN / ARN, tal como (1) exceso de expresión, expresión ectópica, y expresión en otras especies; (2) inactivación génica (genética inversa, inactivación dirigida, silenciamiento de genes inducido por virus (VIGS, véase Baulcombe, 1999); (3) análisis del estado de metilación del gen, en especial de las regiones reguladoras flanqueadoras; y (4) hibridación in situ; (B) análisis de productos génicos, tal como (1) expresión de proteína recombinante; (2) producción de antisuero, (3) inmunolocalización; (4) ensayos bioquímicos de la actividad catalítica u otras actividades; (5) estado de fosforilación; y (6) interacción con otras proteínas por análisis de dos híbridos de levaduras; (C) análisis de ruta, tal como poner un gen o un producto génico en una ruta bioquímica o de señalización particular, basado en el exceso de expresión del fenotipo o por homología de secuencia con genes relacionados; y (D) otros análisis que también se pueden realizar para determinar o confirmar la participación del gen aislado y su producto en una ruta metabólica o de señalización particular, y ayudar a determinar la función génica.

**[0173]** Las endoglucanasas y beta-glucosidasas pueden ser responsables de la producción de disacáridos, tales como soforosa, a partir de celooligosacáridos y glucosa por reacciones de transglicosilación. Se sabe que la soforosa es un inductor muy potente de la expresión del gen celulasa (Ilmen, M. y col., 1997, *Appl. Environ. Microbiol.* 63: 1298 – 1306 y referencias en el mismo). De esta forma, las EG y BGL tienen una función importante en el proceso de inducción de la expresión del gen de celulasa. El exceso de expresión de determinadas EG y BGL en una cepa fúngica puede conducir a una productividad general más alta de la cepa de celulasas.

#### **A. Homología con secuencias conocidas**

**[0174]** La función de una secuencia de ácido nucleico que codifica BGL3 relacionada se puede determinar por homología con genes conocidos que tienen una función particular. Por ejemplo, se usa una comparación de la secuencia codificante de una molécula de ácido nucleico identificada con una base de datos pública de secuencias de ácidos nucleicos, para confirmar la función por homología con genes conocidos o por extensión de la secuencia de ácido nucleico identificada.

**[0175]** La expresión “% de homología” en el presente documento se usa de forma intercambiable con la expresión “% de identidad” y se refiere al nivel de identidad de secuencia del ácido nucleico o de los aminoácidos entre la secuencia de ácido nucleico que codifica BGL3 o la secuencia de aminoácidos de BGL3, cuando se alinean usando un programa de alineamiento de secuencias.

**[0176]** Por ejemplo, como se usa en el presente documento, 80 % de homología significa lo mismo que una identidad de secuencia de 80 % determinada por un algoritmo definido, y por consiguiente un homólogo de una secuencia dada tiene una identidad de secuencia mayor de 80 % a lo largo de la longitud de la secuencia dada.

**[0177]** Los programas de ordenador de ejemplo que se pueden usar para determinar la identidad entre dos secuencias incluyen, pero sin limitar, el conjunto de programas BLAST, p. ej., BLASTN, BLASTX y TBLASTX, BLASTP y TBLASTN, disponibles al público en Internet en <http://www.ncbi.nlm.nih.gov/BLAST/>. Véase también, Altschul, y col., 1990 y Altschul, y col., 1997.

**[0178]** Las búsquedas de secuencias normalmente se llevan a cabo usando el programa BLASTN cuando se evalúa una secuencia de ácido nucleico dada con respecto a secuencias de ácidos nucleicos en GenBank DNA Sequences. Se prefiere el programa BLASTX para la búsqueda de secuencias de ácidos nucleicos que se han traducido en todos los marcos de lectura frente a secuencias de aminoácidos en GenBank Protein Sequences y otras bases de datos públicas. Tanto BLASTN como BLASTX trabajan usando parámetros por defecto de una penalización de apertura de hueco de 11,0, y una penalización de extensión de hueco de 1,0, y utilizan la matriz BLOSUM-62. (Véase, p. ej., Altschul, y col., 1997).

**[0179]** Se lleva a cabo un alineamiento preferido de secuencias seleccionadas con el fin de determinar el “% de identidad” de dos o más secuencias, usando, por ejemplo, el programa CLUSTAL-W en MacVector versión 6.5, trabajando con los parámetros por defecto, que incluyen una penalización de apertura de hueco de 10,0, una penalización de extensión de hueco de 0,1, y una matriz de similitud BLOSUM 30.

**[0180]** En un procedimiento de ejemplo, se puede llevar a cabo la extensión de secuencia de un ácido nucleico que codifica *bg/3* usando procedimientos de extensión de cebador convencionales como se describe en Sambrook y col., véase antes, para detectar precursores de *bg/3* e intermedios del procesamiento del ARNm que pueden no haber experimentado transcripción inversa en ADNc y/o para identificar ORF que codifican la proteína de longitud completa.

**[0181]** En otro aspecto más, se describe la secuencia de nucleótidos entera o parcial de la secuencia del ácido nucleico de *bg/3* para usar como sonda. Dicha sonda se puede usar para identificar y clonar secuencias de ácidos nucleicos homólogas de organismos relacionados.

**[0182]** El cribado de una biblioteca genómica o de ADNc con la sonda seleccionada se puede llevar a cabo usando procedimientos estándar, como se describe en Sambrook y col. (1989). Las condiciones de hibridación, incluyendo la restricción moderada y la restricción alta, se proporcionan en Sambrook y col., véase antes.

**[0183]** También se pueden usar sondas o partes de las mismas en técnicas de PCR para generar un grupo de secuencias para la identificación de secuencias estrechamente relacionadas con *bg/3*. Cuando las secuencias de *bg/3* están dirigidas al uso de sondas, se puede usar una parte particular de una secuencia que codifica BGL3, por ejemplo, una parte altamente conservada de la secuencia codificante.

[0184] Por ejemplo, una secuencia de nucleótidos de *bg/3* se puede usar como una sonda de hibridación para una biblioteca de ADNc para aislar genes, por ejemplo, los que codifican variantes naturales de BGL3 de otras especies de hongos, bacterias o plantas, que tienen un nivel deseado de identidad de secuencia con la secuencia de nucleótidos de *bg/3* descrita en la figura 1 (SEQ ID NO: 1). Las sondas de ejemplo tienen una longitud de aproximadamente 20 a aproximadamente 50 bases.

### **B. Análisis de doble híbrido**

[0185] Las proteínas identificadas por los procedimientos descritos en el presente documento, se pueden usar en el sistema de doble híbrido de levaduras para “capturar” las proteínas de unión a la proteína que son proteínas putativas de la ruta de señales. El sistema de doble híbrido de levaduras se describe en Fields y Song, *Nature* 340: 245 – 246 (1989). Brevemente, en un sistema de doble híbrido, se construye una fusión de un dominio de unión a ADN-*bg/3* (p. ej., fusión GAL4-*bg/3*) y se transfecta en células de levadura. Se puede usar el gen *bg/3* entero, o subregiones del gen *bg/3*. Se cotransfecta una segunda construcción que contiene la biblioteca de potenciales parejas de unión fusionadas al dominio de activación de ADN. Los cotransformantes de levadura que albergan proteínas que se unen a la proteína BGL3 son identificados, por ejemplo, por la producción (un cribado) de beta-galactosidasa o luciferasa, o supervivencia en placas que carecen de un nutriente esencial (una selección), según sea adecuado para los vectores usados.

### **C. Análisis de micromatrices**

[0186] Además, se puede usar el análisis de micromatrices, conocido también como descripción de perfil de expresión o descripción de perfil de transcrito, para evaluar simultáneamente la presencia o expresión de secuencias de ADN dadas, o cambios de la expresión de muchos genes diferentes. En un procedimiento, se forma una matriz de un conjunto grande de secuencias de ADN (sondas), normalmente un conjunto amplio de marcadores de secuencia expresados, ADNc, fragmentos de ADNc, u oligonucleótidos específicos de secuencia, sobre un soporte sólido tal como un portaobjetos de vidrio o membrana de nailon. La diana marcada para hibridar con las sondas se genera aislando ARNm del tejido control e inducido, después marcando cada mezcla de ARNm directamente o a través de un ADNc o ARNc intermedio, con un marcador distinto, normalmente un colorante fluorescente. La micromatriz se hibrida con las sondas complejas y se puede cuantificar la intensidad relativa de la señal de hibridación con cada localización en la matriz para cada colorante marcador. Las diferencias en la expresión entre los estados de control e inducidos se puede medir como una relación de las señales de los dos colorantes marcadores. (Véase Baldwin, D y col., 1999).

[0187] El análisis de micromatrices del organismo fuente del cual se ha obtenido *bg/3* se puede llevar a cabo para facilitar la comprensión de la función del gen identificando otros genes que son regulados de forma coordinada como consecuencia del exceso de expresión de *bg/3*. Para identificar los genes regulados de forma coordinada puede ayudar el poner el gen *bg/3* en una ruta particular. Alternativamente, dicho análisis se puede usar para identificar otros genes implicados en la misma ruta usando análisis de micromatrices.

[0188] Aunque la invención se ha descrito con referencia a procedimientos y realizaciones específicos, se apreciará que se pueden hacer diferentes modificaciones y cambios sin salirse de la invención, como se define en las reivindicaciones.

### **EJEMPLO 1**

[0189] En un procedimiento de ejemplo, se aísla un fragmento de ADNc para usar como una sonda extrayendo el ARN total de micelios de una cepa de *T. reesei* cultivada en condiciones que se sabe que inducen la producción de celulasa, y obteniendo la fracción poliadenilada (poliA) de la misma. El poliA ARN se usa para producir una mezcla de ADNc que después se amplifica usando cebadores específicos basados en la secuencia de ácido nucleico de *bg/3* proporcionada en el presente documento.

[0190] El ARN total se aísla de micelios usando procedimientos conocidos en la materia, por ejemplo, como describen Timberlake y col., 1981; Maniatis, y col., 1989; Ausubel, y col., 1993 y Sambrook y col., 1989. Una vez aislado se llevan a cabo transferencias Northern para confirmar la expresión de la celulasa y seleccionar un tiempo de inducción óptimo para la expresión de la celulasa y el correspondiente aislamiento del ARN.

[0191] El ARN mensajero (ARNm) que tiene una cola de poli (A) en el extremo 3', se puede purificar del ARN total usando procedimientos conocidos en la materia.

[0192] El ARN de *T. reesei* se usa como molde para la RT-PCR usando procedimientos conocidos en la materia (Loftus, J. y col., *Science*, 249: 915 – 918, 1990). Durante este procedimiento se lleva a cabo la transcripción inversa

del ARNm para producir la primera cadena de ADNc. El ADNc posteriormente sirve como molde para la amplificación por PCR de las secuencias de ADNc de *bgl3* usando cebadores oligonucleótidos específicos diseñados de acuerdo con la SEQ ID No. 1 o SEQ ID No. 4.

Tabla 1. Secuencias proporcionadas para apoyar la invención

Descripción	SEQ ID NO.
<p>Secuencia de ácido nucleico de ADNc de <i>bgl3</i> de longitud completa</p> <p><b>CCACGCGTCCGACTAGTTCTAGATCCCGAGTACCTTGGTCCGCGGCCCGT TCATCATGAAGACGTTGTCAGTGTTGCTGCCGCCCTTTGGCGGCCGTA GCTGAGGCCAATCCCTACCCGCTCCTCACTCCAACCAGGCGTACTCGC CTCCTTTCTACCTTCGCCATGGATGGACCCAGTGCTCCAGGCTGGGA GCAAGCCTATGCCCAAGCTAAGGAGTCGTCTCGGGCTTGACTCTTTG GAGAAGGTCAACCTCACCACCGGTGTTGGCTGGATGGGTGAGAAGTGCG TTGGAAACGTTGGTACCGTGCCCTCGCTTGGGCATGCGAAGTCTTTGCATG CAGGACGGCCCCCTGGGTCTCCGATTCAACACGTACAACAGCGCTTTCA GCGTTGGCTTGACGGCCGCCGCCAGCTGGAGCCGACACCTTTGGGTTG ACCGCGGTACCGCTCTGGCTCCGAGGCAAAGGGCAAGGGTGTGATG TTCTTCTCGGACCCCGTGGCTGGCCCTCTCGGTGCGAACCCCAACGGAGG CCGTAACGTCGAGGGTTTTCGGCTCGGATCCCTATCTGGCGGGTTTGGCT CTGGCCGATACCGTGACCGGAATCCAGAACGCGGGCACCATCGCCTGTG CCAAGCACTTCTCCTCAACGAGCAGGAGCATTTCGCCAGGTCCGGCA AGCTAACGTTACGGATACCCCATCACCGAGGCTCTGTCTTCCAACGTTG ATGACAAGACGATTACGAGGTGTACGGCTGGCCCTTCCAGGATGCTGT CAAGGCTGGTGTCCGGTCTTTCATGTGCTCGTACAACCAGGTCAACAAC CGTACGCTTGCCAAAACCTCAAGCTCATCAACGCTTGTCAAGGAGGA GTACGGTTTCCAAGGCTTTGTCATGAGCGACTGGCAGGCCAGCACACG GGTGTGCGCTGTGCTGTTGCCGGTCTCGATATGACCATGCCTGGTGACA CCGCCTTCAACACCGGCGCATCCTACTTTGGAAGCAACCTGACGCTTGT GTTCTCAACGGCACCGTCCCGAGTGGCGCATTGACGACATGGTGATGC GTATCATGGCTCCCTTCTTCAAGGTGGGCAAGACGGTTGACAGCCTCATT GACACCAACTTTGATTCTTGGACCAATGGCGAGTACGGCTACGTTCCAGGC CGCCGTCAATGAGAAGTGGGAGAAGGTCAACTACGGCGTCGATGTCCGC GCTCAACCATGCGAACACATCCGCGAGGTTGGCGCAAGGGAAGTGTCA TCTTCAAGAACACGGCATCCTGCCCTTAAGAAGCCCAAGTTCTGACC GTCATTGGTGAGGATGCTGGCGGCAACCCTGCCGGCCCCAACGGCTGC GGTGACCGCGGCTGTGACGACGGCACTCTTGCCATGGAGTGGGGATCT GGTACTACCAACTTCCCTACCTCGTCACCCCGACGCGGCCCTGCAGA GCCAGGCTCTCCAGGACGGCACCCGCTACGAGAGCATCCTGTCCAAC CGCCATCTCGCAGACCCAGGCGCTCGTCAGCCAGCCCGATGCCATTGCC ATTGCTTTTGCCAACTCGGATAGCGGCGAGGGCTACATCAACGTGATG GCAACGAGGCGACCGCAAGAACCTGACGCTGTGGAAGAACGGCGACG ATCTGATCAAGACTGTTGCTGTGTCACCCCAAGACGATTGTGTCATC CACTCGACCGGCCCGTGATTCTCAAGGACTACGCCAACCAACCCCAACA TCTCTGCCATTCTGTGGGCCGGTCTCCTGGCCAGGAGTCTGGCAACTC GCTGGTGCACATTCTGTACGGCAAGCAGAGCCCGGGCCGCACTCCCTTC ACCTGGGGCCCGTCGCTGGAGAGCTACGGAGTTAGTGTTATGACCACGC CCAACAACGGCAACGGCGCTCCCGAGGATAACTTCAACGAGGGCGCCTT CATCGACTACCGCTACTTTGACAAGGTGGCTCCCGGCAAGCCTCGCAGC TCGGACAAGGCTCCACGTACGAGTTTGGCTTCGGACTGTCGTGGTGA CGTTCAAGTTCTCCAACCTCCACATCCAGAAGAACAATGTCGGCCCCATG AGCCCGCCCAACGGCAAGACGATTGCGGCTCCCTCTCTGGGCAGCTTCA GCAAGAACCTTAAGGACTATGGCTTCCCAAGAACGTTCCGCCATCAAG GAGTTTATCTACCCCTACCTGAGCACCCTACCTCTGGCAAGGAGGCGTC GGGTGACGCTCACTACGGCCAGACTGCGAAGGAGTTCTCCCGCCGG TGCCCTGGACGGCAGCCCTCAGCCTCGCTCTGCGGCCTCTGGCGAAC CGGCGGCAACCGCCAGCTGTACGACATTCTACACCGTGACGGCCACC ATTACCAACACGGGCTCGGTTCATGGACGACGCCGTTCCCGAGCTGTACC</b></p>	<p>1</p>

<p><b>TGAGCCACGGCGGTCCCAACGAGCCGCCCAAGGTGCTGCGTGGCTTCG          ACCGCATCGAGCGCATTGCTCCCGGCCAGAGCGTCACGTTCAAGGCAGA          CCTGACGCGCCGTGACCTGTCCAACCTGGGACACGAAGAAGCAGCAGTG          GGTCATTACCGACTACCCCAAGACTGTGTACGTGGGCAGCTCCTCGCGC          GACCTGCCGCTGAGCGCCCGCCTGCCATGAGGGAGACAAGATGTGACG          CGAATGTTTAGTGTATAGATAAGTATTAGTATTAATCAGATTAATGAAGC          TYTTGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</b></p>	
<p>Secuencia de aminoácidos predicha de BGL3</p> <p><b>EANPYPPPHSNQAYSPPFYPSWMDPSAPGWEQAYAQAKEFVSGTLLEK          VNLTGTVGWMGEKCVGNVGTVPRLGMRSLCMQDGPLGLRFNTYNSAFSVG          LTAAASWSRHLWVDRGTALGSEAKGKGVVLLGPVAGPLGRNPNGGRNVE          GFGSDPYLAGLALADTVTGIQNAGTIACAKHFLNEQEHFRQVGEANGYGY          ITEALSSNVDDKIHEVYGVWPFQDAVKAGVGSFMCYSYNQVNNSYACQNSKLI          NGLLKEEYGFQGFVMSDWQAQHTGVASAVAGLDMTPGDTAFNTGASYFG          SNLTLAVLNGTVPEWRIDDMVMRIMAPFFKVGKTVDSLIDTNFDSWTNGEY          YVQAAVNENWEKVNYGVDVRANHANHIREVGAAGTVIFKNNGILPLKPKFL          TVIGEDAGGNPAGPNGCGDRGCDGTLAMEWGSSTNFPYLVTTPDAALQS          QALQDGTYESILSNYASQTQALVSQPDAAIAIVFANSDSGEGYINVDGNEGD          RKNLTLWKNGDDLKTVAAVNPKTIVVIHSTGPVILKDYANHPNISAILWAGAP          GQESGNSLVDILYKQSPGRTPFTWGPSLESYGVSVMTTPNNGNGAPQDN          FNEGAFIDYRYFDKVAPGKPRSSDKAPTYEFGFGLSWSTFKFSNLHIQKNNV          GPMSPPNGKTIAAPSLGFSKLNLDYGFKNVRRIKEFIYPYLSTTTSGKEAS          GDAHYGQTAKEFLPAGALDGSPQPRSAASGEPGGNRQLYDILYTVTATITNT          GVMDDAVPQLYLSHGPNPEPKVLRGFDRIERIAPGQSVTFKADLTRDLS          NWDTKKQQWVITDYPKTVYVGSSSRDLPLSARLP</b></p>	<p>2</p>
<p>Secuencia señal predicha de la proteína BGL3:</p> <p><b>MKTLVFAAALLAAVA</b></p>	<p>3</p>
<p>Secuencia que codifica el ácido nucleico de <i>bgl3</i></p>	<p>4</p>

<p>ATGAAGACGTTGTCAGTGTTGCTGCCGCCCTTTGGCGGCCGTAGCTG          AGGCCAATCCCTACCCGCCTCCTCACTCCAACCAGGCGTACTCGCCTCC          TTTCTACCCTTCGCCATGGATGGACCCAGTGCTCCAGGCTGGGAGCAA          GCCTATGCCAAGCTAAGGAGTTCGTCTCGGGCTTGACTCTCTTGAGAA          GGTCAACCTCACCACCGGTGTTGGCTGGATGGGTGAGAAGTGCGTTGGA          AACGTTGGTACCGTGCCTCGCTTGGGCATGCGAAGTCTTTGCATGCAGG          ACGGCCCCCTGGGTCTCCGATTCAACACGTACAACAGCGCTTTTCAGCGT          TGGCTTGACGGCCCGCCAGCTGGAGCCGACACCTTTGGGTTGACCG          CGGTACCGCTCTGGGCTCCGAGGCAAAGGGCAAGGGTGTTCGATGTTCTT          CTCGGACCCGTGGCTGGCCCTCTCGGTGCGCAACCCCAACGGAGGCCGT          AACGTCGAGGGTTTCGGCTCGGATCCCTATCTGGCGGGTTTGGCTCTGG          CCGATACCGTGACCGGAATCCAGAACGCGGGCACCATCGCCTGTGCCAA          GCACTTCCTCCTCAACGAGCAGGAGCATTTCGCGCAGGTCGGCGAAGCT          AACGGTTACGGATACCCCATCACCGAGGCTCTGTCTTCCAACGTTGATGA          CAAGACGATTACGAGGTGTACGGCTGGCCCTTCCAGGATGCTGTCAAG          CGTGGTGTGCGGGTCTTCATGTGCTCGTACAACCAGGTCAACAACCTCGTA          CGTTGCCAAAACCTCAAGCTCATCAACGGCTTGCTCAAGGAGGAGTAC          GGTTTCCAAGGCTTTGTCATGAGCGACTGGCAGGCCAGCACACGGGTG          TCGCGTCTGCTGTTGCCGGTCTCGATATGACCATGCCTGGTGACACCGC          CTTCAACACCGGCGCATCCTACTTTGGAAGCAACCTGACGCTTGCTGTTT          TCAACGGCACCGTCCCCGAGTGGCGCATTGACGACATGGTGATGCGTAT          CATGGCTCCCTTCTCAAGGTGGGCAAGACGGTTGACAGCCTCATTGACA          CCAACTTTGATTCTTGACCAATGGCGAGTACGGCTACGTTACAGGCCGC          CGTCAATGAGAAGTGGGAGAAGGTCAACTACGGCGTCGATGTCCGCGCC          AACCATGCGAACCACATCCGCGAGGTTGGCGCCAAGGGAAGTGTATCT          TCAAGAACAACGGCATCCTGCCCTTAAGAAGCCCAAGTTCTGACCGTC</p>	
<p>ATTGGTGAGGATGCTGGCGGCAACCCTGCCGGCCCCAACGGCTGCGGT          GACCGCGGCTGTGACGACGGCACTCTTGCCATGGAGTGGGGATCTGGTA          CTACCAACTTCCCCTACCTCGTCAACCCCGACGCGGCCCTGCAGAGCCA          GGCTCTCAGGACGGCACCCGCTACGAGAGCATCCTGTCCAACCTACGCC          ATCTCGCAGACCCAGGCGCTCGTCAAGCCAGCCGATGCCATTGCCATTG          TCTTTGCCAACTCGGATAGCGGCGAGGGCTACATCAACGTCGATGGCAA          CGAGGGCGACCGCAAGAACCTGACGCTGTGGAAGAACGGCGACGATCT          GATCAAGACTGTTGCTGCTGTCAACCCCAAGACGATTGTGTCATCCACT          CGACCCGGCCCCGTGATTCTCAAGGACTACGCCAACCCACCCCAACATCTC          TGCCATTCTGTGGGCGGTGCTCCTGGCCAGGAGTCTGGCAACTCGCTG          GTCGACATTCTGTACGGCAAGCAGAGCCCGGGCCGCACTCCCTTCACT          GGGGCCCGTGCCTGGAGAGCTACGGAGTTAGTGTATGACCACGCCCAA          CAACGGCAACGGCGCTCCCCAGGATAACTTCAACGAGGGCGCCTTCATC          GACTACCGTACTTTGACAAGGTGGCTCCCGGCAAGCCTCGCAGCTCGG          ACAAGGCTCCCACGTACGAGTTTGGCTTCGGACTGTCGTGGTTCGACGTT          CAAGTTCTCAACCTCCACATCCAGAAGAACAATGTGCGCCCCATGAGCC          CGCCCAACGGCAAGACGATTGCGGCTCCCTCTCTGGGCAGCTTCAGCAA          GAACCTTAAGGACTATGGCTTCCCAAGAAGCTTCGCCGCATCAAGGAGT          TTATCTACCCCTACCTGAGCACCCTACCTCTGGCAAGGAGGCGTCGGG          TGACGCTCACTACGGCCAGACTGCGAAGGAGTTCTCCCGCCGGTGCC          CTGGACGGCAGCCCTCAGCCTCGCTCTGCGGCCTCTGGCGAACC CGGC          GGCAACCGCCAGCTGTACGACATTCTTACACCGTGACGGCCACCATTA          CCAACACGGGCTCGGTCATGGACGACGCGTTCCCGAGCTGTACCTGAG          CCACGGCGGTCCCAACGAGCCGCCAAGGTGCTGCGTGGCTTCGACCG          CATCGAGCGCATTGCTCCCGGCCAGAGCGTCACGTTCAAGGCAGACCTG          ACGCGCCGTGACCTGTCCAACCTGGGACACGAAGAAGCAGCAGTGGGTCA          TTACCGACTACCCCAAGACTGTGTACGTGGGCAGCTCCTCGCGCGACCT          GCCGCTGAGCGCCCCGCTGCCATGA</p>	

**REIVINDICACIONES**

1. Un polinucleótido aislado seleccionado del grupo que consiste en:
- 5 (a) una secuencia de ácido nucleico que codifica o es complementaria de una secuencia que codifica un polipéptido BGL3, que tiene una identidad de secuencia de al menos 98 % con la secuencia de aminoácidos presentada como SEQ ID NO: 21 en la tabla 1.
- (b) una secuencia de ácido nucleico que codifica o es complementaria de una secuencia que codifica un polipéptido BGL3 que tiene la secuencia de aminoácidos presentada como SEQ ID NO: 2 en la tabla 1; y
- 10 (c) una secuencia de ácido nucleico presentada como SEQ ID NO: 4 en la tabla 1 o la complementaria de la misma;
- en el que el % de identidad se calcula como la identidad de la secuencia de aminoácidos de BGL3 presentada como SEQ ID NO: 2 en la tabla 1 a lo largo de su longitud entera, usando el programa CLUSTAL-W de MacVector versión 15 6.5, utilizando los parámetros por defecto que incluyen una penalización de apertura de hueco de 10,0, una penalización de extensión de hueco de 0,1 y una matriz de similitud BLOSUM 30.
2. El polinucleótido aislado de la reivindicación 1, en el que dicho polinucleótido es una molécula de ARN.
- 20 3. Un polinucleótido aislado de acuerdo con una cualquiera de las reivindicaciones 1 y 2, que codifica una enzima que tiene actividad de  $\beta$ -glucosidasa, en el que la enzima deriva de una fuente de *Trichoderma*.
- 25 4. El polinucleótido aislado de la reivindicación 3, en el que la enzima deriva de *Trichoderma reesei*.
5. Una construcción de expresión que incluye una secuencia de polinucleótido (i) que codifica o es complementaria de una secuencia que codifica un polipéptido BGL3 que tiene una identidad de secuencia de al menos 98 % con la secuencia de aminoácidos presentada como SEQ ID NO: 2 en la tabla 1, o (ii) que es como se define en la reivindicación 2.
- 30 6. Un vector que incluye la construcción de expresión de la reivindicación 5.
7. Un vector que comprende un polinucleótido aislado de la reivindicación 1 o reivindicación 2, operativamente unido a secuencias de control reconocidas por una célula huésped transformada con el vector.
- 35 8. Una célula huésped transformada con el vector de la reivindicación 7 o la reivindicación 8.
9. Una célula huésped recombinante que comprende un polinucleótido de la reivindicación 1 o la reivindicación 2.
- 40 10. La célula huésped de la reivindicación 9 o la reivindicación 10, que es una célula procariota.
11. La célula huésped de la reivindicación 9 o la reivindicación 10, que es una célula eucariota.
- 45 12. Un polipéptido BGL3 purificado con la actividad biológica de una  $\beta$ -glucosidasa, que comprende una secuencia seleccionada del grupo que consiste en:
- (a) una secuencia de aminoácidos que tiene una identidad de secuencia de al menos 98 % con la secuencia de aminoácidos presentada como SEQ ID NO: 2 en la tabla 1;
- 50 (b) una secuencia de aminoácidos presentada como SEQ ID NO: 2 en la tabla 1;
- (c) un fragmento biológicamente activo purificado de la secuencia de aminoácidos presentada como SEQ ID NO: 2 en la tabla 1;
- en el que el % de identidad se calcula como la identidad con la secuencia de aminoácidos de BGL3 presentada como SEQ ID NO: 2 en la tabla 1 a lo largo de toda su longitud.
- 60 13. Un procedimiento de producción de una enzima que tiene actividad de  $\beta$ -glucosidasa, que comprende:
- (a) transformar establemente una célula huésped con un vector de expresión que comprende un polinucleótido

como se define en la reivindicación 1 o la reivindicación 2;

(b) cultivar dicha célula huésped transformada en condiciones adecuadas para que dicha célula huésped produzca dicha  $\beta$ -glucosidasa; y

(c) recuperar dicha  $\beta$ -glucosidasa.

14. El procedimiento de la reivindicación 14, en el que la célula huésped es una célula de hongo filamentoso o de levadura.

15. Una enzima purificada que tiene actividad de  $\beta$ -glucosidasa preparada por el procedimiento de la reivindicación 14.

16. Una célula huésped recombinante que comprende una eliminación o inserción u otra alteración en el gen *bg/3* que inactiva el gen y previene la producción del polipéptido BGL3, en la que el polipéptido BGL3 tiene una identidad de secuencia de al menos 98 % con la secuencia del polipéptido BGL3 presentado como SEQ ID NO: 2 en la tabla 1, en el que el % de identidad se calcula como la identidad de la secuencia de aminoácidos de BGL3 presentada como SEQ ID NO: 2 en la tabla 1 a lo largo de toda su longitud.

17. Un oligonucleótido de sentido contrario complementario de un ARN mensajero que codifica un polipéptido BGL3 que tiene la secuencia presentada como SEQ ID NO: 2 en la tabla 1, en el que tras exposición a una célula huésped que produce  $\beta$ -glucosidasa, dicho oligonucleótido disminuye o inhibe la producción de  $\beta$ -glucosidasa por dicha célula huésped.

18. El oligonucleótido de sentido contrario de la reivindicación 18, en el que la célula huésped es un hongo filamentoso.

19. Una composición de detergente, comprendiendo dicha composición un polipéptido seleccionado del grupo que consiste en:

(a) una secuencia de aminoácidos que tiene una identidad de secuencia de al menos 98 % con la secuencia de aminoácidos presentada como SEQ ID NO: 2 en la tabla 1;

(b) una secuencia de aminoácidos presentada como SEQ ID NO: 2 en la tabla 1;

(c) un fragmento biológicamente activo purificado de la secuencia de aminoácidos presentada como SEQ ID NO: 2 en la tabla 1,

en el que el % de identidad se calcula como la identidad con la secuencia de aminoácidos de BGL3 presentada como SEQ ID NO: 2 en la tabla 1 a lo largo de toda su longitud.

20. Un procedimiento para mejorar las características de una masa de levadura o producto horneado hecho con dicha masa, que consiste esencialmente en las etapas de:

(a) mezclar al menos aproximadamente 10 ppm de un BGL3 de acuerdo con la reivindicación 12 con ingredientes de masa para formar una mezcla de masa, y

(b) hornear dicha mezcla de masa para formar un producto horneado.

21. Un procedimiento para mejorar las características de la masa de pan de levadura o masa de bollo de levadura o pan de levadura o bollo de levadura, que consiste esencialmente en las etapas de:

(a) mezclar al menos aproximadamente 10 ppm de un BGL3 de acuerdo con la reivindicación 12 con ingredientes de masa de pan o de bollo, para formar una mezcla de masa, y

(b) dar forma o moldear la mezcla de masa;

(c) levar la mezcla de masa, y

(d) hornear la mezcla de masa para formar pan o bollos.

22. Un procedimiento para expresar un polipéptido heterólogo que tiene actividad de  $\beta$ -glucosidasa en una

especie de *Aspergillus*, que comprende:

- 5 (a) proporcionar un *Aspergillus* huésped con un vector de expresión que comprende un polinucleótido que codifica una secuencia señal de  $\beta$ -glucosidasa de *Aspergillus* unida a un polinucleótido que codifica una  $\beta$ -glucosidasa heteróloga 3 que comprende la SEQ ID NO: 2 en la tabla 1, codificando así un polipéptido quimérico;
- (b) cultivar dicho *Aspergillus* huésped en condiciones adecuadas para que dicho *Aspergillus* produzca dicho polipéptido quimérico, en el que se produce dicho polipéptido quimérico.

CCACGCGTCCGACTAGTTCTAGATCCCCGAGTACCTTGGTCGCGGCCCGTTCATCATGAAGA  
CGTTGTCAAGTGTGCTGCGCCCTTTTGGCGGCCGCTAGCTGAGGCCAATCCCTACCCGCC  
TCTCACTCCAACCAGGCGTACTCGCCTCCTTTCTACCCTTCGCCATGGATGGACCCAGT  
GCTCCAGGCTGGGAGCAAGCCTATGCCCAAGCTAAGGAGTTCGTCFCGGGCTTGACTCTCT  
TGGAGAAGGTCAACCTCACCACCGGTGTTGGCTGGATGGGTGAGAAGTGCCTTGGAAACGT  
TGGTACCCTGCGCTTGGGCATGCGAAGTCTTTGCATGCAGGACGGCCCCCTGGGTCTC  
CGATTC AACACGTACAACAGCGCTTTCAGCGTTGGCTTGACGGCCGCCAGCTGGAGCC  
GACACCTTTGGGTTGACCGCGGTACCCTCTGGGCTCCGAGGCAAAGGGCAAGGGTGTCTGA  
TGTCTCTTCGGACCCGTGGCTGGCCCTCTCGGTGCGAACCCCAACGGAGGCCGTAACGTC  
GAGGGTTTCGGCTCGGATCCCTATCTGGCGGGTTTGGCTCTGGCCGATACCGTGACCGGAA  
TCCAGAACGCGGGCACCATCGCCTGTGCCAAGCACTTCTCCTCAACGAGCAGGAGCATTT  
CCGCCAGGTGGCGAAGCTAACGGTTACGGATACCCATCACCAGGCTCTGTCTTCCAAC  
GTTGATGACAAGACGATTACAGGTTTACGGCTGGCCCTTCCAGGATGCTGTCAAGGCTG  
GTGTGGGTCCTTCATGTGCTCGTACAACCAGGTCAACAACCTGTACGCTTGCCAAAACCTC  
CAAGCTCATCAACGGCTTGTCAAGGAGGAGTACGGTTTCCAAGGCTTTGTTCATGAGCGAC  
TGGCAGGCCAGCACACGGGTGTCGCTCTGTCTGTTGCCGGTCTCGATATGACCATGCCTG  
GTGACACCGCCTTCAACACCGGCGCATCCTACTTTGGAAGCAACCTGACGCTTGTCTTCT  
CAACGGCACCGTCCCGAGTGGCGCATGACGACATGGTGTATGCGTATCATGGCTCCCTTC  
TTCAAGGTGGGCAAGACGGTTGACAGCTCATTTGACACCAACTTTGATTCTTGGACCAATG  
GCGAGTACGGCTACGTTTACGGCCGCCGTCATGAGAAGTGGGAGAAGGTCAACTACGGCGT  
CGATGTCCGCGCCAACCATGCGAACACATCCGCGAGGTTGGCGCCAAGGGAACTGTCTATC  
TTCAAGAACAACGGCATCCTGCCCTTAAGAAGCCCAAGTTCTTGACCGTCAATGGTGAGG  
ATGCTGGCGGCAACCTGCGGCCCAACGGCTGCGGTGACCGCGGCTGTGACGACGGCAC  
TCTTGCCATGGAGTGGGGATCTGGTACTACCAACTTCCCTTACCTCGTCACCCCCGACGCG  
GCCCTGCAGAGCCAGGCTCTCCAGGACGGCACCCGCTACGAGAGCATCTGTCCAACACTAG  
CCATCTCGCAGACCCAGGCGCTCGTCAGCCAGCCCGATGCCATTGCCATTGTCTTTGCCAA  
CTCGGATAGCGGCGAGGGCTACATCAACGTGATGGCAACGAGGGCGACCGCAAGAACCTG  
ACGCTGTGGAAGAACGGCGACGATCTGATCAAGACTGTTGCTGCTGTCAACCCCCAAGACGA  
TTGTGCTCATCCACTCGACCGGCCCGGTGATTCTCAAGGACTACGCCAACCACCCCAACAT  
CTCTGCCATTCTGTGGGCCGGTGTCTCTGGCCAGGAGTCTGGCAACTCGCTGGTTCGACATT  
CTGTACGGCAAGCAGAGCCCGGGCCGCACTCCCTTACCTGGGGCCCGTTCGCTGGAGAGCT  
ACGGAGTTAGTGTATGACCACGCCCAACAACGGCAACGGCGCTCCCCAGGATAACTTCAA  
CGAGGGCCGCTTCATCGACTACCGCTACTTTGACAAGGTGGCTCCCGGCAAGCCTCGCAGC  
TCGGACAAGGCTCCACGTACGAGTTTGGCTTCGGACTGTGCTGGTTCGACGTTCAAGTTCT  
CCAACCTCCACATCCAGAAGAACAATGTGCGCCCATGAGCCCGCCCAACGGCAAGACGAT  
TGCGGCTCCCTCTCTGGGCAGCTTCAGCAAGAACCCTAAGGACTATGGCTTCCCCAAGAAC  
GTTTCGCCGATCAAGGAGTTTATCTACCCCTACCTGAGCACCCTACCTCTGGCAAGGAGG  
CGTCGGGTGACGCTCACTACGGCCAGACTGCGAAGGAGTTCCCTCCCCGCGGTGCCCTGGA  
CGGCAGCCCTCAGCTCGCTCTGCGGCCCTCTGGCGAACCCGGCGGCAACCGCCAGCTGTAC  
GACATTCTTACACCGTGACGGCCACCAATTACCAACACGGGCTCGGTCAATGGACGACCGCG  
TTCCCCAGCTGTACCTGAGCCACGGCGGTCCCAACGAGCCGCCAAGGTGCTGCGTGGCTT  
CGACCGCATCGAGCGCATTGCTCCCGGCCAGAGCGTACGTTCAAGGCAGACCTGACGCGC  
CGTGACCTGTCCAACCTGGGACACGAAGAAGCAGCAGTGGGTCAATACCGACTACCCCAAGA  
CTGTGTACGTGGGCAGCTCCTCGCGCGACCTGCCGCTGAGCGCCCGCCTGCCATGAGGGAG  
ACAAGATGTGACGCGAATGTTTAGTGTATAGATAAGTATTAGTATTAATCAGATTAATGAA  
GCTYTTGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

**FIG. 1**

**MKTL**SVFAAALLAAVAEANPYPPPHSNQAYSPPFYPSPWMDPSAPGWEQAYAQAKEFVSGL  
TLLEKVNLTGTVGWMGEKCVGNVGTVPRLGMRSLCMQDGPLGLRFNTYNSAFSVGLTAAAS  
WSRHLWVDRGTALGSEARKGKGV DVLLGVPVAGPLGRNPNGGRNVEGFGSDPYLAGLALADTV  
TGIQNAGTIACAKHFLNEQEHRQVGEANGYGYPI TEALSSNVDDKTIHEVYGWPFQDAV  
KAGVGSFMCSYNQVNNYSYACQNSKLLINGLLKEEYGFQGFVMSDWQAQHTGVASAVAGLDMT  
MPGDTAFNTGASYFGSNLTLAVLN GTVPEWRIDDMVMRIMAPFFKVGKTVDSLIDTNFDSW  
TNGEYGYVQAAVNENWEKVNYGVDVVRANHANHIREVGA KGTVIFKNNGILPLKKPKFLT VI  
GEDAGGNPAGPNGCGDRGCDDGTLAMEWGS GTTNFPYLVT PDAALQS QALQDGTRYESILS  
NYAISQTQALVSQPDAIAIVFANS DSGEGYINVDGNEGDRKNLTLWKNGDDL IKTVA AVNP  
KTIVVIHSTGVPVILKDYANHPNISA I LWAGAPGOESGNSLVDILY GKQSPGRTPFTWGP SL  
ESYGVSVM TTPNNGNGAPQDNFN EGAFIDYRYFDK VAPGKPRSSDKAPTYEFGFGLSWSTF  
KFSNLHIQKNNVGPMSPPNGK TIAAPSLG SF SKNLKDYGF PKNVRR I KEFIYPYLSTTTSG  
KEASGDAHYGQTAKEFLPAGALDGSPQPRSAASGE PGGNRQLYDILYTVTATITNTGSVMD  
DAVPQLYLSHGGPN EPPKVL RGFDR IERIA PGQSVTFKADL TRRDLSNWDTKKQQWVITDY  
PKTVYVGS SSRDLPLSARLP

La secuencia de aminoácidos predicha (SEQ ID NO: 2) y secuencia señal (SEQ ID NO: 3) basadas en la secuencia de nucleótidos proporcionada en la figura 1, en la que la secuencia señal se indica en negrilla.

## **FIG.\_2**