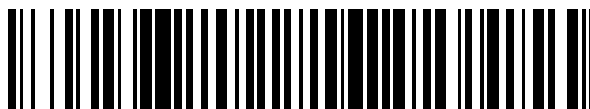


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 601 145**

51 Int. Cl.:

C12Q 1/68 (2006.01)
C12N 9/42 (2006.01)
C12N 1/20 (2006.01)
C12N 15/00 (2006.01)
C07H 21/04 (2006.01)
C07K 1/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **05.11.2003 PCT/US2003/035672**

87 Fecha y número de publicación internacional: **27.05.2004 WO04043980**

96 Fecha de presentación y número de la solicitud europea: **05.11.2003 E 03768790 (2)**

97 Fecha y número de publicación de la concesión europea: **15.06.2016 EP 1556512**

54 Título: **Beta-glucosidasa BGL6 y ácidos nucleicos que codifican la misma**

30 Prioridad:

07.11.2002 US 424784 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
14.02.2017

73 Titular/es:

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

72 Inventor/es:

**DUNN-COLEMAN, NIGEL y
WARD, MICHAEL**

74 Agente/Representante:

RIZZO, Sergio

ES 2 601 145 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 BGL6 y ácidos nucleicos que codifican la misma

Apoyo gubernamental

- 5 **[0001]** Partes de este trabajo fueron financiadas por el subcontrato n.º ZCO-30017-01 con el Laboratorio Nacional de Energía Renovable de EE.UU. según el contrato principal n.º DE-AC36-99GC310337 con el Departamento de Energía de los EE.UU.

Campo

- 10 **[0002]** La presente exposición se refiere a secuencias de ácido nucleico *bgl6* aisladas que codifican polipéptidos que tienen actividad de beta-glucosidasa. La exposición se refiere también a constructos de ácido nucleico, vectores y células huésped que comprenden las secuencias de ácido nucleico así como a métodos para la producción de polipéptidos de BGL6 recombinantes.

Referencias

[0003]

- 15 Altschul, S. F., *et al.*, *J. Mol. Biol.* 215:403-410, 1990.
 Altschul, S. F., *et al.*, *Nucleic Acids Res.* 26:3389-3402, 1997.
 Aro, N., *et al.*, *J. Biol. Chem.*, 10.1074/ M003624200, 13 Abril 2001.
 Aubert, *et al.*, *Ed.*, p 11 *et seq.*, Academic Press, 1988.
 Ausubel G. M., *et al.* *CURRENT PROTOCOLS IN MOLECULAR BIOLOGY*, John Wiley & Sons, Nueva York, N.Y., 1993.
 20 Baldwin, D., *et al.*, *Curr. Opin. Plant Biol.* 2(2):96-103, 1999.
 Baulcombe, D., *Arch. Virol. Suppl.* 15:189-209, 1999.
 Bhikhabhai, R. *et al.*, *J. Appl. Biochem.* 6:336, 1984.
 Brumbauer, A. *et al.*, *Bioseparation* 7:287-295, 1999.
 Carter *et al.*, *Nucl. Acids Res.* 13:4331, 1986.
 25 Chen *et al.*, *Biochem. Biophys. Acta.* 1121:54-60, 1992.
 Coligan, J. E. *et al.*, eds., *CURRENT PROTOCOLS IN IMMUNOLOGY*, 1991.
 Collen, A., *et al.*, *Journal of Chromatography A* 910:275-284, 2001.
 Coughlan, *et al.*, *BIOCHEMISTRY AND GENETICS OF CELLULOSE DEGRADATION*.
 Cummings y Fowler, *Curr. Genet.* 29:227-233; 1996.
 30 Dayhoff *et al.* en *Atlas of Protein Sequence and Structure*, Volumen 5, Suplemento 3, Capítulo 22, pp. 345-382, 1978.
 Deutscher, M.P., *Methods Enzymol.* 182:779-80, 1990.
 Doolittle, R. F., *OF URFs AND ORFs*, University Science Books, CA, 1986.
 Ellouz, S. *et al.*, *J. Chromatography* 396:307, 1987.
 35 Fields y Song, *Nature* 340:245-246, 1989.
 Filho, *et al.* *Can. J. Microbiol.* 42:1-5, 1996.
 Fliess, A., *et al.*, *Eur. J. Appl. Microbiol. Biotechnol.* 17:314, 1983.
 Freer, *et al.* *J. Biol. Chem.* 268:9337-9342, 1993.
 Freshney, R. I., ed., *ANIMAL CELL CULTURE*, 1987.
 40 Goyal, A. *et al.* *Bioresource Technol.* 36:37, 1991.
 Halldorsdottir, S *et al.*, *Appl Microbiol Biotechnol.* 49(3):277-84, 1998.
 Hu *et al.*, *Mol Cell Biol.* 11:5792-9, 1991.
 Hemmpel, W.H. *ITB Dyeing/Printing/Finishing* 3:5-14, 1991.
 Herr *et al.*, *Appl. Microbiol. Biotechnol.* 5:29-36, 1978.
 45 Jakobovits, A, *et al.*, *Ann N Y Acad Sci* 764:525-35, 1995.
 Jakobovits, A, *Curr Opin Biotechnol* 6(5):561-6, 1995.
 Jones *et al.*, *Nature* 321:522-525, 1986.
 Kawaguchi, T *et al.*, *Gene* 173(2):287-8, 1996.
 Knowles, J. *et al.*, *TIBTECH* 5, 255-261, 1987.
 50 Kohler y Milstein, *Nature* 256:495, 1975.
 Krishna, S. *et al.*, *Bioresource Tech.* 77:193-196, 2001.
 Kumar, A., *et al.*, *Textile Chemist and Colorist* 29:37-42, 1997.
 Lehtio, J. *et al.*, *FEMS Microbiology Letters* 195:197-204, 2001.
 Li y Ljungdahl *Appl. Environ. Microbiol.* 62:209-213, 1996.
 55 Linder, M. y Teeri, T.T., *Biotechnol.* 57:15-28, 1997.
 Medve, J. *et al.*, *J. Chromatography A* 808:153, 1998.
 Ohmiya *et al.*, *Biotechnol. Gen. Engineer. Rev.* 14:365-414, 1997.
 Ooi *et al.*, *Nucleic Acids Res.* 18(19):5884, 1990.
 Ortega *et al.*, *International Biodeterioration and Biodegradation* 47:7-14, 2001.

- Penttilä *et al.*, *Yeast* 3:175-185, 1987.
 Penttilä *et al.*, *Gene* 63: 103-112, 1988.
 Pere, J., *et al.*, *In Proc. Tappi Pulping Conf.*, Nashville, TN, 27-31, pp. 693-696, 1996.
 Riechmann *et al.*, *Nature* 332:323-327, 1988.
 5 Rothstein *et al.*, *Gene* 55:353-356, 1987.
 Saarihtti *et al.*, *Gene* 90:9-14, 1990.
 Sakamoto *et al.*, *Curr. Genet.* 27:435-439, 1995.
 Saloheimo M, *et al.*, *Gene* 63:11-22, 1988.
 Sambrook *et al.*, *MOLECULAR CLONING: A LABORATORY MANUAL* (Segunda edición), Cold Spring
 10 Harbor Press, Plainview, N.Y., 1989.
 Schulein, *Methods Enzymol.*, 160,25, pp. 234 *et seq.*, 1988.
 Scopes, *Methods Enzymol.* 90 Pt E:479-90, 1982.
 Spilliaert R, *et al.*, *Eur J Biochem.* 224(3):923-30, 1994.
 Stahlberg, J. *et al.*, *Bio/Technol.* 9:286-290,1991.
 15 Strathern *et al.*, eds. (1981) *The Molecular Biology of the Yeast Saccharomyces*. Suurnakki, A. *et al.*,
Cellulose 7:189-209, 2000.
 Te'o, J. *et al.*, *FEMS Microbiology Letters* 190:13-19, 2000.
 Tilbeurgh, H. *et al.*, *FEBS Lett.* 16:215, 1984.
 Timberlake *et al.*, *Cell* 1:29-37, 1981.
 20 Tomaz, C. y Queiroz, J., *J. Chromatography A* 886:123-128, 1999.
 Tomme, P. *et al.*, *Eur. J. Biochem.* 170:575-581, 1988.
 Tormo, J. *et al.*, *EMBO J.* 15:5739-5751, 1996.
 Tyndall, R.M., *Textile Chemist and Colorist* 24:23-26, 1992.
 Van Rensburg *et al.*, *Yeast* 14:67-76, 1998.
 25 Van Tilbeurgh, H. *et al.*, *FEBS Lett.* 204:223-227, 1986.
 Verhoeyen *et al.*, *Science* 239:1534-1536, 1988.
 Warrington, *et al.*, *Genomics* 13:803-808, 1992.
 Wells *et al.*, *Gene* 34:315, 1985.
 Wells *et al.*, *Philos. Trans. R. Soc. London SerA* 317:415, 1986.
 30 Wood, *Biochem. Soc. Trans.*, 13, pp. 407-410,1985.
 Wood *et al.*, *METHODS IN ENZYMOLOGY*, 160, 25, p. 87 *et seq.*, Academic Press, Nueva York, 1988.
 Zoller *et al.*, *Nucl. Acids Res.* 10:6487, 1987.

Antecedentes

[0004] La celulosa y la hemicelulosa son los materiales vegetales más abundantes producidos mediante
 35 fotosíntesis. Pueden ser degradados y utilizados como fuente de energía por numerosos microorganismos,
 incluyendo bacterias, levaduras y hongos, que producen enzimas extracelulares capaces de hidrolizar los
 sustratos poliméricos en azúcares monoméricos (Aro *et al.*, 2001). Puesto que se acercan los límites de los
 recursos no renovables, el potencial de la celulosa para convertirse en un importante recurso de energía
 40 renovable es enorme (Krishna *et al.*, 2001). La utilización efectiva de la celulosa a través de procesos biológicos
 es un enfoque para superar la escasez de alimentos, piensos y combustibles (Ohmiya *et al.*, 1997).

[0005] Las celulasas son enzimas que hidrolizan celulosa (enlaces beta-1,4-glucano o beta D-glucosídicos)
 dando lugar a la formación de glucosa, celobiosa, celooligosacáridos y similares. 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 ([beta] -D-glucósido glucohidrolasa; EC 3.2.1.21)
 45 ("BG"). (Knowles *et al.*, 1987; Schülein, 1988). Las endoglucanasas actúan principalmente en las partes amorfas
 de la fibra de celulosa, mientras que las celobiohidrolasas son también capaces de degradar celulosa cristalina
 (Nevalainen y Penttilä, 1995). De esta manera, es necesaria la presencia de una celobiohidrolasa en un sistema
 de celulasas para la solubilización eficaz de la celulosa cristalina (Suurnakki, *et al.* 2000). La beta-glucosidasa
 50 actúa para liberar unidades de D-glucosa a partir de 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 celulosa, de manera
 que las celulasas se producen fácilmente en grandes cantidades a través de la fermentación. Los hongos
 filamentosos desempeñan un papel especial, puesto que muchas levaduras, tales como *Saccharomyces*
 55 *cerevisiae*, carecen de la capacidad para hidrolizar celulosa. Véase, p. ej., Aro *et al.*, 2001; Aubert *et al.*, 1988;
 Wood *et al.*, 1988 y Coughlan, *et al.*.

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

[0008] Con el fin de convertir de manera eficaz la celulosa cristalina en glucosa, se necesita el sistema de celulasas completo que comprende componentes de cada una de las clasificaciones CBH, EG y BG, con componentes aislados menos eficaces en la hidrólisis de celulosa cristalina (Filho *et al.*, 1996). Se ha observado una relación sinérgica entre componentes de celulasa de diferentes clasificaciones. En concreto, las celulasas de tipo EG y las celulasas de tipo CBH interactúan sinérgicamente para degradar celulosa de manera más eficiente. Véase, por ejemplo, Wood, 1985.

[0009] Las celulasas se conocen en la técnica por ser útiles en el tratamiento de tejidos con el fin de mejorar la capacidad de limpieza de las composiciones detergentes, para el uso como agente suavizante, para mejorar el tacto y la apariencia de los tejidos de algodón, y similares (Kumar *et al.*, 1997).

[0010] Se han descrito composiciones detergentes que contienen celulasa con rendimiento de limpieza mejorado (Patente estadounidense n.º 4 435 307; solicitudes GB n.º 2 095 275 y 2 094 826) y para su utilización en el tratamiento de telas para mejorar el tacto y la apariencia del tejido (Patentes estadounidenses n.º 5 648 263, 5 691 178, y 5 776 757; solicitud GB n.º 1 358 599; *The Shizuoka Prefectural Hamamatsu Textile Industrial Research Institute Report*, Vol. 24, pp. 54-61, 1986).

[0011] Por lo tanto, las celulasas producidas en hongos y bacterias han recibido una atención significativa. En concreto, se ha demostrado que la fermentación de *Trichoderma spp.* (p. ej., *Trichoderma longibrachiatum* o *Trichoderma reesei*) produce un sistema de celulasas completo capaz de degradar formas cristalinas de la celulosa. La patente estadounidense n.º 5 475 101 describe la purificación y la clonación molecular de una enzima particularmente útil designada EGIII que se deriva de *Trichoderma longibrachiatum*.

[0012] Foreman P. *et al.* (2003), *J. Biol. Chem.*, vol. 278, n.º 34, pp. 31988-31997 describe la secuenciación parcial de más de 5100 clones de ADNc de *T. reesei* aleatorios. Se identificaron doce secuencias que codificaban enzimas previamente desconocidas que probablemente funcionan en la degradación de biomasa. Los análisis de los niveles de expresión de las micromatrices indicaron que varias de las secuencias se sobreexpresaron de manera coordinada en una cepa que producía celulasa en exceso.

[0013] WO99/46362 hace referencia a la modificación genética de un microbio para mejorar su producción de una enzima, beta-glucosidasa. Se comunica la transformación con un constructo que comprende un promotor, una señal de secreción de xilanas y una región de codificación de beta-glucosidasa madura para aumentar la cantidad de beta-glucosidasa producida con relación a los microbios no transformados.

[0014] US6184018 describe una β -glucosidasa de *Orpinomyces sp.* PC2, secuencias de nucleótidos que codifican la proteína madura y la proteína precursora y métodos para la producción recombinante de la β -glucosidasa.

[0015] US6022725 describe un proceso para la expresión de β -glucosidasa extracelular en un hongo filamentoso mediante la expresión de una secuencia de ADN fúngica que codifica una β -glucosidasa mejorada, delecionada o alterada en un microorganismo huésped recombinante. También se describen las composiciones de celulasas fúngicas recombinantes que contienen una expresión mejorada, delecionada o alterada de β -glucosidasa.

[0016] Aunque las composiciones de celulasas se han descrito anteriormente, se siguen necesitando composiciones de celulasas nuevas y mejoradas para utilizarse en detergentes domésticos, composiciones para lavado a la piedra o detergentes para ropa, etc. Son de especial interés las celulasas que muestran resistencia a los surfactantes (p. ej., alquilsulfonatos lineales, LAS (por sus siglas en inglés)), rendimiento mejorado en condiciones de estrés térmico, capacidad celulolítica aumentada o disminuida y/o alto nivel de expresión *in vitro*.

Sumario

[0017] En el presente documento se describen una proteína de celulasa aislada, identificada en el presente documento como BGL6 y ácidos nucleicos que codifican BGL6.

[0018] En un aspecto, las proteínas o polipéptidos de BGL6 comprenden una secuencia que tiene al menos 80 %, 85 %, 90 %, 95 %, 98 % o más de identidad de secuencia con la secuencia presentada como SEQ ID NO:2.

[0019] Un aspecto relacionado incluye (i) fragmentos de BGL6, preferiblemente al menos aproximadamente 20-100 aminoácidos de largo, más preferiblemente aproximadamente 100-200 aminoácidos de largo y (ii) una composición farmacéutica que comprende BGL6. En varios modos de realización, el fragmento corresponde con el dominio N-terminal de BGL6 o el dominio C-terminal de BGL6.

[0020] Otro aspecto incluye un polinucleótido aislado que tiene una secuencia que codifica BGL6, una secuencia complementaria a la secuencia de codificación de *bgl6* y una composición que comprende el polinucleótido. El polinucleótido puede ser ARNm, ADN, ADNc, ADN genómico o un antisentido análogo de los mismos.

[0021] Un polinucleótido de *bgl6* puede comprender una molécula de ácido nucleico aislada que se hibrida al complemento del ácido nucleico presentado como SEQ ID NO: 1 en condiciones de moderada a alta restricción, donde la molécula de ácido nucleico codifica un polipéptido BGL6 que muestra actividad de beta-glucosidasa.

5 **[0022]** El polinucleótido puede codificar una proteína BGL6 que tenga al menos 80 %, 85 %, 90 %, 95 %, 98 % o más de identidad de secuencia con la secuencia presentada como SEQ ID NO:1. En un modo de realización específico, el polinucleótido comprende una secuencia sustancialmente idéntica a la SEQ ID NO:1. También se contemplan fragmentos del polinucleótido, preferiblemente al menos aproximadamente 15-30 nucleótidos de largo.

10 **[0023]** Se describen los vectores de expresión recombinantes que contienen una secuencia de ácido nucleico que codifica BGL6 o un fragmento o variante de corte y empalme de los mismos, operativamente unidos a elementos reguladores eficaces para la expresión de la proteína en un huésped seleccionado. Un aspecto relacionado incluye una célula huésped que contiene el vector.

15 **[0024]** Se incluye un método para la producción de BGL6 mediante técnicas recombinantes, cultivando células huésped procariontes o eucariotas recombinantes que comprenden una secuencia de ácidos nucleicos que codifica BGL6 en condiciones eficaces para promover la expresión de la proteína incluida, y la posterior recuperación de la proteína de la célula huésped o del medio de cultivo celular.

20 **[0025]** Otro aspecto proporciona una composición enzimática útil en la conversión de celulosa a etanol. En un modo de realización preferido, la composición enzimática comprende BGL6. La composición puede comprender además enzimas de celulosa adicionales tales como endoglucanasas y/o celobiohidrolasas. La composición puede enriquecerse en BGL6.

[0026] Otro aspecto adicional incluye un anticuerpo específicamente inmunorreactivo con BGL6.

[0027] Los métodos analíticos para la detección de los ácidos nucleicos *bgl6* y las proteínas BGL6 también forman parte de la exposición.

Breve descripción de las figuras

25 **[0028]**

La Figura 1 es una representación monocatenaria de la secuencia de ácidos nucleicos (SEQ ID NO:1), de la *bgl6* de *T. reesei*, donde la secuencia no codificante se indica como subrayada.

La Figura 2 muestra la secuencia de aminoácidos prevista (SEQ ID NO:2) basada en la secuencia de nucleótidos proporcionada en la Figura 1, donde se utiliza el primer codón de inicio.

30 La Figura 3 muestra la secuencia de aminoácidos prevista (SEQ ID NO:4) basada en la secuencia de nucleótidos proporcionada en la Figura 1, donde se utiliza el segundo codón de inicio.

La Figura 4 es la secuencia de codificación *bgl6*, donde los dos codones de inicio alternativos están subrayados.

Descripción detallada

35 **I. Definiciones.**

[0029] A menos que se indique lo contrario, todos los términos técnicos y científicos utilizados en el presente documento tienen el mismo significado que tendrían para un experto en la materia de la presente invención. Se dirige a los profesionales en concreto a Sambrook *et al.*, 1989, y Ausubel FM *et al.*, 1993, para definiciones y términos de la técnica. Ha de entenderse que la presente invención no está limitada a la metodología, los protocolos y los reactivos concretos descritos, puesto que estos pueden variar.

40 **[0030]** El término "polipéptido" como se utiliza en el presente documento se refiere a un compuesto formado por una sola cadena de residuos de aminoácidos unidos por enlaces peptídicos. El término "proteína" como se utiliza en el presente documento puede ser sinónimo del término "polipéptido" o puede referirse, además, a un compuesto de dos o más polipéptidos.

45 **[0031]** El término "molécula de ácido nucleico" incluye moléculas de ARN, ADN y ADNc. Se entenderá que, como resultado de la degeneración del código genético, se puede producir una multitud de secuencias de nucleótidos que codifican una proteína determinada, tal como BGL6. La presente invención contempla todas las posibles variantes de secuencias de nucleótidos, que codifican BGL6, las cuales son todas posibles debido a la degeneración del código genético.

50 **[0032]** Un constructo o secuencia de ácido nucleico "heteróloga" tiene una parte de la secuencia que no es nativa a la célula en la que se expresa. Heteróloga, con respecto a una secuencia de control, se refiere a una secuencia de control (es decir, promotor o potenciador) que no sirve por naturaleza para regular el mismo gen cuya expresión está regulando actualmente. En general, las secuencias de ácidos nucleicos heterólogas no son

endógenas a la célula o parte del genoma en el que están presentes y se han añadido a la célula mediante infección, transfección, transformación, microinyección, electroporación o similar. Un constructo de ácido nucleico "heterólogo" puede contener una combinación de secuencia de control/secuencia de codificación de ADN que sea la misma que, o diferente a, una combinación de secuencia de control/secuencia de codificación de ADN encontrada en la célula nativa.

[0033] Como se utiliza en el presente documento, el término "vector" se refiere a un constructo de ácido nucleico diseñado para la transferencia entre células huésped diferentes. Un "vector de expresión" se refiere a un vector que tiene la capacidad de incorporar y expresar fragmentos de ADN heterólogo en una célula extraña. Muchos vectores de expresión procariotas y eucariotas están comercialmente disponibles. La selección de vectores de expresión adecuados está dentro del conocimiento de los expertos en la materia.

[0034] Por consiguiente, un "casete de expresión" o "vector de expresión" es un constructo de ácido nucleico generado de forma recombinante o sintética, con una serie de elementos de ácido nucleico especificados que permiten la transcripción de un ácido nucleico concreto en una célula diana. El casete de expresión recombinante se puede incorporar en un plásmido, cromosoma, ADN mitocondrial, ADN plastidial, virus, o fragmento de ácido nucleico. Normalmente, la parte del casete de expresión recombinante de un vector de expresión incluye, entre otras secuencias, una secuencia de ácido nucleico a transcribir y un promotor.

[0035] Como se utiliza en el presente documento, el término "plásmido" se refiere a un constructo de ADN bicatenario circular utilizado como un vector de clonación y que forma un elemento genético autorreplicante extracromosómico en muchas bacterias y algunos eucariotas.

[0036] Como se utiliza en el presente documento, el término "secuencia de nucleótidos que codifica un marcador seleccionable" se refiere a una secuencia de nucleótidos que es capaz de expresión en células y donde la expresión del marcador seleccionable confiere a las células que contienen el gen expresado la capacidad de crecer en presencia de un agente selectivo correspondiente, o en condiciones de crecimiento selectivo correspondientes.

[0037] Como se utiliza en el presente documento, el término "promotor" se refiere a una secuencia de ácido nucleico que sirve para dirigir la transcripción de un gen secuencia abajo. El promotor será en general adecuado para la célula huésped en la que se está expresando el gen diana. El promotor junto con otras secuencias de ácido nucleico reguladoras de la transcripción y la traducción (también denominadas "secuencias de control") son necesarios para expresar un gen determinado. En general, las secuencias reguladoras de la transcripción y la traducción incluyen, pero sin carácter limitativo, secuencias promotoras, sitios de unión al ribosoma, secuencias de inicio y de detención de la transcripción, secuencias de inicio y detención de la traducción y secuencias potenciadoras o activadoras.

[0038] "Gen quimérico" o "constructo de ácido nucleico heterólogo", como se define en el presente documento, se refiere a un gen no nativo (es decir, uno que se ha introducido en un huésped) que puede estar compuesto por partes de genes diferentes, incluyendo elementos reguladores. Un constructo de gen quimérico para la transformación de una célula huésped se compone normalmente de una región reguladora de la transcripción (promotor) operativamente unida a una secuencia de codificación de proteína heteróloga o, en un gen quimérico marcador seleccionable, a un gen marcador seleccionable que codifica una proteína que confiere resistencia a los antibióticos a las células transformadas. Un gen quimérico típico, para la transformación en una célula huésped, incluye una región reguladora de la transcripción que es constitutiva o inducible, una secuencia de codificación de proteína y una secuencia de terminación. Un constructo de gen quimérico puede incluir también una segunda secuencia de ADN que codifica un péptido señal si se desea la secreción de la proteína diana.

[0039] Un ácido nucleico está "operativamente unido" cuando se coloca en una relación funcional con otra secuencia de ácido nucleico. Por ejemplo, el ADN que codifica un líder secretor está operativamente unido al ADN para un polipéptido si se expresa como una preproteína que participa en la secreción del polipéptido; un promotor o potenciador está operativamente unido a una secuencia de codificación si afecta a la transcripción de la secuencia; o un sitio de unión al ribosoma está operativamente unido a una secuencia de codificación si está colocado de manera que facilite la traducción. En general, "operativamente unido" significa que las secuencias de ADN que están unidas son contiguas, y, en el caso de un líder secretor, contiguas y en el marco de lectura. Sin embargo, los potenciadores no tienen que ser contiguos. La unión se consigue mediante ligadura en sitios de restricción convenientes. Si dichos sitios no existen, los adaptadores oligonucleótidos sintéticos, ligadores o cebadores para PCR se utilizan de conformidad con la práctica convencional.

[0040] Como se utiliza en el presente documento, el término "gen" significa el segmento de ADN que participa en la producción de una cadena de polipéptidos, que puede o no incluir regiones anteriores y posteriores a la región de codificación, p. ej., secuencias 5' no traducidas (5' UTR) o secuencias "líder" y secuencias 3' UTR o secuencias "tráiler", así como secuencias intercaladas (intrones) entre los segmentos de codificación individuales (exones).

[0041] En general, las moléculas de ácido nucleico que codifican BGL6 o un análogo u homólogo de las mismas se hibridarán, en condiciones de moderada a alta restricción a la secuencia natural proporcionada en el presente documento como SEQ ID NO: 1. Sin embargo, en algunos casos se emplea una secuencia de nucleótidos que codifica BGL6 que posee un uso de codón sustancialmente diferente, mientras que la proteína codificada por la secuencia de nucleótidos que codifica BGL6 tiene la secuencia de aminoácidos igual o sustancialmente igual que la proteína nativa. Por ejemplo, la secuencia de codificación se puede modificar para facilitar la expresión más rápida de BGL6 en un sistema concreto de expresión procariota o eucariota, de conformidad con la frecuencia con la que el huésped utiliza un codón concreto. Te'o, *et al.* (2000), por ejemplo, describe la optimización de genes para la expresión en hongos filamentosos.

[0042] Una secuencia de ácido nucleico se considera "selectivamente hibridable" a una secuencia de ácido nucleico de referencia si las dos secuencias se hibridan específicamente entre sí en condiciones de lavado y de hibridación de moderada a alta restricción. Las condiciones de hibridación se basan en la temperatura de fusión (T_f) de la sonda o del compuesto de unión a ácido nucleico. Por ejemplo, la "restricción máxima" ocurre normalmente a aproximadamente T_f-5 °C (5 °C por debajo de la T_f de la sonda); la "restricción alta" a aproximadamente 5 °C-10 °C por debajo de la T_f; la "restricción intermedia" a aproximadamente 10 °C-20 °C por debajo de la T_f de la sonda; y la "restricción baja" a aproximadamente 20 °C-25 °C por debajo de la T_f. Funcionalmente, las condiciones de restricción máxima se pueden utilizar para identificar secuencias que tienen identidad estricta o identidad casi estricta con la sonda de hibridación; mientras que las condiciones de restricción alta se utilizan para identificar secuencias que tienen aproximadamente 80 % o más de identidad de secuencia con la sonda.

[0043] Las condiciones de hibridación de moderada y alta restricción se conocen bien en la técnica (véase, por ejemplo, Sambrook, *et al.*, 1989, Capítulos 9 y 11, y en Ausubel, F.M., *et al.*, 1993, expresamente incorporado por referencia en el presente documento). Un ejemplo de condiciones de restricción alta incluye hibridación a aproximadamente 42 °C en formamida al 50 %, SSC 5X, solución de Denhardt 5X, SDS al 0,5 % y 100 µg/ml de ADN portador desnaturalizado seguido del lavado dos veces en SSC 2X y SDS al 0,5 % a temperatura ambiente y dos veces adicionales en SSC 0,1X y SDS al 0,5 % a 42 °C.

[0044] Como se utiliza en el presente documento, el término "recombinante", incluye referencia a una célula o vector que se ha modificado por la introducción de una secuencia de ácido nucleico heteróloga o que la célula se deriva de una célula modificada de esta manera. Por lo tanto, por ejemplo, las células recombinantes expresan genes que no se encuentran en forma idéntica en la forma nativa (no recombinante) de la célula o expresan genes nativos que de otra manera se expresan anormalmente, se subexpresan o no se expresan en absoluto como resultado de la intervención humana deliberada.

[0045] Como se utilizan en el presente documento, los términos "transformada", "establemente transformada" o "transgénica" con referencia a una célula significan que la célula tiene una secuencia de ácido nucleico no nativa (heteróloga) integrada en su genoma o como un plásmido episomal que se mantiene a través de múltiples generaciones.

[0046] Como se utiliza en el presente documento, el término "expresión" se refiere al proceso por el cual se produce un polipéptido basándose en la secuencia de ácido nucleico de un gen. El proceso incluye tanto la transcripción como la traducción.

[0047] El término "introducido", en el contexto de la inserción de una secuencia de ácido nucleico en una célula significa "transfección" o "transformación" o "transducción" e incluye referencia a la incorporación de una secuencia de ácido nucleico en una célula eucariota o procariota en la que la secuencia de ácido nucleico puede incorporarse al genoma de la célula (por ejemplo, cromosoma, plásmido, plástido o ADN mitocondrial), convertirse en un replicón autónomo o expresarse transitoriamente (por ejemplo, ARNm transfectado).

[0048] Se deduce que el término "expresión de BGL6" se refiere a la transcripción y traducción del gen *bgl6*, cuyos productos incluyen ARN precursor, ARNm, polipéptido, polipéptidos procesados tras la traducción y derivados de los mismos, que incluyen BGL6 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 BGL6 incluyen transferencia Western para proteína BGL6, análisis de transferencia Northern y ensayos de reacción en cadena de la polimerasa con transcriptasa inversa (RT-PCR, por sus siglas en inglés) para ARNm de BGL6, y ensayos de actividad de glucosidasa como se describe en Chen *et al.* (1992) y Herr *et al.* (1978).

[0049] El término "corte y empalme alternativo" se refiere al proceso por el cual se generan múltiples isoformas de polipéptido a partir de un solo gen, e implica el corte y empalme junto de exones no consecutivos durante el procesamiento de algunas, pero no todas, las transcripciones del gen. De esta manera, se puede conectar un exón concreto a cualquiera de los varios exones alternativos para formar ARN mensajeros. Los ARNm cortados y empalmados de manera alternativa producen polipéptidos ("variantes de corte y empalme") en los que algunas partes son comunes, mientras que otras partes son diferentes.

[0050] El término "secuencia señal" se refiere a una secuencia de aminoácidos en la parte N-terminal de una proteína que facilita la secreción de la forma madura de la proteína fuera de la célula. La forma madura de la proteína extracelular carece de la secuencia señal que se escinde durante el proceso de secreción.

5 **[0051]** Por el término "célula huésped" se entiende una célula que contiene un vector y soporta la replicación y/o la transcripción o la transcripción y la traducción (expresión) del constructo de expresión. Las células huésped para su uso en la presente invención pueden ser células procariotas, tales como *E. coli*, o células eucariotas, tales como células de levaduras, plantas, insectos, anfibios o mamíferos. En general, las células huésped son hongos filamentosos.

10 **[0052]** El término "hongos filamentosos" significa todos y cada uno de los hongos filamentosos reconocidos por los expertos en la materia. Se selecciona un hongo preferido del grupo consistente en *Aspergillus*, *Trichoderma*, *Fusarium*, *Chrysosporium*, *Penicillium*, *Humicola*, *Neurospora*, o formas sexuales alternativas de los mismos tales como *Emericella*, *Hypocrea*.

[0053] El término "celooligosacárido" se refiere a grupos oligosacáridos que contienen de 2 a 8 unidades de glucosa y tienen enlaces β -1,4, p. ej., celobiosa.

15 **[0054]** 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 glucosa. Se han obtenido numerosos ejemplos de celulasas, tales como exoglucanasas, exocelobiohidrolasas, endoglucanasas y glucosidasas, a partir de organismos celulolíticos, incluyendo particularmente hongos, plantas y bacterias.

20 **[0055]** El término "dominio de unión a celulosa" como se utiliza en el presente documento se refiere a la parte de la secuencia de aminoácidos de una celulasa o una región de la enzima que interviene en la actividad de unión a celulosa de una celulasa o un derivado de la misma. Los dominios de unión a celulosa generalmente funcionan uniéndose de manera no covalente a la celulosa, un derivado de celulosa u otro polisacárido equivalente de la misma. Los dominios de unión a celulosa permiten o facilitan la hidrólisis de fibras de celulosa por la región de núcleo catalítico estructuralmente distinto y funcionan normalmente independientes del núcleo catalítico. De esta manera, un dominio de unión a celulosa no poseerá la actividad hidrolítica significativa atribuible a un núcleo catalítico. Dicho de otro modo, un dominio de unión a celulosa es un elemento estructural de la estructura terciaria de la proteína enzimática de celulasa que es distinto del elemento estructural que posee actividad catalítica.

30 **[0056]** Como se utiliza en el presente documento, el término "surfactante" se refiere a cualquier compuesto generalmente reconocido en la técnica por tener cualidades activas de superficie. De esta manera, por ejemplo, los surfactantes comprenden surfactantes aniónicos, catiónicos y no iónicos, tales como los que se encuentran habitualmente en los detergentes. Los surfactantes aniónicos incluyen alquilbencenosulfonatos lineales o ramificados; sulfatos de éter de alquilo o alqueno que tienen grupos alquilo o grupos alqueno lineales o ramificados; sulfatos de alquilo o alqueno; olefinsulfonatos; y alcanosulfonatos. Los surfactantes anfólicos incluyen sulfonatos de sal de amonio cuaternario y surfactantes anfólicos de tipo betaína. Dichos surfactantes anfólicos tienen tanto grupos de carga positiva como negativa en la misma molécula. Los surfactantes no iónicos pueden comprender éteres de polioxialquilo, así como alcanolamidas de ácidos grasos superiores o aductos de óxido de alquilo de los mismos, monoésteres de glicerina de ácidos grasos, y similares.

40 **[0057]** Como se utiliza en el presente documento, el término "tejido que contiene celulosa" se refiere a cualquier tejido cosido o no cosido, hilo o fibra de algodón o no de algodón que contienen celulosa o de algodón o no de algodón que contienen mezclas de celulosa, incluyendo celulósicos naturales y celulósicos sintéticos (tales como yute, lino, ramio, rayón y liocel).

45 **[0058]** Como se utiliza en el presente documento, el término "tejido que contiene algodón" se refiere a tejidos cosidos o no cosidos, hilos o fibras de algodón puro o de mezclas de algodón, incluyendo tejidos de algodón tejido, prendas de punto de algodón, vaqueros de algodón, hilos de algodón, algodón en bruto y similares.

50 **[0059]** Como se utiliza en el presente documento, el término "composición para lavado a la piedra" se refiere a una formulación para utilizarse en tejidos que contienen celulosa lavados a la piedra. Las composiciones para lavado a la piedra se utilizan para modificar tejidos que contienen celulosa antes de su venta, es decir, durante el proceso de fabricación. Por el contrario, las composiciones detergentes están destinadas a la limpieza de ropa sucia y no se utilizan durante el proceso de fabricación.

55 **[0060]** Como se utiliza en el presente documento, el término "composición detergente" se refiere a una mezcla que está destinada a utilizarse en un medio de lavado para el lavado de tejidos sucios que contienen celulosa. En el contexto de la presente invención, tales composiciones pueden incluir, además de celulasas y surfactantes, enzimas hidrolíticas adicionales, mejoradores, agentes blanqueadores, activadores de blanqueo, agentes azules y tintes fluorescentes, inhibidores del apelmazamiento, agentes enmascarantes, activadores de celulasa, antioxidantes y solubilizantes.

[0061] Como se utiliza en el presente documento, el término “reducción o eliminación de la expresión del gen *bgl6*” significa que el gen *bgl6* ha sido deletado del genoma y, por tanto, no puede ser expresado por el microorganismo huésped recombinante; o que el gen *bgl6* ha sido modificado de tal manera que el microorganismo huésped recombinante no produce una enzima BGL6 funcional.

5 **[0062]** El término “*bgl6* alterada” o “gen *bgl6* alterado” significa que la secuencia de ácido nucleico del gen se ha alterado mediante la eliminación, adición y/o manipulación de la secuencia de codificación o que la secuencia de aminoácidos de la proteína expresada se ha modificado.

10 **[0063]** Como se utiliza en el presente documento, el término “purificar” generalmente se refiere a someter ácidos nucleicos transgénicos o células que contienen proteínas a purificación bioquímica y/o cromatografía en columna.

[0064] Como se utiliza en el presente documento, los términos “activo” y “biológicamente activo” se refieren a una actividad biológica asociada a una determinada proteína, tal como la actividad enzimática asociada a una proteasa. Se deduce que la actividad biológica de una proteína determinada se refiere a cualquier actividad biológica que los expertos en la materia atribuyen normalmente a esa proteína.

15 **[0065]** Como se utiliza en el presente documento, el término “enriquecida” significa que la BGL6 se encuentra en una concentración que es mayor en relación con la concentración de BGL6 encontrada en una composición de celulasa fúngica natural, o de origen natural.

20 **[0066]** Una composición de celulasa fúngica natural es una producida por una fuente fúngica de origen natural y que comprende uno o más componentes BG, CBH y EG, donde cada uno de estos componentes se encuentra en la proporción producida por la fuente fúngica. De esta manera, una composición de BGL6 enriquecida tendría BGL6 en una proporción alterada, donde la proporción de BGL6 con respecto a otros componentes de celulasa (es decir CBH y endoglucanasas) es elevada. Esta proporción se puede incrementar ya sea aumentando la BGL6 o disminuyendo (o eliminando) al menos algún otro componente mediante cualquier medio conocido en la técnica.

25 **[0067]** De esta manera, como ilustración, un sistema de celulasas de origen natural se puede purificar en componentes sustancialmente puros mediante técnicas de separación reconocidas publicadas en la literatura, incluyendo cromatografía de intercambio iónico a un pH adecuado, cromatografía de afinidad, exclusión de tamaño y similares. Por ejemplo, en la cromatografía de intercambio iónico (normalmente cromatografía de intercambio aniónico) es posible separar los componentes de celulasa mediante elución con un gradiente de pH, 30 o un gradiente de sal, o ambos, un gradiente de pH y de sal. La BGL6 purificada se puede añadir entonces a la solución enzimática dando lugar a una solución de BGL6 enriquecida.

35 **[0068]** Las celulasas fúngicas pueden contener más de un componente BG. Los componentes diferentes generalmente tienen puntos isoeléctricos diferentes que permiten su separación mediante cromatografía de intercambio iónico y similares. En una solución enzimática se puede emplear un único componente BG o una combinación de componentes BG.

40 **[0069]** Cuando se emplea en soluciones enzimáticas, el componente BG se añade generalmente en una cantidad suficiente para evitar la inhibición por la celobiosa de cualquier componente de CBH y endoglucanasa encontrados en la composición de celulasa. La cantidad de componente BG añadida depende de la cantidad de celobiosa producida durante el proceso de sacarificación de biomasa, que puede determinarse fácilmente por el experto en la materia. Sin embargo, cuando se utiliza, el porcentaje en peso del componente de BGL6 en 45 relación con cualquier componente del tipo endoglucanasa o CBH presente en la composición de celulasa es preferiblemente de aproximadamente 1, preferiblemente aproximadamente 5, preferiblemente aproximadamente 10, preferiblemente aproximadamente 15 o preferiblemente aproximadamente 20 por ciento en peso a preferiblemente aproximadamente 25, preferiblemente aproximadamente 30, preferiblemente aproximadamente 35, preferiblemente aproximadamente 40, preferiblemente aproximadamente 45 o preferiblemente 50 por ciento en peso. Además, los intervalos preferidos pueden ser aproximadamente 0,5 a 50 aproximadamente 15 por ciento en peso, aproximadamente 0,5 a aproximadamente 20 por ciento en peso, de 55 aproximadamente 1 a aproximadamente 10 por ciento en peso, de aproximadamente 1 a aproximadamente 15 por ciento en peso, de aproximadamente 1 a aproximadamente 20 por ciento en peso, de aproximadamente 1 a aproximadamente 25 por ciento en peso, de aproximadamente 5 a aproximadamente 20 por ciento en peso, de aproximadamente 5 a aproximadamente 25 por ciento en peso, de aproximadamente 5 a aproximadamente 30 por ciento en peso, de aproximadamente 5 a aproximadamente 35 por ciento en peso, de aproximadamente 5 a aproximadamente 40 por ciento en peso, de aproximadamente 5 a aproximadamente 45 por ciento en peso, de aproximadamente 5 a aproximadamente 50 por ciento en peso, de aproximadamente 10 a aproximadamente 20 por ciento en peso, de aproximadamente 10 a aproximadamente 25 por ciento en peso, de aproximadamente 10 a aproximadamente 30 por ciento en peso, de aproximadamente 10 a aproximadamente 35 por ciento en peso, de aproximadamente 10 a aproximadamente 40 por ciento en peso, de aproximadamente 10 a aproximadamente 45 por ciento en peso, de aproximadamente 10 a aproximadamente 50 por ciento en peso, de aproximadamente 15 a aproximadamente 20 por ciento en peso, de aproximadamente 15 a aproximadamente 25 por ciento en

peso, de aproximadamente 15 a aproximadamente 30 por ciento en peso, de aproximadamente 15 a aproximadamente 35 por ciento en peso, de aproximadamente 15 a aproximadamente 30 por ciento en peso, de aproximadamente 15 a aproximadamente 45 por ciento en peso, de aproximadamente 15 a aproximadamente 50 por ciento en peso.

5 II. Organismos diana

A. Hongos filamentosos

[0070] Los hongos filamentosos incluyen todas las formas filamentosas de la subdivisión Eumycota y Oomycota. Los hongos filamentosos se caracterizan por un micelio vegetativo que tiene una pared celular compuesta por quitina, glucano, quitosano, manano y otros polisacáridos complejos, con crecimiento vegetativo mediante elongación hifal y catabolismo de carbono que es obligadamente aeróbico.

[0071] La célula parental de hongo filamentoso puede ser una célula de una especie de, pero sin carácter limitativo, *Trichoderma*, p. ej., *Trichoderma longibrachiatum* (reesei), *Trichoderma virile*, *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 utiliza en el presente documento, el término "*Trichoderma*" o "*Trichoderma* sp." se refiere a cualquier cepa fúngica que ha sido previamente clasificada como *Trichoderma* o que está actualmente clasificada como *Trichoderma*.

[0072] En un modo de realización preferido, la célula parental de hongo filamentoso es una célula de *Aspergillus niger*, *Aspergillus awamori*, *Aspergillus aculeatus*, o *Aspergillus nidulans*.

[0073] En otro modo de realización preferido, la célula parental de hongo filamentoso es una célula de *Trichoderma reesei*.

III. Celulasas

[0074] Las celulasas se conocen en la técnica como enzimas que hidrolizan celulosa (enlaces beta-1,4-glucano o beta D-glucosídicos) dando lugar a la formación de glucosa, celobiosa, celooligosacáridos y similares. Como se ha descrito anteriormente, las celulasas se han dividido tradicionalmente en tres clases principales: endoglucanasas (EC 3.2.1.4) ("EG"), exoglucanasas o celobiohidrolasas (EC 3.2.1.91) ("CBH") y beta-glucosidasas (EC 3.2.1.21) ("BG"). (Knowles *et al.*, 1987; Schulein, 1988).

[0075] Algunos hongos producen sistemas de celulasas completos que incluyen exocelobiohidrolasas o celulasas de tipo CBH, endoglucanasas o celulasas de tipo EG y beta-glucosidasas o celulasas de tipo BG (Schulein, 1988). Sin embargo, a veces estos sistemas carecen de celulasas de tipo CBH y las celulasas bacterianas también incluyen normalmente pocas o ninguna celulasa de tipo CBH. Además, se ha demostrado que los componentes EG y los componentes CBH interactúan sinérgicamente para degradar celulosa de manera más eficiente. Véase, p. ej., Wood, 1985. Los componentes diferentes, es decir, las diversas endoglucanasas y exocelobiohidrolasas en un sistema de celulasas completo o multicomponente, generalmente tienen propiedades diferentes, tales como punto isoelectrónico, peso molecular, grado de glicosilación, especificidad de sustrato y patrones de acción enzimática.

[0076] Se cree que las celulasas de tipo endoglucanasa hidrolizan los enlaces internos beta-1,4-glucosídicos en regiones de baja cristalinidad de la celulosa y que las celulasas de tipo exocelobiohidrolasa hidrolizan la celobiosa a partir del extremo reductor o no reductor de la celulosa. Se deduce que la acción de los componentes de endoglucanasa pueden facilitar en gran medida la acción de las exocelobiohidrolasas mediante la creación de nuevos extremos de cadena que son reconocidos por los componentes de exocelobiohidrolasa. Además, se ha demostrado que las celulasas de tipo beta-glucosidasa catalizan la hidrólisis de alquil y/o aril β -D-glucósidos, tales como metil β -D-glucósido y p-nitrofenil glucósido, así como glicósidos que contienen únicamente residuos de carbohidratos, tales como celobiosa. Esto produce glucosa como el único producto para el microorganismo y reduce o elimina la celobiosa que inhibe las celobiohidrolasas y las endoglucanasas.

[0077] Por consiguiente, las celulasas del tipo β -glucosidasa se consideran una parte esencial del sistema de celulasa porque conducen la reacción general a glucosa. Se ha demostrado que la expresión aumentada de la BG en *T. reesei* mejora la degradación de celulosa a glucosa. Véase EP0562003. Además, las β -glucosidasas pueden catalizar la hidrólisis de un número de sustratos diferentes y, por lo tanto, son útiles en una variedad de aplicaciones diferentes. Se pueden añadir algunas β -glucosidasas a las uvas durante la producción del vino para mejorar el aroma potencial del producto de vino acabado. Otra aplicación adicional puede ser la utilización de la β -glucosidasa en la fruta para mejorar el aroma de la misma. De manera alternativa, la β -glucosidasa puede utilizarse directamente en los aditivos alimenticios o en la elaboración del vino para mejorar el sabor y el aroma.

[0078] Las celulasas también encuentran un número de usos en composiciones detergentes, incluyendo la mejora de la capacidad de limpieza, como un agente suavizante y la mejora del tacto de los tejidos de algodón (Hemmpel, 1991; Tyndall, 1992; Kumar *et al.*, 1997). Aunque el mecanismo no forma parte de la invención, las

propiedades suavizantes y de restauración de color de la celulasa se han atribuido a los componentes de endoglucanasa alcalina en composiciones de celulasa, como lo demuestran las patentes estadounidenses n.º 5 648 263, 5 691 178 y 5 776 757, que exponen que las composiciones detergentes que contienen una composición de celulasa enriquecida en un componente de endoglucanasa alcalina especificado imparten restauración de color y suavizado mejorado a las prendas tratadas en comparación con las composiciones de celulasa no enriquecidas en dicho componente. Además, se ha demostrado que la utilización de dichos componentes de endoglucanasa alcalina en composiciones detergentes complementa los requisitos de pH de la composición detergente (p. ej., exhibiendo actividad máxima en un pH alcalino de 7.5 a 10, como se describe en las patentes estadounidenses n.º 5 648 263, 5 691 178 y 5 776 757).

[0079] También se ha demostrado que las composiciones de celulasa degradan tejidos que contienen algodón, contribuyendo a una pérdida de resistencia reducida en el tejido (patente estadounidense n.º 4 822 516), contribuyendo a la retención a utilizar composiciones de celulasa en aplicaciones detergentes comerciales. Se ha sugerido que las composiciones de celulasa que comprenden componentes de endoglucanasa presentan una pérdida de resistencia reducida para tejidos que contienen algodón en comparación con las composiciones que comprenden un sistema de celulosas completo.

[0080] También se ha demostrado que las celulosas son útiles en la degradación de biomasa de celulosa en etanol (donde la celulosa degrada celulosa en glucosa y levadura u otros microbios y además fermenta la glucosa en etanol), en el tratamiento de pulpa mecánica (Pere *et al.*, 1996), para su utilización como aditivo para alimentación animal (WO 91/04673) y en la molienda en húmedo de grano.

[0081] Se han descrito numerosas celulosas en la literatura científica, ejemplos de las cuales incluyen: de *Trichoderma reesei*: Shoemaker, S. *et al.*, *Bio/Technology*, 1:691-696, 1983, que da a conocer CBHI; Teeri, T. *et al.*, *Gene*, 51:43-52, 1987, que da a conocer CBHII. Penttilä, M. *et al.*, *Gene*, 45:253-263, 1986, que da a conocer EGI; Saloheimo, M. *et al.*, *Gene*, 63:11-22, 1988, que da a conocer EGII; Okada, M. *et al.*, *Appl. Environ. Microbiol.*, 64:555-563, 1988, que da a conocer EGIII; Saloheimo, M. *et al.*, *Eur. J. Biochem.*, 249:584-591, 1997, que da a conocer EGIV; Saloheimo, A. *et al.*, *Molecular Microbiology*, 13:219-228, 1994, que da a conocer EGV; Barnett, C. C., *et al.*, *Bio/Technology*, 9:562-567, 1991, que da a conocer BGL1, y Takashima, S. *et al.*, *J. Biochem.*, 125:728-736, 1999, que da a conocer BGL2. También se han descrito celulosas de especies distintas a *Trichoderma*, p. ej., Ooi *et al.*, 1990, que da a conocer la secuencia de ADNc que codifica la endoglucanasa F1-CMC producida por *Aspergillus aculeatus*; Kawaguchi T *et al.*, 1996, que da a conocer la clonación y la secuenciación del ADNc que codifica la beta-glucosidasa 1 de *Aspergillus aculeatus*; Sakamoto *et al.*, 1995, que da a conocer la secuencia de ADNc que codifica la endoglucanasa CMCasa-1 de *Aspergillus kawachii* IFO 4308; Saarihtti *et al.*, 1990, que da a conocer una endoglucanasa de *Erwinia carotovora*; Spilliaert R, *et al.*, 1994, que da a conocer la clonación y la secuenciación de bglA, que codifica una beta-glucanasa termoestable a partir de *Rhodothermus marinus*; y Halldorsdottir S *et al.*, 1998, que da a conocer la clonación, la secuenciación y la sobreexpresión de un gen de *Rhodothermus marinus* que codifica una celulasa termoestable de la familia 12 de glicosil hidrolasa. Sin embargo, todavía resulta necesaria la identificación y caracterización de nuevas celulosas, con propiedades mejoradas, tales como rendimiento mejorado en condiciones de estrés térmico o en presencia de surfactantes, actividad específica aumentada, patrón de escisión de sustrato alterado, y/o alto nivel de expresión *in vitro*.

[0082] El desarrollo de nuevas y mejoradas composiciones de celulasa que comprenden cantidades diversas de celulosas de tipo CBH, de tipo EG y de tipo BG es de interés para utilizarse: (1) en composiciones detergentes que presentan capacidad de limpieza mejorada, funcionan como un agente suavizante y/o mejoran el tacto de los tejidos de algodón (p. ej., "lavado a la piedra" o "biopulido"); (2) en composiciones para la degradación de pulpa de madera u otra biomasa en azúcares (p. ej., para la producción de bioetanol); y/o (3) en composiciones para pienso.

IV. Métodos de identificación de secuencias novedosas

[0083] Los marcos de lectura abiertos (ORF, por sus siglas en inglés) se analizan después de la secuenciación total o parcial del genoma de *T. reesei* o de clones de genotecas de ADNc derivadas de ARNm de *T. reesei* y se analizan además utilizando *software* de análisis de secuencias y mediante la determinación de la homología con secuencias conocidas en bases de datos (públicas/privadas).

V. Ácidos nucleicos *bgl6* y polipéptidos BGL6.

A. Ácidos nucleicos *bgl6*

[0084] Las moléculas de ácido nucleico incluyen la secuencia de codificación nativa, la secuencia de ADNc para *bgl6* presentada en el presente documento como SEQ. ID. NO:1 y homólogos de las mismas en otras especies, variantes de corte y empalme y alélicas de origen natural, fragmentos de ácido nucleico y derivados biológicamente activos (funcionales) de los mismos, tales como, variantes de secuencias de aminoácidos de la molécula nativa y secuencias que codifican las proteínas de fusión. El gen *bgl6* tiene dos supuestos codones de

inicio. Los dos codones de inicio están subrayados en la Figura 4. Las secuencias se denominan colectivamente en el presente documento “secuencias de ácido nucleico que codifican BGL6”.

[0085] El 1 de octubre de 2002 se llevó a cabo una búsqueda Basic BLASTN (<http://www.ncbi.nlm.nih.gov/BLAST>) de la base de datos de secuencias de ácido nucleico no redundantes, con la secuencia del gen *bgl6* presentada en la Figura 1 (SEQ ID NO:1), que indicó que no había secuencias que produjeran alineamientos significativos (es decir, con un valor E de 10^{-5} o menos).

[0086] Una secuencia de ácido nucleico *bgl6* de la presente invención puede ser una secuencia de ADN o de ARN, derivada de ADN genómico, ADNc, ARNm o puede sintetizarse en su totalidad o en parte. El ADN puede ser bicatenario o monocatenario y si es monocatenario puede ser la cadena codificante o la cadena no codificante (antisentido, complementaria). La secuencia de ácido nucleico puede clonarse, por ejemplo, aislando el ADN genómico de una fuente adecuada y amplificando y clonando la secuencia de interés utilizando una reacción en cadena de la polimerasa (PCR, por sus siglas en inglés). De manera alternativa, la secuencia de ácido nucleico puede sintetizarse, por completo o en parte, especialmente donde se desea proporcionar secuencias preferidas por el huésped para la expresión óptima. Por lo tanto, todo el gen estructural deseado o una parte del mismo (la parte del gen que codifica un polipéptido o proteína) puede sintetizarse utilizando codones preferidos por un huésped seleccionado.

[0087] Debido a la degeneración inherente del código genético, las secuencias de ácido nucleico distintas a la forma nativa que codifican sustancialmente la misma secuencia de aminoácidos o una funcionalmente equivalente pueden utilizarse para clonar y/o expresar secuencias de ácidos nucleicos que codifican BGL6. De esta manera, para una secuencia de ácidos nucleicos determinada que codifica BGL6, se entiende que como resultado de la degeneración del código genético, puede producirse un número de secuencias de codificación que codifican una proteína que tiene la misma secuencia de aminoácidos. Por ejemplo, el triplete CGT codifica el aminoácido arginina. La arginina es codificada de manera alternativa por CGA, CGC, CGG, AGA y AGG. Por lo tanto, se entiende que dichas sustituciones en la región de codificación entran dentro de las variantes de secuencia de ácido nucleico cubiertas. Todas y cada una de estas variantes de secuencia pueden utilizarse de la misma manera que se describe en el presente documento para la forma nativa de una secuencia de ácido nucleico que codifica BGL6.

[0088] Una “variante” de secuencia de ácido nucleico que codifica BGL6 puede codificar una “variante” de secuencia de aminoácidos BGL6 que es alterada por uno o más aminoácidos de la secuencia de polipéptidos nativa o puede truncarse por la eliminación de uno o más aminoácidos de cualquier extremo de la secuencia de polipéptidos, ambos incluidos. De manera similar, el término “forma modificada de”, en relación con BGL6, significa un derivado o forma variante de la secuencia de ácido nucleico que codifica la proteína BGL6 nativa o la secuencia de aminoácidos BGL6 nativa.

[0089] De manera similar, los polinucleótidos que se puede utilizar incluyen secuencias que codifican proteínas BGL6 nativas y variantes de corte y empalme de las mismas, secuencias complementarias a la secuencia de codificación de la proteína nativa y fragmentos novedosos de polinucleótidos que codifican BGL6. Una secuencia de ácido nucleico que codifica BGL6 puede contener una o más secuencias de intrones si es una secuencia de ADN genómico.

[0090] En un modo de realización general, una secuencia de nucleótidos que codifica BGL6 tiene al menos 70 %, preferiblemente 80 %, 85 %, 90 %, 95 %, 98 %, o más identidad de secuencia con la secuencia de codificación de *bgl6* presentada en el presente documento como SEQ ID NO:1.

[0091] En otro modo de realización, una secuencia de nucleótidos que codifica BGL6 se hibridará en condiciones de moderada a alta restricción a una secuencia de nucleótidos que codifica una proteína BGL6. En un modo de realización relacionado, una secuencia de nucleótidos que codifica BGL6 se hibridará en condiciones de moderada a alta restricción a la secuencia de nucleótidos presentada como SEQ ID NO:1.

[0092] Se entiende que algunas variantes de secuencias de ácidos nucleicos que codifican BGL6 pueden o no hibridarse de manera selectiva a la secuencia parental. A modo de ejemplo, en situaciones en las que la secuencia de codificación se ha optimizado basándose en la degeneración del código genético, puede producirse una variante de secuencia de codificación que codifica una proteína BGL6, pero no se hibrida a una secuencia de ácidos nucleicos que codifican BGL6 en condiciones de moderada a alta restricción. Esto ocurriría, por ejemplo, cuando la variante de secuencia incluye un codón diferente para cada uno de los aminoácidos codificados por el nucleótido parental.

[0093] Como entenderán de manera adicional los expertos en la materia, en algunos casos puede resultar ventajoso producir secuencias de nucleótidos que posean codones que no sean de origen natural, p. ej., inosina u otro análogo de nucleótido que no sea de origen natural. Los codones preferidos por un huésped eucariota particular pueden seleccionarse, por ejemplo, para aumentar el índice de expresión de la proteína BGL6 o para producir transcritos de ARN recombinante que tienen propiedades deseables, tales como una vida media más larga que los transcritos producidos a partir de la secuencia de origen natural. Por consiguiente, una secuencia

nativa de nucleótidos que codifica BGL6 puede crearse con el fin de alterar la secuencia de codificación para una variedad de razones, incluyendo, pero sin carácter limitativo, alteraciones que modifican la clonación, el procesamiento y/o la expresión de la proteína BGL6 por una célula.

[0094] Se prefieren particularmente las sustituciones, adiciones y deleciones de ácidos nucleicos que son silenciosas de manera que no alteran las propiedades o actividades del polinucleótido o polipéptido nativo.

[0095] Las variaciones pueden realizarse utilizando métodos conocidos en la técnica tales como mutagénesis (de sitio dirigido) mediada por oligonucleótidos y mutagénesis por PCR. La mutagénesis de sitio dirigido (Carter *et al.*, 1986; Zoller *et al.*, 1987), la mutagénesis por inserción de un casete (Wells *et al.*, 1985), la mutagénesis por restricción-selección (Wells *et al.*, 1986) u otras técnicas conocidas pueden realizarse en el ADN clonado para producir la variante de ADN que codifica polipéptido BGL6.

[0096] Sin embargo, en algunos casos puede resultar ventajoso expresar variantes de *bgl6* que carezcan de las propiedades o actividades del polipéptido BGL6 o del polinucleótido *bgl6* nativo. En tales casos, las formas mutantes o modificadas de la secuencia nativa de ácidos nucleicos que codifican BGL6 pueden generarse utilizando técnicas empleadas de manera rutinaria por los expertos en la materia.

B. Polipéptidos BGL6

[0097] En un modo de realización preferido, se describe un polipéptido BGL6 que tiene una secuencia de polipéptidos BGL6 nativa madura o completa que comprende la secuencia presentada en la Figura 2 (SEQ ID NO:2). Un polipéptido BGL6 puede ser el polipéptido BGL6 maduro, parte de una proteína de fusión o un fragmento o variante de la secuencia de polipéptido BGL6 presentada en la Figura 2 (SEQ ID NO:2).

[0098] Generalmente, un polipéptido BGL6 tiene al menos 80 % de identidad con una secuencia de aminoácidos BGL6 en toda su longitud. Son más preferibles las secuencias de polipéptidos BGL6 que comprenden una región que tiene al menos 80 %, 85 %, 90 %, 95 %, 98 % o más de identidad de secuencia con la secuencia de polipéptidos BGL6 de la Figura 2 (SEQ ID NO:2), utilizando un programa de alineamiento de secuencias, como se detalla en el presente documento.

[0099] Normalmente, una “forma modificada de” una proteína BGL6 nativa o una “variante” de proteína BGL6 tiene una secuencia derivada que contiene al menos una sustitución, adición, deleción o inserción de aminoácidos, respectivamente.

[0100] En la técnica se sabe que se pueden realizar ciertas sustituciones de aminoácidos en las secuencias de proteína sin afectar a la función de la proteína. En general, las sustituciones de aminoácidos conservadoras o las sustituciones de aminoácidos similares se toleran sin afectar a la función de la proteína. Aminoácidos similares pueden ser los que son similares en tamaño y/o propiedades de carga, por ejemplo, el aspartato y el glutamato, y la isoleucina y la valina, son ambos pares de aminoácidos similares. La similitud entre los pares de aminoácidos se ha evaluado en la técnica de numerosas maneras. Por ejemplo, Dayhoff *et al.* (1978), proporciona tablas de frecuencias para sustituciones de aminoácidos que pueden emplearse como una medida de similitud de aminoácidos. Las tablas de frecuencias de Dayhoff *et al.* se basan en comparaciones de secuencias de aminoácidos para proteínas que tienen la misma función de una variedad de fuentes evolutivamente diferentes.

[0101] Se consideran los fragmentos y variantes de la secuencia de polipéptidos BGL6 de la Figura 2 (SEQ ID NO:2). Un fragmento es una variante de polipéptido que tiene una secuencia de aminoácidos que es completamente igual como parte pero no toda la secuencia de aminoácidos de los polipéptidos descritos anteriormente. Los fragmentos pueden ser “independientes” o estar comprendidos dentro de un polipéptido más grande del cual el fragmento forma una parte o una región, más preferiblemente como una única región continua. Los fragmentos preferidos son fragmentos biológicamente activos que son aquellos fragmentos que median las actividades de los polipéptidos de la invención, incluyendo aquellos con actividad similar o actividad mejorada o con una actividad reducida. También se incluyen aquellos fragmentos que son antigénicos o inmunogénicos en un animal, en concreto un humano. En este aspecto, se incluyen (i) fragmentos de BGL6, preferiblemente al menos aproximadamente 20-100 aminoácidos en longitud, más preferiblemente aproximadamente 100-200 aminoácidos en longitud, y (ii) una composición farmacéutica que comprende BGL6. En varios modos de realización, el fragmento corresponde al dominio N-terminal de BGL6 o al dominio C-terminal de BGL6.

[0102] Los polipéptidos BGL6 también incluyen polipéptidos que varían de la secuencia de polipéptidos BGL6 de la Figura 2 (SEQ ID NO:2). Estas variantes pueden ser variantes de sustitución, de inserción o de deleción. Las variantes normalmente muestran la misma actividad biológica cualitativa que los análogos de origen natural, aunque también pueden seleccionarse variantes que tienen características modificadas como se describe de manera adicional a continuación.

[0103] Una “sustitución” resulta del reemplazo de uno o más nucleótidos o aminoácidos por nucleótidos o aminoácidos diferentes, respectivamente.

[0104] Una “inserción” o “adición” es ese cambio en una secuencia de nucleótidos o de aminoácidos que ha dado lugar a la adición de uno o más residuos de nucleótidos o de aminoácidos, respectivamente, en comparación con la secuencia de origen natural.

5 **[0105]** Una “delección” se define como un cambio en una secuencia de nucleótidos o de aminoácidos en la que uno o más residuos de nucleótidos o de aminoácidos, respectivamente, están ausentes.

[0106] Las sustituciones de aminoácidos son normalmente de residuos solos; las inserciones generalmente serán del orden de aproximadamente 1 a aproximadamente 20 aminoácidos, aunque pueden tolerarse inserciones considerablemente más grandes. Las delecciones varían de aproximadamente 1 a aproximadamente 20 residuos, aunque en algunos casos las delecciones pueden ser mucho más grandes.

10 **[0107]** Las sustituciones, delecciones, inserciones o cualquier otra combinación de las mismas puede utilizarse para llegar a un derivado final. En general, estos cambios se realizan en unos pocos aminoácidos para minimizar la alteración de la molécula. Sin embargo, pueden tolerarse cambios más grandes en determinadas circunstancias.

15 **[0108]** Las sustituciones de aminoácidos pueden ser el resultado del reemplazo de un aminoácido por otro aminoácido con propiedades estructurales y/o químicas similares, tal como el reemplazo de una isoleucina por una valina, es decir, reemplazos de aminoácidos conservadores. Las inserciones o delecciones pueden opcionalmente estar en el intervalo de 1 a 5 aminoácidos.

20 **[0109]** Las sustituciones se realizan generalmente de conformidad con “sustituciones conservadoras” conocidas. Una “sustitución conservadora” se refiere a la sustitución de un aminoácido en una clase por un aminoácido en la misma clase, donde una clase se define por propiedades de cadena lateral de aminoácidos fisicoquímicas comunes y frecuencias de sustitución altas en proteínas homólogas encontradas en la naturaleza (como se determina, p. ej., por una matriz de intercambio de frecuencias de Dayhoff estándar o una matriz BLOSUM). (Véase generalmente, Doolittle, R.F., 1986.)

25 **[0110]** Una “sustitución no conservadora” se refiere a la sustitución de un aminoácido en una clase por un aminoácido de otra clase.

30 **[0111]** Las variantes de polipéptido BGL6 normalmente muestran la misma actividad biológica cualitativa que el análogo de origen natural, aunque las variantes también se seleccionan para modificar las características del polipéptido BGL6, según sea necesario. Por ejemplo, los sitios de glicosilación, y más en concreto uno o más sitios de O-glicosilación o de N-glicosilación pueden alterarse o eliminarse. Los expertos en la materia entenderán que los cambios en los aminoácidos pueden alterar los procesos posteriores a la traducción del polipéptido BGL6, tal como cambiar el número o la posición de los sitios de glicosilación o alterar las características de anclaje de la membrana o las características de secreción u otras características de localización celular.

35 **[0112]** En la definición de polipéptidos BGL6 también se incluyen otros polipéptidos BGL6 relacionados. Por lo tanto, las secuencias de cebador de reacción en cadena de la polimerasa (PCR, por sus siglas en inglés) degeneradas o de sonda pueden utilizarse para encontrar otros polipéptidos relacionados. Las secuencias de cebador o de sonda útiles pueden diseñarse para: toda o parte de la secuencia de polipéptidos BGL6, o secuencias fuera de la región codificante. Como se conoce de manera general en la técnica, los cebadores PCR preferidos son de aproximadamente 15 a aproximadamente 35 nucleótidos de longitud, prefiriéndose de
40 aproximadamente 20 a aproximadamente 30, y pueden contener inosina según sea necesario. Las condiciones para la reacción PCR se conocen por lo general en la técnica.

45 **[0113]** También se incluyen las modificaciones covalentes de polipéptidos BGL6. Por ejemplo, se proporcionan polipéptidos BGL6 que son una proteína madura y pueden comprender aminoácidos amino o carboxilo terminal adicionales, o aminoácidos dentro del polipéptido maduro (por ejemplo, cuando la forma madura de la proteína tiene más de una cadena de polipéptidos). Estas secuencias pueden, por ejemplo, desempeñar un papel en el procesamiento de la proteína de una forma precursora a una madura, permitir el transporte de la proteína, acortar o alargar la vida media de la proteína o facilitar la manipulación de la proteína en los ensayos o en la producción. Como ejemplo, se cree que el presente polipéptido BGL6 novedoso es una proteína intracelular. Por lo tanto, con el fin de exportarse al medio extracelular una señal de secreción que se elimina posteriormente
50 puede ser deseable.

[0114] También se contemplan modificaciones dirigidas a la alteración de un sitio activo, la alteración del pH óptimo, la temperatura óptima y/o la afinidad del sustrato de la enzima BGL6.

55 **[0115]** La Figura 2 muestra la secuencia de aminoácidos prevista (SEQ ID NO:2) de un polipéptido BGL6 de ejemplo basada en la secuencia de nucleótidos proporcionada en la Figura 1. El peso molecular previsto del polipéptido BGL6 codificado es 92 kDa. Ninguna secuencia que se parezca a un péptido de señal (Nielsen, H.,

Engelbrecht, J., Brunak, S., von Heijne, G., *Protein Engineering*, 10:1-6, 1997) está presente en el amino terminal de BGL6 sugiriendo que el polipéptido BGL6 no está secretado.

[0116] Una búsqueda Basic BLASTP (<http://www.ncbi.nlm.nih.gov/BLAST>) de la base de datos de proteínas no redundantes, llevada a cabo el 1 de octubre de 2002 con la secuencia de aminoácidos BGL6 indicó un 42 % de identidad de secuencia con el número de acceso GenBank P07337 (precursor de beta-glucosidasa de *marxianus* de la variedad *Kluyveromyces marxianus*), 43 % de identidad de secuencia con el número de acceso GenBank AL355920 (precursor de beta-glucosidasa de *Schizosaccharomyces pombe*), 38 % de identidad de secuencia con el número de acceso GenBank AF329731 (beta-glucosidasa de *Volvariella volvacea*) y 38 % de identidad de secuencia con el número de acceso GenBank AJ293760 (supuesta beta-glucosidasa de *Agaricus bisporus*). Las diez secuencias con la identidad más alta pero inferior a 43 % de identidad con BGL6 se anotaron todas como beta-glucosidasas. Estas similitudes de secuencias indican que la BGL6 es un miembro de la familia 3 de glicosil hidrolasas (Henrissat, B. y Bairoch, A. (1993) *Biochem. J.* 293:781-788).

C. Anticuerpos anti-BGL6.

[0117] La presente exposición proporciona además anticuerpos anti-BGL6. Los anticuerpos pueden ser anticuerpos policlonales, monoclonales, humanizados, biespecíficos o heteroconjugados.

[0118] El experto en la materia conoce métodos para la preparación de anticuerpos policlonales. El agente inmunizante puede ser un polipéptido BGL6 o una proteína de fusión del mismo. Puede resultar útil conjugar el antígeno a una proteína que se sabe que es inmunogénica en el mamífero que se está inmunizando. El protocolo de inmunización puede determinarse por un experto en la materia basándose en protocolos estándar o en experimentación rutinaria.

[0119] De manera alternativa, los anticuerpos anti-BGL6 pueden ser anticuerpos monoclonales. Los anticuerpos monoclonales pueden ser producidos por células inmunizadas en un animal o utilizando métodos de ADN recombinante. (Véase, p. ej., Kohler *et al.*, 1975; la patente estadounidense N.º 4 816 567).

[0120] Un anticuerpo anti-BGL6 puede comprender además un anticuerpo humano o humanizado. El término "anticuerpo humanizado" se refiere a formas humanizadas de anticuerpos no humanos (p. ej., murinos) que son anticuerpos quiméricos, cadenas de inmunoglobulina o fragmentos de los mismos (tales como Fv, Fab, Fab', F(ab')₂ u otras secuencias de anticuerpos parciales de unión al antígeno) que contienen alguna parte de la secuencia derivada de un anticuerpo no humano. En la técnica se conocen métodos para la humanización de anticuerpos no humanos, como se detalla de manera adicional en Jones *et al.*, 1986; Riechmann *et al.*, 1988; y Verhoeyen *et al.*, 1988. En la técnica también se conocen métodos para la producción de anticuerpos humanos. Véase, p. ej., Jakobovits, A. *et al.*, 1995 y Jakobovits, A., 1995.

VI. Expresión de BGL6 recombinante

[0121] Los métodos se basan en la utilización de células para expresar BGL6, sin que se requiera ningún método concreto de expresión de la BGL6.

[0122] Se proporcionan células huésped que se han transducido, transformado o transfectado con un vector de expresión que comprende una secuencia de ácido nucleico que codifica BGL6. Las condiciones de cultivo, tales como la temperatura, el pH y similares, son las utilizadas anteriormente para la célula huésped parental antes de la transducción, transformación o transfección y resultarán evidentes para los expertos en la materia.

[0123] En un enfoque, se transfecta una célula de hongo filamentoso o una célula de levadura con un vector de expresión que tiene un promotor o un fragmento de promotor biológicamente activo o uno o varios (por ejemplo, una serie) de los potenciadores que funcionan en la línea de la célula huésped, operativamente unido a un segmento de ADN que codifica BGL6, de manera que la BGL6 se expresa en la línea celular.

A. Constructos de ácido nucleico/Vectores de expresión.

[0124] Se pueden incorporar fragmentos de polinucleótidos naturales o sintéticos que codifican BGL6 ("secuencias de ácido nucleico que codifican BGL6") en vectores o constructos de ácido nucleico heterólogos, capaces de la introducción y la replicación en una célula de levadura o de hongo filamentoso. Los vectores y métodos expuestos en el presente documento son adecuados para utilizarse en células huésped para la expresión de BGL6. Se puede utilizar cualquier vector siempre que sea replicable y viable en las células en las que se introduce. Los expertos en la materia conocen un gran número de vectores y promotores adecuados, y estos están disponibles comercialmente. Los vectores de clonación y expresión también se describen en Sambrook *et al.*, 1989, Ausubel FM *et al.*, 1989, y Strathern *et al.*, 1981. Se describen vectores de expresión adecuados para hongos en van den Hondel, C.A.M.J.J. *et al.* (1991) En: Bennett, J.W. y Lasure, L.L. (eds.) *More Gene Manipulations in Fungi*. Academic Press, pp. 396-428. La secuencia de ADN adecuada se puede insertar en un plásmido o vector (denominados colectivamente en el presente documento como "vectores") mediante una variedad de procedimientos. En general, la secuencia de ADN se inserta en un sitio o sitios de endonucleasa de restricción adecuado(s) mediante procedimientos estándar. Se considera que dichos procedimientos y los

procedimientos de subclonación relacionados están dentro del alcance del conocimiento de los expertos en la materia.

[0125] Se pueden producir hongos filamentosos recombinantes que comprenden la secuencia de codificación para BGL6 introduciendo un constructo de ácido nucleico heterólogo que comprende la secuencia de codificación de BGL6 en las células de una cepa seleccionada de los hongos filamentosos.

[0126] Una vez que se obtiene la forma deseada de una secuencia de ácido nucleico de *bgl6*, un homólogo, variante o fragmento de la misma, se puede modificar de diversas maneras. Cuando la secuencia implica regiones flanqueantes no codificantes, las regiones flanqueantes se pueden someter a resección, mutagénesis, etc. De esta manera, se pueden realizar transiciones, transversiones, delecciones e inserciones en la secuencia de origen natural.

[0127] Puede insertarse una secuencia de codificación de *bgl6* seleccionada en un vector adecuado de conformidad con técnicas recombinantes bien conocidas y utilizarse para transformar los hongos filamentosos capaces de la expresión de BGL6. Debido a la degeneración inherente del código genético, se pueden utilizar otras secuencias de ácido nucleico que codifican una secuencia de aminoácidos sustancialmente igual o funcionalmente equivalente para clonar y expresar BGL6. Por lo tanto, se entiende que dichas sustituciones en la región de codificación están dentro de las variantes de secuencias cubiertas por la presente invención. Se pueden utilizar todas y cada una de estas variantes de secuencia de la misma manera que se describe en el presente documento para una secuencia de ácido nucleico parental que codifica BGL6.

[0128] Los constructos de ácido nucleico recombinante que comprenden una o más de las secuencias de ácido nucleico que codifican BGL6 como se describe anteriormente están incluidos. Los constructos comprenden un vector, tal como un vector viral o plásmido, en el que se ha insertado una secuencia de la invención, en una orientación directa o inversa.

[0129] Los constructos de ácido nucleico heterólogos pueden incluir la secuencia de codificación para *bgl6*, o una variante, fragmento o variante de corte y empalme de la misma: (i) de manera aislada; (ii) en combinación con secuencias de codificación adicionales; tales como secuencias de codificación de proteína de fusión o de péptido señal, en las que la secuencia que codifica *bgl6* es la secuencia de codificación 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 de codificación en un huésped adecuado; y/o (iv) en un vector o medio huésped en el que la secuencia que codifica *bgl6* es un gen heterólogo.

[0130] En un aspecto, se emplea un constructo de ácido nucleico heterólogo para transferir una secuencia de ácido nucleico que codifica BGL6 en una célula *in vitro*, con líneas de levadura y de hongos filamentosos establecidas preferidas. Para la producción a largo plazo y de alto rendimiento de BGL6, se prefiere la expresión estable. Se deduce que se puede utilizar cualquier método efectivo para generar transformantes estables.

[0131] Los vectores adecuados están normalmente equipados con una secuencia de ácido nucleico que codifica un marcador seleccionable, sitios de inserción y elementos de control adecuados, tales como secuencias de promotor y de terminación. El vector puede comprender secuencias reguladoras, que incluyen, por ejemplo, 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 de codificación en células huésped (y/o en un medio de célula huésped o vector en el que normalmente no se expresa una secuencia de codificación de antígeno de proteína soluble modificada), operativamente unidas a la secuencia de codificación. Los expertos en la materia conocen un gran número de vectores y promotores adecuados, muchos de los cuales están disponibles comercialmente y/o se describen en Sambrook, *et al.*, (*supra*).

[0132] Promotores de ejemplo incluyen tanto promotores constitutivos como promotores inducibles, ejemplos de los cuales incluyen un promotor CMV, un promotor SV40 temprano, un promotor RSV, un promotor EF-1 α , un promotor que contiene el elemento de respuesta tet (TRE, por sus siglas en inglés) en el sistema tet-on o tet-off como se describe (ClonTech y BASF), el promotor beta actina y el promotor metalotionina que se pueden regular por incremento mediante la adición de ciertas sales de metales. Una secuencia promotora es una secuencia de ADN que es reconocida por el hongo filamentoso concreto para propósitos de expresión. Está operativamente unida a una secuencia de ADN que codifica un polipéptido de BGL6. Dicha unión comprende el posicionamiento del promotor con respecto al codón de iniciación de la secuencia de ADN que codifica el polipéptido de BGL6 en los vectores de expresión expuestos. La secuencia del promotor contiene la secuencia de control de la transcripción y la traducción que media en la expresión del polipéptido de BGL6. 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 proteinasa aspártica de *A. niger* o *Rhizomucor miehei*; los genes que codifican *cbh1*, *cbh2*, *egl1*, *egl2*, u otras celulasas de *T. reesei*.

[0133] La elección del marcador seleccionable adecuado dependerá de la célula huésped, y en la técnica se conocen bien los marcadores adecuados para diferentes huéspedes. Los genes marcadores seleccionables típicos incluyen *argB* de *A. nidulans* o *T. reesei*, *amdS* de *A. nidulans*, *pyr4* de *Neurospora crassa* o *T. reesei*, *pyrG* de *Aspergillus niger* o *A. nidulans*. Ejemplos adicionales de marcadores seleccionables incluyen, pero sin carácter limitativo, *trpC*, *trp1*, *oliC31*, *niaD* o *leu2*, que se incluyen en constructos de ácido nucleico heterólogo utilizados para transformar una cepa mutante tal como *trp*-, *pyr*-, *leu*- y similares.

[0134] Dichos marcadores seleccionables confieren a los transformantes la capacidad de utilizar un metabolito que normalmente no es metabolizado por los hongos filamentosos. Por ejemplo, el gen *amdS* de *T. reesei* que codifica la enzima acetamidasa que permite que las células transformantes crezcan en acetamida como una fuente de nitrógeno. El marcador seleccionable (por ejemplo, *pyrG*) puede restaurar la capacidad de una cepa mutante auxotrófica para crecer en un medio mínimo selectivo o el marcador seleccionable (por ejemplo, *oliC31*) puede conferir a los transformantes la capacidad de crecer en presencia de un antibiótico o fármaco inhibidor.

[0135] La secuencia de codificación del marcador seleccionable se clona en cualquier plásmido adecuado utilizando métodos generalmente empleados en la técnica. Ejemplos de plásmidos incluyen pUC18, pBR322 y pUC100.

[0136] La práctica de la presente descripción empleará, a menos que se indique lo contrario, técnicas convencionales de biología molecular, microbiología, ADN recombinante e inmunología, que están dentro de la técnica. Dichas técnicas se explican con detalle en la literatura. Véase, por ejemplo, Sambrook *et al.*, 1989; Freshney, 1987; Ausubel, *et al.*, 1993; y Coligan *et al.*, 1991.

B. Células huésped y condiciones de cultivo para la producción de BGL6 mejorada

(i) Hongos filamentosos

[0137] Se proporcionan hongos filamentosos que comprenden células que han sido modificadas, seleccionadas y cultivadas de manera efectiva para dar lugar a la producción o expresión de BGL6 mejorada con respecto a los correspondientes hongos parentales no transformados.

[0138] Ejemplos de especies de hongos filamentosos parentales que se pueden tratar y/o modificar para la expresión de BGL6 mejorada incluyen, pero sin carácter limitativo, *Trichoderma*, por ejemplo, *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.*

[0139] Las células que expresan BGL6 se cultivan en condiciones normalmente empleadas para cultivar la línea fúngica parental. Generalmente, las células se cultivan en un medio estándar que contiene sales fisiológicas y nutrientes, tal como se describe en Pourquie, J. *et al.*, *Biochemistry and Genetics of Cellulose Degradation*, eds. Aubert, J. P. *et al.*, Academic Press, pp. 71-86, 1988 e Ilmen, M. *et al.*, *Appl. Environ. Microbiol.* 63:1298-1306, 1997. Las condiciones de cultivo también son estándar, por ejemplo, los cultivos se incuban a 28 °C en agitadores de cultivos o fermentadores hasta que se consiguen niveles de expresión de BGL6 deseados.

[0140] Las condiciones de cultivo preferidas para un hongo filamentoso determinado se pueden encontrar en la literatura científica y/o de la fuente de los hongos, tal como la American Type Culture Collection (ATCC; "www.atcc.org/"). Después de establecer el crecimiento fúngico, las células se exponen a condiciones efectivas para causar o permitir la sobreexpresión de BGL6.

[0141] En casos en los que una secuencia que codifica BGL6 está bajo el control de un promotor inducible, el agente inductor, por ejemplo, un azúcar, una sal de metal o antibióticos, se añade al medio en una concentración efectiva para inducir la expresión de BGL6 a alto nivel.

(ii) Levadura

[0142] También se contempla la utilización de levadura como célula huésped para la producción de BGL6. Varios otros genes que codifican enzimas hidrolíticas se han expresado en varias cepas de la levadura *S. cerevisiae*. Estos incluyen secuencias que codifican dos endoglucanasas (Penttilä *et al.*, 1987), dos celobiohidrolasas (Penttilä *et al.*, 1988) y una beta-glucosidasa de *Trichoderma reesei* (Cummings y Fowler, 1996), una xilanasa de *Aureobasidium pullulans* (Li y Ljungdahl, 1996), una alfa-amilasa de trigo (Rothstein *et al.*, 1987), etc. Además, un casete del gen de celulasa que codifica la endo-[beta]-1,4-glucanasa (END1) de *Butyrivibrio fibrisolvens*, celobiohidrolasa (CBH1) de *Phanerochaete chrysosporium*, la celodextrinasa (CEL1) de *Ruminococcus flavefaciens* y la celobiasa de *Endomyces fibrilizer* (Bgl1) se expresó satisfactoriamente en una cepa de laboratorio de *S. cerevisiae* (Van Rensburg *et al.*, 1998).

C. Introducción de una secuencia de ácido nucleico que codifica BGL6 en células huésped.

[0143] Se proporcionan células y composiciones celulares que han sido modificadas genéticamente para comprender una secuencia de ácido nucleico que codifica BGL6 proporcionada de manera exógena. Se puede

modificar genéticamente una célula parental o línea celular (es decir, transducirse, transformarse o transfectarse) con un vector de clonación o un vector de expresión. El vector puede estar, por ejemplo, en forma de un plásmido, una partícula viral, un fago, etc., como se describe de manera adicional anteriormente.

[0144] Pueden emplearse varios métodos para entregar un vector de expresión en células *in vitro*. Después de construir un vector adecuado, este se utiliza para transformar cepas de hongos o levadura. Los expertos en la materia conocen los métodos generales para la introducción de ácidos nucleicos en células para la expresión de secuencias de ácidos nucleicos heterólogas. Dichos métodos incluyen, pero sin carácter limitativo, electroporación; microinyección nuclear o microinyección directa en células individuales; fusión de protoplastos bacteriana con células intactas; utilización de policones, p. ej., polibreno o poliornitina; fusión de membrana con liposomas, transfección mediada por lipofección o lipofectamina; bombardeo de alta velocidad con microproyectiles recubiertos de ADN; incubación con precipitado de ADN-fosfato de calcio; transfección mediada por DEAE-Dextrano; infección con ácidos nucleicos virales modificados; y similares.

[0145] Métodos preferidos para la introducción de un constructo de ácido nucleico heterólogo (vector de expresión) en hongos filamentosos (por ejemplo, *T. reesei*) incluyen, pero sin carácter limitativo, la utilización de una pistola de genes o partículas, permeabilización de las paredes celulares de hongos filamentosos antes del proceso de transformación (por ejemplo, mediante la utilización de altas concentraciones de álcali, por ejemplo, 0,05 M a 0,4 M de CaCl_2 o acetato de litio), fusión de protoplastos o transformación mediada por agrobacterium. Se describe un método de ejemplo para la transformación de hongos filamentosos mediante el tratamiento de protoplastos o esferoplastos con polietilenglicol y CaCl_2 en Campbell, E.I. *et al.*, *Curr. Genet.* 16:53-56, 1989 y Penttilä, M. *et al.*, *Gene*, 63:11-22, 1988.

[0146] Además, los constructos de ácido nucleico heterólogos que comprenden una secuencia de ácido nucleico que codifica BGL6 se pueden transcribir *in vitro*, y el ARN resultante se puede introducir en la célula huésped mediante métodos conocidos, por ejemplo, mediante inyección.

[0147] Después de la introducción de un constructo de ácido nucleico heterólogo que comprende la secuencia de codificación para *bgl6*, las células modificadas genéticamente se pueden cultivar en medios de nutrientes convencionales modificados según convenga para activar los promotores, seleccionar los transformantes o aumentar la expresión de una secuencia de ácido nucleico que codifica BGL6. Las condiciones del cultivo, tales como la temperatura, el pH y similares, son las utilizadas anteriormente para la célula huésped seleccionada para la expresión, y resultarán evidente para los expertos en la materia.

[0148] Generalmente, se considera que la progenie de las células en las que se han introducido dichos constructos de ácidos nucleicos heterólogos comprende la secuencia de ácidos nucleicos que codifica BGL6 encontrada en el constructo de ácido nucleico heterólogo.

[0149] Se incluyen, además, transformantes nuevos y útiles de hongos filamentosos tales como *Trichoderma reesei* para utilizarse en la producción de composiciones de celulasa fúngicas. La invención incluye transformantes de hongos filamentosos, especialmente hongos que comprenden la secuencia que codifica *bgl6*, que comprenden una forma modificada de la secuencia que codifica *bgl6* o la delección de la secuencia que codifica la *bgl6*.

[0150] En general, se pueden diferenciar los transformantes estables de hongos filamentosos de los transformantes inestables por su índice de crecimiento más rápido y la formación de colonias circulares con un contorno liso en lugar de irregular en el medio de cultivo sólido. De manera adicional, en algunos casos, se puede realizar una prueba de estabilidad adicional mediante el crecimiento de los transformantes en un medio sólido no selectivo, la recogida de las esporas de este medio de cultivo y la determinación del porcentaje de estas esporas que germinarán posteriormente y crecerán en el medio selectivo.

VII. Análisis para secuencias codificantes de ácido nucleico de BGL6 y/o expresión de proteínas.

[0151] Con el fin de evaluar la expresión de BGL6 mediante una línea celular que ha sido transformada con un constructo de ácido nucleico que codifica BGL6, se pueden realizar ensayos al nivel de la proteína, al nivel del ARN o mediante la utilización de bioensayos funcionales concretos con respecto a la actividad y/o la producción de glucosidasa.

[0152] En una aplicación de ejemplo de las secuencias de proteínas y ácido nucleico de *bgl6* descritas en el presente documento, se diseña una cepa de hongos filamentosos modificada genéticamente, por ejemplo, *Trichoderma reesei*, para producir una mayor cantidad de BGL6. Dichos hongos filamentosos genéticamente modificados serían útiles para producir un producto de celulasa con mayor capacidad celulolítica incrementada. En un enfoque, esto se consigue introduciendo la secuencia de codificación para *bgl6* en un huésped adecuado, por ejemplo, un hongo filamentoso, tal como *Trichoderma reesei*.

[0153] Por consiguiente, se incluyen métodos para la expresión de BGL6 en un hongo filamentoso u otro huésped adecuado introduciendo un vector de expresión que contiene la secuencia de ADN que codifica BGL6 en las células del hongo filamentoso u otro huésped adecuado.

[0154] En otro aspecto, se incluyen métodos para la modificación de la expresión de BGL6 en un hongo filamentoso u otro huésped adecuado. Dicha modificación incluye una reducción o eliminación de la expresión, o la expresión de una forma alterada de BGL6. Una forma alterada de BGL6 puede tener una secuencia de aminoácidos alterada o una secuencia de ácidos nucleicos alterada.

[0155] En general, los ensayos empleados para analizar la expresión de BGL6 incluyen, la transferencia Northern, la transferencia puntual (análisis de ADN o ARN), RT-PCR (reacción en cadena de la polimerasa con transcriptasa inversa) o hibridación *in situ*, utilizando una sonda debidamente marcada (basada en la secuencia de codificación de ácido nucleico) y la autoradiografía y la transferencia Southern convencional.

[0156] Además, la producción y/o la expresión de BGL6 puede medirse en una muestra directamente, por ejemplo, mediante ensayos para la actividad, expresión y/o producción de glucosidasa. Dichos ensayos se describen, por ejemplo, en Chen *et al.* (1992), Herr *et al.* (1978) y la patente estadounidense 6 184 018 (Li *et al.*; 2001). La capacidad de BGL6 para hidrolizar sustratos aislados solubles e insolubles puede medirse utilizando ensayos descritos en Suurnakki *et al.* (2000) y Ortega *et al.* (2001). Los sustratos útiles para los ensayos de las actividades de celobiohidrolasa, endoglucanasa o β -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 glicósido. Los tres últimos son particularmente útiles para ensayar β -glucosidasas. Los ensayos de β -glucosidasa se conocen bien en la técnica. Véase Cummings y Fowler (1996).

[0157] Además, la expresión de proteínas se puede evaluar mediante métodos inmunológicos, tales como tinción inmunohistoquímica de células, secciones de tejido o inmunoensayo del medio de cultivo de tejidos, por ejemplo, mediante transferencia Western o ELISA. Dichos inmunoensayos se pueden utilizar para evaluar de manera cualitativa y cuantitativa la expresión de BGL6. Los expertos en la materia conocen los detalles de tales métodos y muchos reactivos para practicar dichos métodos están disponibles comercialmente.

[0158] Se puede utilizar una forma purificada de BGL6 para producir anticuerpos monoclonales o policlonales específicos de la proteína expresada para utilizarse en diferentes inmunoensayos. (Véase, por ejemplo, Hu *et al.*, 1991). Ejemplos de ensayos incluyen ELISA, inmunoensayos competitivos, radioinmunoensayos, transferencia Western, ensayos de inmunofluorescencia indirecta y similares. En general, se pueden utilizar anticuerpos y/o kits disponibles comercialmente para el inmunoensayo cuantitativo del nivel de expresión de las proteínas de glucosidasa.

VIII. Aislamiento y purificación de proteína BGL6 recombinante.

[0159] En general, una proteína BGL6 producida en un cultivo celular se secreta en el medio y se puede purificar o aislar, por ejemplo, eliminando los componentes no deseados del medio de cultivo celular. No obstante, en algunos casos, se puede producir una proteína BGL6 en una forma celular que requiere la recuperación de un lisado celular. En tales casos, la proteína BGL6 se purifica de las células en las que se produjo utilizando técnicas empleadas de forma rutinaria por los expertos en la materia. Los ejemplos incluyen, pero sin carácter limitativo, cromatografía de afinidad (Tilbeurgh *et al.*, 1984), métodos cromatográficos de intercambio iónico (Goyal *et al.*, 1991; Fliess *et al.*, 1983; Bhikhabhai *et al.*, 1984; Ellouz *et al.*, 1987), incluyendo intercambio iónico utilizando materiales con alto poder de resolución (Medve *et al.*, 1998), cromatografía de interacción hidrofoba (Tomaz y Queiroz, 1999), y separación de dos fases (Brumbauer, *et al.*, 1999).

[0160] Normalmente, la proteína BGL6 se fracciona para segregar proteínas que tienen propiedades seleccionadas, tales como afinidad de unión a agentes de unión concretos, por ejemplo, anticuerpos o receptores; o que tienen un intervalo de peso molecular seleccionado, o un intervalo de puntos isoeléctricos.

[0161] Una vez se consigue la expresión de una proteína BGL6 determinada, la proteína BGL6 producida de esta manera se purifica a partir de las células o cultivo celular. Ejemplos de procedimientos adecuados para tal purificación incluyen los siguientes: cromatografía en columna de afinidad de anticuerpos, cromatografía de intercambio iónico; precipitación con etanol; HPLC de fase inversa; cromatografía sobre sílice o sobre una resina de intercambio catiónico tal como DEAE; cromatografía de exclusión; SDS-PAGE; precipitación con sulfato de amonio; y filtración en gel utilizando, por ejemplo, Sephadex G-75. Se pueden emplear diferentes métodos de purificación de proteínas y tales métodos se conocen en la técnica y se describen, por ejemplo, en Deutscher, 1990; Scopes, 1982. La(s) etapa(s) de purificación seleccionada(s) dependerá(n), por ejemplo, de la naturaleza del proceso de producción utilizado y de la proteína concreta producida.

IX. Utilidad de *bgl6* y BGL6

[0162] Se puede entender que el nucleótido *bgl6*, la proteína BGL6 y las composiciones que comprenden actividad de proteína BGL6 resultan útiles en una amplia variedad de aplicaciones, algunas de las cuales se describen a continuación.

[0163] Las nuevas y mejoradas composiciones de celulasa que comprenden cantidades distintas de celulasas de tipo BG, de tipo EG y de tipo CBH son útiles en composiciones detergentes que muestran capacidad de limpieza mejorada, funcionan como un agente suavizante y/o mejoran el tacto de los tejidos de algodón (por ejemplo, "lavado a la piedra" o "biopulido"), en las composiciones para degradar pulpa de madera en azúcares (por ejemplo, para la producción de bioetanol), y/o en composiciones para pienso. El aislamiento y caracterización de la celulasa de cada tipo proporciona la capacidad para controlar los aspectos de tales composiciones.

[0164] En un enfoque preferido, la celulasa de la invención es útil en composiciones detergentes o en el tratamiento de tejidos para mejorar el tacto y la apariencia.

[0165] Las β -glucosidasas inventivas pueden utilizarse en una variedad de aplicaciones diferentes. Por ejemplo, la β -glucosidasa puede añadirse a las uvas durante la producción del vino para mejorar el aroma potencial del producto de vino acabado. Otra aplicación adicional puede ser la utilización de la β -glucosidasa en la fruta para mejorar el aroma de la misma. De manera alternativa, el producto de fermentación recombinante aislado que contiene la β -glucosidasa mejorada puede utilizarse directamente en los aditivos alimenticios o en la elaboración del vino para mejorar el sabor o el aroma.

[0166] Puesto que la velocidad de hidrólisis de los productos celulósicos se puede aumentar utilizando un transformante que tenga al menos una copia adicional del gen *bgl6* insertado en el genoma, los productos que contienen celulosa o heteroglicanos se pueden degradar a un ritmo más rápido y en mayor medida. Los productos fabricados con celulosa, tales como papel, algodón, pañales de celulosa, y similares se pueden degradar de manera más eficiente en un vertedero. De esta manera, el producto de la fermentación obtenible de los transformantes o los transformantes solos se pueden utilizar en composiciones para ayudar a degradar mediante licuefacción una variedad de productos de celulosa que se añaden a los vertederos llenos.

[0167] La sacarificación y fermentación por separado es un proceso por el que la celulosa presente en la biomasa, por ejemplo, rastrojos de maíz, se convierte en glucosa y, posteriormente, las cepas de levadura convierten la glucosa en etanol. La sacarificación y fermentación simultánea es un proceso por el que la celulosa presente en la biomasa, por ejemplo, rastrojos de maíz, se convierte en glucosa y, al mismo tiempo y en el mismo reactor, las cepas de levadura convierten la glucosa en etanol. De esta manera, en otro enfoque preferido, la celulasa de tipo glucosidasa de la invención resulta útil en la degradación de la biomasa en etanol. La producción de etanol a partir de las fuentes de celulosa fácilmente disponibles proporciona una fuente de combustible renovable estable.

[0168] Las materias primas a base de celulosa se componen de residuos agrícolas, pastos y maderas y otra biomasa de bajo valor, tal como residuos municipales (por ejemplo, papel reciclado, hierba cortada, etc.). El etanol se puede producir a partir de la fermentación de cualquiera de estas materias primas celulósicas. Sin embargo, la celulosa debe convertirse primero en azúcares antes de que pueda haber una conversión en etanol.

[0169] Se puede utilizar una gran variedad de materias primas con la β -glucosidasa inventiva y la seleccionada para utilizarse puede depender de la región en la que se realiza la conversión. Por ejemplo, en el medio oeste de Estados Unidos pueden predominar residuos agrícolas tales como paja de trigo, rastrojo de maíz y bagazo, mientras que en California puede predominar la paja de arroz. Sin embargo, debe entenderse que cualquier biomasa celulósica disponible puede utilizarse en cualquier región.

[0170] Una composición de celulasa que contiene una cantidad mejorada de β -glucosidasa resulta útil en la producción de etanol. El etanol de este proceso también se puede utilizar como un potenciador de octanaje o directamente como combustible en lugar de la gasolina, lo cual es ventajoso puesto que el etanol como fuente de combustible es más ecológico que los productos derivados del petróleo. Se sabe que la utilización de etanol mejorará la calidad del aire y posiblemente reducirá los niveles locales de ozono y el smog. Además, la utilización de etanol en lugar de gasolina puede ser de importancia estratégica en la amortiguación del impacto de los cambios repentinos en la energía no renovable y los suministros petroquímicos.

[0171] El etanol se puede producir a través de 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 importante encontrado en este proceso es la falta de β -glucosidasa en el sistema para convertir celobiosa en glucosa. Se sabe que la celobiosa actúa como un inhibidor de celobiohidrolasas y endoglucanasas y, de esta manera, reduce el índice de hidrólisis para todo el sistema de celulasas. Por lo tanto, la utilización de actividad de β -glucosidasa aumentada para la conversión rápida de celobiosa en glucosa mejoraría considerablemente la producción de etanol.

[0172] De esta manera, la β -glucosidasa inventiva resulta útil en la hidrólisis de celulosa en sus componentes de azúcar. En un modo de realización, se añade la β -glucosidasa a la biomasa antes de la adición de un organismo

fermentador. En un segundo modo de realización, se añade la β -glucosidasa a la biomasa al mismo tiempo que un organismo fermentador. De manera opcional, puede haber otros componentes de celulasa presentes en cualquier modo de realización.

[0173] En otro modo de realización, la materia prima celulósica se puede pretratar. El pretratamiento puede ser mediante temperatura elevada y la adición de ácido diluido, ácido concentrado o una solución de álcali diluido. La solución de pretratamiento se añade durante un tiempo suficiente para hidrolizar al menos parcialmente los componentes de la hemicelulosa y después neutralizarlos.

[0174] En un enfoque alternativo, se prefiere una composición de celulasa que no contiene o es deficiente en β -glucosidasa. La delección del gen de β -glucosidasa sería particularmente útil en la preparación de composiciones de celulasa para su utilización en detergentes. De manera adicional, tales composiciones son útiles para la producción de celobiosa y otros celooligosacáridos. La delección del gen de *bgl6* de las cepas de *T. reesei* sería particularmente útil en la preparación de composiciones de celulasa para su utilización en los detergentes y en el aislamiento de celobiosa. Las enzimas de celulasa se han utilizado en una variedad de composiciones detergentes para limpiar ropa de manera enzimática. Sin embargo, en esta técnica se sabe que la utilización de enzimas de celulasa puede impartir degradación de las fibras de celulosa en la ropa. Una posibilidad para reducir el efecto de degradación es producir un detergente que no contenga β -glucosidasa. De esta manera, la delección de esta proteína haría que el sistema de celulasa inhibiera los otros componentes mediante acumulación de celobiosa. Los microorganismos modificados son particularmente adecuados para la preparación de tales composiciones porque el gen de *bgl6* puede deleccionarse dejando los componentes de EG y CBH restantes dando lugar a una limpieza mejorada y beneficios suavizantes en la composición sin efectos degradativos.

[0175] Las composiciones detergentes pueden emplear, además de la composición de celulasa (independientemente del contenido de β -glucosidasa, es decir sin β -glucosidasa, sustancialmente sin β -glucosidasa, o β -glucosidasa mejorada), un surfactante, incluyendo surfactantes aniónicos, no iónicos y anfólicos, una hidrolasa, agentes mejoradores, agentes blanqueadores, agentes azulantes y tintes fluorescentes, inhibidores del apelmazamiento, solubilizantes, surfactantes catiónicos y similares. Todos estos componentes se conocen en la técnica de los detergentes. La composición de celulasa como se describe anteriormente puede añadirse a la composición detergente ya sea en un diluyente líquido, en gránulos, en emulsiones, en geles, en pastas y similares. El experto en la materia conoce bien dichas formas. Cuando se emplea una composición detergente sólida, la composición de celulasa se formula preferiblemente en forma de gránulos. Preferiblemente, los gránulos se pueden formular de manera que contengan un agente protector de celulasa. Para una exposición más completa, véase la patente estadounidense número 6 162 782 titulada "Detergent compositions containing cellulase compositions deficient in CBH I type components".

[0176] En otro modo de realización adicional, las composiciones detergentes también pueden contener niveles mejorados de beta-glucosidasa o beta-glucosidasa alterada. En este sentido, depende realmente del tipo de producto que se desea utilizar en las composiciones detergentes para proporcionar el efecto apropiado.

[0177] Preferiblemente, las composiciones de celulasa se emplean desde aproximadamente 0,00005 por ciento en peso hasta aproximadamente 5 por ciento en peso en relación con la composición detergente total. Más preferiblemente, las composiciones de celulasa se emplean desde aproximadamente 0,0002 por ciento en peso hasta aproximadamente 2 por ciento en peso en relación con la composición detergente total.

[0178] La delección del gen de *bgl6* también proporcionaría la acumulación de celobiosa en el sistema de celulasa, que se puede purificar a partir de la misma. En este sentido, la presente exposición presenta la posibilidad de aislar celobiosa a partir de microorganismos de una manera fácil y efectiva.

[0179] Partes de la secuencia de ácido nucleico de *bgl6* que son capaces de unión a celulosa se pueden utilizar para generar proteínas de superficie quimérica bacteriana, permitiendo la inmovilización de toda la célula sobre filtros de celulosa u otros soportes sólidos fibrosos como se describe en Lehtio *et al.*, 2001.

[0180] Además, la secuencia de ácidos nucleicos de *bgl6* resulta útil en la identificación y caracterización de secuencias de ácido nucleico relacionadas. Varias técnicas útiles para determinar (predecir o confirmar) la función de los genes o productos génicos relacionados incluyen, pero sin carácter limitativo, (A) análisis de ADN/ARN, tal como (1) sobreexpresión, expresión ectópica y expresión en otras especies; (2) bloqueo de genes (genética inversa, bloqueo dirigido, silenciamiento génico inducido por virus (VIGS, por sus siglas en inglés, véase Baulcombe, 1999); (3) análisis del estado de metilación del gen, especialmente las regiones reguladoras flanqueantes; y (4) hibridación *in situ*; (B) análisis de producto génico, tal como (1) expresión de proteína recombinante; (2) producción de antisueros, (3) inmunofocalización; (4) ensayos bioquímicos para actividad catalítica u otra actividad; (5) estado de fosforilación; y (6) interacción con otras proteínas a través de análisis de dos híbridos de levadura; (C) análisis de sistema, tal como la colocación de un gen o producto génico en el interior de un sistema bioquímico o de transducción de señales concreto basado en su fenotipo de sobreexpresión o mediante homología de secuencia con genes relacionados; y (D) otros análisis que también pueden llevarse a cabo para determinar o confirmar la participación del gen aislado y su producto en un sistema metabólico o de transducción de señales concreto, y ayudar a determinar la función del gen.

[0181] Las endoglucanasas y las beta-glucosidasas pueden ser responsables de la producción de disacáridos, tales como sofrosa, a partir de celooligosacáridos y glucosa mediante reacciones de transglicosilación. Se sabe que la sofrosa es un inductor de la expresión del gen de celulasa muy potente (Ilmen, M. *et al.*, 1997, *Appl. Environ. Microbiol.* 63:1298-1306 y referencias en el mismo). De esta manera, las EG y las BGL pueden desempeñar un papel importante en el proceso de inducción de la expresión del gen de celulasa. La sobreexpresión de ciertas EG o BGL en una cepa fúngica puede llevar a una mayor productividad total de celulasa por esa cepa.

A. Homología con secuencias conocidas

[0182] La función de una secuencia de ácido nucleico que codifica BGL6 relacionada puede determinarse mediante homología con genes conocidos que tienen una función particular. Por ejemplo, una comparación de la secuencia de codificación de una molécula de ácido nucleico identificada con bases de datos de secuencias de ácidos nucleicos públicas se utiliza para confirmar la función por homología con genes conocidos o por extensión de la secuencia de ácido nucleico identificada.

[0183] El término “% de homología” se utiliza de manera indistinta en el presente documento con el término “% de identidad” en el presente documento y se refiere al nivel de identidad de secuencias de aminoácidos o ácidos nucleicos entre la secuencia de ácidos nucleicos que codifica BGL6 o la secuencia de aminoácidos de BGL6, cuando se alinean utilizando un programa de alineamiento de secuencias.

[0184] Por ejemplo, como se utiliza en el presente documento, un 80 % de homología significa lo mismo que un 80 % de identidad de secuencia determinado por un algoritmo definido y, por consiguiente, un homólogo de una secuencia dada tiene más de un 80 % de identidad de secuencia en una longitud de la secuencia dada. Niveles de identidad de secuencia de ejemplo incluyen, pero sin carácter limitativo, 80 %, 85 %, 90 %, 95 %, 98 % o más de identidad de secuencia con una secuencia dada, por ejemplo, la secuencia codificante para *bgl6*, como se describe en el presente documento.

[0185] Los programas informáticos de ejemplo que pueden utilizarse para determina la identidad entre dos secuencias incluyen, pero sin carácter limitativo, la serie de programas BLAST, p. ej., BLASTN, BLASTX, y TBLASTX, BLASTP y TBLASTN, a disposición del público en Internet en <http://www.ncbi.nlm.nih.gov/BLAST/>. Véase también Altschul, *et al.*, 1990 y Altschul, *et al.*, 1997.

[0186] Las búsquedas de secuencias se llevan a cabo normalmente utilizando el programa BLASTN cuando se evalúa una secuencia de ácido nucleico dada en relación con secuencias de ácido nucleico de las secuencias de ADN de GenBank y otras bases de datos públicas. Se prefiere el programa BLASTX para buscar secuencias de ácidos nucleicos que se han traducido en todos los marcos de lectura frente a las secuencias de aminoácidos de las secuencias de proteínas de GenBank y otras bases de datos públicas. Tanto BLASTN como BLASTX funcionan utilizando parámetros por defecto de una penalización por hueco abierto 11.0 y una penalización por hueco extendido de 1.0 y utilizan la matriz BLOSUM-62. (Véase, p.ej., Altschul, *et al.*, 1997.)

[0187] Un alineamiento preferido de secuencias seleccionadas con el fin de determinar el “% de identidad” entre dos o más secuencias se lleva a cabo utilizando, por ejemplo, el programa CLUSTAL-W en MacVector versión 6.5, que funciona con parámetros por defecto, incluyendo una penalización por hueco abierto de 10.0, una penalización por hueco extendido de 0.1 y una matriz de similitud BLOSUM 30.

[0188] En un enfoque de ejemplo, la extensión de la secuencia de un ácido nucleico que codifica *bgl6* puede llevarse a cabo utilizando procedimientos de extensión de cebador convencionales como se describe en Sambrook *et al.*, *supra*, para detectar precursores de *bgl6* e intermediarios de procesamiento de ARNm que puede que no se hayan transcrito de manera inversa en ADNc y/o para identificar ORF que codifican una proteína completa.

[0189] Otro aspecto adicional incluye la secuencia de nucleótidos completa o parcial de la secuencia de ácidos nucleicos de *bgl6* para utilizarse como una sonda. Dicha sonda puede utilizarse para identificar y clonar secuencias de ácidos nucleicos homólogas a partir de organismos relacionados.

[0190] El rastreo de una genoteca genómica o de ADNc con la sonda seleccionada puede llevarse a cabo utilizando procedimientos estándar tales como los descritos en Sambrook *et al.*, (1989). Las condiciones de hibridación, incluyendo la restricción moderada y la restricción alta, se proporcionan en Sambrook, *et al.*, *supra*.

[0191] Las sondas o partes de las mismas también se pueden emplear en técnicas de PCR para generar un grupo de secuencias para la identificación de secuencias de *bgl6* estrechamente relacionadas. Cuando las secuencias de *bgl6* se destinan a la utilización como sondas, se puede utilizar una parte concreta de una secuencia de codificación de BGL6, por ejemplo, una parte altamente conservada de la secuencia de codificación.

[0192] Por ejemplo, se puede utilizar una secuencia de nucleótidos de *bgl6* como sonda de hibridación para una genoteca de ADNc con el fin de aislar genes, por ejemplo, los que codifican las variantes de BGL6 de origen

natural a partir de otras especies fúngicas, bacterianas o vegetales, que tienen un nivel deseado de identidad de secuencia con la secuencia de nucleótidos de *bgl6* 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 dos híbridos

- 5 [0193] Las proteínas identificadas pueden utilizarse en el sistema de dos híbridos de levadura para "capturar" proteínas de unión a proteínas que son supuestas proteínas de sistema de transducción de señales. El sistema de dos híbridos de levadura se describe en Fields y Song, *Nature* 340:245-246 (1989). En resumen, en un sistema de dos híbridos, una fusión de un dominio de unión al ADN-*bgl6* (p.ej., fusión de GAL4-*bgl6*) se construye y se transfecta en células de levadura. Puede utilizarse el gen *bgl6* entero, o subregiones del gen *bgl6*.
10 Se cotransfecta un segundo constructo que contiene la genoteca de ligandos potenciales fusionados al dominio de activación del ADN. Los cotransformantes de levadura que albergan proteínas que se unen a la proteína BGL6 se identifican mediante, por ejemplo, producción de beta-galactosidasa o luciferasa (un muestreo), o supervivencia en placas que carecen de un nutriente esencial (una selección), según proceda para los vectores utilizados.

C. Análisis de micromatrices

- 15 [0194] Además, los análisis de micromatrices, también conocidos como perfiles de expresión o perfiles de transcripción, pueden utilizarse para evaluar de manera simultánea la presencia o expresión de secuencias de ADN dadas, o cambios en la expresión de muchos genes diferentes. En un enfoque, un gran conjunto de secuencias de ADN (sondas), normalmente un conjunto amplio de marcadores de secuencia expresada, ADNc, fragmentos de ADNc u oligonucleótidos de secuencia específica, se dispone en un soporte sólido tal como un portaobjetos o una membrana de nailon. El objetivo marcado para la hibridación en las sondas se genera aislando ARNm de tejidos inducidos y de control, marcando entonces cada grupo de ARNm directamente o mediante un intermediario de ADNc o ARNc, con un marcador distinto, normalmente un tinte fluorescente. La micromatriz se hibrida con las sondas de compuestos y la intensidad de la señal de hibridación relativa asociada a cada localización en la matriz puede cuantificarse para cada tinte marcador. Las diferencias de expresión entre los estados inducidos y de control pueden medirse como una relación de la señal de los dos tintes marcadores. (Véase Baldwin, D *et al.*, 1999.)
20

- [0195] El análisis de micromatrices del organismo fuente del que se derivó la *bgl6* puede llevarse a cabo para facilitar la comprensión de la función del gen identificando otros genes que se regulan de manera coordinada como consecuencia de la sobreexpresión de la *bgl6*. La identidad de los genes regulados coordinados puede ayudar a colocar el gen de *bgl6* en un sistema concreto. De manera alternativa, dicho análisis puede utilizarse para identificar otros genes involucrados en el mismo sistema utilizando análisis de micromatrices.
30

- [0196] Aunque la invención se ha descrito con referencia a modos de realización y métodos específicos, se entenderá que pueden realizarse varios cambios y modificaciones sin alejarse de la invención como se define en las reivindicaciones adjuntas.
35

EJEMPLO 1

- [0197] En un enfoque de ejemplo, un fragmento de ADNc para su utilización como una sonda se aísla extrayendo el ARN total de los micelios de una cepa de *T. reesei* que ha crecido en condiciones que se sabe que inducen la producción de celulasa y la obtención de la fracción poliadenilada (polyA) del mismo. El ARN poliA se utiliza para producir un grupo de ADNc que se amplifica entonces utilizando cebadores específicos basados en la secuencia de ácido nucleico de *bgl6* proporcionada en el presente documento.
40

- [0198] El ARN total se aísla de los micelios utilizando métodos conocidos en la técnica, por ejemplo como se describe en Timberlake *et al.*, 1981; Maniatis, *et al.*, 1989; Ausubel, *et al.*, 1993 y Sambrook *et al.*, 1989. Una vez aislado, se realizan las transferencias Northern para confirmar la expresión de celulasa y seleccionar un tiempo de inducción óptimo para la expresión de la celulasa y el correspondiente aislamiento del ARN.
45

[0199] El ARN mensajero (ARNm), que tienen una cola de poli (A) en el extremo 3', puede purificarse del ARN total utilizando métodos conocidos en la técnica.

- [0200] El ARN de *T. reesei* se utiliza como modelo para la RT-PCR utilizando métodos conocidos en la técnica (Loftus, J. *et al.*, *Science*, 249:915-918, 1990). Durante este procedimiento, el ARNm se transcribe de manera inversa para producir la primera cadena de ADNc. Posteriormente, el ADNc sirve como modelo para la amplificación de la PCR de las secuencias de ADNc de *bgl6* utilizando cebadores de oligonucleótidos específicos diseñados de conformidad con la SEQ ID No. 1 o la SEQ ID No. 3.
50

Tabla 1. Secuencias proporcionadas

Descripción	SEQ. ID NO.
secuencia de ácidos nucleicos de ADN de <i>bgl6</i> completa	1
<u>GATCACACCCCTCCACCCCTTCTCTTTTCAAGGTTGTCCCTTCTCCACGG</u> <u>CTTTATGTACTTCCCACTCTMTAATTCGCTCTTCCATTCCAAGCCAAGCAA</u> <u>CATCTGTGAGCAGCTCATCTTCCCAATATGGGCGAATGGCAGGAGCAGAT</u> GATGGGTTTTGACGTGGAGGATGTTCTGTCTCAGCTGAGCCAAAATGAGAA GATTGCTCTCTTGTCCGGCATTGATTTCTGGCATACTTATCCCATAACAAAG TACAACGTCCCTTCAGTCCGCCTAACGGACGGTCCTAACGGCATAACGAGGC ACAAAGTTTTTGTCTGGCATTCTGCTGCCTGCCTGCCATGTGGGACGGCC CTGGCCTCTACCTGGGATAAGCAGCTGCTGAAGAAGGCTGGGAAGCTGCT CGGTGATGAGTGCAATCGCAAAGGGCGCCACTGCTGGCTGGGCCCAACAA TCAATACTCCCCGATCTCCTCTGGGGGGGGCGGGCTTCGAGTCATTTTCGG AAGATCCGTACCTGTCCGGCATCTTGTCTGCATCTATGATTCTCGGCTGTG AAAGCACAGGTGTCATCTCTGCCGTCAAACACTTTGTGCGCAACGACCAGG AGCACGAGCGGCGAGCGGTGACTGTCTCATCACCCAGCGGGCTCTCCGG GAGGTCTATCTGCGACCCCTCCAGATCGTAGCCCGAGATGCAAGGCCCGGC GCATTGATGACATCCTACAACAAGGTCAATGGCAAGCACGTCGCTGACAG CGCCGAGTTCCCTCAGGGCATTCTCCGGACTGAGTGGAATTGGGATCCTCT CATTGTGACGCGACTGGTACGGCACCTACACCACTATTGATGCCATCAAAGC CGGCCTTGATCTCGAGATGCCGGGCGTTTCACGATATCGCGGCAAATACAT CGAGTCTGCTCTGCAGGCCCCGTTTGTCTGAAGCAGTCCACTATCGATGAGCG CGCTCGCCGCGTGCTCAGGTTCGCCCAGAAGGCCAGCCATCTCAAGGTCTC CGAGGTAGAGCAAGGCCGTGACTTCCCAGAGGATCGCGTCCTCAACCGTC AGATCTGCGGCAGCAGCATTGTCCTACTGAAGAATGAGAACTCCATCTTAC CTCTCCCCAAGTCCGTCAAGAAGGTGCGCCCTTGTGGATCCCACGTGCGTC TACCGGCTATCTCGGGAGGAGGCAGCGCCTCTCTTGTCCCTTACTATGCCA TATCTCTATACGATGCCGTCTCTGAGGTACTAGCCGGTGCCACGATCACGC	

Descripción	SEQ. ID NO.
ACGAGGTCGGTGCCTATGCCACCAAATGCTGCCCGTCATCGACGCAATGA TCAGCAACGCCGTAATCCACTTCTACAACGACCCCATCGATGTCAAAGACA GAAAGCTCCTTGGCAGTGAGAACGTATCGTCGACATCGTTCCAGCTCATGG ATTACAACAACATCCCAACGCTCAACAAGGCCATGTTCTGGGGTACTCTCG TGGGCGAGTTTATCCCTACCGCCACGGAATTTGGGAATTTGGCCTCAGTG TCTTTGGCACTGCCGACCTTTATATTGATAATGAGCTCGTGATTGAAAATA CAACACATCAGACGCGTGGTACCGCCTTTTTCGGAAAGGGAACGACGGAA AAAGTCGCTACCAGGAGGATGGTGGCCGGCAGCACCTACAAGCTGCGTCT CGAGTTTGGGTCTGCCAACACGACCAAGATGGAGACGACCGGTGTTGTCA ACTTTGGCGGCGGTGCCGTACACCTGGGTGCCTGTCTCAAGGTCGACCCAC AGGAGATGATTGCGCGGGCCGTCAAGGCCGCAGCCGATGCCGACTACACC ATCATCTGCACGGGACTCAGCGGCGAGTGGGAGTCTGAGGGTTTTGACCG GCCTCACATGGACCTGCCCCCTGGTGTGGACACCATGATCTCGCAAGTTCT TGACGCCGCTCCCAATGCTGTAGTCGTCAACCAGTCAGGCACCCCAGTGAC AATGAGCTGGGCTCATAAAGCAAAGGCCATTGTGCAGGCTTGGTATGGTG GTAACGAGACAGGCCACGGAATCTCCGATGTGCTCTTTGGCAACGTCAACC CGTCGGGGAAACTCTCCCTATCGTGGCCAGTCGATGTGAAGCACAACCCA GCATATCTCAACTACGCCAGCGTTGGTGGACGGGTCTTGTATGGCGAGGAT GTTTACGTTGGCTACAAGTTCTACGACAAAACGGAGAGGGAGGTTCTGTTT CCTTTTGGGCATGGCCTGTCTTACGCTACCTTCAAGCTCCCAGATTCTACCG TGAGGACGGTCCCCGAAACCTTCCACCCGACCAGCCACAGTAGCCATT GTCAAGATCAAGAACACGAGCAGTGTCCCGGGCGCCCAGGTCTCGCAGCT ATACATTTTCGGCCCCAAACTCGCCTACACATCGCCCCGGTCAAGGAGCTGCA CGGATTCGAAAAGGTGTATCTTGAAGCTGGCGAGGAGAAGGAGGTACAAA TACCCATTGACCAGTACGCTACTAGCTTCTGGGACGAGATTGAGAGCATGT GGAAGAGCGAGAGGGGCATTTATGATGTGCTTGTAGGATTCTCGAGTCAG GAAATCTCGGGCAAGGGGAAGCTGATTGTGCCTGAAACGCGATTCTGGAT GGGGCTGTAGATTCAACACGTGAGCAAAAGCGATTGCGGAAAGTACCAGA AAAGCCAAGGGAGTCAAAGGATGGGAACTTGTGTCAATAGAAGATATGCA TAGATGGGCATTCTGGGATGGTGGTTTGGCATTAATGCAAAGAAGACAAA GATGGATGTGATAAAAAAAAAAAAAAAAAAAAAA	
Secuencia de aminoácidos prevista de BGL6	2
MGEWQEQMMMGFDVEDVLSQLSQNEKIALLSGIDFWHTYPIPKYNVPSVRLTD GPNGIRGTFKFFAGIPAAACLPCTALASTWDKQLLKKAGKLLGDECIAGAH WLGPINTPRSPLGGRGFESFSEDPYLSGILAAASMLGCESTGVISAVKHVAND QEHERRAVDCLITQRALREVYLRPFQIVARDARPGALMTSYNKNVNGKHVADS AEFLQGILRTEWNWDPLIVSDWYGYTTIDAIAKAGLDLEMPGVSRYRGKYIES ALQARLLKQSTIDERARRVLRFAQKASHLKVSEVEQGRDFPEDRVLNRIQCS SIVLLKNENSILPLPKSVKKVALVGSHVRLPAISGGGSASLVPYYAISLYDAVSE VLGATITHEVGAYAHQMLPVIDAMISNAVIHFYNDPIDVKDRKLLGSENVSS TSFQLMDYNNIPTLNKAMFWGTLVGEFIPATGIWEFGLSVFGTADLYIDNEL VIENTTHQTRGTAFGKGTTEKVATRRMVAGSTYKLRLEFGSANTTKMETTG VYNFGGGA VHLGACLKVDPEMIARAVKAAADADYTICTGLSGEWESEGFD RPHMDLPPGVDTMISQVLDAAPNAV VVNQSGTPVTMSWAHKAKAIVQAWY GGNETGHGISDVLFGNVNPSGKLSLWPDVKHNPAYLNYASVGGRVLYGE DVYVGYKYFDKTEREVLFPFGHLSYATFKLPDSTVRTVPETFHPDQPTVAIV KIKNTSSVPGAQVLQLYISAPNSPTHRVPKELHGFEKVYLEAGEEKEVQIPIDQ	
<u>YATSFWDEIESMWKSERGIYDVLVGFSSQEISGKGKLIPETRFWMGL</u>	

Descripción	SEQ. ID NO.
Secuencia de aminoácidos prevista de BGL6 con inicio alternativo MMGFĐVEDVLSQLSQNEKIÁLLSGIDFWHTYPIPKYNVPSVRLTDGPNNGIRGT KFFAGIPAACLPCGTALASTWDKQLLKKAGKLLGDECIAGAHWCWLGPNTINTP RSPLGGRGFESFSEDPLYSGILAASMLGCESTGVISAVKHVANDQEHERRAV DCLITQRALREVYLRPFQIVARDARPGALMTSYNKVNGKHVADSAEFLQGILR TEWNWDPLIVSDWYGYTTTIDAIKAGLDLEMPGVSRYRGKYIESALQARLLK QSTIDERARRVLRFAQKASHLKVSEVEQGRDFPEDRVLRQICGSSIVLLKNEN SILPLPKSVKKVALVGSHVRLPAISGGGSASLVPYYAISLYDAVSEVLAGATIT HEVGAYAHQMLPVIDAMISNAVIHFYNDPIDVKDRKLLGSENVSSTSFQLMDY NNIPTLNKAMFWGTLVGEFIPTATGIWEFGLSVFGTADLYIDNELVIENTHQT RGTAFFGKGTEK VATRRMVAGSTYKLRLEFGSANTTKMETTG VVNFGGGA VHLGACLKVDPQEMIARAVKAAADADYTICTGLSGEWESEGFDRPHMDLPP GVDTMISQVLDAAPNAV VVNQSGTPVTMSWAHKAKAIVQAWYGGNETGHG ISDVLFGNVNPSGKLSLWPDVKHNPAYLNYASVGGRVLYGEDVYVGYKF YDKTEREVLFPFGHGLSYATFKLPDSTVRTVPETFHPDQPTVAIVKIKNTSSVP GAQVLQLYISAPNSPTHRVPKELHGFKEKVYLEAGEEKEVQIPIDQYATSFWDEI ESMWKSERGIYDVLVGFSSQEISGKGKLVIPETRFWMGL	4
Secuencia codificante de ácido nucleico de <i>bgl6</i> ATGGGCGAATGGCAGGAGCAGATGATGGGTTTTGACGTGGAGGATGTTCT GTCTCAGCTGAGCCAAAATGAGAAGATTGCTCTCTTGTCCGGCATTGATTT CTGGCATACTTATCCCATACCAAAGTACAACGTCCCTTCAGTCCGCCTAAC GGACGGTCCTAACGGCATAACGAGGCACAAAGTTTTTTGCTGGCATTCTCTGC TGCCTGCTGCCATGTGGGACGGCCCTGGCCTCTACCTGGGATAAGCAGCT GCTGAAGAAGGCTGGGAAGCTGCTCGGTGATGAGTGCATCGCAAAGGCG CCCACTGCTGGCTGGGCCCAACAATCAATACTCCCCGATCTCCTCTGGGGG GGCGCGGCTTCGAGTCATTTTCGGAAGATCCGTACCTGTCCGGCATCCTTG CTGCATCTATGATTCTCGGCTGTGAAAGCACAGGTGTCATCTCTGCCGTCA AACACTTTGTGCGCAACGACCAGGAGCACGAGCGCGAGCGGTGCGACTGT CTCATCACCCAGCGGGCTCTCCGGGAGGTCTATCTGCGACCCTTCCAGATC GTAGCCCGAGATGCAAGGCCCGGCGCATTGATGACATCCTACAACAAGGT CAATGGCAAGCACGTCGCTGACAGCGCCGAGTTCCTTCAGGGCATTCTCCG GACTGAGTGGAATTGGGATCCTCTCATTGTGACGCACTGGTACGGCACCTA CACCATAATTGATGCCATCAAAGCCGGCCTTGATCTCGAGATGCCGGGCGT TTCACGATATCGCGCAAATACATCGAGTCTGCTCTGCAGGCCCGTTTGCT GAAGCAGTCCACTATCGATGAGCGCGCTCGCCGCGTGCTCAGGTTGCCCCA GAAGGCCAGCCATCTCAAGGTCTCCGAGGTAGAGCAAGGCCGTGACTTCC CAGAGGATCGCGTCTCAACCGTCAGATCTGCGGCAGCAGCATTGTCCTAC TGAAGAATGAGAACTCCATCTTACCTCTCCCCAAGTCCGTCAAGAAGGTG CCCTTGTTGGATCCACGTGCGTCTACCGGCTATCTCGGGAGGAGGCAGCG CCTCTCTTGTCCCTTACTATGCCATATCTCTATACGATGCCGTCTCTGAGGT ACTAGCCGGTGCCACGATCACGCACGAGGTGCGTGCTATGCCACCAA TGCTGCCCCGTCATCGACGCAATGATCAGCAACGCCGTAATCCAATTCTACA ACGACCCCATCGATGTCAAAGACAGAAAGCTCCTTGGCAGTGAGAACGTA TCGTGACATCGTTCCAGCTCATGGATTACAACAACATCCCAACGCTCAAC AAGGCCATGTTCTGGGGTACTCTCGTGGGCGAGTTTATCCCTACCGCCACG GGAATTTGGGAATTTGGCCTCAGTGTCTTTGGCACTGCCGACCTTTATATTG ATAATGAGCTCGTGATTGAAAATACAACACATCAGACGCGTGGTACCGCC TTTTTCGGAAGGGAACGACGGAAAAAGTCGCTACCAGGAGGATGGTGGC CGGCAGCACCTACAAGCTGCGTCTCGAGTTTGGGTCTGCCAACACGACCAA GATGGAGACGACCGGTGTTGTCAACTTTGGCGGGCGGTGCCGTACACCTGG GTGCCTGTCTCAAGGTCGACCCACAGGAGATGATTGCGCGGGCCGTCAAG GCCGACCGGATGCCGACTACACCATCATCTGCACGGGACTCAGCGGCGA GTGGGAGTCTGAGGGTTTTGACCGGCCCTCACATGGACCTGCCCCCTGGTGT GGACACCATGATCTCGCAAGTTCTTGACGCCGCTCCCAATGCTGTAGTCGT CAACCACTCAGGCACCCAGTGACAATGAGCTGGGCTCATAAAGCAAAGG CCATTGTGACAGGCTTGGTATGGTGGTAACGAGACAGGCCACGGAATCTCC GATGTGCTCTTTGGCAACGTCAACCCGTCGGGGAAACTCTCCCTATCGTGG CCAGTCGATGTGAAGACAACCCAGCATATCTCAACTACGCCAGCGTTGGT GGACGGGTCTTGATGGCGAGGATGTTTACGTTGGCTACAAGTTCTACGAC	3

Descripción	SEQ. ID NO.
AAAACGGAGAGGGAGGTTCTGTTTCCTTTTGGGCATGGCCTGTCTTACGCT ACCTTCAAGCTCCCAGATTCTACCGTGAGGACGGTCCCCGAAACCTTCCAC CCGGACCAGCCCACAGTAGCCATTGTCAAGATCAAGAACACGAGCAGTGT CCCGGGCGCCCAGGTCCTGCAGCTATACATTTCTGGCCCCAAACTCGCCTAC ACATCGCCCGGTCAAGGAGCTGCACGGATTGAAAAAGGTGTATCTTGAAG CTGGCGAGGAGAAGGAGGTACAAATACCCATTGACCAGTACGCTACTAGC TTCTGGGACGAGATTGAGAGCATGTGGAAGAGCGAGAGGGGCATTTATGA TGTGCTTGTAGGATTCTCGAGTCAGGAAATCTCGGGCAAGGGGAAGCTGA TTGTGCCTGAAACGCGATTCTGGATGGGGCTGTAG	

REIVINDICACIONES

1. Un polinucleótido aislado derivado de una fuente fúngica, dicho polinucleótido comprende una secuencia de nucleótidos que codifica una enzima que tiene actividad de β -glucosidasa donde dicho polinucleótido aislado se selecciona de entre el grupo consistente en:
 - 5 (a) una secuencia de ácido nucleico que codifica o es complementaria a una secuencia que codifica un polipéptido BGL6 que tiene al menos 85 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
 - (b) una secuencia de ácido nucleico que codifica o es complementaria a una secuencia que codifica un polipéptido BGL6 que tiene al menos 90 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
 - 10 (c) una secuencia de ácido nucleico que codifica o es complementaria a una secuencia que codifica un polipéptido BGL6 que tiene al menos 95 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
 - (d) una secuencia de ácido nucleico que codifica o es complementaria a una secuencia que codifica un polipéptido BGL6 que tiene la secuencia de aminoácidos presentada como SEQ ID NO:2; y
 - 15 (e) una secuencia de ácido nucleico presentada como SEQ ID NO:3 o el complemento de la misma.
2. Polinucleótido aislado de la reivindicación 1, en el que el % de identidad se calcula utilizando el programa CLUSTAL-W en MacVector versión 6.5, que funciona con parámetros por defecto, que incluyen una penalización por hueco abierto de 10,0, una penalización por hueco extendido de 0,1 y una matriz de similitud BLOSUM 30.
- 20 3. Polinucleótido aislado de la reivindicación 1, en el que dicho polinucleótido es una molécula de ARN.
4. Polinucleótido aislado de la reivindicación 2 que codifica una enzima que tiene actividad de β -glucosidasa, en el que la enzima se deriva de una fuente de *Trichoderma*.
5. Polinucleótido aislado de la reivindicación 3, en el que la enzima se deriva de *Trichoderma reesei*.
6. Constructo de expresión que incluye una secuencia de polinucleótidos (i) que codifica un polipéptido BGL6 que tiene al menos 85 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2, o (ii) que es complementaria a una secuencia de nucleótidos que tiene al menos 85 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
- 25 7. Vector que incluye el constructo de expresión de la reivindicación 6.
8. Vector que comprende un polinucleótido aislado de la reivindicación 1, operativamente unido a secuencias de control reconocidas por una célula huésped transformada con el vector.
- 30 9. Célula huésped transformada con el vector de la reivindicación 7.
10. Célula huésped transformada con el vector de la reivindicación 8.
11. Célula huésped de la reivindicación 10, que es una célula procariota.
12. Célula huésped de la reivindicación 10, que es una célula eucariota.
- 35 13. Célula huésped recombinante procariota que comprende un polinucleótido de la reivindicación 1.
14. Polipéptido BGL6 aislado con la actividad biológica de una β -glucosidasa, que comprende una secuencia seleccionada de entre el grupo consistente en:
 - (a) una secuencia de aminoácidos que tiene al menos 85 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
 - 40 (b) una secuencia de aminoácidos que tiene al menos 90 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
 - (c) una secuencia de aminoácidos que tiene al menos 95 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
 - (d) una secuencia de aminoácidos presentada como SEQ ID NO:2; y
 - 45 (e) un fragmento aislado de la secuencia de aminoácidos presentada como SEQ ID NO:2 que tiene la actividad biológica de una β -glucosidasa.
15. Método para la producción de un polipéptido BGL6 que tiene actividad de β -glucosidasa, que comprende:
 - (a) la transformación de manera estable de una célula huésped con un vector de expresión que comprende un polinucleótido como se define en la reivindicación 1;
 - 50 (b) el cultivo de dicha célula huésped transformada en una condición adecuada para que dicha célula huésped produzca dicha β -glucosidasa; y

(c) la recuperación de dicha β -glucosidasa.

16. Método de la reivindicación 15 en el que la célula huésped es un hongo filamentoso o una célula de levadura.

17. Oligonucleótido antisentido complementario a la secuencia codificante bgl6 de un ARN mensajero que codifica un polipéptido BGL6 que tiene la secuencia presentada como SEQ ID NO:2, en el que tras la exposición a una célula huésped que produce β -glucosidasa, dicho oligonucleótido reduce o inhibe la producción de β -glucosidasa por dicha célula huésped.

18. Oligonucleótido antisentido de la reivindicación 17, en el que la célula huésped es un hongo filamentoso.

19. Composición detergente, comprendiendo dicha composición un polipéptido con la actividad biológica de una β -glucosidasa seleccionada de entre el grupo consistente en:

- 10 (a) una secuencia de aminoácidos que tiene al menos 85 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
- (b) una secuencia de aminoácidos que tiene al menos 90 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
- 15 (c) una secuencia de aminoácidos que tiene al menos 95 % de identidad de secuencia con la secuencia de aminoácidos presentada como SEQ ID NO:2;
- (d) una secuencia de aminoácidos presentada como SEQ ID NO:2;
- (e) un fragmento aislado de la secuencia de aminoácidos presentada como SEQ ID NO:2 que tiene la actividad biológica de una β -glucosidasa.

20. Método para la mejora de las características de una masa de levadura o un producto horneado hecho a partir de dicha masa que consiste esencialmente en las etapas de:

- (a) mezclado de al menos aproximadamente 10 ppm de una BGL6 de conformidad con la reivindicación 14 con los ingredientes de la masa para formar una mezcla de masa, y
- (b) horneado de dicha mezcla de masa para formar un producto horneado.

21. Método para la mejora de las características de una masa de pan con levadura o de una masa de panecillo con levadura o de un pan con levadura o de un panecillo con levadura que consiste esencialmente en las etapas de:

- (a) mezclado de al menos aproximadamente 10 ppm de una BGL6 de conformidad con la reivindicación 14 con los ingredientes de la masa de panecillo o de pan para formar una mezcla de masa;
- (b) formado o moldeado de la mezcla de masa;
- 30 (c) fermentación de la mezcla de masa, y
- (d) horneado de la mezcla de masa para formar pan o panecillos.

GATCACACCC	CTCCCACCCT	TCTCTTTTCA	AGGTTGTCCC	CTTCTCCCAC	50
GGCTTTATGT	ACTTCCCAC	CTMTAATTG	CTCTTTCCAT	TCCAAGCCAA	100
GCAACATCTG	TGAGCAGCTC	ATCCTTCCCA	ATATGGGCGA	ATGGCAGGAG	150
CAGATGATGG	GTTTTGACGT	GGAGGATGTT	CTGTCTCAGC	TGAGCCAAAA	200
TGAGAAGATT	GCTCTCTTGT	CCGGCATTGA	TTTCTGGCAT	ACTTATCCCA	250
TACCAAAGTA	CAACGTCCCT	TCAGTCCGCC	TAACGGACGG	TCCTAACGGC	300
ATACGAGGCA	CAAAGTTTTT	TGCTGGCATT	CCTGCTGCCT	GCCTGCCATG	350
TGGGACGGCC	CTGGCCTCTA	CCTGGGATAA	GCAGCTGCTG	AAGAAGGCTG	400
GGAAGCTGCT	CGGTGATGAG	TGCATCGCAA	AAGGCGCCCA	CTGCTGGCTG	450
GGCCCAACAA	TCAATACTCC	CCGATCTCCT	CTGGGGGGGC	GCGGCTTCGA	500
GTCATTTTCG	GAAGATCCGT	ACCTGTCCGG	CATCCTTGCT	GCATCTATGA	550
TTCTCGGCTG	TGAAAGCACA	GGTGTCTCT	CTGCCGTCAA	ACACTTTGTC	600
GCCAACGACC	AGGAGCACGA	GCGGCGAGCG	GTCGACTGTC	TCATCACCCA	650
GCGGGCTCTC	CGGGAGGTCT	ATCTGCGACC	CTTCCAGATC	GTAGCCCGAG	700
ATGCAAGGCC	CGGCGCATTG	ATGACATCCT	ACAACAAGGT	CAATGGCAAG	750
CACGTCGCTG	ACAGCGCCGA	GTTCTTCAG	GGCATTCTCC	GGACTGAGTG	800
GAATTGGGAT	CCTCTCATTG	TCAGCGACTG	GTACGGCACC	TACACCACTA	850
TTGATGCCAT	CAAAGCCGGC	CTTGATCTCG	AGATGCCGGG	CGTTTTCACGA	900
TATCGCGGCA	AATACATCGA	GTCTGCTCTG	CAGGCCCGTT	TGCTGAAGCA	950
GTCCACTATC	GATGAGCGCG	CTCGCCGCGT	GCTCAGGTTT	GCCCAGAAGG	1000
CCAGCCATCT	CAAGGTCTCC	GAGGTAGAGC	AAGGCCGTGA	CTTCCAGAG	1050
GATCGCGTCC	TCAACCGTCA	GATCTGCGGC	AGCAGCATTG	TCCTACTGAA	1100
GAATGAGAAC	TCCATCTTAC	CTCTCCCCAA	GTCCGTCAAG	AAGGTCGCCC	1150
TTGTTGGATC	CCACGTGCGT	CTACCGGCTA	TCTCGGGAGG	AGGCAGCGCC	1200
TCTCTTGTC	CTTACTATGC	CATATCTCTA	TACGATGCCG	TCTCTGAGGT	1250
ACTAGCCGGT	GCCACGATCA	CGCACGAGGT	CGGTGCCTAT	GCCCACCAAA	1300
TGCTGCCCCG	CATCGACGCA	ATGATCAGCA	ACGCCGTAAT	CCACTTCTAC	1350
AACGACCCCA	TCGATGTCAA	AGACAGAAAG	CTCCTTGGCA	GTGAGAACGT	1400
ATCGTCGACA	TCGTTCCAGC	TCATGGATTA	CAACAACATC	CCAACGCTCA	1450
ACAAGGCCAT	GTTCTGGGGT	ACTCTCGTGG	GCGAGTTTAT	CCCTACCGCC	1500
ACGGGAATTT	GGGAATTTGG	CCTCAGTGTC	TTTGGCACTG	CCGACCTTTA	1550
TATTGATAAT	GAGCTCGTGA	TTGAAAATAC	AACACATCAG	ACGCGTGGTA	1600
CCGCCTTTTT	CGGAAAGGGA	ACGACGGAAA	AAGTCGCTAC	CAGGAGGATG	1650
GTGGCCGGCA	GCACCTACAA	GCTGCGTCTC	GAGTTTGGGT	CTGCCAACAC	1700
GACCAAGATG	GAGACGACCG	GTGTTGTCAA	CTTTGGCGGC	GGTGCCGTAC	1750
ACCTGGGTGC	CTGTCTCAAG	GTCGACCCAC	AGGAGATGAT	TGCGCGGGCC	1800
GTCAAGGCCG	CAGCCGATGC	CGACTACACC	ATCATCTGCA	CGGGACTCAG	1850
CGGCGAGTGG	GAGTCTGAGG	GTTTTGACCG	GCCTCACATG	GACCTGCCCC	1900

FIG. _1A

CTGGTGTGGA	CACCATGATC	TCGCAAGTTC	TTGACGCCGC	TCCCAATGCT	1950
GTAGTCGTCA	ACCAGTCAGG	CACCCCAGTG	ACAATGAGCT	GGGCTCATAA	2000
AGCAAAGGCC	ATTGTGCAGG	CTTGGTATGG	TGGTAACGAG	ACAGGCCACG	2050
GAATCTCCGA	TGTGCTCTTT	GGCAACGTCA	ACCCGTCGGG	GAAACTCTCC	2100
CTATCGTGGC	CAGTCGATGT	GAAGCACAAC	CCAGCATATC	TCAACTACGC	2150
CAGCGTTGGT	GGACGGGTCT	TGTATGGCGA	GGATGTTTAC	GTTGGCTACA	2200
AGTTCTACGA	CAAAACGGAG	AGGGAGGTTT	TGTTTCCTTT	TGGGCATGGC	2250
CTGTCTTACG	CTACCTTCAA	GCTCCCAGAT	TCTACCGTGA	GGACGGTCCC	2300
CGAAACCTTC	CACCCGGACC	AGCCCACAGT	AGCCATTGTC	AAGATCAAGA	2350
ACACGAGCAG	TGTCCCGGGC	GCCCAGGTCC	TGCAGCTATA	CATTTCGGCC	2400
CCAAACTCGC	CTACACATCG	CCCGGTCAAG	GAGCTGCACG	GATTTCGAAAA	2450
GGTGTATCTT	GAAGCTGGCG	AGGAGAAGGA	GGTACAAATA	CCCATTGACC	2500
AGTACGCTAC	TAGCTTCTGG	GACGAGATTG	AGAGCATGTG	GAAGAGCGAG	2550
AGGGGCATTT	ATGATGTGCT	TGTAGGATTC	TCGAGTCAGG	AAATCTCGGG	2600
CAAGGGGAAG	CTGATTGTGC	CTGAAACGCG	ATTCTGGATG	GGGCTGTAGA	2650
<u>TTCAACACGT</u>	<u>GAGCAAAAGC</u>	<u>GATTGCGGAA</u>	<u>AGTACCAGAA</u>	<u>AAGCCAAGGG</u>	2700
<u>AGTCAAAGGA</u>	<u>TGGGAACTTG</u>	<u>TGTCAATAGA</u>	<u>AGATATGCAT</u>	<u>GATGGGCATT</u>	2750
<u>TGGGATGGTG</u>	<u>TTTGGCATT</u>	<u>TGCAAAGAAG</u>	<u>CAAAGATGGA</u>	<u>GTGATAAAAA</u>	2800
<u>AAAAAAAAAA</u>	<u>AA</u>				2812

FIG._1B

MGEWQEQMMG	FDVEDVLSQL	SQNEKIALLS	GIDFWHTYPE	PKYNVPSVRL	50
TDGPNGIRGT	KFFAGIPAAC	LPCGTALAST	WDKQLLKKAG	KLLGDECIAC	100
GAHCWLGPIT	NTPRSPLGGR	GFESFSEDPI	LSGILAAAMI	LGCESTGVIS	150
AVKHVFVANDQ	EHERRAVDCL	ITQRALREVI	LRPFQIVARD	ARPGALMTSY	200
NKVNKGHVAD	SAEFLQGILR	TEWNWDPLIV	SDWYGTYTTI	DAIKAGLDLE	250
MPGVSRVRGK	YIESALQARL	LKQSTIDERA	RRVLRFAQKA	SHLKVSEVEQ	300
GRDFPEDRVL	NRQICGSSIV	LLKNENSILP	LPKSVKKVAL	VGSHVRLPAI	350
SGGGSASLVP	YYAISLYDAV	SEVLAGATIT	HEVGAYAHQM	LPVIDAMISN	400
AVIHFYNDPI	DVKDRKLLGS	ENVSSSTFQL	MDYNNIPTLN	KAMFWGTLVG	450
EFIPTATGIW	EFGLSVFGTA	DLYIDNELVI	ENTTHQTRGT	AFFGKGTTEK	500
VATRRMVAGS	TYKLRLEFGS	ANTTKMETTG	VVNFGGGAVH	LGACLKVDPO	550
EMIRARVAAA	ADADYTIICT	GLSGEWESEG	FDRPHMDLPP	GVDTMISQVL	600
DAAPNAVNVN	QSGTPVTMSW	AHKAKAIVQA	WYGGNETGHG	ISDVLFGNVN	650
PSGKLSLSWP	VDVKHNPAYL	NYASVGGRVL	YGEDVYVGYK	FYDKTEREVL	700
FPFGHGLSYA	TFKLPDSTVR	TVPETFHPDQ	PTVAIVKIKN	TSSVPGAQVL	750
QLYISAPNSP	THRPVKELHG	FEKVYLEAGE	EKEVQIPIDQ	YATSFWDEIE	800
SMWKSERGIY	DVLVGFSSQE	ISGKGKLIVP	ETRFWMGL		838

Figura 2

MMGFDVEDVL	SQLSQNEKIA	LLSGIDFWHT	YPIPKYNVPS	VRLTDGPNGI	50
RGTKFFAGIP	AACLPCGTAL	ASTWDKQLLK	KAGKLLGDEC	IAKGAHCWLG	100
PTINTPRSPL	GGRGFESFSE	DPYLSGILAA	SMILGCESTG	VISAVKHFWA	150
NDQEHERRAV	DCLITQRALR	EVYLRPFQIV	ARDARPGALM	TSYNKVNGKH	200
VADSAEFLQG	ILRTEWNWDP	LIVSDWYGTY	TTIDAIKAGL	DLEMPGVSRY	250
RGKYIESALQ	ARLLKQSTID	ERARRVLRFA	QKASHLKVSE	VEQGRDFPED	300
RVLNRQICGS	SIVLLKNENS	ILPLPKSVKK	VALVGSHVRL	PAISGGGSAS	350
LVPYYAISLY	DAVSEVLAGA	TITHEVGAYA	HQMLPVIDAM	ISNAVIHFYN	400
DPIDVKDRKL	LGSENVSSTS	FQLMDYNNIP	TLNKAMFWGT	LVGEFIPTAT	450
GIWEFGLSVF	GTADLYIDNE	LVIENTTHQT	RGTAFFGKGT	TEKVATRMRV	500
AGSTYKLRLE	FGSANTTKME	TTGVVNFGGG	AVHLGACLKV	DPQEMIARAV	550
KAAADADYTI	ICTGLSGEWE	SEGFDRPHMD	LPPGVDTMIS	QVLDAAPNAV	600
VVNQSGTPVT	MSWAHKAKAI	VQAWYGGNET	GHGISDVLFG	NVNPSGKLSL	650
SWPVDVKHNP	AYLNYASVGG	RVLYGEDVYV	GYKFYDKTER	EVLFPFGHGL	700
SYATFKLPDS	TVRTVPETFH	PDQPTVAIVK	IKNTSSVPGA	QVLQLYISAP	750
NSPTHRPVKE	LHGFEEKVYLE	AGEEKEVQIP	IDQYATSFWD	EIESMWKSER	800
GIYDVLVGFS	SQEISGKGKL	IVPETREWMG	L		831

Figura 3

ATGGGCGAAT	GGCAGGAGCA	GATGATGGGT	TTTGACGTGG	AGGATGTTCT	50
GTCTCAGCTG	AGCCAAAATG	AGAAGATTGC	TCTCTTGTC	GGCATTGATT	100
TCTGGCATA	TTATCCCATA	CCAAAGTACA	ACGTCCCTTC	AGTCCGCCTA	150
ACGGACGGTC	CTAACGGCAT	ACGAGGCACA	AAGTTTTTTG	CTGGCATTCC	200
TGCTGCCTGC	CTGCCATGTG	GGACGGCCCT	GGCCTCTACC	TGGGATAAGC	250
AGCTGCTGAA	GAAGGCTGGG	AAGCTGCTCG	GTGATGAGTG	CATCGCAAAA	300
GGCGCCCACT	GCTGGCTGGG	CCCAACAATC	AATACTCCCC	GATCTCCTCT	350
GGGGGGGCGC	GGCTTCGAGT	CATTTTCGGA	AGATCCGTAC	CTGTCCGGCA	400
TCCTTGCTGC	ATCTATGATT	CTCGGCTGTG	AAAGCACAGG	TGTCATCTCT	450
GCCGTCAAAC	ACTTTGTGCG	CAACGACCAG	GAGCACGAGC	GGCGAGCGGT	500
CGACTGTCTC	ATCACCAGC	GGGCTCTCCG	GGAGGTCTAT	CTGCGACCCT	550
TCCAGATCGT	AGCCCGAGAT	GCAAGGCCCG	GCGCATTGAT	GACATCCTAC	600
AACAAGGTCA	ATGGCAAGCA	CGTCGCTGAC	AGCGCCGAGT	TCCTTCAGGG	650
CATTCTCCGG	ACTGAGTGGA	ATTGGGATCC	TCTCATTGTC	AGCGACTGGT	700
ACGGCACCTA	CACCACTATT	GATGCCATCA	AAGCCGGCCT	TGATCTCGAG	750
ATGCCGGGCG	TTTCACGATA	TCGCGGCAAA	TACATCGAGT	CTGCTCTGCA	800
GGCCCGTTTG	CTGAAGCAGT	CCACTATCGA	TGAGCGCGCT	CGCCGCGTGC	850
TCAGGTTTCG	CCAGAAGGCC	AGCCATCTCA	AGGTCTCCGA	GGTAGAGCAA	900
GGCCGTGACT	TCCCAGAGGA	TCGCGTCCTC	AACCGTCAGA	TCTGCGGCAG	950
CAGCATTGTC	CTACTGAAGA	ATGAGAACTC	CATCTTACCT	CTCCCCAAGT	1000
CCGTCAAGAA	GGTCGCCCTT	GTTGGATCCC	ACGTGCGTCT	ACCGGCTATC	1050
TCGGGAGGAG	GCAGCGCCTC	TCTTGTCCTT	TACTATGCCA	TATCTCTATA	1100
CGATGCCGTC	TCTGAGGTAC	TAGCCGGTGC	CACGATCACG	CACGAGGTCT	1150
GTGCCTATGC	CCACCAAATG	CTGCCCCTCA	TCGACGCAAT	GATCAGCAAC	1200
GCCGTAAATC	ACTTCTACAA	CGACCCCATC	GATGTCAAAG	ACAGAAAGCT	1250
CCTTGGCAGT	GAGAACGTAT	CGTCGACATC	GTTCCAGCTC	ATGGATTACA	1300
ACAACATCCC	AACGCTCAAC	AAGGCCATGT	TCTGGGGTAC	TCTCGTGGGC	1350
GAGTTTATCC	CTACCGCCAC	GGGAATTTGG	GAATTTGGCC	TCAGTGTCTT	1400
TGGCACTGCC	GACCTTTATA	TTGATAATGA	GCTCGTGATT	GAAAATACAA	1450
CACATCAGAC	GCGTGGTACC	GCCTTTTTTC	GAAAGGGAAC	GACGGAAAAA	1500
GTCGCTACCA	GGAGGATGGT	GGCCGGCAGC	ACCTACAAGC	TGCGTCTCGA	1550
GTTTGGGTCT	GCCAACACGA	CCAAGATGGA	GACGACCGGT	GTTGTCAACT	1600
TTGGCGGGCG	TGCCGTACAC	CTGGGTGCCT	GTCTCAAGGT	CGACCCACAG	1650
GAGATGATTG	CGCGGGCCGT	CAAGGCCGCA	GCCGATGCCG	ACTACACCAT	1700
CATCTGCACG	GGACTCAGCG	GCGAGTGGGA	GTCTGAGGGT	TTTGACCGGC	1750
CTCACATGGA	CCTGCCCCCT	GGTGTGGACA	CCATGATCTC	GCAAGTTCTT	1800
GACGCCGCTC	CCAATGCTGT	AGTCGTCAAC	CAGTCAGGCA	CCCCAGTGAC	1850
AATGAGCTGG	GCTCATAAAG	CAAAGGCCAT	TGTGCAGGCT	TGGTATGGTG	1900
GTAACGAGAC	AGGCCACGGA	ATCTCCGATG	TGCTCTTTGG	CAACGTCAAC	1950
CCGTCGGGGA	AACTCTCCCT	ATCGTGGCCA	GTCGATGTGA	AGCACAAACC	2000
AGCATATCTC	AACTACGCCA	GCGTTGGTGG	ACGGGTCTTG	TATGGCGAGG	2050

FIG. 4A

ES 2 601 145 T3

ATGTTTACGT	TGGCTACAAG	TTCTACGACA	AAACGGAGAG	GGAGGTTCTG	2100
TTTCCTTTTG	GGCATGGCCT	GTCTTACGCT	ACCTTCAAGC	TCCCAGATTC	2150
TACCGTGAGG	ACGGTCCCCG	AAACCTTCCA	CCCGGACCAG	CCCACAGTAG	2200
CCATTGTCAA	GATCAAGAAC	ACGAGCAGTG	TCCCGGGCGC	CCAGGTCCTG	2250
CAGCTATACA	TTTCGGCCCC	AAACTCGCCT	ACACATCGCC	CGGTCAAGGA	2300
GCTGCACGGA	TTGAAAAGG	TGTATCTTGA	AGCTGGCGAG	GAGAAGGAGG	2350
TACAAATACC	CATTGACCAG	TACGCTACTA	GCTTCTGGGA	CGAGATTGAG	2400
AGCATGTGGA	AGAGCGAGAG	GGGCATTTAT	GATGTGCTTG	TAGGATTCTC	2450
GAGTCAGGAA	ATCTCGGGCA	AGGGGAAGCT	GATTGTGCCT	GAAACGCGAT	2500
TCTGGATGGG	GCTGTAG				2517

FIG. _4B