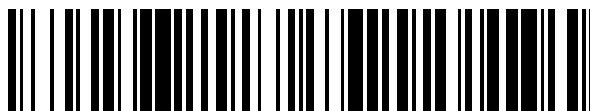


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 468 366**

51 Int. Cl.:

C07K 14/37 (2006.01)

C12N 9/42 (2006.01)

C11D 3/386 (2006.01)

D21C 1/00 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **28.01.2005 E 10180399 (7)**

97 Fecha y número de publicación de la concesión europea: **19.03.2014 EP 2301958**

54 Título: **Polipéptidos con actividad de mejora celulolítica y polinucleótidos que los codifican**

30 Prioridad:

30.01.2004 US 540661 P

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

16.06.2014

73 Titular/es:

**NOVOZYMES INC. (100.0%)
1445 Drew Avenue
Davis, CA 95616, US**

72 Inventor/es:

**BROWN, KIMBERLY;
HARRIS, PAUL;
ZARETSKY, ELIZABETH;
RE, EDWARD;
VLASENKO, ELENA;
MCFARLAND, KEITH y
LOPES DE LEON, ALFREDO**

74 Agente/Representante:

TOMAS GIL, Tesifonte Enrique

ES 2 468 366 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

Polipéptidos con actividad de mejora celulolítica y polinucleótidos que los codifican

5 **Declaración sobre Derechos de Invencciones realizadas por medio de Investigación y Desarrollo patrocinados federalmente**

10 [0001] Esta invención se realizó con el apoyo del Gobierno bajo subcontrato de NREL N° ZCO-30017-02, Contrato Principal DE-AC36-98GO10337 adjudicado por el Departamento de Energía. El gobierno tiene ciertos derechos sobre esta invención.

Antecedentes de la invención

15 **Campo de la invención**

[0002] La presente invención se refiere a polipéptidos aislados con actividad de mejora celulolítica y polinucleótidos aislados que codifican los polipéptidos. La invención también se refiere a construcciones de ácidos nucleicos, vectores y células huéspedes que comprenden los polinucleótidos al igual que métodos para producir y usar los polipéptidos.

20 **Descripción del Estado de la técnica relacionado**

25 [0003] La celulosa es un polímero de la glucosa de azúcar simple unida de manera covalente por enlaces beta-1,4. Muchos microorganismos producen enzimas que hidrolizan glucanos beta-unidos. Estas enzimas incluyen endoglucanasas, celobiohidrolasas y beta-glucosidasas. Las endoglucanasas digieren el polímero de celulosa en lugares aleatorios, abriéndolo para atacar con celobiohidrolasas. Las celobiohidrolasas liberan secuencialmente moléculas de celobiosa de los extremos del polímero de celulosa. La celobiosa es un dímero de glucosa hidrosoluble beta-1,4-unido. Las beta-glucosidasas hidrolizan celobiosa a glucosa.

30 [0004] La conversión de materias primas celulósicas en etanol tiene las ventajas de la disponibilidad inmediata de cantidades grandes de materia prima, la conveniencia de evitar quema o vertido de los materiales y la limpieza del combustible de etanol. Madera, residuos agrícolas, brotes herbáceos y desperdicios sólidos municipales se han considerado como materias primas para producción de etanol. Estos materiales principalmente consisten en celulosa, hemicelulosa y lignina. Una vez la celulosa se convierte en glucosa, la glucosa es fácilmente fermentada por levadura en etanol.

35 [0005] Sería ventajoso en la técnica para mejorar la conversión de materias primas celulósicas.

40 [0006] Es un objeto de la presente invención proporcionar polipéptidos aislados con actividad de mejora celulolítica y secuencias de ácidos nucleicos aisladas que codifican los polipéptidos para mejorar la conversión de materias primas celulósicas.

Resumen de la invención

45 [0007] La presente invención se refiere a un polipéptido aislado con actividad de mejora celulolítica y una secuencia de aminoácidos que tiene al menos 85%, más preferiblemente al menos 90%, y de la forma más preferible al menos 97% de identidad con los aminoácidos 19 a 226 de la SEC ID n°: 8 .

50 [0008] La presente invención también se refiere a polinucleótidos aislados que codifican un polipéptido según la reivindicación 1 con actividad de mejora celulolítica.

[0009] La presente invención también se refiere a construcciones de ácidos nucleicos, vectores de expresión recombinantes y células huéspedes recombinantes comprendiendo los polinucleótidos.

55 [0010] La presente invención también se refiere a métodos para producir tal polipéptido con actividad de mejora celulolítica comprendiendo (a) cultivo de una célula huésped recombinante comprendiendo una construcción de ácidos nucleicos comprendiendo un polinucleótido que codifica el polipéptido bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.

60 [0011] La presente invención también se refiere a métodos para degradar o convertir un material celulósico, que comprende: tratar el material celulósico con una cantidad eficaz de una proteína celulolítica en presencia de una cantidad eficaz de un polipéptido que tiene actividad de mejora celulolítica, en el que la presencia del polipéptido que tiene actividad de mejora celulolítica aumenta la degradación del material celulósico en comparación con la ausencia del polipéptido que tiene actividad de mejora celulolítica.

65 [0012] La presente invención se refiere además a métodos para producir una sustancia orgánica, que comprende: (a) sacarificar un material celulósico con una cantidad eficaz de una proteína celulolítica en presencia de una cantidad

eficaz de un polipéptido que tiene actividad de mejora celulolítica, en el que la presencia del polipéptido que tiene actividad de mejora celulolítica aumenta la degradación de material celulósico en comparación con la ausencia del polipéptido que tiene actividad de mejora celulolítica;

(b) fermentar el material celulósico sacarificado del paso (a) con uno o más microorganismos fermentadores; y

5 (c) recuperar la sustancia orgánica de la fermentación.

Breve descripción de las figuras

[0013]

10 Las figuras 1A y 1 B muestran la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un polipéptido de *Thielavia terrestris* NRRL 8126 GH61B con actividad de mejora celulolítica (SEC ID n°: 1 y 2, respectivamente). Intrones predichos están en cursiva. El péptido señal predicho está subrayado.

15 La figura 2 muestra la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un polipéptido de *Thielavia terrestris* NRRL 8126 GH61C con actividad de mejora celulolítica (SEC ID n°: 3 y 4, respectivamente). Intrones predichos están en cursiva. El péptido señal predicho está subrayado.

La figura 3 muestra la secuencia de ADN genómico y la secuencia de aminoácidos deducida de un polipéptido de *Thielavia terrestris* NRRL 8126 GH61 D con actividad de mejora celulolítica (SEC ID n°: 5 y 6, respectivamente). Intrones predichos están en cursiva. El péptido señal predicho está subrayado.

20 La figura 4 muestra la secuencia de ADNc y la secuencia de aminoácidos deducida de un polipéptido de *Thielavia terrestris* NRRL 8126 GH61 E con actividad de mejora celulolítica (SEC ID n°: 7 y 8, respectivamente). El péptido señal predicho está subrayado.

La figura 5 muestra la secuencia de ADNc y la secuencia de aminoácidos deducida de un polipéptido *Thielavia terrestris* NRRL 8126 GH61 G con actividad de mejora celulolítica (SEC ID n°: 9 y 10, respectivamente). El péptido señal predicho está subrayado.

25 La Figura 6 muestra un mapa de restricción de pAILo1.

La Figura 7 muestra un mapa de restricción de pBANE10.

La Figura 8 muestra un mapa de restricción de pAILo2.

La Figura 9 muestra un mapa de restricción de pPH27.

La Figura 10 muestra un mapa de restricción de pPH29.

30 La Figura 11 muestra un mapa de restricción de pPH28.

La Figura 12 muestra un mapa de restricción de pEJG120.

La Figura 13 muestra un mapa de restricción de pEJG111.

La Figura 14 muestra un mapa de restricción de pEJG112.

La Figura 15 muestra un mapa de restricción de pTter61C.

35 La Figura 16 muestra un mapa de restricción de pTter61D.

La Figura 17 muestra un mapa de restricción de pTter61E.

La Figura 18 muestra un mapa de restricción de pTter61G.

La Figura 19 muestra un mapa de restricción de pEJG108.

La Figura 20 muestra un mapa de restricción de pAILo24.

40 La Figura 21 muestra un mapa de restricción de pAILo28.

La Figura 22 muestra un mapa de restricción de pMJ04.

La Figura 23 muestra un mapa de restricción de pMJ06.

La Figura 24 muestra un mapa de restricción de pMJ09.

La Figura 25 muestra un mapa de restricción de pEmY10.

45 La Figura 26 muestra un mapa de restricción de pSMAI155.

La Figura 27 muestra un mapa de restricción de pEJG110.

La Figura 28 muestra un mapa de restricción de pEJG114.

La Figura 29 muestra un mapa de restricción de pCW076.

La Figura 30 muestra un mapa de restricción de pCW078.

50 La Figura 31 muestra un mapa de restricción de pEJG97.

Las Figuras 32A y 32B muestran la secuencia de ADN genómico y la secuencia de aminoácidos deducida de una beta-glucosidasa de *Aspergillus fumigatus* (SEC ID n°: 75 y 76, respectivamente). El péptido señal predicho está subrayado e intrones predichos están en cursiva.

La Figura 33 muestra un mapa de restricción de pEJG107.

55 La Figura 34 muestra el efecto de un polipéptido *Thielavia terrestris* GH61 B con actividad de mejora celulolítica (1 mg/g PCS) en la hidrólisis de PCS (1% p/v) por caldo de fermentación de *Trichoderma reesei* que expresa actividad y una beta-glucosidasa de *Aspergillus fumigatus* (5 mg/g PCS) a 50-60°C.

60 La Figura 35 muestra el efecto de un polipéptido de *Thielavia terrestris* GH61B con actividad de mejora celulolítica (1 mg/g PCS) en la hidrólisis de PCS (1% p/v) por caldo de fermentación de *Trichoderma reesei* que expresa actividad celulolítica de caldo de fermentación y una beta-glucosidasa de *Aspergillus fumigatus* (5 mg/g PCS) a 50°C.

La Figura 36 muestra conversión de celulosa por Tr/AoBG sola o Tr/AoBG suplementada con polipéptidos de *Thielavia terrestris* GH61 con actividad celulolítica a 50°C, pH 5,0 durante 115 horas.

Definiciones

65

[0014] El término "actividad de mejora celulolítica" se define aquí como una actividad biológica que aumenta la hidrólisis de un material celulósico por proteínas con actividad celulolítica. Para fines de la presente invención, actividad de mejora celulolítica se determina por medición del aumento en azúcares reductores de la hidrólisis de un material celulósico por proteína celulolítica bajo las siguientes condiciones: 5,0 mg de proteína celulolítica/g de celulosa en PCS durante 5-7 días a 50°C en presencia y ausencia de 0,01-2,5 mg de actividad de mejora celulolítica por g de celulosa en PCS en comparación con una hidrólisis de control con carga igual de proteína total sin actividad de mejora celulolítica (5,01-7,5 mg de proteína celulolítica/g de celulosa en PCS). En un aspecto preferido, una mezcla de muestra de Celluclast® 1,5L (Novozymes A/S, Bagsværd, Dinamarca) en presencia de 3% de beta-glucosidasa de *Aspergillus oryzae* (producida de manera recombinante en *Aspergillus oryzae* según la WO 02/095014) o 3% de beta-glucosidasa de *Aspergillus fumigatus* (producida de manera recombinante en *Aspergillus oryzae* según el ejemplo 22) de carga de proteína celulosa se usa como la fuente de la actividad celulolítica.

[0015] El término "actividad celulolítica" se define aquí como una actividad biológica que hidroliza un material celulósico. Para fines de la presente invención, actividad celulolítica se determina por medición del aumento en la hidrólisis de un material celulósico por una mezcla celulolítica bajo las siguientes condiciones: 1-10 mg de proteína celulolítica/g de celulosa en PCS durante 5-7 días a 50°C en comparación con una hidrólisis de control sin adición de proteína celulolítica. En un aspecto preferido, una mezcla de muestra de Celluclast® 1,5L (Novozymes A/S, Bagsværd, Dinamarca) en presencia de 3% de beta-glucosidasa de *Aspergillus oryzae* (producida de manera recombinante en *Aspergillus oryzae* según la WO 02/095014) o 3% de beta glucosidasa de *Aspergillus fumigatus* (producida de manera recombinante en *Aspergillus oryzae* según el ejemplo 22) de carga de proteína celulosa como la fuente de la actividad celulolítica.

[0016] El término "PCS" o "restos de maíz pretratado" se define aquí como un material celulósico derivado de restos de maíz por tratamiento con calor y ácido diluido. Para fines de la presente invención, PCS se hace por el método descrito en el ejemplo 24, o variaciones del mismo en tiempo, temperatura y cantidad de ácido.

[0017] El término "glucósido hidrolasa de la Familia 61" o "Familia GH61" se define aquí como un polipéptido que entra en la Familia 61 de glucósido hidrolasa según Henrissat B., 1991, *A classification of glycosyl hydrolases based on amino- acid sequence similarities*, *Biochem. J.* 280: 309-316 y Henrissat B., y Bairoch A., 1996, *Updating the sequence-based classification of glycosyl hydrolases*, *Biochem. J.* 316: 695-696. Actualmente, Henrissat enumera la Familia GH61 como sin clasificar indicando que propiedades tales como mecanismo, nucleófilo catalítico/base, donantes de protones catalíticos y estructura 3-D no se conocen para polipéptidos pertenecientes a esta familia.

[0018] El material celulósico puede ser cualquier material con celulosa. La celulosa se encuentra generalmente, por ejemplo, en los tallos, hojas, vainas, cáscaras y mazorcas de plantas o hojas, derivaciones y madera de árboles. El material celulósico puede también ser, pero no está limitado a, material herbáceo, residuos agrícolas, residuos de silvicultura, residuos de sólidos municipales, papel usado y pulpa y residuos de fábrica de papel. Se entiende aquí que la celulosa puede estar en forma de lignocelulosa, un material de pared celular vegetal con lignina, celulosa y hemicelulosa en una matriz mezclada.

[0019] En una forma de realización preferida, el material celulósico es restos de maíz. En otra forma de realización preferida, el material celulósico es fibra de maíz. En otra forma de realización preferida, el material celulósico es paja de arroz. En otra forma de realización preferida, el material celulósico es residuos de tratamiento de papel y de pulpa. En otra forma de realización preferida, el material celulósico es plantas boscosas y herbáceas. En otra forma de realización preferida, el material celulósico es bagazo.

[0020] El material celulósico se puede utilizar como tal o se puede someter a pretratamiento, usando métodos convencionales conocidos en la técnica. Por ejemplo, técnicas de pretratamiento físicas puede incluir varios tipos de fresado, irradiación, explosión por vapor e hidrotérmolisis; técnicas de pretratamiento químicas pueden incluir ácido diluido, solvente alcalino orgánico, amoníaco, dióxido de azufre, dióxido de carbono y hidrotérmolisis con pH controlado; y técnicas de pretratamiento biológicas pueden implicar aplicación de microorganismos de solubilización de lignina (ver, por ejemplo, Hsu, T.-A., 1996, *Pretreatment of biomass*, en *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212; Ghosh, P., y Singh, A., 1993, *Physicochemical and biological treatments for enzymatic/microbial conversion of lignocellulosic biomass*, *Adv. Appl. Microbiol.* 39: 295-333; McMillan, J. D., 1994, *Pretreating lignocellulosic biomass: una revisión*, en *Enzymatic Conversion of Biomass for Fuels Production*, Himmel, M. E., Baker, J. O., y Overend, R. P., eds., ACS Symposium Series 566, American Chemical Society, Washington, DC, capítulo 15; Gong, C. S., Cao, N. J., Du, J., y Tsao, G. T., 1999, *Ethanol production from renewable resources*, in *Advances in Biochemical Engineering/Biotechnology*, Scheper, T., ed., Springer-Verlag Berlin Heidelberg, Germany, 65: 207-241; Olsson, L., y Hahn-Hagerdal, B., 1996, *Fermentation of lignocellulosic hydrolysates for ethanol production*, *Enz. Microb. Tech.* 18: 312-331; y Vallander, L., y Eriksson, K.-E. L., 1990, *Production of ethanol from lignocellulosic materials: Estado de la técnica*, *Adv. Biochem. Eng./Biotechnol.* 42: 63-95).

[0021] Los polipéptidos de la presente invención tienen al menos 20%, preferiblemente al menos 40%, más preferiblemente al menos 50%, más preferiblemente al menos 60%, más preferiblemente al menos 70%, más preferiblemente al menos 80%, incluso más preferiblemente al menos 90%, de forma más preferida al menos 95%, e

incluso de la forma más preferida al menos 100% de la actividad de mejora celulolítica del polipéptido que consiste en la secuencia de aminoácidos mostrada como los aminoácidos 19 a 226 de la SEC ID n°: 8.

5 [0022] **Polipéptido aislado:** el término "polipéptido aislado" como se usa aquí se refiere a un polipéptido que es al menos 20% puro, preferiblemente al menos 40% puro, más preferiblemente al menos 60% puro, incluso más preferiblemente al menos 80% puro, de la forma más preferida al menos 90% puro, e incluso de la forma más preferida al menos 95% puro, como se determina por SDS-PAGE.

10 [0023] **Polipéptido substancialmente puro:** el término "polipéptido substancialmente puro" denota aquí una preparación de polipéptido que contiene como mucho 10%, preferiblemente como mucho 8%, más preferiblemente como mucho 6%, más preferiblemente como mucho 5%, más preferiblemente como mucho 4%, más preferiblemente como mucho 3%, incluso más preferiblemente como mucho 2%, de forma más preferida como mucho 1%, e incluso de forma más preferida como mucho 0,5% en peso de otro material de polipéptido con el cual se asocia originalmente. Se prefiere, por lo tanto, que el polipéptido substancialmente puro sea al menos 92% puro, preferiblemente al menos 94% puro, más preferiblemente al menos 95% puro, más preferiblemente al menos 96% puro, más preferiblemente al menos 96% puro, más preferiblemente al menos 97% puro, más preferiblemente al menos 98%, puro incluso más preferiblemente al menos 99%, de forma más preferida al menos 99,5% puro, e incluso de forma más preferida 100% puro en peso del material de polipéptido total presente en la preparación.

20 [0024] Los polipéptidos de la presente invención están preferiblemente en una forma substancialmente pura. En particular, se prefiere que los polipéptidos estén en "forma esencialmente pura", es decir, que la preparación de polipéptido esté esencialmente libre de otro material de polipéptido con el cual se asocia originalmente. Esto se puede lograr, por ejemplo, preparando el polipéptido mediante métodos recombinantes o por métodos de purificación tradicionales bien conocidos.

25 [0025] Aquí, el término "polipéptido substancialmente puro" es sinónimo de los términos "polipéptido aislado" y "polipéptido en forma aislada".

30 [0026] **Identidad:** la relación entre dos secuencias de aminoácidos o entre dos secuencias de nucleótidos se describe por el parámetro "identidad".

35 [0027] Para fines de la presente invención, el grado de identidad entre dos secuencias de aminoácidos se puede determinar por el método Clustal (Higgins, 1989, *CABIOS* 5: 151-153) usando el software LASERGENE™ MEGALIGN™ (DNASTAR, Inc., Madison, WI) con una tabla de identidad y los siguientes parámetros de alineamiento múltiples: penalización de espacio de 10 y penalización de longitud del espacio de 10. Parámetros de alineamiento por parejas son Ktuple=1, penalización de espacio=3, ventanas=5 y diagonales=5. El grado de identidad entre dos secuencias de aminoácidos pueden ser también determinado por el algoritmo de Smith-Waterman (Waterman et al., 1976, *Adv.Math.* 20: 367) con una penalización abierta de espacio de 11, una penalización de extensión del espacio de 1 y la matriz BLOSUM62.

40 [0028] Para fines de la presente invención, el grado de identidad entre dos secuencias de nucleótidos se determina por el método Wilbur-Lipman (Wilbur and Lipman, 1983, *Proceedings of the National Academy of Science USA* 80: 726-730) usando el software LASERGENE™ MEGALIGN™ (DNASTAR, Inc., Madison, WI) con una tabla de identidad y los siguientes parámetros de alineamiento múltiples: penalización de espacio de 10 y penalización de longitud del espacio de 10. Parámetros de alineamiento por parejas son Ktuple=3, penalización de espacio=3 y ventanas=20.

50 [0029] **Fragmento de polipéptido:** el término "fragmento de polipéptido" se define aquí como un polipéptido con uno o más aminoácidos eliminados del amino y/o carboxi terminal de los aminoácidos 19 a 226 de la SEC ID n°: 8 o una secuencia homóloga de esta, donde el fragmento tiene actividad de mejora celulolítica. Preferiblemente, un fragmento de los aminoácidos 19 a 226 de la SEC ID n°: 8 contiene al menos 175 residuos de aminoácidos, más preferiblemente al menos 185 residuos de aminoácidos y de forma más preferida al menos 195 residuos de aminoácidos.

55 [0030] **Subsecuencia:** el término "subsecuencia" se define aquí como una secuencia de nucleótidos con uno o más nucleótidos eliminados del extremo 5' y/o 3' de los nucleótidos 55 a 678 de la SEC ID n°7; o una secuencia homóloga de la misma, en la que la subsecuencia codifica un fragmento de polipéptido que tiene actividad de mejora celulolítica. Preferiblemente, una subsecuencia de nucleótidos 55-678 de la SEQ ID N °: 7 contiene al menos 525 nucleótidos, más preferiblemente al menos 555 nucleótidos, y más preferiblemente al menos 585 nucleótidos.

60 [0031] **Variante alélica:** el término "variante alélica" denota aquí cualquiera de dos o más formas alternativas de un gen que ocupa el mismo locus cromosómico. La variación alélica surge naturalmente a través de mutación y puede suponer polimorfismo dentro de poblaciones. Las mutaciones genéticas pueden ser silenciosas (ningún cambio en el polipéptido codificado) o pueden codificar polipéptidos con secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

65 [0032] **Polinucleótido aislado:** el término "polinucleótido aislado" como se usa aquí se refiere a un polinucleótido que es al menos 20% puro, preferiblemente al menos 40% puro, más preferiblemente al menos 60% puro, incluso más

preferiblemente al menos 80% puro, de forma más preferida al menos 90% puro, e incluso de forma más preferida al menos 95% puro, como determinado por electroforesis de agarosa.

5 [0033] **Polinucleótido substancialmente puro:** el término "polinucleótido substancialmente puro" como se usa aquí se refiere a una preparación polinucleótida libre de otros nucleótidos indeseados o extraños y en una forma adecuada para uso dentro de sistemas de producción de proteína creada genéticamente. Así, un polinucleótido substancialmente puro contiene como mucho 10%, preferiblemente como mucho 8%, más preferiblemente como mucho 6%, más preferiblemente como mucho 5%, más preferiblemente como mucho 4%, más preferiblemente como mucho 3%, incluso más preferiblemente como mucho 2%, de forma más preferida como mucho 1%, e incluso de forma más preferida como mucho 0,5% en peso de otro material polinucleótido con el cual se asocia originalmente. Un polinucleótido substancialmente puro puede, no obstante, incluir 5' y 3' regiones no codificantes de origen natural, tales como promotores y terminadores. Se prefiere que el polinucleótido substancialmente puro sea al menos 90% puro, preferiblemente al menos 92% puro, más preferiblemente al menos 94% puro, más preferiblemente al menos 95% puro, más preferiblemente al menos 96% puro, más preferiblemente al menos 97% puro, incluso más preferiblemente al menos 98% puro, de forma más preferida al menos 99%, e incluso de forma más preferida al menos 99,5% puro en peso. Los polinucleótidos de la presente invención están preferiblemente en una forma substancialmente pura. En particular, se prefiere que los polinucleótidos descritos aquí estén en "forma esencialmente pura", es decir, que la preparación polinucleótida esté esencialmente libre de otro material polinucleótido con el cual se asocia originalmente. Aquí, el término "polinucleótido substancialmente puro" es sinónimo de los términos "polinucleótido aislado" y "polinucleótido en forma aislada". Los polinucleótidos pueden ser de origen genómico, ADNc, ARN, semisintético, sintético, o cualquier combinaciones de estos.

25 [0034] **ADNc:** el término "ADNc" se define aquí como una molécula de ADN que se puede preparar por transcripción inversa de una molécula de ARNm maduro dividido obtenido de una célula eucariota. ADNc carece de secuencias de intrones que están presentes normalmente en el ADN genómico correspondiente. La transcripción inicial de ARN primario es un precursor para ARNm que se procesa a través de una serie de pasos antes de aparecer como ARNm maduro dividido. Estos pasos incluyen la eliminación de secuencias de intrones por un proceso llamado empalme. ADNc derivado de ARNm carece, por lo tanto, de cualquier secuencia de intrones.

30 [0035] **Construcción de ácidos nucleicos:** el término "construcción de ácidos nucleicos" como se utiliza en este caso se refiere a una molécula de ácido nucleico, mono o bicatenaria, que se aísla de un gen de origen natural o que se modifica para contener segmentos de ácidos nucleicos de una manera que de lo contrario no existirían en la naturaleza. El término construcción de ácidos nucleicos es sinónimo del término "casete de expresión" cuando la construcción de ácidos nucleicos contiene las secuencias de control requeridas para la expresión de una secuencia codificante de la presente invención.

35 [0036] **Secuencia de control:** el término "secuencias de control" se define aquí para incluir todos los componentes que son necesarios o ventajosos para la expresión de un polinucleótido que codifica un polipéptido de la presente invención. Cada secuencia de control puede ser nativa o extranjera a la secuencia de nucleótidos que codifica el polipéptido o extranjero o nativo entre sí. Tales secuencias de control incluyen, pero de forma no limitativa, una guía, secuencia de poliadenilación, secuencia de propéptido, promotor, secuencia de péptido señal y terminador de transcripción. A un mínimo, las secuencias de control incluyen un promotor y señales de parada traduccionales y transcripcionales. Las secuencias de control se pueden proporcionar con enlazadores con el propósito de introducir sitios de restricción que facilitan ligamiento de las secuencias de control con la región codificante de la secuencia de nucleótidos que codifica un polipéptido.

50 [0037] **Operativamente unido:** el término "operativamente unido" denota aquí una configuración en que una secuencia de control se coloca en una posición apropiada en relación a la secuencia de codificación de la secuencia polinucleótida de tal manera que la secuencia de control dirige la expresión de la secuencia codificante de un polipéptido.

55 [0038] **Secuencia codificante:** cuando se usa aquí el término "secuencia codificante" significa una secuencia de nucleótidos que especifica directamente la secuencia de aminoácidos de su producto proteínico. Los bordes de la secuencia codificante son generalmente determinados por un marco de lectura abierto, que empieza normalmente con el codón de inicio ATG o codones de inicio alternativos tales como TTG y GTG y extremos con un codón de terminación tal como TAA, TAG y TGA. La secuencia codificante puede ser una secuencia de ADN, ADNc, o de nucleótidos recombinante.

60 [0039] **Expresión:** el término "expresión" incluye cualquier paso implicado en la producción de un polipéptido incluyendo, pero no limitado a, transcripción, modificación postranscripcional, traducción, modificación postraduccional y secreción.

65 [0040] **Vector de expresión:** el término "vector de expresión" se define aquí como una molécula de ADN circular o lineal que comprende un polinucleótido que codifica un polipéptido de la invención y que se une operativamente a nucleótidos adicionales que proporcionan su expresión.

[0041] **Célula huésped:** el término "célula huésped", como se utiliza en este caso, incluye cualquier tipo de célula que es susceptible de transformación, transfección, transducción y similares con una construcción de ácidos nucleicos o vector de expresión comprendiendo un polinucleótido de la presente invención.

5 [0042] **Modificación:** el término "modificación" significa aquí cualquier modificación química del polipéptido que comprende o que consiste en los aminoácidos 19 a 226 de la SEC ID n°: 8 o una secuencia homóloga de esta, al igual que manipulación genética del ADN que codifica este polipéptido. La modificación puede ser sustituciones, deleciones y/o inserciones de uno o más aminoácidos al igual que reemplazos de una o más cadenas laterales de aminoácidos.

10 [0043] **Variante artificial:** cuando se usa aquí, el término "variante artificial" significa un polipéptido con actividad de mejora celulolítica producido por un organismo que expresa una secuencia de nucleótidos modificada de la SEC ID n°: 7 o una secuencia homóloga de esta, o la región codificante madura de esta. La secuencia de nucleótidos modificada se obtiene a través de intervención humana por modificación de la secuencia de nucleótidos descrita en la SEC ID n°: 7 o una secuencia homóloga de esta, o la región codificante madura de esta.

15

Descripción detallada de la invención

Polipéptidos con actividad de mejora celulolítica

20 [0044] En un aspecto, la presente descripción se refiere a polipéptidos aislados que tienen actividad de mejora celulolítica, que comprende los siguientes motivos:

[ILMV]-P-X(4,5)-G-X-Y-[ILMV]-X-R-X-[EQ]-X(4)-[HNQ] y [FW]-[TF]-K- [AIV],

en donde X es cualquier aminoácido, X(4,5) es cualquier aminoácido en 4 o 5 posiciones contiguas, y X(4) es cualquier aminoácido en 4 posiciones contiguas.

25 [0045] El polipéptido aislado que comprende los motivos indicados anteriormente puede comprender además:

H-X (1,2)-G-P-X (3)-[YW]-[AILMV],

[EQ]-X-Y-X (2)-C-X-[EHQN]-[FILV]-X-[ILV], o

HX (1,2)-GPX (3)-[YW]-[AILMV] y [EQ]-XYX (2)-CX-[EHQN]-[FILV]-X-[ILV],

30 en donde X es cualquier aminoácido, X(1,2) es cualquier aminoácido en la posición 1 o 2 posiciones contiguas, X(3) es cualquier aminoácido en 3 posiciones contiguas, y X(2) es cualquier aminoácido en 2 posiciones contiguas. En los motivos anteriores, se emplea la abreviatura de aminoácidos de una sola letra de la IUPAC aceptada.

[0046] En una forma de realización preferida, el polipéptido aislado que tiene actividad de mejora celulolítica comprende además H-X(1,2)-G-P-X(3)-[YW]-[AILMV]. En otra forma de realización preferida, el polipéptido aislado que tiene actividad de mejora celulolítica comprende además [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV]. En otra forma de realización preferida, el polipéptido aislado que tiene actividad de mejora celulolítica comprende además H-X(1,2)-G-P-X(3)-[YW]-[AILMV] y [EQ]-X-Y-X(2)-C-X-[EHQN]-[FILV]-X-[ILV].

40 [0047] La presente invención se refiere a polipéptidos aislados con actividad de mejora celulolítica, y con una secuencia de aminoácidos que tiene un grado de identidad a los aminoácidos 19 a 226 de la SEC ID n°: 8 (es decir, el polipéptido maduro) de al menos 85%, más preferiblemente al menos 90%, de la forma más preferible al menos 95%, e incluso de forma más preferida al menos 97%, 98%, o 99%, que tienen actividad de mejora celulolítica (de ahora en adelante "polipéptidos homólogos"). En un aspecto preferido, los polipéptidos homólogos tienen una secuencia de aminoácidos que difiere por diez aminoácidos, preferiblemente por cinco aminoácidos, más preferiblemente por cuatro aminoácidos, incluso más preferiblemente por tres aminoácidos, de forma más preferida por dos aminoácidos, e incluso de forma más preferida por un aminoácido de los aminoácidos 19 a 226 de la SEC ID n°: 8.

[0048] Un polipéptido de la presente invención comprende preferiblemente la secuencia de aminoácidos de la SEC ID N°: 8 o una variante alélica de la misma; o un fragmento de la misma que tiene actividad de mejora celulolítica. En un aspecto preferido, un polipéptido comprende la secuencia de aminoácidos de la SEC ID N°: 8. En otro aspecto preferido, un polipéptido comprende los aminoácidos 19 a 226 de la SEC ID N°: 8, o una variante alélica de la misma; o un fragmento de la misma que tiene actividad de mejora celulolítica. En otro aspecto preferido, un polipéptido comprende los aminoácidos 19 a 226 de la SEC ID N°: 8. En otro aspecto preferido, un polipéptido consiste en la secuencia de aminoácidos de la SEC ID N°: 8 o una variante alélica de la misma; o un fragmento de la misma que tiene actividad de mejora celulolítica. En otro aspecto preferido, un polipéptido consiste en la secuencia de aminoácidos de la SEC ID N°: 8. En otro aspecto preferido, un polipéptido consiste en los aminoácidos 19 a 226 de la SEC ID N°: 8 o una variante alélica de la misma; o un fragmento de la misma que tiene actividad de mejora celulolítica. En otro aspecto preferido, un polipéptido consiste en los aminoácidos 19 a 226 de la SEC ID N°: 8.

60 [0049] En otro aspecto, la presente descripción se refiere a polipéptidos aislados con actividad de mejora celulolítica que son codificados por polinucleótidos que hibridan bajo condiciones de astringencia muy baja, preferiblemente condiciones de astringencia baja, condiciones de astringencia más preferiblemente medias, más preferiblemente condiciones de astringencia medias-altas, incluso más preferiblemente condiciones de astringencia altas, y de la forma más preferida condiciones de astringencia muy altas con (i) los nucleótidos 55 a 678 de la SEC ID n°: 7, (ii) la secuencia de ADNc contenida en los nucleótidos 98 a 821 de la SEC ID n°: 3, (iii) una subsecuencia de (i) o (ii), o (iv) una cadena complementaria de (i), (ii), o (iii) (J. Sambrook, E.F. Fritsch, and T. Maniatus, 1989, *Molecular Cloning, A Laboratory*

65

Manual, 2d edition, Cold Spring Harbor, New York). Una subsecuencia de la SEC ID N°: 7, contiene al menos 100 nucleótidos contiguos o preferiblemente al menos 200 nucleótidos contiguos. Por otra parte, la subsecuencia puede codificar un fragmento de polipéptido que tiene actividad de mejora celulolítica.

5 [0050] La secuencia de nucleótidos de la SEC ID n°: 7; o una subsecuencia de ésta, así como la secuencia de aminoácidos de la SEC ID n°: 8 o un fragmento de esta, se puede utilizar para diseñar una sonda de ácidos nucleicos para identificar y clonar ADN que codifica polipéptidos con actividad de mejora celulolítica de cepas de diferentes géneros o especies según métodos bien conocidos en la técnica. En particular, tales sondas se pueden usar para hibridación con el genómico o ADNc del género o especie de interés, seguido de procedimientos de transferencia de Southern estándar, para identificar y aislar el gen correspondiente en este. Tales sondas pueden ser considerablemente más cortas que la secuencia entera, pero deberían ser al menos 14, preferiblemente al menos 25, más preferiblemente al menos 35 y de forma más preferida al menos 70 nucleótidos en longitud. Se prefiere, no obstante, que la sonda de ácidos nucleicos sea al menos 100 nucleótidos en longitud. Por ejemplo, la sonda de ácidos nucleicos puede ser al menos 200 nucleótidos, preferiblemente al menos 300 nucleótidos, más preferiblemente al menos 400 nucleótidos, o de forma más preferida al menos 500 nucleótidos en longitud. Sondas incluso más largas pueden ser utilizadas, por ejemplo, sondas de ácidos nucleicos que son al menos 600 nucleótidos, al menos preferiblemente al menos 700 nucleótidos, más preferiblemente al menos 800 nucleótidos, o de forma más preferida al menos 900 nucleótidos en longitud. Ambas sondas de ADN y ARN pueden usarse. Las sondas se marcan típicamente para detectar el gen correspondiente (por ejemplo, con ³²P, ³H, ³⁵S, biotina, o avidina). Tales sondas están comprendidas por la presente invención.

25 [0051] Un ADN genómico o genoteca de ADNc obtenido a partir de estos otros organismos puede, por lo tanto, seleccionarse para ADN que hibridiza con las sondas anteriormente descritas y que codifica un polipéptido con actividad de mejora celulolítica. Genómico u otro ADN de estos otros organismos se pueden separar por agarosa o electroforesis en gel de poliacrilamida, u otras técnicas de separación. ADN de las bibliotecas o el ADN separado se puede transferir e inmovilizar en nitrocelulosa u otro material portador adecuado. Para identificar un clon o ADN que es homólogo a la SEC ID n°: 7, o una subsecuencia de esta, el material portador se usa en una transferencia de Southern.

30 [0052] Para fines de la presente divulgación, hibridación indica que la secuencia de nucleótidos hibrida a una sonda de ácidos nucleicos marcada correspondiente a la secuencia de nucleótidos mostrada en la SEC ID n°: 7, la secuencia genómica que comprende la SEC ID n°: 7, su cadena complementaria, o una subsecuencia de esta, bajo condiciones de astringencia de muy bajas a muy altas. Moléculas a la que la sonda de ácidos nucleicos hibrida bajo estas condiciones pueden detectarse usando película radiográfica.

35 [0053] La sonda de ácidos nucleicos es una secuencia de ácidos nucleicos que codifica el polipéptido de la SEC ID n°: 8, o una subsecuencia de esta. En otra forma de realización, la sonda de ácidos nucleicos es la SEC ID N°: 7. En otra forma de realización la sonda de ácidos nucleicos es la región codificante del polipéptido maduro de la SEC ID n°: 7. En otra forma de realización, la sonda de ácidos nucleicos es la secuencia de ácidos nucleicos contenida en el plásmido PTter61E que está contenido en *Escherichia coli* NRRL B-30814, donde la secuencia de ácidos nucleicos codifica un polipéptido con actividad de mejora celulolítica. En otra forma de realización, la sonda de ácidos nucleicos es la región codificante del polipéptido maduro contenida en el plásmido PTter61E que está contenido en *Escherichia coli* NRRL B-30814.

45 [0054] Para sondas largas de al menos 100 nucleótidos en longitud, condiciones de astringencia de muy bajas a muy altas se definen como prehibridación e hibridación a 42°C en 5X SSPE, 0,3% de SDS, 200 µg/ml de ADN de esperma de salmón desnaturalizado y cortado, y bien 25% de formamida para astringencias bajas y muy bajas, 35% de formamida para astringencias medias y medias-altas, o 50% de formamida para astringencias altas y muy altas, seguido de procedimientos de transferencia de Southern estándar durante 12 a 24 horas óptimamente.

50 [0055] Para sondas largas de al menos 100 nucleótidos en longitud, el material portador se lava finalmente tres veces cada durante 15 minutos usando 2X SSC, 0,2% de SDS preferiblemente al menos a 45°C (astringencia muy baja), más preferiblemente al menos a 50°C (astringencia baja), más preferiblemente al menos a 55°C (astringencia media), más preferiblemente al menos a 60°C (astringencia media-alta), incluso más preferiblemente al menos a 65°C (astringencia alta), y más preferiblemente al menos a 70°C (astringencia muy alta).

55 [0056] Para sondas cortas que son aproximadamente de 15 nucleótidos a aproximadamente 70 nucleótidos en longitud, condiciones de astringencia se definen como prehibridación, hibridación y post-hibridación de lavado de aproximadamente 5°C a aproximadamente 10°C por debajo de la T_m calculada usando el cálculo según Bolton y McCarthy (1962, *Proceedings of the National Academy of Sciences USA* 48:1390) en 0,9 M de NaCl, 0,09 M de Tris-HCl pH 7,6, 6 mM de EDTA, 0,5% de NP-40, solución de Denhardt 1X, 1 mM de pirofosfato de sodio, 1 mM de fosfato monobásico de sodio, 0,1 mM de ATP y 0,2 mg de levadura ARN por ml seguido de procedimientos de transferencia de Southern estándar durante 12 a 24 horas óptimamente.

65 [0057] Para sondas cortas que son aproximadamente de 15 nucleótidos a aproximadamente 70 nucleótidos en longitud, el material portador se lava una vez en 6X SCC más 0,1% de SDS durante 15 minutos y dos veces cada durante 15 minutos usando 6X SSC a de 5°C a 10°C por debajo de la T_m calculada.

- [0058] La divulgación se refiere a variantes artificiales comprendiendo una sustitución conservadora, eliminación, y/o inserción de uno o más aminoácidos. Preferiblemente, cambios de aminoácidos son de naturaleza secundaria, que es sustituciones de aminoácido conservadoras o inserciones que significativamente no afectan al pliegue y/o actividad de la proteína; deleciones pequeñas, típicamente de uno a aproximadamente 30 aminoácidos; pequeñas extensiones amino- o carboxi- terminales, tales como un residuo de metionina aminoterminal; un pequeño péptido de enlace de hasta aproximadamente 20-25 residuos; o una pequeña extensión que facilita purificación cambiando la carga de red u otra función, tal como un tracto de polihistidina, un epítipo antigénico o un dominio de unión.
- [0059] Ejemplos de sustituciones conservadoras están dentro del grupo de aminoácidos básicos (arginina, lisina e histidina), aminoácidos ácidos (ácido glutámico y ácido aspártico), aminoácidos polares (glutamina y asparagina), aminoácidos hidrofóbicos (leucina, isoleucina y valina), aminoácidos aromáticos (fenilalanina, triptófano y tirosina), y aminoácidos pequeños (glicina, alanina, serina, treonina y metionina). Sustituciones de aminoácidos que generalmente no alteran actividad específica se conocen en la técnica y se describen, por ejemplo, por H. Neurath y R.L. Hill, 1979, en *The Proteins, Academic Press, New York*. Los intercambios que ocurren más frecuentemente son Ala/Ser, Val/Ile, Asp/Glu, Thr/Ser, Ala/Gly, Ala/Thr, Ser/Asn, Ala/Val, Ser/Gly, Tyr/Phe, Ala/Pro, Lys/Arg, Asp/Asn, Leu/Ile, Leu/Val, Ala/Glu y Asp/Gly.
- [0060] Además de los 20 aminoácidos estándar, aminoácidos no estándar (tales como 4-hidroxi prolina, 6-N-metilo lisina, ácido 2-aminoisobutírico, isovalina, y alfa-metil serina) se pueden sustituir por residuos de aminoácidos de un polipéptido de tipo salvaje. Un número limitado de aminoácidos no conservadores, aminoácidos que no se codifican por el código genético y aminoácidos no naturales se pueden sustituir por residuos de aminoácidos. "Aminoácidos no naturales" han sido modificados después de síntesis de proteína, y/o tienen una estructura química en su cadena(s) lateral diferente de los aminoácidos estándar. Aminoácidos no naturales pueden sintetizarse químicamente y, preferiblemente, están disponibles comercialmente, e incluyen ácido pipercolico, tiazolidina ácido carboxílico, dehidroprolina, 3- y 4-metilprolina y 3,3-dimetilprolina.
- [0061] Alternativamente, los cambios de aminoácidos son de tal naturaleza que las propiedades físico-químicas de los polipéptidos son alteradas. Por ejemplo, cambios de aminoácidos pueden mejorar la termoestabilidad del polipéptido, alterar la especificidad del sustrato, cambiar el pH óptimo y similares.
- [0062] Aminoácidos esenciales en el polipéptido progenitor se pueden identificar según procedimientos conocidos en la técnica, tales como mutagénesis dirigida o mutagénesis de escaneado de alanina (Cunningham and Wells, 1989, *Science* 244: 1081- 1085). En esta última técnica, mutaciones de alanina únicas se introducen en cada residuo en la molécula y las moléculas resultantes mutantes se evalúan para actividad biológica (es decir, actividad de mejora celuloítica) para identificar residuos de aminoácidos que son muy importantes para la actividad de la molécula. Ver también, Hilton *et al.*, 1996, *J. Biol. Chem.* 271: 4699-4708. El sitio activo de la enzima u otra interacción biológica puede también determinarse por análisis físico de estructura, como determinado por tales técnicas como resonancia magnética nuclear, cristalografía, difracción electrónica, o marcaje por fotoafinidad, conjuntamente con mutación de aminoácidos de sitio de contactos putativo. Ver, por ejemplo, de Vos *et al.*, 1992, *Science* 255: 306- 312; Smith *et al.*, 1992, *J. Mol. Biol.* 224: 899-904; Wlodaver *et al.*, 1992, *FEBS Lett.* 309: 59-64. Las identidades de aminoácidos esenciales puede también inferirse de análisis de identidades con polipéptidos que se relacionan con un polipéptido según la invención.
- [0063] Sustituciones de aminoácidos múltiples o únicas pueden hacerse y evaluarse usando métodos conocidos de mutagénesis, recombinación, y/o redistribución, seguidos de un procedimiento de selección pertinente, tal como aquellos descritos por Reidhaar-Olson and Sauer, 1988, *Science* 241: 53-57; Bowie and Sauer, 1989, *Proc. Natl. Acad. Sci. USA* 86: 2152-2156; WO 95/17413; o WO 95/22625. Otros métodos que pueden usarse incluyen PCR con tendencia al error, visualización de fago (p. ej., Lowman *et al.*, 1991, *Biochem.* 30: 10832-10837; U.S. Patent nº 5,223,40; WO 92/06204) y mutagénesis dirigida por región (Derbyshire *et al.*, 1986, *Gene* 46: 145; Ner *et al.*, 1988, *DNA* 7: 127).
- [0064] Métodos de mutagénesis/redistribución se pueden combinar con alto rendimiento, métodos de selección automatizados para detectar actividad de polipéptidos clonados mutagenizados expresados por células huéspedes (Ness *et al.*, 1999, *Nature Biotechnology* 17: 893-896). Moléculas de ADN mutagenizadas que codifican polipéptidos activos se pueden recuperar de las células huéspedes y rápidamente secuenciar usando métodos estándar en la técnica. Estos métodos permiten la determinación rápida de la importancia de residuos de aminoácidos individuales en un polipéptido de interés y se pueden aplicar a polipéptidos de estructura desconocida.
- [0065] El número total de sustituciones de aminoácido, deleciones y/o inserciones de los aminoácidos 19 a 226 de la SEC ID nº: 8 es 10, preferiblemente 9, más preferiblemente 8, más preferiblemente 7, más preferiblemente como mucho 6, más preferiblemente 5, más preferiblemente 4, incluso más preferiblemente 3, de forma más preferida 2, e incluso de forma más preferida 1.
- [0066] En un aspecto preferido, el polipéptido es un polipéptido de *Thielavia terrestris*. En una forma de realización más preferida, el polipéptido es un polipéptido de *Thielavia terrestris* NRRL 8126, por ejemplo, el polipéptido con la secuencia de aminoácidos de la SEC ID nº: 8, o fragmentos de ésta, por ejemplo, la proteína madura.

[0067] Cepas de estas especies son accesibles fácilmente al público en un número de colecciones de cultivo, tal como la American Type Culture Collection (ATCC), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSM), Centraalbureau Voor Schimmelcultures (CBS) y Agricultural Research Service Patent Culture Collection, Northern Regional Research Center (NRRL).

[0068] Polipéptidos de la presente invención también incluyen polipéptidos fundidos o polipéptidos de fusión divisibles en los que otro polipéptido se funde al N-terminal o el C-terminal del polipéptido o fragmento de este. Un polipéptido fusionado se produce por fundición de una secuencia de nucleótidos (o una parte de esta) que codifica otro polipéptido a una secuencia de nucleótidos (o una parte de esta) de la presente invención. Técnicas para producir polipéptidos de fusión se conocen en la técnica, e incluyen ligar las secuencias de codificación que codifican los polipéptidos de modo que estas están en marco y que expresión del polipéptido fusionado está bajo control del mismo promotor(es) y terminador.

15 Polinucleótidos

[0069] La presente descripción también se refiere a polinucleótidos aislados con secuencias de nucleótidos que codifican polipéptidos de la presente invención.

La secuencia de nucleótidos se expone en la SEC ID n°: 7. En otro aspecto más preferido, la secuencia de nucleótidos es la secuencia contenida en el plásmido pTter61E que está contenido en *Escherichia coli* NRRL B-30814. La secuencia de nucleótidos es la región codificante del polipéptido maduro de la SEC ID n°: 8. En otro aspecto más preferido, la secuencia de nucleótidos es la región codificante del polipéptido maduro contenido en el plásmido pTter61E que está contenido en *Escherichia coli* NRRL B-30814. La presente invención también abarca secuencias de nucleótidos que codifican un polipéptido con la secuencia de aminoácidos de la SEC ID n°: 8 o el polipéptido maduro de ésta, que difiere de la SEC ID n°: 7 en virtud de la degeneración del código genético. Las técnicas usadas para aislar o clonar un polinucleótido que codifica un polipéptido se conocen en la técnica e incluyen aislamiento de ADN genómico, preparación de ADNc, o una combinación de estos. La clonación de los polinucleótidos de la presente invención de este ADN genómico puede ser efectuada, por ejemplo, usando reacción en cadena de polimerasa (PCR) bien conocida o seleccionando anticuerpos de bibliotecas de expresión para detectar fragmentos de ADN clonados con características estructurales compartidas. Ver, por ejemplo, Innis *et al.*, 1990, *PCR: A Guide to Methods and Application*, Academic Press, New York. Otros procedimientos de amplificación de ácido nucleico tal como reacción en cadena de la ligasa (LCR), transcripción activada ligada (LAT) y amplificación basada en secuencias de nucleótidos (NASBA) pueden utilizarse. Los polinucleótidos se pueden clonar de una cepa de *Thielavia*, u otra u organismo relacionado y así, por ejemplo, puede ser una variante alélica o de especies de la región codificante de polipéptido de la secuencia de nucleótidos.

[0070] La presente divulgación también se refiere a polinucleótidos con secuencias de nucleótidos que tienen un grado de identidad a la secuencia codificante del polipéptido maduro de la SEC ID n°: 7 (es decir, los nucleótidos 55 a 678) de al menos 75%, preferiblemente al menos 80%, más preferiblemente al menos 85%, incluso más preferiblemente al menos 90%, de forma más preferida al menos 95%, e incluso de forma más preferida al menos 97% de identidad, que codifica un polipéptido activo.

[0071] Modificación de una secuencia de nucleótidos que codifica un polipéptido de la presente invención puede ser necesaria para la síntesis de polipéptidos sustancialmente similares al polipéptido. El término "sustancialmente similar" al polipéptido se refiere a formas que no ocurren de manera natural del polipéptido. Estos polipéptidos pueden diferir en alguna manera creada genéticamente del polipéptido aislado de su fuente nativa, por ejemplo, variantes artificiales que difieren en la actividad específica, termostabilidad, pH óptimo, o similares. La secuencia variante puede construirse basándose en la secuencia de nucleótidos presentada como la región codificante del polipéptido de la SEC ID n°: 7, por ejemplo, una subsecuencia de esta, y/o por introducción de sustituciones de nucleótido que no dan lugar a otra secuencia de aminoácidos del polipéptido codificadas por la secuencia de nucleótidos, pero que corresponden al uso del codón del organismo huésped destinado para la producción de la enzima, o por introducción de sustituciones de nucleótidos que pueden dar lugar a una secuencia de aminoácidos diferente. Para una descripción general de sustitución de nucleótidos, ver, por ejemplo, Ford *et al.*, 1991, *Protein Expression and Purification 2: 95-107*.

[0072] Será aparente a los expertos en la técnica que estas sustituciones pueden hacerse fuera de las regiones más importantes a la función de la molécula y todavía suponer un polipéptido activo. Residuos de aminoácidos esenciales a la actividad del polipéptido codificados por un polinucleótido aislado de la invención y, por lo tanto, preferiblemente no sujeto a sustitución, se puede identificar según procedimientos conocidos en la técnica, tal como mutagénesis dirigida o mutagénesis de escaneado de alanina (ver, por ejemplo, Cunningham and Wells, 1989, *Science 244: 1081-1085*). En esta técnica, mutaciones se introducen en cada residuo positivamente cargado en la molécula y las moléculas mutantes resultantes se evalúan para actividad de mejora celulolítica para identificar residuos de aminoácidos que son muy importantes para la actividad de la molécula. Sitios de interacción sustrato-enzima pueden también ser determinados por análisis de la estructura tridimensional como determinado por tales técnicas como análisis de resonancia magnética nuclear, cristalografía o etiquetado de fotoafinidad (ver, por ejemplo, de Vos *et al.*, 1992, *Science 255: 306-312*; Smith *et al.*, 1992, *Journal of Molecular Biology 224: 899-904*; Wlodaver *et al.*, 1992, *FEBS Letters 309: 59-64*).

[0073] La presente divulgación también se refiere a polinucleótidos aislados que codifican un polipéptido de la presente invención, que se hibridan bajo condiciones de astringencia muy baja, preferiblemente condiciones de astringencia baja, más preferiblemente condiciones de astringencia media, más preferiblemente condiciones de astringencia media-alta, incluso más preferiblemente condiciones de astringencia alta, y de la forma más preferible condiciones de astringencia muy alta con (i) los nucleótidos 55-678 de la SEQ ID N°: 7 (ii) la secuencia de ADNc contenida en los nucleótidos 98-821 de la SEQ ID N°: 3, o (iii) una cadena complementaria de (i) o (ii); o variantes alélicas y subsecuencias de las mismas (Sambrook et al., 1989, supra), como se define aquí.

[0074] La presente divulgación también se refiere a polinucleótidos aislados obtenidos por (a) hibridación de una población de ADN bajo condiciones de astringencia muy baja, baja, media, media-alta, alta o muy alta con (i) los nucleótidos 55-678 de la SEQ ID NO: 7 (ii) la secuencia de ADNc contenida en los nucleótidos 126-978 de la SEC ID N°: 5 o (iii) una cadena complementaria de (i) o (ii); y (b) aislamiento del polinucleótido de hibridación, que codifica un polipéptido que tiene actividad de mejora celulolítica.

15 Construcciones de ácidos nucleicos

[0075] La presente invención también se refiere a construcciones de ácidos nucleicos comprendiendo un polinucleótido aislado de la presente invención operativamente unido a una o más secuencias de control que dirigen la expresión de la secuencia codificante en una célula huésped adecuada bajo condiciones compatibles con las secuencias de control.

[0076] Un polinucleótido aislado que codifica un polipéptido de la presente invención se puede manipular en una variedad de maneras para proveer expresión del polipéptido. Manipulación de la secuencia del polinucleótido antes de su inserción en un vector puede ser deseable o necesaria dependiendo del vector de expresión. Las técnicas para modificar secuencias polinucleótidas utilizando métodos de ADN recombinante se conocen bien en la técnica.

[0077] La secuencia de control puede ser una secuencia del promotor apropiada, una secuencia de nucleótidos que es reconocida por una célula huésped para la expresión de un polinucleótido que codifica un polipéptido de la presente invención. La secuencia del promotor contiene secuencias de control transcripcionales que median la expresión del polipéptido. El promotor puede ser cualquier secuencia de nucleótidos que muestra actividad transcripcional en la célula huésped de elección, incluyendo promotores mutantes, truncados e híbridos, y se pueden obtener de genes que codifican polipéptidos intracelulares o extracelulares bien homólogos o heterólogos a la célula huésped.

[0078] Ejemplos de promotores adecuados para dirigir la transcripción de las construcciones de ácidos nucleicos de la presente invención, especialmente en una célula huésped bacteriana, son los promotores obtenidos del operón lac de *E. coli*, gen de agarasa de *Streptomyces coelicolor* (dagA), gen de levansucrasa de *Bacillus subtilis* (sacB), gen de alfa-amilasa de *Bacillus licheniformis* (amyL), gen de amilasa maltogénica de *Bacillus stearothermophilus* (amyM), gen de alfa-amilasa de *Bacillus amyloliquefaciens* (amyQ), gen de penicilinas de *Bacillus licheniformis* (penP), genes xylA y xylB de *Bacillus subtilis* y gen de beta-lactamasa procariótica (Villa-Kamaroff et al., 1978, *Proceedings of the National Academy of Sciences USA* 75: 3727-3731), así como el promotor tac (DeBoer et al., 1983, *Proceedings of the National Academy of Sciences USA* 80: 21-25). Además promotores son descritos en "Useful proteins from recombinant bacteria" en *Scientific American*, 1980, 242: 74-94; y en Sambrook et al., 1989, supra.

[0079] Ejemplos de promotores adecuados para dirigir la transcripción de las construcciones de ácidos nucleicos de la presente invención en una célula huésped filamentosa fúngica son promotores obtenidos de los genes para TAKA amilasa de *Aspergillus oryzae*, proteinasa aspártica de *Rhizomucor miehei*, alfa-amilasa neutra de *Aspergillus niger*, alfa-amilasa estable en ácido de *Aspergillus niger*, glucoamilasa (glaA) de *Aspergillus niger* o *Aspergillus awamori*, lipasa de *Rhizomucor miehei*, proteasa alcalina de *Aspergillus oryzae*, triosa fosfato isomerasa de *Aspergillus oryzae*, acetamidasa de *Aspergillus nidulans*, amiloglucosidasa de *Fusarium venenatum* (WO 00/56900), *Fusarium venenatum Daria* (WO 00/56900), *Fusarium venenatum Quinn* (WO 00/56900), proteasa de tipo tripsina de *Fusarium oxysporum* (WO 96/00787), beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I de *Trichoderma reesei*, celobiohidrolasa II de *Trichoderma reesei*, endoglucanasa I de *Trichoderma reesei*, endoglucanasa II de *Trichoderma reesei*, endoglucanasa III de *Trichoderma reesei*, endoglucanasa IV de *Trichoderma reesei*, endoglucanasa V de *Trichoderma reesei*, xilanasa I de *Trichoderma reesei*, xilanasa II de *Trichoderma reesei*, beta-xilosidasa de *Trichoderma reesei*, al igual el promotor NA2-tpi (un híbrido de los promotores de los genes para alfa-amilasa neutra de *Aspergillus niger* e isomerasa de fosfato de triosa de *Aspergillus oryzae*); y promotores híbridos truncados mutantes de estos.

[0080] En un huésped de levadura, promotores útiles se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), galactocinasa de *Saccharomyces cerevisiae* (GAL1), alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH1,ADH2/GAP), triosa fosfato isomerasa de *Saccharomyces cerevisiae* (TPI), metalotioneína de *Saccharomyces cerevisiae* (CUP1) y 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para células huéspedes de levadura son descritos por Romanos et al., 1992, *Yeast* 8: 423-488.

[0081] La secuencia de control puede también ser una secuencia de terminación de transcripción adecuada, una secuencia reconocida por una célula huésped para terminar la transcripción. La secuencia de terminación se une

operativamente al 3' terminal de la secuencia de nucleótidos que codifica el polipéptido. Cualquier terminador que es funcional en la célula huésped de elección se puede utilizar en la presente invención.

5 [0082] Terminadores preferidos para células huéspedes filamentosas fúngicas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae*, glucoamilasa de *Aspergillus niger*, antranilato sintasa de *Aspergillus nidulans*, alfa-glucosidasa de *Aspergillus niger* y proteasa de tipo tripsina de *Fusarium oxysporum*.

10 [0083] Terminadores preferidos para células huéspedes de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae*, citocromo de *Saccharomyces cerevisiae* C (CYC1) y gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae*. Otros terminadores útiles para células huéspedes de levadura son descritos por Romanos *et al.*, 1992, *supra*.

15 [0084] La secuencia de control puede también ser una secuencia líder adecuada, una región no traducida de un ARNm que es importante para traducción por la célula huésped. La secuencia líder se une operativamente al 5' terminal de la secuencia de nucleótidos que codifica el polipéptido. Cualquier secuencia líder que es funcional en la célula huésped de elección se puede utilizar en la presente invención.

20 [0085] Líderes preferidas para células huéspedes filamentosas fúngicas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae* y triosa fosfato isomerasa de *Aspergillus nidulans*.

25 [0086] Líderes adecuadas para células huéspedes de levadura se obtienen de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), 3-fosfoglicerato quinasa de *Saccharomyces cerevisiae*, alfa-factor de *Saccharomyces cerevisiae* y alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH2/GAP).

30 [0087] La secuencia de control puede también ser una secuencia de poliadenilación, una secuencia unida operativamente al 3' terminal de la secuencia de nucleótidos y que, cuando se transcribe, es reconocida por la célula huésped como una señal para añadir residuos de poliadenosina a ARNm transcrito. Cualquier secuencia de poliadenilación que es funcional en la célula huésped de elección se puede utilizar en la presente invención.

[0088] Secuencias de poliadenilación preferidas para células huéspedes filamentosas fúngicas se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae*, glucoamilasa de *Aspergillus niger*, antranilato sintasa de *Aspergillus nidulans*, proteasa de tipo tripsina de *Fusarium oxysporum* y alfa-glucosidasa de *Aspergillus niger*.

35 [0089] Secuencias de poliadenilación útiles para células huéspedes de levadura son descritas por Guo y Sherman, 1995, *Molecular Cellular Biology* 15: 5983-5990.

40 [0090] La secuencia de control puede también ser una región codificante de péptido señal que codifica a una secuencia de aminoácidos unida al amino terminal de un polipéptido y dirige al polipéptido codificado en la vía secretora de la célula. El extremo 5' de la secuencia codificante de la secuencia de nucleótidos puede intrínsecamente contener una región codificante de péptido señal naturalmente unida en el marco de lectura de traducción con el segmento de la región codificante que codifica el polipéptido segregado. Alternativamente, el extremo 5' de la secuencia codificante puede contener una región codificante de péptido señal que es extranjero a la secuencia codificante. La región codificante de péptido señal extranjero puede requerirse donde la secuencia codificante no contiene naturalmente una región codificante de péptido señal. Alternativamente, la región codificante del péptido señal extranjero puede simplemente reemplazar la región codificante del péptido señal natural para mejorar secreción del polipéptido. No obstante, cualquier región codificante del péptido señal que dirige al polipéptido expresado en la vía secretora de una célula huésped de elección, es decir, segregado en un medio de cultivo, se puede utilizar en la presente invención.

50 [0091] Regiones codificantes del péptido señal eficaces para células huéspedes bacterianas son las regiones codificantes del péptido señal obtenidas de los genes para amilasa maltogénica de *Bacillus* NCIB 11837, alfa-amilasa de *Bacillus stearothermophilus*, dubtilisina de *Nacillus licheniformis*, beta-lactamasa de *Bacillus licheniformis*, proteasas neutras de *Bacillus stearothermophilus* (nprT, nprS, nprM) y *Bacillus subtilis* prsA. Además, péptidos señal son descritos por Simonen y Palva, 1993, *Microbiological Reviews* 57: 109-137.

55 [0092] Regiones codificantes de péptido señal eficaces para células huéspedes filamentosas fúngicas son las regiones codificantes del péptido señal obtenido de los genes para TAKA amilasa de *Aspergillus oryzae*, amilasa neutra de *Aspergillus niger*, glucoamilasa de *Aspergillus niger*, proteinasa aspártica de *Rhizomucor miehei*, celulasa de *Humicola insolens*, endoglucanasa V de *Humicola insolens* y lipasa de *Humicola lanuginosa*.

60 [0093] La región codificante del péptido señal es los nucleótidos 1 a 54 de la SEC ID n°: 7 que codifica los aminoácidos 1 a 18 de la SEC ID n°: 8. Péptidos señal útiles para células huéspedes de levadura se obtienen de los genes para alfa-factor de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*. Otras regiones codificantes de péptido señal útiles son descritas por Romanos *et al.*, 1992, *supra*.

65

[0094] La secuencia de control puede también ser una región codificante de propéptido que codifica a una secuencia de aminoácidos situada en el amino terminal de un polipéptido. El polipéptido resultante es conocido como una proenzima o propolipéptido (o un zimógeno en algunos casos). Un propolipéptido es generalmente inactivo y se puede convertir a un polipéptido maduro activo por escisión autocatalítica o catalítica del propéptido del propolipéptido. La región codificante del propéptido se puede obtener de los genes para proteasa alcalina de *Bacillus subtilis* (aprE), proteasa neutra de *Bacillus subtilis* (nprT), alfa-factor de *Saccharomyces cerevisiae*, proteinasa aspártica de *Rhizomucor miehei* y lacasa de *Myceliophthora thermophila* (WO 95/33836).

[0095] Donde ambos péptido señal y regiones de propéptido están presentes en el amino terminal de un polipéptido, la región de propéptido está situada junto al amino terminal de un polipéptido y la región de péptido señal está situada junto al amino terminal de la región de propéptido.

[0096] Puede también ser deseable añadir secuencias reguladoras que permiten la regulación de la expresión del polipéptido en relación al crecimiento de la célula huésped. Ejemplos de sistemas reguladores son los que causan que la expresión del gen cambie de ON u OFF en respuesta a un estímulo físico o químico, incluyendo la presencia de un compuesto regulador. Sistemas reguladores en sistemas procarióticos incluyen los sistemas operadores lac, tac y trp. En la levadura, el sistema ADH2 o sistema GAL1 pueden ser utilizados. En hongos filamentosos, el promotor de TAKA alfa-amilasa, el promotor de glucoamilasa de *Aspergillus niger* y el promotor de glucoamilasa de *Aspergillus oryzae* se pueden utilizar como secuencias reguladoras. Otros ejemplos de secuencias reguladoras son los que permiten amplificación génica. En sistemas eucariotas, estos incluyen el gen de dihidrofolato-reductasa que se amplifica en presencia de metotrexato, y los genes de metalotioneína que se amplifican con metales pesados. En estos casos, la secuencia de nucleótidos que codifica el polipéptido sería unida operativamente con la secuencia reguladora.

Vectores de expresión

[0097] La presente invención también se refiere a vectores de expresión recombinantes comprendiendo un polinucleótido de la presente invención, un promotor y señales de parada traduccionales y transcripcionales. Varios ácidos nucleicos y secuencias de control descritos aquí pueden juntarse para producir un vector de expresión recombinante que puede incluir uno o más sitios de restricción convenientes para permitir inserción o sustitución de la secuencia de nucleótidos que codifica el polipéptido en estos sitios. Alternativamente, una secuencia de nucleótidos de la presente invención se puede expresar por inserción de la secuencia de nucleótidos o una construcción de ácidos nucleicos comprendiendo la secuencia en un vector apropiado para la expresión. En la creación del vector de expresión, la secuencia codificante se localiza en el vector de modo que la secuencia codificante se une operativamente con las secuencias de control apropiadas para la expresión.

[0098] El vector de expresión recombinante puede ser cualquier vector (p. ej., un plásmido o virus) que puede ser convenientemente sometido a procedimientos de ADN recombinante y puede provocar expresión de la secuencia de nucleótidos. La elección del vector típicamente dependerá de la compatibilidad del vector con la célula huésped en la que el vector debe ser introducida. Los vectores pueden ser plásmidos circulares cerrados o lineales.

[0099] El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, la replicación del cual es independiente de replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un minicromosoma, o un cromosoma artificial. El vector puede contener cualquier medio para asegurar autorreplicación. Alternativamente, el vector puede ser uno que, cuando se introduce en la célula huésped, se integra en el genoma y se replica con el cromosoma(s) en el que ha sido integrado. Además, un único vector o plásmido o dos o más vectores o plásmidos que juntos contienen el ADN total para ser introducido en el genoma de la célula huésped, o un transposón pueden ser utilizados.

[0100] Los vectores de la presente invención preferiblemente contienen uno o más marcadores seleccionables que permiten selección fácil de células transformadas, modificadas, transducidas, o similares. Una etiqueta seleccionable es un gen, el producto del cual proporciona resistencia biocida o vírica, resistencia a metales pesados, prototrofia a auxótrofos y similares.

[0101] Ejemplos de marcadores bacterianos seleccionables son los genes dal de *Bacillus subtilis* o *Bacillus licheniformis*, o marcadores que confieren resistencia antibiótica tal como ampicilina, kanamicina, cloranfenicol, o resistencia a la tetraciclina. Marcadores adecuados para células huéspedes de levadura son ADE2, HIS3, LEU2, LYS2, MET3, TRP1 y URA3. Marcadores seleccionables para uso en una célula huésped filamentosa fúngica incluyen, pero de forma no limitativa, *amdS* (acetamidasa), *argB* (ornitina-carbamoyltransferasa), *bar* (fosfonitricina acetiltransferasa), *hph* (higromicina fosfotransferasa), *niaD* (nitrito-reductasa), *pyrG* (orotidina-5'-fosfato-descarboxilasa), *sC* (adeniltransferasa de sulfato), y *trpC* (sintasa de antranilato), al igual que equivalentes de estos. Preferido para uso en una célula de *Aspergillus* son los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces hygroscopicus*.

[0102] Los vectores de la presente invención preferiblemente contienen un elemento(s) que permite integración del vector en el genoma de la célula huésped o replicación autónoma del vector en la celular independiente del genoma.

[0103] Para integración en el genoma de la célula huésped, el vector puede depender de la secuencia del polinucleótido que codifica el polipéptido o cualquier otro elemento del vector para integración en el genoma por recombinación homóloga o no homóloga. Alternativamente, el vector puede contener secuencias de nucleótidos adicionales para dirigir integración por recombinación homóloga en el genoma de la célula huésped en una ubicación(es) precisa en el cromosoma(s). Para aumentar la probabilidad de integración en una ubicación precisa, los elementos integracionales deberían contener preferiblemente un número suficiente de ácidos nucleicos, tal como de 100 a 10.000 pares de bases, preferiblemente de 400 a 10.000 pares de bases y de forma más preferida de 800 a 10.000 pares de bases, que tienen un alto grado de identidad con la secuencia blanco correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos integracionales pueden ser cualquier secuencia que es homóloga a la secuencia blanco en el genoma de la célula huésped. Además, los elementos integracionales pueden ser secuencias codificantes o no codificantes de nucleótidos. Por otro lado, el vector se puede integrar en el genoma de la célula huésped por recombinación no homóloga.

[0104] Para replicación autónoma, el vector puede comprender además un origen de replicación que permite al vector replicar de manera autónoma en la célula huésped en cuestión. El origen de replicación puede ser cualquier replicador plásmido que media replicación autónoma que funciona en una célula. El término "origen de replicación" o "replicador plásmido" se define aquí como una secuencia de nucleótidos que habilita a un plásmido o vector para replicar *in vivo*.

[0105] Ejemplos de orígenes bacterianos de replicación son los orígenes de replicación de plásmidos pBR322, pUC19, pACYC177 y pACYC184 permitiendo replicación en *E. coli* y pUB110, pE194, pTA1060 y pAMB1 permitiendo replicación en *Bacillus*.

[0106] Ejemplos de orígenes de replicación para uso en una célula huésped de levadura son el origen de replicación de 2 micras, ARS1, ARS4, la combinación de ARS1 y CEN3 y la combinación de ARS4 y CEN6.

[0107] Ejemplos de orígenes de replicación útiles en una célula filamentosa fúngica son AMA1 y ANS1 (Gems et al., 1991, *Gene* 98: 61-67; Cullen et al., 1987, *Nucleic Acids Research* 15: 9163-9175; WO 00/24883). Aislamiento del gen AMA1 y construcción de plásmidos o vectores comprendiendo el gen, lo que se puede realizar según los métodos descritos en la WO 00/24883.

[0108] Más de una copia de un polinucleótido de la presente invención se puede insertar en la célula huésped para aumentar producción del producto genético. Un aumento en el número de copias del polinucleótido se puede obtener integrando al menos una copia adicional de la secuencia en el genoma de la célula huésped o incluyendo un gen marcador amplificable seleccionable con el polinucleótido donde células con copias amplificadas del gen marcador seleccionable, y así copias adicionales del polinucleótido, se pueden seleccionar para cultivar las células en presencia del agente seleccionable apropiado.

[0109] Los procedimientos usados para unir los elementos anteriormente descritos para construir los vectores de expresión recombinantes de la presente invención se conocen por un experto en la técnica (ver, por ejemplo, Sambrook et al., 1989, *supra*).

Células huéspedes

[0110] La presente invención también se refiere a células huéspedes recombinantes, comprendiendo un polinucleótido de la presente invención, que es ventajosamente usado en la producción recombinante de los polipéptidos. Un vector comprendiendo un polinucleótido de la presente invención se introduce en una célula huésped de modo que el vector ese mantiene como un integrante cromosómico o como un vector extracromosomal que se duplica como se describe anteriormente. El término "célula huésped" abarca cualquier descendiente de una célula madre que no es idéntico a la célula madre debido a mutaciones que ocurren durante la replicación. La elección de una célula huésped en gran parte depende del gen que codifica el polipéptido y su fuente.

[0111] La célula huésped puede ser un microorganismo unicelular, por ejemplo, un procarionta, o un microorganismo no unicelular, por ejemplo, un eucariota.

[0112] Microorganismos unicelulares útiles son células bacterianas tales como bacterias gram positivas incluyendo, pero no limitado a, una célula de *Bacillus*, por ejemplo, *Bacillus alkalophilus*, *Bacillus amyloliquefaciens*, *Bacillus brevis*, *Bacillus circulans*, *Bacillus clausii*, *Bacillus coagulans*, *Bacillus lautus*, *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus megaterium*, *Bacillus stearothermophilus*, *Bacillus subtilis* y *Bacillus thuringiensis* o una célula de *Streptomyces*, por ejemplo, *Streptomyces lividans* y *Streptomyces murinus*, o bacterias gram negativas tales como *E. coli* y *Pseudomonas sp.* En un aspecto preferido, la célula huésped bacteriana es una célula de *Bacillus lentus*, *Bacillus licheniformis*, *Bacillus stearothermophilus*, o *Bacillus subtilis*. En otro aspecto preferido, la célula de *Bacillus* es un *Bacillus* alcalofílico.

[0113] La introducción de un vector en una célula huésped bacteriana puede, por ejemplo, ser efectuada por transformación de protoplasto (ver, por ejemplo, Chang and Cohen, 1979, *Molecular General Genetics* 168: 111-115), usando células competentes (ver, por ejemplo, Young and Spizizen, 1961, *Journal of Bacteriology* 81: 823-829, o Dubnau and Davidoff-Abelson, 1971, *Journal of Molecular Biology* 56: 209-221), electroporación (ver, por ejemplo,

Shigekawa and Dower, 1988, *Biotechniques* 6: 742-751), o conjugación (ver, por ejemplo, Koehler and Thorne, 1987, *Journal of Bacteriology* 169: 5771-5278).

[0114] La célula huésped puede también ser un eucariota, tal como un mamífero, insecto, planta, o célula fúngica.

[0115] En un aspecto preferido, la célula huésped es una célula fúngica. "Hongo", como se utiliza en este caso, incluye los fila *Ascomycota*, *Basidiomycota*, *Chytridiomycota* y *Zygomycota* (como se define por Hawksworth *et al.*, En *Einsworth and Bisby's Dictionary of The Fungi*, 8th edition, 1995, CAB International, University Press, Cambridge, UK) al igual que el *Oomycota* (como se cita en Hawksworth *et al.*, 1995, *supra*, page 171) y todos los hongos mitospóricos (Hawksworth *et al.*, 1995, *supra*). [0119] En un aspecto más preferido, la célula huésped fúngica es una célula de levadura. "Levadura", como se utiliza en este caso, incluye levadura ascoesporógena (Endomicetales), levadura *basidiosporogenous* y levadura de los hongos Imperfecti (*Blastomycetes*). Ya que la clasificación de levadura puede cambiar en el futuro, para los fines de esta invención, la levadura debe ser definida como se describe en *Biology and Activities of Yeast* (Skinner, F.A., Passmore, S.M., and Davenport, R.R., eds, *Soc. App. Bacteriol. Symposium Series No. 9*, 1980).

[0116] En un aspecto incluso más preferido, la célula huésped de levadura es una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, o *Yarrowia*.

[0117] En un aspecto más preferido, la célula huésped de levadura es una célula de *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, o *Saccharomyces oviformis*. En otro aspecto más preferido, la célula huésped de levadura es una célula de *Kluyveromyces lactis*. En otro aspecto más preferido, la célula huésped de levadura es una célula de *Yarrowia lipolytica*.

[0118] En otro aspecto más preferido, la célula huésped fúngica es una célula micótica filamentosa. "Hongos filamentosos" incluye todas las formas filamentosas de la subdivisión *Eumycota* y *Oomycota* (como se define por Hawksworth *et al.*, 1995, *supra*). Los hongos filamentosos se caracterizan generalmente por una pared micelial compuesta por quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos. Crecimiento vegetativo es por alargamiento hifal y catabolismo de carbono es estrictamente aeróbico. En cambio, crecimiento vegetativo por levaduras tales como *Saccharomyces cerevisiae* es por injerto de un talo unicelular y catabolismo de carbono puede ser fermentativo.

[0119] En un aspecto incluso más preferido, la célula huésped filamentosa fúngica es una célula de *Acremonium*, *Aspergillus*, *Aureobasidium*, *Bjerkandera*, *Ceriporiopsis*, *Coprinus*, *Coriolus*, *Cryptococcus*, *Filibasidium*, *Fusarium*, *Humicola*, *Magnaporthe*, *Mucor*, *Myceliophthora*, *Neocallimastix*, *Neurospora*, *Paecilomyces*, *Penicillium*, *Phanerochaete*, *Phlebia*, *Piromyces*, *Pleurotus*, *Schizophyllum*, *Talaromyces*, *Thermoascus*, *Thielavia*, *Tolypocladium*, *Trametes*, o *Trichoderma*.

[0120] En un aspecto más preferido, la célula huésped filamentosa fúngica es una célula de *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* o *Aspergillus oryzae*. En otro aspecto más preferido, la célula huésped filamentosa fúngica es una célula de *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium graminum*, *Fusarium heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium reticulatum*, *Fusarium roseum*, *Fusarium sambucinum*, *Fusarium sarcochromum*, *Fusarium sporotrichioides*, *Fusarium sulphureum*, *Fusarium torulosum*, *Fusarium trichothecioides*, o *Fusarium venenatum*. En otro aspecto más preferido, la célula huésped filamentosa fúngica es una células de *Bjerkandera adusta*, *Ceriporiopsis aneirina*, *Ceriporiopsis aneirina*, *Ceriporiopsis caregiea*, *Ceriporiopsis gilvescens*, *Ceriporiopsis pannocinta*, *Ceriporiopsis rivulosa*, *Ceriporiopsis subrufa*, *Ceriporiopsis subvermispora*, *Coprinus cinereus*, *Coriolus hirsutus*, *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Phanerochaete chrysosporium*, *Phlebia radiata*, *Pleurotus eryngii*, *Thielavia terrestris*, *Trametes villosa*, *Trametes versicolor*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.

[0121] Células fúngicas se pueden transformar por un proceso implicando formación de protoplasto, transformación de los protoplastos y regeneración de la pared celular en una manera conocida *per se*. Procedimientos adecuados para transformación de células huéspedes de *Aspergillus* y *Trichoderma* se describen en EP 238 023 y Yelton *et al.*, 1984, *Proceedings of the National Academy of Sciences USA* 81: 1470-1474. Métodos adecuados para transformación de especies de *Fusarium* son descritos por Malardier *et al.*, 1989, *Gene* 78: 147-156 y WO 96/00787. La levadura puede ser transformada usando los procedimientos descritos por Becker y Guarente, en *Abelson, J.N. and Simon, M.I., editors, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, pp 182-187, Academic Press, Inc., New York*; Ito *et al.*, 1983, *Journal of Bacteriology* 153: 163; y Hinnen *et al.*, 1978, *Proceedings of the National Academy of Sciences USA* 75: 1920.

Métodos de producción

[0122] La presente divulgación también se refiere a métodos para producir un polipéptido de la presente invención, comprendiendo (a) cultivo de una célula, que en su forma de tipo salvaje es capaz de producir el polipéptido, bajo condiciones propicias para la producción del polipéptido; y (b) recuperación el polipéptido. Preferiblemente, la célula es del género *Thielavia*, y más preferiblemente *Thielavia terrestris*.

[0123] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención, comprendiendo: (a) cultivo de una célula huésped bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.

[0124] En los métodos de producción de la presente invención, las células se cultivan en un medio nutritivo adecuado para la producción del polipéptido usando métodos bien conocidos en la técnica. Por ejemplo, la célula se puede cultivar por cultivo en matraz de agitación y fermentación a gran escala o a pequeña escala (incluyendo fermentaciones continua, por lotes, por lote alimentado, o de estado sólido) en el laboratorio o fermentadores industriales realizadas en un medio adecuado y bajo condiciones que permiten al polipéptido expresarse y/o aislarse. El cultivo se desarrolla en un medio nutritivo adecuado comprendiendo fuentes de nitrógeno y carbono y sales inorgánicas, usando procedimientos conocidos en la técnica. Medios adecuados están disponibles de proveedores comerciales o se pueden preparar según composiciones publicadas (p. ej., en catálogos de la American Type Culture Collection). Si el polipéptido se segrega en el medio nutritivo, el polipéptido se puede recuperar directamente del medio. Si el polipéptido no se segrega en el medio, este se puede recuperar de lisatos celulares.

[0125] Los polipéptidos con actividad de mejora celulolítica son detectados usando los métodos descritos aquí.

[0126] El polipéptido resultante puede ser recuperado usando métodos conocidos en la técnica. Por ejemplo, el polipéptido puede ser recuperado del medio nutritivo por procedimientos convencionales incluyendo, pero no limitado a, centrifugado, filtración, extracción, secado por pulverización, evaporación, o precipitación.

[0127] Los polipéptidos de la presente invención se pueden purificar por una variedad de procedimientos conocidos en la técnica incluyendo, pero no limitado a, cromatografía (p. ej., intercambio de iones, afinidad, cromatografía de exclusión por tamaño), procedimientos electroforéticos (p. ej., isoelectroenfoque preparativo), solubilidad diferencial (p. ej., precipitación de sulfato amónico), SDS-PAGE, o extracción (ver, por ejemplo, *Protein Purification*, J.-C. Janson and Lars Ryden, editors, VCH Publishers, New York, 1989) para obtener polipéptidos substancialmente puros.

Plantas

[0128] La presente divulgación también se refiere a una planta transgénica, parte de planta, o célula vegetal que ha sido transformada con una secuencia de nucleótidos que codifica un polipéptido con actividad de mejora celulolítica de la presente invención para expresar y producir el polipéptido en cantidades recuperables. El polipéptido se puede recuperar de la planta o parte de planta. Alternativamente, la planta o parte de planta con el polipéptido recombinante se puede utilizar como tal para mejorar la calidad de un alimento o pienso, por ejemplo, mejora de valor nutritivo, palatabilidad y propiedades reológicas, o para destruir un factor antinutritivo.

[0129] La planta transgénica puede ser *dicotyledonous* (un dicotiledóneo) o *monocotyledonous* (un monocotiledóneo). Ejemplos de plantas monocotiledóneas son hierbas, tales como césped de pradera (pasto azul, Poa), hierba forrajera tal como *Festuca*, *Lolium*, césped templado, tal como *Agrostis* y cereales, por ejemplo, trigo, avena, centeno, cebada, arroz, sorgo y maíz.

[0130] Ejemplos de plantas dicotiledóneas son tabaco, leguminosas, tales como altramuces, patata, remolacha azucarera, guisante, judía y semilla de soja, y plantas crucíferas (familia *Brassicaceae*), tal como coliflor, semilla de colza y el modelo cercanamente relacionado con organismo *Arabidopsis thaliana*.

[0131] Ejemplos de partes de planta son vástago, callo, hojas, raíz, frutas, semillas y tubérculos al igual que los tejidos individuales comprendiendo estas partes, por ejemplo, epidermis, mesófilo, parénquima, tejidos vasculares, meristemas. Compartimentos de célula vegetal específicos, tales como cloroplastos, apoplastos, mitocondria, vacuolas, peroxisomas y citoplasma son también considerados como una parte de planta. Además, cualquier célula vegetal, cualquiera que sea el origen del tejido, se considera como una parte de planta. Asimismo, partes de planta tales como tejidos específicos y células aisladas para facilitar la utilización de la invención son también consideradas partes de planta, por ejemplo, embriones, endospermas, aleurona y revestimientos de semillas.

[0132] También incluido dentro del campo de la presente invención está la progenie de tales plantas, partes de planta y células vegetales.

[0133] La planta transgénica o célula vegetal que expresa un polipéptido de la presente invención se puede construir conforme a métodos conocidos en la técnica. En resumen, la planta o célula vegetal se construye por incorporación de una o más construcciones de expresión que codifican un polipéptido de la presente invención en el genoma huésped de

planta o genoma de cloroplasto y propagan la planta modificada resultante o célula vegetal en una planta transgénica o célula vegetal.

[0134] La construcción de expresión es convenientemente una construcción de ácidos nucleicos que comprende un polinucleótido que codifica un polipéptido de la presente invención operativamente unido con secuencias reguladoras apropiadas requeridas para la expresión de la secuencia de nucleótidos en la planta o parte de planta de elección. Además, la construcción de expresión puede comprender una etiqueta seleccionable útil para identificación de células huéspedes en las que la construcción de expresión ha sido integrada y secuencias de ADN necesarias para introducción de la construcción en la planta en cuestión (esta última depende del método de introducción de ADN que se usa).

[0135] La elección de secuencias reguladoras, tales como promotor y secuencias de terminación y opcionalmente secuencias señal o de tránsito son determinadas, por ejemplo, basándose en cuándo, dónde y cómo se desea el polipéptido que se exprese. Por ejemplo, la expresión del gen que codifica un polipéptido de la presente invención puede ser constitutiva o inducible, o puede ser desarrollable, fase o tejido específico, y el producto genético puede estar dirigido a un tejido específico o parte de planta tal como semillas o hojas. Secuencias reguladoras son, por ejemplo, descritas por Tague *et al.*, 1988, *Plant Physiology* 86: 506.

[0136] Para expresión constitutiva, el 35S-CaMV, la ubiquitina de maíz 1, y el promotor de actina de arroz 1 se pueden usar (Franck *et al.*, 1980, *Cell* 21: 285-294, Christensen *et al.*, 1992, *Plant Mo. Biol.* 18: 675-689; Zhang *et al.*, 1991, *Plant Cell* 3: 1155-1165). Promotores específicos a un órgano pueden ser, por ejemplo, un promotor de tejidos sumidero de almacenamiento tal como semillas, tubérculos de patata y frutas (Edwards & Coruzzi, 1990, *Ann. Rev. Genet.* 24: 275-303), o de tejidos sumidero metabólicos tales como meristemas (Ito *et al.*, 1994, *Plant Mol. Biol.* 24: 863-878), un promotor específico de semilla tal como la glutelina, prolamina, globulina, o promotor de albúmina de arroz (Wu *et al.*, 1998, *Plant and Cell Physiology* 39: 885-889), un promotor de *Vicia faba* de la legúmina B4 y el gen de proteína de semilla desconocida de *Vicia faba* (Conrad *et al.*, 1998, *Journal of Plant Physiology* 152: 708-711), un promotor de una proteína del cuerpo de aceite de semilla (Chen *et al.*, 1998, *Plant and Cell Physiology* 39: 935- 941), el promotor napA de la proteína de almacenamiento de *Brassica napus*, o cualquier otro promotor específico de semilla conocido en la técnica, por ejemplo, como se describe en la WO 91/14772. Además, el promotor puede ser un promotor específico de hoja tal como el promotor de rbcS de arroz o tomate (Kyoizuka *et al.*, 1993, *Plant Physiology* 102: 991-1000), el promotor de gen de metiltransferasa de adenina de virus de chlorella (Mittra and Higgins, 1994, *Plant Molecular Biology* 26: 85-93), o el promotor aldP de gen de arroz (Kagaya *et al.*, 1995, *Molecular and General Genetics* 248: 668-67), o un promotor inducible dañado tal como el promotor pin2 de patata (Xu *et al.*, 1993, *Plant Molecular Biology* 22: 573-588). Asimismo, el promotor puede ser inducible por tratamientos abióticos tales como temperatura, sequía, o alteraciones en la salinidad o inducido por sustancias aplicadas exógenamente que activan el promotor, por ejemplo, etanol, estrógenos, hormonas de planta tal como etileno, ácido abscísico y ácido giberélico y metales pesados.

[0137] Un elemento promotor intensificador puede también ser usado para conseguir expresión mayor de un polipéptido de la presente invención en la planta. Por ejemplo, el elemento promotor intensificador puede ser un intrón que es colocado entre el promotor y la secuencia de nucleótidos que codifica un polipéptido de la presente invención. Por ejemplo, Xu *et al.*, 1993, *supra*, revela el uso del primer intrón del gen de actina de arroz 1 para mejorar expresión.

[0138] El gen marcador seleccionable y cualquier otra parte de la construcción de expresión se puede elegir de aquellas disponibles en la técnica.

[0139] La construcción de ácidos nucleicos se incorpora en el genoma de planta según técnicas convencionales conocidas en la técnica, incluyendo transformación mediada por *Agrobacterium*, transformación mediada por virus, microinyección, bombardeo de partícula, transformación biolística y electroporación (Gasser *et al.*, 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto *et al.*, 1989, *Nature* 338: 274).

[0140] Actualmente, transferencia de gen mediada por *Agrobacterium tumefaciens* es el método de elección para generación de dicotiledóneas transgénicas (para una revisión, ver Hooykas and Schilperoort, 1992, *Plant Molecular Biology* 19: 15-38) y también puede usarse para transformación de monocotiledóneas, aunque otros métodos de transformación se usan frecuentemente usado para estas plantas. Actualmente, el método de elección para generación de monocotiledóneas transgénicas es bombardeo de partículas (oro microscópico o partículas de tungsteno revestido con la transformación ADN) de callos embrionarios o embriones en desarrollo (Christou, 1992, *Plant Journal* 2: 275-281; Shimamoto, 1994, *Current Opinion Biotechnology* 5: 158-162; Vasil *et al.*, 1992, *Bio/Technology* 10: 667-674). Un método alternativo para transformación de monocotiledóneas se basa en transformación de protoplasto como se describe por Omirulleh *et al.*, 1993, *Plant Molecular Biology* 21: 415-428.

[0141] Después de transformación, los transformantes que han sido incorporados a la construcción de expresión se seleccionan y regeneran en plantas enteras según métodos bien conocidos en la técnica. Con frecuencia, el procedimiento de transformación se diseña para la eliminación selectiva de selección de genes bien durante regeneración o en las siguientes generaciones usando, por ejemplo, cotransformación con dos construcciones de T-ADN separados o escisión específica de sitio del gen de selección por una recombinasa específica.

[0142] La presente invención también se refiere a métodos para producir un polipéptido de la presente invención comprendiendo: (a) cultivo de una planta transgénica o una célula vegetal comprendiendo un polinucleótido que codifica un polipéptido con actividad de mejora celulolítica de la presente invención bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.

Composiciones

[0143] La presente invención también se refiere a composiciones que comprenden un polipéptido de la presente invención. Preferiblemente, las composiciones se enriquecen en un polipéptido tal. El término "enriquecido" indica que la actividad de mejora celulolítica de la composición ha sido aumentada, por ejemplo, con un factor de enriquecimiento de al menos 1.1.

[0144] La composición puede comprender un polipéptido de la presente invención como el componente principal, por ejemplo, una composición monocomponente. Alternativamente, la composición puede comprender además una o más actividades enzimáticas, tales como una aminopeptidasa, amilasa, carbohidrasa, carboxipeptidasa, catalasa, celobiohidrolasa, celulasa, quitinasa, cutinasa, glicosiltransferasa ciclodextrina, desoxirribonucleasa, endoglucanasa, esterasa, alfa-galactosidasa, beta-galactosidasa, glucoamilasa, alfa-glucosidasa, beta-glucosidasa, haloperoxidasa, invertasa, lacasa, lipasa, manosidasa, oxidasa, enzima pectinolítica, peptidoglutaminasa, peroxidasa, fitasa, polifenoloxidasas, enzima proteolítica, ribonucleasa, transglutaminasa, o xilanasas. La(s) enzima(s) adicional(es) se puede(n) producir, por ejemplo, por un microorganismo que pertenece al género *Aspergillus*, preferiblemente *Aspergillus aculeatus*, *Aspergillus awamori*, *Aspergillus fumigatus*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger*, o *Aspergillus oryzae*; *Fusarium*, preferiblemente *bactridioides Fusarium*, *Fusarium*, *Fusarium cerealis crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*, *Fusarium*, *Fusarium graminum heterosporum*, *Fusarium negundi*, *Fusarium oxysporum*, *Fusarium* *Fusarium roseum*, *reticulatum*, *Fusarium*, *Fusarium sambucinum sarcochroum*, *Fusarium*, *Fusarium sulphureum toruloseum*, *trichothecioides Fusarium*, o *Fusarium venenatum*; *Humicola*, preferiblemente *Humicola insolens* o *Humicola lanuginosa*; o *Trichoderma*, preferiblemente *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei*, o *Trichoderma viride*.

[0145] Las composiciones de polipéptido se pueden preparar de acuerdo con métodos conocidos en la técnica y pueden estar en forma de un líquido o una composición seca. Por ejemplo, la composición de polipéptido puede estar en forma de un granulado o un microgranulado. El polipéptido para ser incluido en la composición puede ser estabilizado de acuerdo con métodos conocidos en la técnica.

[0146] Se dan ejemplos más abajo de usos preferidos de las composiciones de polipéptidos de la invención. La dosificación de la composición de polipéptido de la invención y otras condiciones bajo las cuales se usa la composición puede ser determinada sobre la base de los métodos conocidos en la técnica.

Degradación o conversión de biomasa a monosacáridos, disacáridos y polisacáridos

[0147] La presente invención también se refiere a métodos para degradar o convertir un material celulósico, comprendiendo: tratamiento del material celulósico con una cantidad eficaz de una proteína celulolítica en presencia de una cantidad eficaz del polipéptido con actividad de mejora celulolítica, donde la presencia del polipéptido con actividad de mejora celulolítica aumenta la degradación del material celulósico en comparación con la ausencia del polipéptido con actividad de mejora celulolítica.

[0148] Los polipéptidos y células huéspedes de la presente invención se pueden utilizar en la producción de monosacáridos, disacáridos y polisacáridos como materias primas químicas o de fermentación a partir de biomasa para la producción de etanol, plástico, otros productos o productos intermedios. En particular, los polipéptidos y células huéspedes de la presente invención se pueden utilizar para aumentar el valor de residuos de tratamiento (grano de destiladores secos, granos consumidos de elaboración, bagazo de caña de azúcar, etc.) por solubilización completa o parcial de celulosa o hemicelulosa. En el impulso del tratamiento por proteínas celulolíticas de material celulósico para glucosa, xilosa, manosa, galactosa y arabinosa, sus polímeros, o productos derivados de estos como se describe debajo. Los polipéptidos pueden estar en forma de un caldo de fermentación crudo con o sin las células o en forma de una preparación semi-purificada o purificada enzimática. La proteína de mejora celulolítica puede ser una preparación monocomponente, por ejemplo, una proteína de la familia 61, una preparación de proteína de varios componentes, por ejemplo, varias proteínas de la familia 61 o una combinación de preparaciones de proteína monocomponente y de varios componentes. Las proteínas de mejora celulolíticas pueden impulsar la actividad de proteínas celulolíticas, bien en el intervalo de pH ácido, neutro o alcalino. Alternativamente, una célula huésped de la presente invención se puede utilizar como una fuente del polipéptido en un proceso de fermentación con la biomasa. La célula huésped puede también contener genes heterólogos o nativos que codifican proteína celulolítica al igual que otras enzimas útiles en el procesamiento de biomasa.

[0149] La biomasa puede incluir, pero no está limitado a, recursos de madera, desperdicios sólidos municipales, papel usado, cultivos y residuos de cultivo (ver, por ejemplo, Wiseloge *et al.*, 1995, en *Handbook on Bioethanol* (Charles E. Wyman, editor), pp.105- 118, Taylor & Francis, Washington D.C.; Wyman, 1994, *Bioresource Technology* 50: 3-16; Lynd, 1990, *Applied Biochemistry and Biotechnology* 24/25: 695-719; Mosier *et al.*, 1999, Recent Progress en

Bioconversion of Lignocellulosics, in Advances in Biochemical Engineering/Biotechnology, T. Scheper, managing editor, Volume 65, pp.23-40, Springer-Verlag, New York).

[0150] El polisacárido predominante en la pared celular primaria de biomasa es celulosa, el segundo más abundante es hemicelulosa y el tercero es pectina. La pared celular secundaria, producida después de que la célula ha detenido el crecimiento, también contiene polisacáridos y se refuerza por lignina polimérica reticulada de manera covalente a hemicelulosa. Celulosa es un homopolímero de anhidrocelobiosa y así un beta-(1-4)-D-glucano lineal, mientras hemicelulosas incluyen una variedad de compuestos, tales como xilanos, xiloglucanos, arabinoxilanas y mananos en las estructuras ramificadas de complejo con un espectro de sustituyentes. Aunque generalmente celulosa polimorfa se encuentra en tejido vegetal principalmente como una matriz de cristalina insoluble de cadenas de glucano paralelas. Hemicelulosas normalmente une hidrógeno a celulosa, al igual que a otras hemicelulosas, que ayudan estabilizan a la matriz de la pared celular.

[0151] En los métodos de la presente invención, la proteína celulolítica puede ser cualquier proteína implicada en el procesamiento de material celulósico a glucosa, o hemicelulosa a xilosa, manosa, galactosa y arabinosa, sus polímeros, o productos derivados de estos como se describe abajo. La proteína celulolítica puede ser una preparación monocomponente, por ejemplo, una celulasas, una preparación de varios componentes, por ejemplo, endoglucanasa, celobiohidrolasa, glucohidrolasa, beta-glucosidasa, tal y como se define por debajo o una combinación de preparaciones de proteína monocomponente y de varios componentes. Las proteínas celulolíticas pueden tener actividad, es decir, hidrolizar celulosa, bien en el intervalo de pH ácido, neutro, o alcalino.

[0152] La proteína celulolítica puede ser de origen bacteriano o fúngico, que puede ser obtenible o aislado y purificado de microorganismos que se conocen por ser capaces de producir enzimas celulolíticas, por ejemplo, especies de *Bacillus*, *Pseudomonas*, *Humicola*, *Coprinus*, *Thielavia*, *Fusarium*, *Myceliophthora*, *Acremonium*, *Cephalosporium*, *Scytalidium*, *Penicillium* o *Aspergillus* (ver, por ejemplo, EP 458162), especialmente aquellas producidas por una cepa seleccionada de la especie *Humicola insolens* (reclasificada como *Scytalidium thermophilum*, ver por ejemplo, patente US n°. 4,435,307), *Coprinus cinereus*, *Fusarium oxysporum*, *Myceliophthora thermophila*, *Meripilus giganteus*, *Thielavia terrestris*, *Acremonium sp.*, *Acremonium persicinum*, *Acremonium acremonium*, *Acremonium brachyphenium*, *Acremonium dichromosporum*, *Acremonium obclavatum*, *Acremonium pinkertoniae*, *Acremonium roseogriseum*, *Acremonium incoloratum* y *Acremonium furatum*; preferiblemente de la especie *Humicola insolens* DSM 1800, *Fusarium oxysporum* DSM 2672, *Myceliophthora thermophila* CBS 117.65, *Cephalosporium sp.* RYM-202, *Acremonium sp.* CBS 478.94, *Acremonium sp.* CBS 265.95, *Acremonium persicinum* CBS 169.65, *Acremonium acremonium* AHU 9519, *Cephalosporium sp.* CBS 535.71, *Acremonium brachyphenium* CBS 866.73, *Acremonium dichromosporum* CBS 683.73, *Acremonium obclavatum* CBS 311.74, *Acremonium pinkertoniae* CBS 157.70, *Acremonium roseogriseum* CBS 134.56, *Acremonium incoloratum* CBS 146.62 y *Acremonium furatum* CBS 299.70H. Proteínas celulolíticas se pueden obtener también de *Trichoderma* (particularmente *Trichoderma viride*, *Trichoderma reesei* y *Trichoderma koningii*), *Bacillos* alcalofílicos (ver, por ejemplo, patente US n°. 3,844,890 y EP 458162) y *Streptomyces* (ver, por ejemplo, EP 458162). Mutantes de proteína químicamente modificados o creados genéticamente se incluyen.

[0153] Proteínas celulolíticas especialmente adecuadas son las celulasas neutras o alcalinas. Ejemplos de tales celulasas son celulasas descritas en EP 495,257, EP 531,372, WO 96/11262, WO 96/29397, WO 98/08940. Otros ejemplos son variantes de celulasas tales como los descritos en WO 94/07998, EP 531,315, patente US n°. 4,435,307, patente US n°. 5,457,046, patente US n°. 5,648,263, patente US n°. 5,686,593, patente US n°. 5,691,178, patente US n°. 5,763,254, patente US n°. 5,776,757, WO 89/09259, WO 95/24471, WO 98/12307 y PCT/DK98/00299.

[0154] Las proteínas celulolíticas y proteínas de mejora celulolíticas usadas en los métodos de la presente invención se pueden producir por fermentación de las cepas microbianas mencionadas anteriormente en un medio nutritivo con fuentes de nitrógeno y carbono adecuadas y sales inorgánicas, usando procedimientos conocidos en la técnica (ver, por ejemplo, Bennett, J.W. and LaSure, L. (eds.), *More Gene Manipulations in Fungi*, Academic Press, CA, 1991). Medios adecuados están disponibles de proveedores comerciales o se pueden preparar según composiciones publicadas (p. ej., en catálogos de la American Type Culture Collection). Intervalos de temperatura y otras condiciones adecuadas para crecimiento y producción de proteína celulolítica se conocen en la técnica (ver, por ejemplo, Bailey, J.E., and Ollis, D.F., *Biochemical Engineering Fundamentals*, McGraw-Hill Book Company, NY, 1986).

[0155] La fermentación puede ser cualquier método de cultivo de una célula dando como resultado la expresión o aislamiento de una proteína celulolítica o proteína de mejora celulolítica. Fermentación puede, por lo tanto, ser entendida como comprendiendo cultivo en matriz de agitación, fermentación a pequeña o gran escala (incluyendo fermentaciones continuas, por lotes, por lote alimentado, o de estado sólido) en el laboratorio o fermentadores industriales realizados en un medio adecuado y bajo condiciones que permiten a la proteína celulolítica o proteína de mejor celulolítica expresarse o aislarse.

[0156] Las proteínas resultantes celulolíticas o proteínas de mejora celulolíticas producidas por los métodos anteriormente descritos pueden ser recuperadas del medio de fermentación por procedimientos convencionales incluyendo, pero no limitado a, centrifugado, filtración, secado por pulverización, evaporación, o precipitación. La proteína recuperada puede luego ser además purificada por una variedad de procedimientos cromatográficos, por ejemplo, cromatografía de intercambio de iones, cromatografía de filtración en gel, cromatografía de afinidad, o similar.

[0157] La proteína celulolítica puede hidrolizar o hidroliza carboximetilcelulosa (CMC), disminuyendo así la viscosidad de la mezcla de incubación. La reducción resultante en la viscosidad se puede determinar por un viscosímetro de vibración (p. ej., MIVI 3000, de Sofraser Francia). Determinación de actividad de celulasa, medida en cuanto a unidad de viscosidad de celulasa (CEVU), cuantifica la cantidad de actividad catalítica presente en prueba A por medición de la capacidad de la muestra para reducir la viscosidad de una solución de carboximetilcelulosa (CMC). El ensayo se realiza a la temperatura y el pH adecuados para la proteína celulolítica y sustrato. Para Celluclast™ (Novozymes A/S, Bagsværd, Dinamarca) el ensayo se realiza a 40°C en 0,1 M de tampón fosfato pH 9,0 durante 30 minutos con CMC como sustrato (33,3 g/L de carboximetilcelulosa Hercules 7 LFD) y una concentración enzimática de aproximadamente 3,3-4,2 CEVU/ml. La actividad CEVU se calcula en relación a un estándar de enzima declarado, tal como CELLUZYME™ estándar 17-1194 (obtenido de Novozymes A/S).

[0158] Ejemplos de preparaciones celulolíticas adecuadas para uso en la presente invención incluyen, por ejemplo, CELLUCLAST™ (disponible de Novozymes A/S) y NOVOZYM™ 188 (disponible de Novozymes A/S). Otras preparaciones disponibles comercialmente que comprenden celulasa que se puede usar incluyen CELLUZYME™, CEREFLO™ y ULTRAFLO™ (Novozymes A/S), LAMINEX™ y SPEZYME™ CP (Genencor Int.) y ROHAMENT™ a 7069 W (Röhm GmbH). Las enzimas de celulasa se agregan en cantidades efectivas de aproximadamente 0,001% a aproximadamente 5,0 % en peso de sólidos, más preferiblemente de aproximadamente 0,025% a aproximadamente 4,0% en peso de sólidos y de forma más preferida de aproximadamente 0,005% a aproximadamente 2,0% en peso de sólidos.

[0159] Como se ha mencionado anteriormente, las proteínas celulolíticas o proteínas de mejora celulolíticas usadas en los métodos de la presente invención pueden ser preparaciones monocomponente, es decir, un componente esencialmente libre de otros componentes celulolíticos. El único componente puede ser un componente recombinante, es decir, producido por clonación de una secuencia de ADN que codifica el único componente y célula posterior transformadas con la secuencia de ADN y expresada en un huésped (ver, por ejemplo, WO 91/17243 y WO 91/17244). Otros ejemplos de proteínas celulolíticas de monocomponente incluyen, pero de forma no limitativa, aquellas descritos en JP-07203960-A y WO-9206209. El huésped es preferiblemente un huésped heterólogo (la enzima es extranjera al huésped), pero el huésped puede bajo condiciones determinadas también ser un huésped homólogo (la enzima es nativa al huésped). Proteínas monocomponente celulolíticas puede también ser preparadas por purificación de tal proteína de un caldo de fermentación.

[0160] Ejemplos de proteínas celulolíticas monocomponente útiles en la práctica de los métodos de la presente invención incluyen, pero de forma no limitativa, endoglucanasa, celobiohidrolasa, glucohidrolasa y beta-glucosidasa.

[0161] El término "endoglucanasa" se define aquí como una endo-1,4-(1,3,1,4)-beta-D-glucano 4-glucanohidrolasa (E.C. N°. 3.2.1.4), que cataliza endohidrólisis de enlaces 1,4-beta-D-glicosídicos en la celulosa, derivados de celulosa (tales como carboximetilcelulosa y hidroxietil celulosa), liquenina, enlaces beta-1,4 en mezcla con beta-1,3 glucanos tales como beta-D-glucanos de cereal o xiloglucanos, y otro material vegetal con componentes celulósicos. Para fines de la presente invención, la actividad de endoglucanasa es determinada usando la hidrólisis de carboximetil celulosa (CMC) según el procedimiento de Ghose, 1987, *Pure and Appl. Chem.* 59: 257-268.

[0162] Las exo-1,4-beta-D-glucanasas incluyen ambas celobiohidrolasas y glucohidrolasas.

[0163] El término "celobiohidrolasa" se define aquí como una 1,4-beta-D-glucano celobiohidrolasa (E.C. 3.2.1.91), que cataliza la hidrólisis de enlaces 1,4-beta-D-glicosídicos en la celulosa, celooligosacáridos, o cualquier beta-1,4-glucosa unida con polímero, liberando celobiosa de los extremos reductores o no reductores de la cadena. Para fines de la presente invención, actividad de celobiohidrolasa se determina según los procedimientos descritos por Lever *et al.*, 1972, *Anal. Biochem.* 47: 273-279 y por van Tilbeurgh *et al.*, 1982, *FEBS Letters* 149: 152-156; van Tilbeurgh and Claeysens, 1985, *FEBS Letters* 187: 283-288. En la presente invención, el método de Lever *et al.* fue empleado para valorar hidrólisis de celulosa en los restos de maíz, mientras el método de van Tilbeurgh *et al.* fue usado para determinar la actividad de celobiohidrolasa en un derivado de disacárido fluorescente.

[0164] El término "glucohidrolasa" se define aquí como una 1,4-beta-D-glucano glucohidrolasa (E.C. 3.2.1.74), que cataliza la hidrólisis de enlaces 1,4 (enlaces O-glicosil) en 1,4-beta-D-glucanos para eliminar unidades de glucosa sucesivas. Para fines de la presente invención, actividad de exoglucanasa se determina según el procedimiento descrito por Himmel *et al.*, 1986, *J. Biol. Chem.* 261: 12948-12955.

[0165] El término "beta-glucosidasa" se define aquí como una beta-D-glucósido glucohidrolasa (E.C. 3.2.1.21) que cataliza la hidrólisis de residuos de beta-D-glucosa terminal no-reductores con la liberación de beta-D-glucosa. Para fines de la presente invención, actividad de beta-glucosidasa se determina según el procedimiento básico descrito por Venturi *et al.*, 2002, *J. Basic Microbiol.* 42: 55-66, excepto que condiciones diferentes fueron empleadas como se describe en este caso. Una unidad de actividad de beta-glucosidasa es definida como 1,0 µmol de p-nitrofenol producido por minuto a 50°C, pH 5 de 4 mM de P-nitrofenil-beta-d-glucopiranosido como sustrato en 100 mM de citrato sódico, 0,01% de Tween-20.

[0166] Los polipéptidos de la presente invención se usan conjuntamente con proteínas celulolíticas para degradar el componente celulósico del sustrato de biomasa, (ver, por ejemplo, Brigham *et al.*, 1995, en *Handbook on Bioethanol* (Charles E. Wyman, editor), pp.119-141, Taylor & Francis, Washington D.C.; Lee, 1997, *Journal of Biotechnology* 56: 1-24).

[0167] Las cantidades óptimas de un polipéptido con actividad de mejora celulolítica y de proteínas celulolíticas depende de los diferentes factores incluyendo, pero no limitado a, la mezcla de componentes de proteínas celulolíticas, el sustrato celulósico, la concentración de sustrato celulósico, el pretratamiento(s) del sustrato celulósico, temperatura, tiempo, pH, e inclusión de organismo de fermentación (p. ej., levadura para sacarificación y fermentación simultáneas). El término "proteínas celulolíticas" se define aquí como aquellas proteínas o mezclas de proteínas que muestran ser capaces de hidrolización o conversión o degradación de celulosa bajo las condiciones evaluadas. Sus cantidades son normalmente medidas por un ensayo común tal como BCA (ácido bicinónico, P.K. Smith *et al.*, 1985, *Anal. Biochem.* 150: 76), y la cantidad preferida añadida en proporción a la cantidad de biomasa que se hidroliza. La cantidad de polipéptido con actividad de mejora celulolítica por g de material celulósico puede ser de aproximadamente 0,01 a aproximadamente 2,0 mg, preferiblemente de aproximadamente 0,025 a aproximadamente 1,5 mg, más preferiblemente de aproximadamente 0,05 a aproximadamente 1,25 mg, más preferiblemente de aproximadamente 0,075 a aproximadamente 1,25 mg, más preferiblemente de aproximadamente 0,01 a aproximadamente 1,25 mg, incluso más preferiblemente de aproximadamente 0,15 a aproximadamente 1,25 mg, y de forma más preferida aproximadamente de 0,25 a aproximadamente 1,0 mg por g de material celulósico.

[0168] La cantidad de proteínas celulolíticas por g de material celulósico puede ser de aproximadamente 0,5 a aproximadamente 50 mg, preferiblemente de aproximadamente 0,5 a aproximadamente 40 mg, más preferiblemente de aproximadamente 0,5 a aproximadamente 25 mg, más preferiblemente de aproximadamente 0,75 a aproximadamente 20 mg, más preferiblemente aproximadamente 0,75 a aproximadamente 15 mg, incluso más preferiblemente de aproximadamente 0,5 a aproximadamente 10 mg, y de forma más preferida de aproximadamente 2,5 a aproximadamente 10 mg por g de material celulósico.

[0169] La cantidad de polipéptido con actividad de mejora celulolítica por g de proteínas celulolíticas puede ser de aproximadamente 0,005 a aproximadamente 1,0 g, preferiblemente de aproximadamente 0,01 a aproximadamente 1,0 g, más preferiblemente de aproximadamente 0,15 a aproximadamente 0,75 g, más preferiblemente de aproximadamente 0,15 a aproximadamente 0,5 g, más preferiblemente de aproximadamente 0,1 a aproximadamente 0,5 g, incluso más preferiblemente de aproximadamente 0,1 a aproximadamente 0,5 g, y de forma más preferida de aproximadamente 0,05 a aproximadamente 0,2 g por g de proteínas celulolíticas.

[0170] Los métodos de la presente invención se pueden utilizar para procesar un material celulósico para muchos productos orgánicos, productos químicos y combustibles útiles. Además de etanol, algunos productos químicos de consumo diario y de especialidad que se pueden producir a partir de celulosa incluyen xilosa, acetona, acetato, glicina, lisina, ácidos orgánicos (p. ej., ácido láctico), 1,3-propanodiol, butanodiol, glicerol, etilenglicol, furfuro, polihidroxialcanoatos, cis, ácido cis-mucónico, y pienso para animales (Lynd, L. R., Wyman, C. E., y Gerngross, T. U., 1999, *Biocommodity Engineering, Biotechnol. Prog.*, 15: 777-793; Philippidis, G. P., 1996, *Cellulose bioconversion technology*, en *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212; y Ryu, D. D. Y., y Mandels, M., 1980, *Cellulases: biosynthesis and applications*, *Enz. Microb. Technol.*, 2: 91-102). Beneficios de coproducción potenciales se extienden más allá de la síntesis de productos múltiples orgánicos de carbohidrato fermentable. Residuos ricos en lignina restantes después de tratamiento biológico pueden convertirse en productos químicos derivados de lignina, o usarse para producción de energía.

[0172] Métodos convencionales usados para procesar el material celulósico conforme a los métodos de la presente invención son bien entendidos por los expertos en la técnica. Los métodos de la presente invención pueden ser implementados usando cualquier aparato de tratamiento de biomasa convencional configurado para operar conforme a la invención.

[0173] Tal aparato puede incluir un reactor agitado por lotes, un reactor agitado de flujo continuo con ultrafiltración, un reactor de columna de flujo de tapón continua (Gusakov, A. V., and Sinitsyn, A. P., 1985, *Kinetics of the enzymatic hydrolysis of cellulose: 1. A mathematical model for a batch reactor process*, *Enz. Microb. Technol.* 7: 346-352), un reactor de fricción (Ryu, S. K., and Lee, J. M., 1983, *Bioconversion of waste cellulose by using an attrition bioreactor*, *Biotechnol. Bioeng.* 25: 53-65), o un reactor con agitación intensiva inducido por un campo electromagnético (Gusakov, A. V., Sinitsyn, A. P., Davydkin, I. Y., Davydkin, V. Y., Protas, O. V., 1996, *Enhancement of enzymatic cellulose hydrolysis using a novel type of bioreactor with intensive stirring induced by electromagnetic field*, *Appl. Biochem. Biotechnol.* 56: 141-153).

[0174] Los métodos convencionales incluyen, pero de forma no limitativa, sacarificación, fermentación, hidrólisis y fermentación separadas (SHF), sacarificación y fermentación simultáneas (SSF), sacarificación y cofermentación simultáneas (SSCF), hidrólisis híbrida y fermentación (HHF) y conversión directa microbiana (DMC).

[0175] SHF usa pasos de proceso separados para primero hidrolizar enzimáticamente celulosa a glucosa y luego fermentar glucosa a etanol. En SSF, la hidrólisis enzimática de celulosa y la fermentación de glucosa a etanol se

combina en un paso (Philippidis, G. P., 1996, *Cellulose bioconversion technology*, en *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212). SSCF incluye la cofermentación de azúcares múltiples (Sheehan, J., and Himmel, M., 1999, *Enzymes, energy and the environment: A strategic perspective on the U.S. Department of Energy's research and development activities for bioethanol*, *Biotechnol. Prog.* 15: 817-827). HHF incluye dos pasos separados realizados en el mismo reactor pero a temperaturas diferentes, es decir, sacarificación enzimática de alta temperatura seguida de SSF a una temperatura inferior que la cepa de fermentación puede tolerar. DMC combina todos los tres procesos (producción de celulasa, hidrólisis de celulosa y fermentación) en un paso (Lynd, L. R., Weimer, P. J., van Zyl, W. H., and Pretorius, I. S., 2002, *Microbial cellulose utilization: Fundamentals and biotechnology*, *Microbiol. Mol. Biol. Reviews* 66: 506-577).

[0176] "Fermentación" o "proceso de fermentación" se refieren a cualquier proceso de fermentación o cualquier proceso comprendiendo un paso de fermentación. Un proceso de fermentación incluye, sin limitación, procesos de fermentación usados para producir productos de fermentación incluyendo alcoholes (p. ej., arabinitol, butanol, etanol, glicerol, metanol, 1,3-propanediol, sorbitol, y xilitol); ácidos orgánicos (p. ej., ácido acético, ácido acetónico, ácido adípico, ácido ascórbico, ácido cítrico, ácido 2,5-diceto-D-glucónico, ácido fórmico, ácido fumárico, ácido glucárico, ácido glucónico, ácido glucurónico, ácido glutárico, ácido 3-hidroxipropiónico, ácido itacónico, ácido láctico, ácido málico, ácido malónico, ácido oxálico, ácido propiónico, ácido succínico y ácido xilónico); cetonas (p. ej., acetona); aminoácidos (p. ej., ácido aspártico, ácido glutámico, glicina, lisina, serina y treonina); gases (p. ej., metano, hidrógeno (H₂), dióxido de carbono (CO₂) y monóxido de carbono (CO)). Procesos de fermentación también incluyen procesos de fermentación usados en la industria del alcohol consumible (p. ej., cerveza y vino), industria de lechería (p. ej., productos lácteos fermentados), industria de cuero e industria del tabaco.

[0177] La presente invención además se refiere a métodos para producir una sustancia orgánica, comprendiendo: (a) sacarificación de un material celulósico con una cantidad eficaz de una proteína celulolítica en presencia de una cantidad eficaz de un polipéptido con actividad de mejora celulolítica, donde la presencia del polipéptido con actividad de mejora celulolítica aumenta la degradación de material celulósico en comparación con la ausencia del polipéptido con actividad de mejora celulolítica; (b) fermentación del material sacarificado celulósico del paso (a) con uno o más microorganismos fermentadores; y (c) recuperación de la sustancia orgánica de la fermentación. El polipéptido con actividad de mejora celulolítica puede estar en forma de un caldo de fermentación crudo con o sin células o en forma de una preparación enzimática semi-purificada o purificada. La proteína de mejora celulolítica puede ser una preparación monocomponente, por ejemplo, una proteína de la familia 61, una preparación de proteína de varios componentes, por ejemplo, varias proteínas de la familia 61, o una combinación de preparaciones de proteína monocomponente y de varios componentes.

[0178] La sustancia orgánica puede ser cualquier sustancia derivada de la fermentación. En un aspecto preferido, la sustancia orgánica es un alcohol. Se entenderá que el término "alcohol" abarca una sustancia orgánica que contiene una o más fracciones de hidróxilo. En una forma de realización más preferida, el alcohol es arabinitol. En otra forma de realización más preferida, el alcohol es butanol. En otra forma de realización más preferida, el alcohol es etanol. En otra forma de realización más preferida, el alcohol es glicerol. En otra forma de realización más preferida, el alcohol es metanol. En otra forma de realización más preferida, el alcohol es 1,3-propanediol. En otra forma de realización más preferida, el alcohol es sorbitol. En otra forma de realización más preferida, el alcohol es xilitol. Ver, por ejemplo, Gong, C. S., Cao, N. J., Du, J., and Tsao, G. T., 1999, *Ethanol production from renewable resources*, in *Advances in Biochemical Engineering/Biotechnology*, Scheper, T., ed., Springer-Verlag Berlin Heidelberg, Germany, 65: 207-241; Silveira, M. M., and Jonas, R., 2002, *The biotechnological production of sorbitol*, *Appl. Microbiol. Biotechnol.* 59: 400-408; Nigam, P.; and Singh, D., 1995, *Processes for fermentative production of xylitol - a sugar substitute*, *Process Biochemistry* 30 (2): 117-124; Ezeji, T. C., Qureshi, N. and Blaschek, H. P., 2003, *Production of acetone, butanol and ethanol by Clostridium beijerinckii BA101 and in situ recovery by gas stripping*, *World Journal of Microbiology and Biotechnology* 19 (6): 595-603.

[0179] En otra forma de realización preferida, la sustancia orgánica es un ácido orgánico. En otra forma de realización más preferida, el ácido orgánico es ácido acético. En otra forma de realización más preferida, el ácido orgánico es ácido acetónico. En otra forma de realización más preferida, el ácido orgánico es ácido adípico. En otra forma de realización más preferida, el ácido orgánico es ácido ascórbico. En otra forma de realización más preferida, el ácido orgánico es ácido cítrico. En otra forma de realización más preferida, el ácido orgánico es ácido 2,5-diceto-D-glucónico. En otra forma de realización más preferida, el ácido orgánico es ácido fórmico. En otra forma de realización más preferida, el ácido orgánico es ácido fumárico. En otra forma de realización más preferida, el ácido orgánico es ácido glucárico. En otra forma de realización más preferida, el ácido orgánico es ácido glucónico. En otra forma de realización más preferida, el ácido orgánico es ácido glucurónico. En otra forma de realización más preferida, el ácido orgánico es ácido glutárico. En otra forma de realización preferida, el ácido orgánico es ácido 3-hidroxipropiónico. En otra forma de realización más preferida, el ácido orgánico es ácido itacónico. En otra forma de realización más preferida, el ácido orgánico es ácido láctico. En otra forma de realización más preferida, el ácido orgánico es ácido málico. En otra forma de realización más preferida, el ácido orgánico es ácido malónico. En otra forma de realización más preferida, el ácido orgánico es ácido oxálico. En otra forma de realización más preferida, el ácido orgánico es ácido propiónico. En otra forma de realización más preferida, el ácido orgánico es ácido succínico. En otra forma de realización más preferida, el ácido orgánico es ácido xilónico. Ver, por ejemplo, Chen, R., and Lee, Y. Y., 1997, *Membrane-mediated extractive fermentation for lactic acid production from cellulosic biomass*, *Appl. Biochem. Biotechnol.* 63-65: 435-448.

[0180] En otra forma de realización preferida, la sustancia orgánica es una cetona. Se entenderá que el término "cetona" abarca una sustancia orgánica que contiene una o más fracciones de cetona. En otra forma de realización más preferida, la cetona es acetona. Ver, por ejemplo, Qureshi y Blaschek, 2003, *supra*.

[0181] En otra forma de realización preferida, la sustancia orgánica es un aminoácido. En otra forma de realización más preferida, el ácido orgánico es ácido aspártico. En otra forma de realización más preferida, el aminoácido es ácido glutámico. En otra forma de realización más preferida, el aminoácido es glicina. En otra forma de realización más preferida, el aminoácido es lisina. En otra forma de realización más preferida, el aminoácido es serina. En otra forma de realización más preferida, el aminoácido es treonina. Ver, por ejemplo, Richard, A., and Margaritis, A., 2004, *Empirical modeling of batch fermentation kinetics for poly(glutamic acid) production and other microbial biopolymers, Biotechnology and Bioengineering* 87 (4): 501-515.

[0182] En otra forma de realización preferida, la sustancia orgánica es un gas. En otra forma de realización más preferida, el gas es metano. En otra forma de realización más preferida, el gas es metano. En otra forma de realización más preferida, el gas es H₂. En otra forma de realización más preferida, el gas es CO₂. En otra forma de realización más preferida, el gas es CO. Ver, por ejemplo, Kataoka, N., A. Miya, and K. Kiriya, 1997, *Studies on hydrogen production by continuous culture system of hydrogen-producing anaerobic bacteria, Water Science and Technology* 36 (6-7): 41-47; y Gunaseelan V.N. en *Biomass and Bioenergy, Vol. 13 (1-2), pp. 83-114, 1997, Anaerobic digestion of biomass for methane production: A review*.

[0183] Producción de una sustancia orgánica de material celulósico típicamente requiere cuatro pasos importantes. Estos cuatro pasos son pretratamiento, hidrólisis enzimática, fermentación y recuperación. Ejemplificado debajo está un proceso para producir etanol, pero se entenderá que procesos similares pueden utilizarse para producir otras sustancias orgánicas, por ejemplo, las sustancias anteriormente descritas.

[0184] Pretratamiento. En el paso de pretratamiento o de pre-hidrólisis, el material celulósico se calienta para descomponer la lignina y la estructura de carbohidrato, solubilizar la mayor parte de la hemicelulosa y hacer la fracción de celulosa accesible a enzimas celulolíticas. La calefacción es realizada bien directamente con vapor o en el compuesto acuoso donde un catalizador puede también ser añadido al material para acelerar las reacciones. Catalizadores incluyen ácidos fuertes, tales como ácido sulfúrico y SO₂, o álcali, tal como hidróxido sódico. El propósito de la fase de pretratamiento es facilitar la penetración de las enzimas y microorganismos. Biomasa celulósica puede también ser sujeto para un pretratamiento de explosión de vapor hidrotérmico (ver solicitud de patente US n°. 20020164730).

[0185] Sacarificación. En el paso de hidrólisis enzimática, también conocido como sacarificación, enzimas como se describe en este caso se agregan al material pretratado para convertir la fracción de celulosa a glucosa y/o otros azúcares. La sacarificación es generalmente realizada en los reactores de tanque agitado o fermentadores bajo condiciones de pH, temperatura y mezcla controladas. Un paso de sacarificación puede durar hasta 200 horas. La sacarificación se puede llevar a cabo a temperaturas de aproximadamente 30°C a aproximadamente 65°C, en particular, alrededor de 50°C, y a un pH en el intervalo entre aproximadamente 4 y aproximadamente 5, especialmente alrededor de pH 4.5. Para producir glucosa que se puede metabolizar por levadura, la hidrólisis es típicamente realizada en presencia de una beta-glucosidasa.

[0186] Fermentación. En el paso de fermentación, azúcares, liberados del material celulósico como resultado del pretratamiento y pasos de hidrólisis enzimática, se fermentan a etanol por un organismo de fermentación, tal como levadura. La fermentación puede también efectuarse simultáneamente con la hidrólisis enzimática en el mismo recipiente, nuevamente bajo condiciones de pH, temperatura y mezcla controladas. Cuando la sacarificación y la fermentación se realizan simultáneamente en el mismo recipiente, el proceso es generalmente denominado sacarificación y fermentación simultáneas o SSF.

[0187] Cualquier sustrato celulósico o materia prima adecuados se puede utilizar en un proceso de fermentación de la presente invención. El sustrato es generalmente seleccionado basado en el producto de fermentación deseado, es decir, la sustancia orgánica para ser obtenida de la fermentación, y el proceso empleado, como es bien conocido en la técnica. Ejemplos de sustratos adecuados para uso en los métodos de presente invención, incluyen materiales con celulosa, tal como madera o residuos de planta o azúcares de peso molecular bajo DP1-3 obtenidos de material celulósico procesado que se puede metabolizar por el microorganismo fermentador y que se puede suministrar por adición directa al medio de fermentación.

[0188] El término "medio de fermentación" se entenderá para referirse a un medio antes de que el microorganismo(s) fermentador(es) es(son) añadido, tal como, un medio que resulta de un proceso de sacarificación, al igual que un medio usado en una proceso de sacarificación y fermentación simultáneas (SSF).

[0189] "Microorganismo fermentador" se refiere a cualquier microorganismo adecuado para uso en un proceso de fermentación deseado. Microorganismos fermentadores adecuados según la invención son capaces de fermentar, es decir, convertir azúcares, tales como glucosa, xilosa, arabinosa, manosa, galactosa, u oligosacáridos directa o

indirectamente en el producto de fermentación deseado. Ejemplos de microorganismos fermentadores incluyen organismos fúngicos, tal como levadura. Levadura preferida incluye cepas de la *Saccharomyces spp.* y, en particular, *Saccharomyces cerevisiae*. Levadura disponible comercialmente incluye, por ejemplo, Red Star®/™/Lesaffre Ethanol Red (disponible de Red Star/Lesaffre, EEUU) FALI (disponible de Fleischmann's Yeast, una división de Burns Philp Food Inc., EEUU), SUPERSTART (disponible de Alltech), GERT STRAND (disponible de Gert Strand AB, Suecia) y FERMIOL (disponible de DSM Specialties).

[0190] En una forma de realización preferida, la levadura es una *Saccharomyces spp.* En una forma de realización más preferida, la levadura es *Saccharomyces cerevisiae*. En una forma de realización más preferida, la levadura es *Saccharomyces distaticus*. En una forma de realización más preferida, la levadura es *Saccharomyces uvarum*. En una forma de realización más preferida, la levadura es un *Kluyveromyces*. En una forma de realización más preferida, la levadura es *Kluyveromyces marxianus*. En una forma de realización más preferida, la levadura es *Kluyveromyces fragilis*. En una forma de realización más preferida, la levadura es una *Candida*. En una forma de realización más preferida, la levadura es *Candida pseudotropicalis*. En una forma de realización más preferida, la levadura es *Candida brassicae*. En una forma de realización preferida, la levadura es una *Clavispora*. En una forma de realización más preferida, la levadura es *Clavispora lusitaniae*. En una forma de realización más preferida, la levadura es *Clavispora opuntiae*. En una forma de realización preferida, la levadura es un *Pachysolen*. En una forma de realización más preferida, la levadura es *Pachysolen tannophilus*. En una forma de realización preferida, la levadura es un *Bretannomyces*. En una forma de realización más preferida, la levadura es *Bretannomyces clausenii* (Philippidis, G. P., 1996, *Cellulose bioconversion technology*, en *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 179-212).

[0191] Bacterias que pueden fermentar eficazmente glucosa a etanol incluyen, por ejemplo, *Zymomonas mobilis* y *Clostridium thermocellum* (Philippidis, 1996, *supra*).

[0192] Es bien conocido en la técnica que los organismos descritos anteriormente también pueden usarse para producir otras sustancias orgánicas, como se describe en este caso.

[0193] La clonación de genes heterólogos en *Saccharomyces cerevisiae* (*Saccharomyces cerevisiae* (Chen, Z., Ho, N. W. Y., 1993, *Cloning and improving the expression of Pichia stipitis xylose reductase gene in Saccharomyces cerevisiae*, *Appl. Biochem. Biotechnol.* 39-40: 135-147; Ho, N. W. Y., Chen, Z., Brainard, A. P., 1998, *Genetically engineered Saccharomyces yeast capable of effectively cofermenting glucose and xylose*, *Appl. Environ. Microbiol.* 64: 1852-1859), o en bacterias tales como *Escherichia coli* (Beall, D. S., Ohta, K., Ingram, L. O., 1991, *Parametric studies of ethanol production from xylose and other sugars by recombinant Escherichia coli*, *Biotech. Bioeng.* 38: 296-303), *Klebsiella oxytoca* (Ingram, L. O., Gomes, P. F., Lai, X., Moniruzzaman, M., Wood, B. E., Yomano, L. P., York, S. W., 1998, *Metabolic engineering of bacteria for ethanol production*, *Biotechnol. Bioeng.* 58: 204-214), y *Zymomonas mobilis* (Zhang, M., Eddy, C., Deanda, K., Finkelstein, M., and Picataggio, S., 1995, *Metabolic engineering of a pentose metabolism pathway in ethanologenic Zymomonas mobilis*, *Science* 267: 240-243; Deanda, K., Zhang, M., Eddy, C., and Picataggio, S., 1996, *Development of an arabinose-fermenting Zymomonas mobilis strain by metabolic pathway engineering*, *Appl. Environ. Microbiol.* 62: 4465-4470) han llevado a la construcción de organismos capaces de convertir hexosas y pentosas a etanol (cofermentación).

[0194] Levadura u otro microorganismo típicamente se añade a la celulosa degradada o hidrolizada y la fermentación está en curso durante de aproximadamente 24 a aproximadamente 96 horas, tal como de aproximadamente 35 a aproximadamente 60 horas. La temperatura está típicamente entre de aproximadamente 26°C a aproximadamente 40°C, en particular, a aproximadamente 32°C, y de sobre pH 3 a sobre pH 6, en particular, alrededor de pH 4-5.

[0195] En una forma de realización preferida, levadura u otro microorganismo se aplica a la celulosa degradada o hidrolizada y la fermentación está en curso durante de aproximadamente 24 a aproximadamente 96 horas, tal como típicamente 35-60 horas. En unas formas de realización preferidas, la temperatura está generalmente entre de aproximadamente 26 a aproximadamente 40°C, en particular, aproximadamente 32°C y el pH está generalmente de sobre pH 3 a sobre pH 6, preferiblemente alrededor de pH 4-5. Levadura u otro microorganismo se aplica preferiblemente en cantidades de aproximadamente 10⁵ a 10¹², preferiblemente de aproximadamente 10⁷ a 10¹⁰, especialmente aproximadamente 5x10⁷ a recuento viable por ml de caldo de fermentación. Durante una fase de producción de etanol, el recuento de células de levadura debería preferiblemente estar en el intervalo de aproximadamente 10⁷ a 10¹⁰, especialmente alrededor de aproximadamente 2 x 10⁸. Otras orientaciones respecto al uso de levadura para fermentación se pueden encontrar en, por ejemplo, "The Alcohol Textbook" (Editors K. Jacques, T.P. Lyons and D.R. Kelsall, Nottingham University Press, United Kingdom 1999), que es por la presente incorporado como referencia.

[0196] El proceso usado más ampliamente en la técnica es la sacarificación y fermentación simultáneas (SSF) proceso donde no hay fase de retención para la sacarificación, lo que significa que levadura y enzima se agregan juntas.

[0197] Para producción de etanol, después de la fermentación, la masa se destila para extraer el etanol. El etanol obtenido según el proceso de la invención puede ser utilizado como, por ejemplo, etanol de combustible; etanol potable, es decir, licores potables neutros, o etanol industrial.

[0198] Un estimulador de fermentación se puede utilizar en combinación con cualquiera de los procesos enzimáticos descritos aquí para mejorar además el proceso de fermentación y, en particular, el rendimiento del microorganismo fermentador, tal como, mejorar la velocidad y rendimiento del etanol. Un "estimulador de fermentación" se refiere a estimuladores para crecimiento de los microorganismos fermentadores, en particular, levadura. Estimuladores de fermentación preferidos para crecimiento incluyen vitaminas y minerales. Ejemplos de vitaminas incluyen multivitaminas, biotina, pantotenato, ácido nicotínico, meso-inositol, tiamina, piridoxina, ácido para-aminobenzoico, ácido fólico, riboflavina y vitaminas A, B, C D, y E. Ver, por ejemplo, Alfenore *et al.*, *Improving ethanol production and viability of Saccharomyces cerevisiae by a vitamin feeding strategy during fed-batch process*, Springer-Verlag (2002), que se incorpora por la presente como referencia. Ejemplos de minerales incluyen minerales y sales minerales que pueden suministrar nutrientes comprendiendo P, K, Mg, S, Ca, Fe, Zn, Mn y Cu.

[0199] Recuperación. El alcohol es separado del material celulósico fermentado y purificado por métodos convencionales de destilación. Etanol con una pureza de hasta aproximadamente 96 % vol. de etanol puede obtenerse, que se puede usar como, por ejemplo, etanol de combustible, etanol potable, es decir, licores potables neutros, o etanol industrial.

[0200] Para otras sustancias orgánicas, cualquier método conocido en la técnica se puede usar incluyendo, pero no limitado a, cromatografía (p. ej., intercambio de iones, afinidad, cromatografía de exclusión de tamaño), procedimientos electroforéticos (p. ej., isoelectroenfoque preparativo), solubilidad diferencial (p. ej., precipitación de sulfato amónico), SDS-PAGE, destilación, o extracción.

[0201] En los métodos de la presente invención, las proteína(s) celulolíticas y polipéptido(s) de mejora celulolítica se puede complementar por una o más actividades enzimáticas adicionales para mejorar la degradación del material celulósico. Enzimas preferidas adicionales son hemicelulasas, esterasas (p. ej., lipasas, fosfolipasas y/o cutinasas), proteasas, lacasas, peroxidasas, o mezclas derivadas.

[0202] En los métodos de la presente invención, la enzima(s) adicional puede ser añadida antes o durante fermentación, incluso durante o después de la propagación del microorganismo(s) fermentador.

[0203] Las enzimas referenciadas aquí pueden ser derivadas u obtenidas de cualquier origen adecuado, incluyendo origen bacteriano fúngico, de levadura o mamífero. El término "obtenido" significa aquí que la enzima puede haber sido aislada de un organismo que produce naturalmente la enzima como una enzima nativa. El término "obtenido" significa también aquí que la enzima puede haber sido producida de manera recombinante en un organismo huésped, donde la enzima producida de manera recombinante es bien nativa o extranjera al organismo huésped o tiene una secuencia de aminoácidos modificada, por ejemplo, teniendo uno o más aminoácidos que son eliminados, insertados y/o sustituidos, es decir, una enzima producida de manera recombinante que es un mutante y/o un fragmento de una secuencia de aminoácidos nativa o una enzima producida por procesos de redistribución de ácido nucleico conocidos en la técnica. Abarcado en el significado de una enzima nativa son variantes naturales y en el significado de una enzima extranjera son variantes obtenidas de manera recombinante, tal como por mutagénesis dirigida o redistribución.

[0204] Las enzimas pueden también ser purificadas. El término "purificado" como se utiliza en este caso cubre enzimas libres de otros componentes del organismo del que se deriva. El término "purificado" también cubre enzimas libres de componentes del organismo nativo del que es obtenida. Las enzimas puede ser purificadas, con cantidades solo menores de otra proteína que están presentes. La expresión "otras proteínas" se refiere en particular a otras enzimas. El término "purificado" como se utiliza en este caso también se refiere a eliminación de otros componentes, particularmente otras proteínas y más particularmente otras enzimas presentes en la célula de origen de la enzima de la invención. La enzima puede ser "substancialmente pura", que es, libre de otros componentes del organismo en el que es producida, que es, por ejemplo, un organismo huésped para enzimas producidas de manera recombinante. En un aspecto preferido, las enzimas son al menos 75% (p/p), preferiblemente al menos 80%, más preferiblemente al menos 85%, más preferiblemente al menos 90%, más preferiblemente al menos 95%, más preferiblemente al menos 96%, más preferiblemente al menos 97%, incluso más preferiblemente al menos 98%, o de forma más preferida al menos 99% puras. En otro aspecto preferido, la enzima es 100% pura.

[0205] Las enzimas usadas en la presente invención pueden estar en cualquier forma adecuada para uso en los procesos descritos aquí, tal como, por ejemplo, un caldo de fermentación crudo con o sin células, un polvo seco o granulado, un granulado no polvoriento, un líquido, un líquido estabilizado, o una enzima protegida. Granulados pueden ser producidos, por ejemplo, como se describen en las patentes US n.º. 4,106,991 y 4,661,452, y pueden ser revestidos opcionalmente por procesos conocidos en la técnica. Preparaciones enzimáticas líquidas pueden, por ejemplo, ser estabilizadas añadiendo estabilizadores tales como un azúcar, un alcohol de azúcar u otro poliol, y/o ácido láctico u otro ácido orgánico según proceso establecido. Enzimas protegidas se pueden preparar según el proceso descrito en la EP 238,216.

Hemicelulasas

[0206] Hidrólisis enzimática de hemicelulosas se puede realizar por una amplia variedad de hongos y bacterias. Similar a la degradación de celulosa, hidrólisis de hemicelulosa requiere acción coordinada de muchas enzimas. Hemicelulasas se pueden colocar en tres categorías generales: las enzimas de acción endo que atacan enlaces internos en la cadena polisacárida, las enzimas de acción exo que actúan progresivamente de bien el extremo reductor o no reductor de la cadena polisacárida, y las enzimas accesorio, acetilesterasas y esterases que hidrolizan enlaces de lignina glucósido, tales como esterasa de ácido cumárico y esterasa de ácido ferúlico (Wong, K. K. Y., Tan, L. U. L., and Saddler, J. N., 1988, *Multiplicity of β -1,4-xylanase in microorganisms: Functions and applications*, *Microbiol. Rev.* 52: 305-317; Tenkanen, M., and Poutanen, K., 1992, *Significance of esterases in the degradation of xylans*, en *Xylans and Xylanases*, Visser, J., Beldman, G., Kuster-van Someren, M. A., and Voragen, A. G. J., eds., Elsevier, New York, NY, 203-212; Coughlan, M. P., and Hazlewood, G. P., 1993, *Hemicellulose and hemicellulases*, Portland, London, UK; Brigham, J. S., Adney, W. S., and Himmel, M. E., 1996, *Hemicellulases: Diversity and applications*, en *Handbook on Bioethanol: Production and Utilization*, Wyman, C. E., ed., Taylor & Francis, Washington, DC, 119-141).

[0207] Hemicelulasas incluyen xilanasas, arabinofuranosidasas, acetil xilano esterasa, glucuronidasas, endo-galactanasa, mananasas, endo o endo arabinasas, endo-galactanses y sus mezclas derivadas. Ejemplos de hemicelulasas de acción endo y enzimas auxiliares incluyen endoarabinanasa, endoarabinogalactanasa, endoglucanasa, endomananasa, endoxilanasas y feraxan endoxilanasas. Ejemplos de hemicelulasas de acción endo y enzimas auxiliares incluyen β -L-arabinosidasa, α -L-arabinosidasa, α -1,2-L-fucosidasa, α -D-galactosidasa, β -D-galactosidasa, β -D-glucosidasa, β -D-glucuronidasa, β -D-manosidasa, β -D-xilosidasa, exoglucosidasa, exocelobiohidrolasa, exomanobiohidrolasa, exomananasa, exoxilanasas, xilano α -glucuronidasa y coniferin β -glucosidasa. Ejemplos de esterases incluyen acetil esterases (acetilgalactano esterasa, acetilmanano esterasa y acetilxilano esterasa) y aril esterases (esterasa de ácido cumárico y esterasa de ácido ferúlico).

[0208] Preferiblemente, la hemicelulasa es una hemicelulasa de acción exo y, más preferiblemente, una hemicelulasa de acción exo que tiene la capacidad para hidrolizar hemicelulosa bajo condiciones ácidas por debajo de pH 7. Un ejemplo de una hemicelulasa adecuada para uso en la presente invención incluye VISCOZYME™ (disponible de Novozymes A/S).

La hemicelulasa se añade en una cantidad eficaz de aproximadamente 0,001% a aproximadamente 5,0% en peso de sólidos, más preferiblemente de aproximadamente 0,025% a aproximadamente 4,0% en peso de sólidos y de forma más preferida de aproximadamente 0,005% a aproximadamente 2,0% en peso de sólidos.

[0209] Una xilanasas (E.C. 3.2.1.8) se puede obtener de cualquier fuente adecuada, incluyendo organismos bacterianos y fúngicos, tales como *Aspergillus*, *Disporotrichum*, *Penicillium*, *Neurospora*, *Fusarium*, *Trichoderma*, *Hemicola*, *Thermomyces* y *Bacillus*. Preparaciones preferidas disponibles comercialmente comprendiendo xilanasas incluyen SHEARZYME®, BIOFEED WHEAT®, BIOFEED Plus®L, CELLUCLAST®, ULTRAFLO®, VISCOZYME®, PENTOPAN MONO® BG, y PULPZYME® HC (Novozymes A/S) y LAMINEX® and SPEZYME® CP (Genencor Int.).

Esterasas

[0210] Esterasas que se pueden usar para bioconversión de celulosa incluyen acetil esterases tales como acetilgalactano esterasa, acetilmanano esterasa y acetilxilano esterasa, y esterases que hidrolizan enlaces de lignina glucósido, tales como esterasa de ácido cumárico y esterasa de ácido ferúlico.

[0211] Como se utiliza en este caso, una "esterasa" conocida también como éster hidrolasa carboxílica, se refiere a enzimas que actúan en enlaces estéricos e incluye enzimas clasificadas en EC 3.1.1 hidrolasas de éster carboxílico según nomenclatura enzimática (*Enzyme Nomenclature*, 1992, Academic Press, San Diego, California, con Suplemento 1 (1993), Suplemento 2 (1994), Suplemento 3 (1995), Suplemento 4 (1997) y Suplemento 5, en *Eur. J. Biochem.* 223: 1-5, 1994; *Eur. J. Biochem.* 232: 1-6, 1995; *Eur. J. Biochem.* 237: 1-5, 1996; *Eur. J. Biochem.* 250: 1-6, 1997 y *Eur. J. Biochem.* 264: 610-650, 1999; respectivamente). Ejemplos no limitativos de esterases incluyen arilesterasa, triacilglicerol-lipasa, acetilesterasa, acetilcolinesterasa, colinesterasa, tropinesterasa, pectinesterasa, esteroil-esterasa, clorofilasa, L-arabinonolactonasa, gluconolactonasa, uronolactonasa, tanasa, retinil-palmitato esterasa, hidroxibutirato-dímero hidrolasa, acilglicerol lipasa, 3-oxoadipato enol-lactonasa, 1,4-lactonasa, galactolipasa, 4-piridoxolactonasa, acilcarnitina hidrolasa, aminoacil-ARNt hidrolasa, D-arabinonolactonasa, 6-fosfogluconolactonasa, fosfolipasa A1, 6-acetilglucosa deacetilasa, lipoproteína-lipasa, dihidrocoumarina lipasa, limonina-D-anillo-lactonasa, esteroide-lactonasa, triacetato-lactonasa, actinomicina lactonasa, orsellinato-depsido hidrolasa, cefalosporina-C desacetilasa, clorogenato hidrolasa, alfa-aminoácido esterasa, 4-metiloxaloacetato esterasa, carboximetilenobutenolidasas, desoxilimonato A-anillo-lactonasa, 2-acetil-1-alkilglicerofosfocolina esterasa, fusarinina-C omitinesterasa, sinapina esterasa, cera-éster hidrolasa, forbol-diéster hidrolasa, fosfatidilinositol deacilasa, sialato O-acetilesterasa, acetoxibutirilbitiofeno deacetilasa, acetilsalicilato deacetilasa, metilumbeliferil-acetato desacetilasa, 2-pirona 4,6-dicarboxilato lactonasa, N-acetilgalactosaminoglicano desacetilasa, esterasa de hormona juvenil, bis(2-etilhexil) esterasa de ftalato, glutamato metilesterasa de proteína, 11- hidrolasa de cis-retinilo-palmitato, hidrolasa de todos-trans-retinilo-palmitato, L-ramnono-1,4-lactonasa, 5-(3,4-diacetoxibut-1-inil)-2,2'-bitiofeno deacetilasa, etil-éster-sintasa de ácido graso, xilono-1,4-lactonasa, N-acetilglucosaminilfosfatidilinositol desacetilasa, cetraxato bencilsterasa, acetilalquilglicerol acetilhidrolasa y acetilxilano esterasa.

[0212] Esterasas preferidas para uso en la presente invención son enzimas lipolíticas, tales como lipasas (clasificadas como EC 3.1.1.3, EC 3.1.1.23 y/o EC 3.1.1.26) y fosfolipasas (clasificadas como EC 3.1.1.4 y/o EC 3.1.1.32, incluyendo lisofosfolipasas clasificadas como EC 3.1.1.5). Otras esterazas preferidas son cutinasas (clasificadas como EC 3.1.1.74).

[0213] La esteraza se puede añadir en una cantidad eficaz para obtener el beneficio deseado para mejorar el rendimiento del microorganismo fermentador, por ejemplo, para cambiar la composición/concentración lipídica interior y/o exterior del microorganismo fermentador o en la membrana celular del microorganismo fermentador, para suponer una mejora en el movimiento de solutos en y/o fuera de los microorganismos fermentadores durante fermentación y/o para proporcionar fuentes de energía más metabolizables (tales como, por ejemplo, conversión de componentes, tales como aceite del sustrato de maíz, a componentes útiles, el microorganismo fermentador, por ejemplo, ácidos insaturados grasos y glicerol), para aumentar rendimiento de etanol. Ejemplos de cantidades eficaces de esteraza son de aproximadamente 0,01 a aproximadamente 400 LU/g DS (sólidos secos). Preferiblemente, la esteraza se usa en una cantidad de aproximadamente 0,1 a aproximadamente 100 LU/g DS, más preferiblemente de aproximadamente 0,5 a aproximadamente 50 LU/g DS, e incluso más preferiblemente de aproximadamente 1 a aproximadamente 20 LU/g DS. Posterior optimización de la cantidad de esteraza se puede obtener de aquí en adelante utilizando procedimientos estándar conocidos en la técnica.

[0214] Una unidad lipásica (LU) es la cantidad de enzima que libera 1,0 μmol de ácido graso con tributirino por minuto como sustrato y goma arábica como emulsionante a 30°C, pH 7,0 (tampón fosfato).

[0215] En una forma de realización preferida, la esteraza es una enzima lipolítica, más preferiblemente, una lipasa. Como se utiliza en este caso, una "enzima lipolítica" se refiere a lipasas y fosfolipasas (incluyendo lisofosfolipasas). La enzima lipolítica es preferiblemente de origen microbiano, en particular de origen bacteriano, fúngico o de levadura. La enzima lipolítica usada puede derivar de cualquier fuente, incluyendo, por ejemplo, una cepa de *Absidia*, en particular, *Absidia blakesleena* y *Absidia corymbifera*, una cepa de *Achromobacter*, en particular, *Achromobacter iofagus*, una cepa de *Aeromonas*, una cepa de *Alternaria*, en particular, *Alternaria brassiciola*, una cepa de *Aspergillus*, en particular, *Aspergillus niger*, *Aspergillus oryzae*, *Aspergillus fumigatus* y *Aspergillus flavus*, una cepa de *Achromobacter*, en particular, *Achromobacter iofagus*, una cepa de *Aureobasidium*, en particular, *Aureobasidium pullulans*, una cepa de *Bacillus*, en particular *Bacillus pumilus*, *Bacillus stearothermophilus* y *Bacillus subtilis*, una cepa de *Beauveria*, una cepa de *Brochothrix*, en particular, *Brochothrix thermosohata*, una cepa de *Candida*, en particular, *Candida cylindracea* (*Candida rugosa*), *Candida parapolityca* y *Candida antarctica*, una cepa de *Chromobacter*, en particular, *Chromobacter viscosum*, una cepa de *Coprinus*, en particular, *Coprinus cinereus*, una cepa de *Fusarium*, en particular, *Fusarium graminearum*, *Fusarium oxysporum*, *Fusarium solani*, *Fusarium solani pisi*, *Fusarium roseum culmorum* y *Fusarium venenatum*, una cepa de *Geotricum*, en particular, *Geotricum penicillatum*, una cepa de *Hansenula*, en particular, *Hansenula anomala*, una cepa de *Humicola*, en particular, *Humicola brevispora*, *Humicola brevis var. thermoidea* y *Humicola insolens*, una cepa de *Hifozima*, una cepa de *Lactobacillus*, en particular, *Lactobacillus curvatus*, una cepa de *Metarhizium*, una cepa de *Mucor*, una cepa de *Paecilomyces*, una cepa de *Penicillium*, en particular, *Penicillium cyclopium*, *Penicillium crustosum* y *Penicillium expansum*, una cepa de *Pseudomonas*, en particular, *Pseudomonas aeruginosa*, *Pseudomonas alcaligenes*, *Pseudomonas cepacia* (sin. *Burkholderia cepacia*), *Pseudomonas fluorescens*, *Pseudomonas fragi*, *Pseudomonas maltophilia*, *Pseudomonas mendocina*, *Pseudomonas mephitica lipolytica*, *Pseudomonas alcaligenes*, *Pseudomonas plantari*, *Pseudomonas pseudoalcaligenes*, *Pseudomonas putida*, *Pseudomonas stutzeri* y *Pseudomonas wisconsinensis*, una cepa de *Rhizooctonia*, en particular, *Rhizooctonia solani*, una cepa de *Rhizomucor*, en particular, *Rhizomucor miehei*, una cepa de *Rhizopus*, en particular, *Rhizopus japonicus*, *Rhizopus microsporus* y *Rhizopus nodosus*, una cepa de *Rhodosporidium*, en particular, *Rhodosporidium toruloides*, una cepa de *Rhodotorula*, en particular, *Rhodotorula glutinis*, una cepa de *Sporobolomyces*, en particular, *Sporobolomyces shibatanus*, una cepa de *Thermomyces*, en particular, *Thermomyces lanuginosus* (anteriormente *Humicola lanuginosa*), una cepa de *Thiarosporella*, en particular, *Thiarosporella phaseolina*, una cepa de *Trichoderma*, en particular, *Trichoderma harzianum* y *Trichoderma reesei*, y/o una cepa de *Verticillium*.

[0216] En una forma de realización preferida, la enzima lipolítica deriva de una cepa de *Aspergillus*, *Achromobacter*, *Bacillus*, *Candida*, *Chromobacter*, *Fusarium*, *Humicola*, *Hyphozyma*, *Pseudomonas*, *Rhizomucor*, *Rhizopus*, o *Thermomyces*.

[0217] En formas de realización más preferidas, la enzima lipolítica es una lipasa. Las lipasas se pueden aplicar aquí por su capacidad para modificar la estructura y composición de aceites de triglicéridos y grasas en los medios de fermentación (incluyendo levadura de fermentación), por ejemplo, se producen de un sustrato de maíz. Las lipasas catalizan diferentes tipos de conversiones de triglicéridos, tales como hidrólisis, esterificación y transesterificación. Lipasas adecuadas incluyen lipasas ácidas, neutras y básicas, como son bien conocidas en la técnica, aunque lipasas ácidas (tales como, por ejemplo, la lipasa G AMANO 50, disponible de Amano) parecen ser más eficaces a concentraciones inferiores de lipasa en comparación con bien lipasas básicas o neutras. Lipasas preferidas para uso en la presente invención incluyen lipasa de *Candida antarctica* y lipasa de *Candida cylindracea*. Lipasas más preferidas son lipasas purificadas tales como lipasa de *Candida antarctica* (lipasa A), lipasa de *Candida antarctica* (lipasa B), lipasa de *Candida cylindracea* y lipasa de *Penicillium camembertii*.

[0218] La lipasa puede ser la describa en la EP 258,068-A o puede ser una variante de lipasa tal como una variante descrita en la WO 00/60063 o WO 00/32758, incorporadas aquí como referencia. Lipasas comerciales preferidas incluyen LECITASE™, LIPOLASET™ y LIPEX™ (disponibles de Novozymes A/S) y G AMANO™ 50 (disponible de Amano).

5

[0219] Las lipasas se añaden preferiblemente en cantidades de aproximadamente 1 a aproximadamente 400 LU/g DS, preferiblemente de aproximadamente 1 a aproximadamente 10 LU/g DS y más preferiblemente de aproximadamente 1 a aproximadamente 5 LU/g DS.

10

[0220] En otra forma de realización preferida de la presente divulgación, la esterasa es una cutinasa. Las cutinasas son enzimas que son capaces de degradar cutina. La cutinasa puede derivar de cualquier fuente. En un aspecto preferido, la cutinasa deriva de una cepa de *Aspergillus*, en particular, *Aspergillus oryzae*, una cepa de *Alternaria*, en particular, *Alternaria brassiciola*, una cepa de *Fusarium*, en particular, *Fusarium solani*, *Fusarium solani* pisi, *Fusarium roseum culmorum*, o *Fusarium roseum sambucium*, una cepa de *Helminthosporum*, en particular, *Helminthosporum sativum*, una cepa de *Humicola*, en particular, *Humicola insolens*, una cepa de *Pseudomonas*, en particular, *Pseudomonas mendocina* o *Pseudomonas putida*, una cepa de *Rhizooctonia*, en particular, *Rhizooctonia solani*, una cepa de *Streptomyces*, en particular, *Streptomyces scabies*, o una cepa de *Ulocladium*, en particular, *Ulocladium consortiale*. En una forma de realización más preferida la cutinasa deriva de una cepa de *Humicola insolens*, en particular, la cepa de *Humicola insolens* DSM 1800. Cutinasa de *Humicola insolens* se describe en la WO 96/13580, que se incorpora por la presente como referencia. La cutinasa puede ser un variante tal como una de las variantes descritas en las WO 00/34450 y WO 01/92502, que son incorporadas por la presente como referencia. Variantes de cutinasa preferidas incluyen variantes catalogadas en el ejemplo 2 de la WO 01/92502 que es incorporada por la presente específicamente como referencia. Una cantidad eficaz de cutinasa es de aproximadamente 0,01 a aproximadamente 400 LU/g DS, preferiblemente de aproximadamente 0,1 a aproximadamente 100 LU/g DS, y más preferiblemente de aproximadamente 1 a aproximadamente 50 LU/g DS. Optimización adicional de la cantidad de cutinasa puede de aquí en adelante ser obtenida usando procedimientos estándar conocidos en la técnica.

15

20

25

30

35

[0221] En otra forma de realización preferida, la esterasa es una fosfolipasa. Como se utiliza en este caso, el término "fosfolipasa" es una enzima que tiene actividad hacia fosfolípidos, por ejemplo, actividad hidrolítica. Fosfolípidos, tales como lecitina o fosfatidilcolina, consisten en glicerol esterificado con dos ácidos grasos en una posición externa (sn-1) y una en medio (sn-2) y esterificado con ácido fosfórico en la tercera posición. El ácido fosfórico se puede esterificar a un aminoalcohol. Varios tipos de actividad fosfolipásica se pueden distinguir, incluyendo fosfolipasas A1 y A2 que hidrolizan un grupo acilo graso (en las posiciones sn-1 y sn-2, respectivamente) para formar lisofosfolípido; y lisofosfolipasa (o fosfolipasa B) que hidroliza el grupo ácido graso restante en el lisofosfolípido. Fosfolipasa C y fosfolipasa D (fosfodiesterasas) liberan diacil glicerol o ácido fosfatídico respectivamente.

40

[0222] El término "fosfolipasa" incluye enzimas con actividad fosfolipásica, por ejemplo, fosfolipasa A (A1 o A2), actividad fosfolipasa B, actividad fosfolipasa C, o actividad fosfolipasa D. El término "fosfolipasa A" como se utiliza en este caso se destina a cubrir una enzima con actividad fosfolipasa A1 y/o fosfolipasa A2. La actividad fosfolipasa se puede proporcionar por enzimas con otras actividades también, tales como, por ejemplo, una lipasa con actividad fosfolipasa. La actividad fosfolipasa puede, por ejemplo, ser de una lipasa con actividad fosfolipasa lateral. En otros aspectos, la actividad enzimática de fosfolipasa está provista por una enzima con esencialmente solo actividad fosfolipasa y donde la actividad enzimática de fosfolipasa no es una actividad lateral.

45

[0223] La fosfolipasa puede ser de cualquier origen, por ejemplo, de origen animal (p. ej., mamífero, por ejemplo, páncreas bovino o porcino), o veneno de serpiente o veneno de abeja. Alternativamente, la fosfolipasa puede ser de origen microbiano, por ejemplo, de hongos filamentosos, levadura o bacterias, tales como *Aspergillus*, por ejemplo, *A. awamori*, *A. foetidus*, *A. japonicus*, *A. niger*, o *A. oryzae*, *Dictyostelium*, por ejemplo, *D. discoideum*; *Fusarium*, por ejemplo, *F. culmorum*, *F. graminearum*, *F. heterosporum*, *F. solani*, *F. oxysporum*, o *F. venenatum*; *Mucor*, por ejemplo, *M. javanicus*, *M. mucedo*, o *M. subtilissimus*; *Neurospora*, por ejemplo, *N. crassa*; *Rhizomucor*, por ejemplo, *R. pusillus*; *Rhizopus*, por ejemplo, *R. arrhizus*, *R. japonicus*, o *R. stolonifer*; *Sclerotinia*, por ejemplo, *S. libertiana*; *Trichophyton*, por ejemplo, *T. rubrum*; *Whetzelinia*, por ejemplo, *W. sclerotiorum*; *Bacillus*, por ejemplo, *B. megaterium* o *B. subtilis*; *Citrobacter*, por ejemplo, *C. freundii*; *Enterobacter*, por ejemplo, *E. aerogenes* o *E. cloacae*; *Edwardsiella*, *E. tarda*; *Erwinia*, por ejemplo, *E. herbicola*; *Escherichia*, por ejemplo, *E. coli*; *Klebsiella*, por ejemplo, *K. pneumoniae*; *Proteus*, por ejemplo, *P. vulgaris*; *Providencia*, por ejemplo, *P. stuartii*; *Salmonella*, por ejemplo, *S. typhimurium*; *Serratia*, por ejemplo, *S. liquefaciens*, *S. marcescens*; *Shigella*, por ejemplo, *S. flexneri*, *Streptomyces*, por ejemplo, *S. violeceoruber*, o *Yersinia*, por ejemplo, *Y. enterocolitica*.

50

55

60

[0224] Fosfolipasas preferidas comerciales incluyen LECITASE™ y LECITASE™ ULTRA (disponibles de Novozymes A/S).

65

[0225] Una cantidad eficaz de fosfolipasa es de aproximadamente 0,01 a aproximadamente 400 LU/g DS, preferiblemente de aproximadamente 0,1 a aproximadamente 100 LU/g DS, y más preferiblemente de aproximadamente 1 a aproximadamente 50 LU/g DS. Optimización adicional de la cantidad de fosfolipasa puede de aquí en adelante ser obtenida usando procedimientos estándar conocidos en la técnica.

Proteasas

[0226] En otro aspecto preferido de la divulgación, al menos un tensioactivo y al menos una enzima generadora de carbohidrato se usa en combinación con al menos una proteasa. La proteasa puede ser utilizada, por ejemplo, para digerir proteína para producir nitrógeno libre de amino (FAN). Esta función de aminoácidos libres como nutrientes para la levadura, mejora así el crecimiento de la levadura y, consecuentemente, la producción de etanol.

[0227] El microorganismo fermentador para uso en un proceso de fermentación se puede producir mediante la propagación del microorganismo fermentador en presencia de al menos una proteasa. Aunque sin limitarse a ninguna teoría de operación, se cree que la propagación del microorganismo fermentador con una cantidad eficaz de al menos una proteasa reduce el tiempo de retraso del microorganismo fermentador cuando el microorganismo fermentador se usa posteriormente en un proceso de fermentación en comparación con un microorganismo fermentador que fue propagado bajo las mismas condiciones sin la adición de la proteasa. La acción de la proteasa en el proceso de propagación se cree que directa o indirectamente supone la supresión o expresión de genes que son perjudiciales o provechosos, respectivamente, al microorganismo fermentador durante fermentación, así disminuye el tiempo de retardo y da como resultado un ciclo de fermentación más rápido.

[0228] Proteasas se conocen bien en la técnica y se refieren a enzimas que catalizan la escisión de enlaces peptídicos. Proteasas adecuadas incluyen proteasas fúngicas y bacterianas. Proteasas preferidas son proteasas ácidas, es decir, proteasas caracterizadas por la capacidad de hidrolizar proteínas bajo condiciones ácidas por debajo de pH 7. Proteasas fúngicas ácidas adecuadas incluyen proteasas fúngicas derivadas de *Aspergillus*, *Mucor*, *Rhizopus*, *Candida*, *Coriolus*, *Endothia*, *Entomophtra*, *Irpex*, *Penicillium*, *Sclerotium* y *Torulopsis*. Se contemplan especialmente proteasas derivadas de *Aspergillus niger* (ver, por ejemplo, Koaze *et al.*, 1964, *Agr. Biol. Chem. Japan* 28: 216), *Aspergillus saitoi* (ver, por ejemplo, Yoshida, 1954, *J. Agr. Chem. Soc. Japan* 28: 66), *Aspergillus awamori* (Hayashida *et al.*, 1977, *Agric. Biol. Chem.* 42: 927-933), *Aspergillus aculeatus* (WO 95/02044), o *Aspergillus oryzae*; y proteasas ácidas de *Mucor pusillus* o *Mucor miehei*.

[0229] Proteasas bacterianas, que no son proteasas ácidas, incluyen los productos disponibles comercialmente ALCALASE™ y NEUTRASE™ (disponibles de Novozymes A/S). Otras proteasas incluyen GC106 de Genencor internacional, Inc., EEUU y NOVOZYM™ 50006 de Novozymes A/S

[0230] Preferiblemente, la proteasa es una proteasa de ácido aspártico, como se describe, por ejemplo, en *Handbook of Proteolytic Enzymes*, Editado por A.J. Barrett, N.D. Rawlings y J.F. Woessner, Academic Press, San Diego, 1998, Chapter 270. Ejemplos adecuados de proteasas de ácido aspártico incluyen, por ejemplo, aquellos descritos por Berka *et al.*, 1990, *Gene* 96: 313; Berka *et al.*, 1993, *Gene* 125: 195-198; and Gomi *et al.*, 1993, *Biosci. Biotech. Biochem.* 57: 1095-1100.

Peroxidasas

[0231] Otros compuestos que poseen actividad de peroxidasa pueden ser cualquier peroxidasa (EC 1.11.1.7), o cualquier fragmento con actividad de peroxidasa derivado de esta, que muestra actividad de peroxidasa.

[0232] Preferiblemente, la peroxidasa se produce por plantas (p. ej., peroxidasa de alfalfa o de semilla de soja) o microorganismos tales como hongos o bacterias.

[0233] Algunos hongos preferidos incluyen cepas de la subdivisión *Deuteromycotina*, clase *Hyphomycetes*, por ejemplo, *Fusarium*, *Humicola*, *Trichoderma*, *Myrothecium*, *Verticillum*, *Arthomyces*, *Caldariomyces*, *Ulocladium*, *Embellisia*, *Cladosporium*, o *Dreschlera*, en particular, *Fusarium oxysporum* (DSM 2672), *Humicola insolens*, *Trichoderma reesei*, *Myrothecium verrucaria* (IFO 6113), *Verticillum alboatrum*, *Verticillum dahlie*, *Arthomyces ramosus* (FERM P-7754), *Caldariomyces fumago*, *Ulocladium chartarum*, *Embellisia alli*, o *Dreschlera halodes*.

[0234] Otros hongos preferidos incluyen cepas de la subdivisión *Basidiomycotina*, clase *Basidiomycotina*, por ejemplo, *Coprinus*, *Phanerochaete*, *Coriolus*, o *Trametes*, en particular, *Coprinus cinereus f. microsporus* (IFO 8371), *Coprinus macrorhizus*, *Phanerochaete chrysosporium* (e.g. NA-12), o *Trametes* (anteriormente llamado *Polyporus*), por ejemplo, *T. versicolor* (p. ej., PR4 28-A).

[0235] Otros hongos preferidos incluyen cepas de la subdivisión *Zygomycotina*, clase *Mycoraceae*, por ejemplo, *Rhizopus* o *Mucor*, en particular, *Mucor hiemalis*.

[0236] Algunas bacterias preferidas incluyen cepas del orden *Actinomycetales*, por ejemplo, *Streptomyces spheroides* (ATCC 23965), *Streptomyces thermoviolaceus* (IFO 12382), o *Streptovercillum verticillum ssp. verticillum*.

[0237] Otras bacterias preferidas incluyen *Rhodobacter sphaeroides*, *Rhodomonas palustri*, *Streptococcus lactis*, *Pseudomonas purrocina* (ATCC 15958), *Pseudomonas fluorescens* (NRRL B-11) y cepas de *Bacillus*, por ejemplo, *Bacillus pumilus* (ATCC 12905) y *Bacillus stearothermophilus*.

[0238] Otras bacterias preferidas incluyen cepas de *Myxococcus*, por ejemplo, *M. virescens*.

[0239] La peroxidasa puede también ser una que se produce por un método comprendiendo cultivo una célula huésped transformada con un vector de ADN recombinante que lleva una secuencia de ADN que codifica la peroxidasa al igual que secuencias de ADN para la expresión de la secuencia de ADN que codifica la peroxidasa, en un medio de cultivo bajo condiciones que permiten la expresión de la peroxidasa y recuperación la peroxidasa del cultivo.

[0240] En una forma de realización preferida, una peroxidasa producida de manera recombinante es una peroxidasa derivada de un *Coprinus sp.*, en particular, *C. macrorhizus* o *C. cinereus* según la WO 92/16634.

[0241] En la presente divulgación, compuestos que poseen actividad de peroxidasa comprenden enzimas peroxidasa y fragmentos activos de peroxidasa derivados de citocromos, hemoglobina, o enzimas peroxidasa.

[0242] Una unidad de peroxidasa (POXU) es la cantidad de enzima que bajo las siguientes condiciones cataliza la conversión de 1 μ mol de peróxido de hidrógeno por minuto a 30°C en 0,1 M de tampón de fosfato pH 7,0, 0,88 mM de peróxido de hidrógeno y 1,67 mM de 2,2'-azino-bis(3-etilbenzotiazolina-6-sulfonato) (ABTS). La reacción es seguida durante 60 segundos (15 segundos después de mezcla) por el cambio en la absorbancia a 418 nm, que debería estar en el intervalo de 0,15 a 0,30. Para cálculo de actividad, un coeficiente de absorción de ABTS oxidado de 36 mM⁻¹ cm⁻¹ y una estequiometría de un μ mol de H₂O₂ convertida por dos μ mol de ABTS oxidado se usan.

Lacasas

[0243] En la presente divulgación, lacasas y enzimas relacionadas con la lacasa comprenden cualquier enzima lacasa clasificada como EC 1.10.3.2, cualquier enzima catecol oxidasa clasificada como EC 1.10.3.1, cualquier enzima bilirrubina oxidasa clasificada como EC 1.3.3.5, o cualquier enzima de monofenol-monooxigenasa clasificada como EC 1.14.18.1.

[0244] Las enzimas mencionadas arriba pueden ser microbianas, es decir, obtenidas de bacterias u hongos (incluyendo hongos filamentosos y levaduras), o pueden derivar de plantas.

[0245] Ejemplos adecuados de hongos incluyen una lacasa obtenida de una cepa de *Aspergillus*, *Neurospora*, por ejemplo, *N. crassa*, *Podospora*, *Botrytis*, *Collybia*, *Fomes*, *Lentinus*, *Pleurotus*, *Trametes*, por ejemplo, *T. villosa* and *T. versicolor*, *Rhizoctonia*, por ejemplo, *R. solani*, *Coprinus*, e.g., *C. cinereus*, *C. comatus*, *C. friesii*, and *C. plicatilis*, *Psathyrella*, por ejemplo, *P. condelleana*, *Panaeolus*, por ejemplo, *P. papilionaceus*, *Myceliophthora*, por ejemplo, *M. thermophila*, *Schytalidium*, por ejemplo, *S. thermophilum*, *Polyporus*, por ejemplo, *P. pinsitus*, *Pycnoporus*, por ejemplo, *P. cinnabarinus*, *Phlebia*, por ejemplo, *P. radita* (WO 92/01046), o *Coriolus*, por ejemplo, *C. hirsutus* (JP 2-238885).

[0246] Ejemplos adecuados de bacterias incluyen una lacasa obtenida de una cepa de *Bacillus*.

[0247] Una lacasa obtenida de *Coprinus*, *Myceliophthora*, *Polyporus*, *Pycnoporus*, *Scytalidium* o *Rhizoctonia* es preferida; en particular, una lacasa obtenida de *Coprinus cinereus*, *Myceliophthora thermophila*, *Polyporus pinsitus*, *Pycnoporus cinnabarinus*, *Scytalidium thermophilum*, o *Rhizoctonia solani*.

[0248] Lacasas comercialmente disponibles son NS51001 (una lacasa de *Polyporus pinsitus*, disponible de Novozymes A/S) y NS51002 (una lacasa de *Myceliophthora thermophila*, disponible de Novozymes A/S).

[0249] La lacasa o la enzima relacionada con la lacasa puede también ser una que se produce por un método comprendiendo cultivo de una célula huésped transformada con un vector ADN recombinante que lleva una secuencia de ADN que codifica la lacasa al igual que secuencias de ADN para la expresión de la secuencia de ADN que codifica la lacasa, en un medio de cultivo bajo condiciones que permiten la expresión de la enzima lacasa y recuperación de la lacasa del cultivo.

[0250] Actividad de la lacasa (LACU) se determina de la oxidación de siringaldazina bajo condiciones aeróbicas a pH 5,5. El color de violeta producido se fotometra a 530 nm. Las condiciones analíticas son 19 mM de siringaldazina, 23 mM de tampón acetato, pH 5,5, 30°C, 1 minuto de tiempo de reacción. Una unidad de lacasa (LACU) es la cantidad de enzima que cataliza la conversión de 1,0 μ mol de siringaldazina por minuto bajo las condiciones anteriores.

[0251] La actividad de lacasa (LAMU) se determina de la oxidación de siringaldazina bajo condiciones aeróbicas a pH 7,5. El color de violeta producido se fotometra a 530 nm. Las condiciones analíticas son 19 mM de siringaldazina, 23 mM de tris/maleato pH 7,5, 30°C, 1 minuto de tiempo de reacción. Una unidad de lacasa (LAMU) es la cantidad de enzima que cataliza la conversión de 1,0 μ mol de siringaldazina por minuto bajo las condiciones anteriores.

[0252] Los polipéptidos de la presente invención se pueden utilizar conjuntamente con las enzimas mencionadas anteriormente y/o proteínas celulolíticas para además degradar el componente de celulosa del sustrato de biomasa, (ver, por ejemplo, Brigham *et al.*, 1995, en *Handbook on Bioethanol* (Charles E. Wyman, editor), pp.119-141, Taylor & Francis, Washington D.C.; Lee, 1997, *Journal of Biotechnology* 56: 1-24). [0253] Los polipéptidos de la presente

invención se pueden utilizar conjuntamente con las enzimas mencionadas anteriormente y/o proteínas celulolíticas para además degradar el componente de celulosa del sustrato de biomasa, (ver, por ejemplo, Brigham *et al.*, 1995, en *Handbook on Bioethanol* (Charles E. Wyman, editor), pp.119-141, Taylor & Francis, Washington D.C.; Lee, 1997, *Journal of Biotechnology* 56: 1-24).

[0254] Las cantidades óptimas un polipéptido con actividad de mejora celulolítica y de proteínas celulolíticas depende de distintos factores incluyendo, pero no limitado a, la mezcla de enzimas componentes, el sustrato celulósico, concentración de sustrato celulósico, pretratamiento de sustrato celulósico, temperatura, tiempo, pH, e inclusión de organismo de fermentación (p. ej., levadura para SSF).

[0255] En un aspecto preferido, una cantidad eficaz de polipéptido(s) con actividad de mejora celulolítica para material celulósico es aproximadamente de 0,01 a aproximadamente 2,0 mg, preferiblemente de aproximadamente 0,025 a aproximadamente 1,5 mg, más preferiblemente de aproximadamente 0,05 a aproximadamente 1,25 mg, más preferiblemente de aproximadamente 0,075 a aproximadamente 1,25 mg, más preferiblemente de aproximadamente 0,1 a aproximadamente 1,25 mg, incluso más preferiblemente de aproximadamente 0,15 a aproximadamente 1,25 mg y de forma más preferida de aproximadamente 0,25 a aproximadamente 1,0 mg por g de material celulósico.

[0256] En otro aspecto preferido, una cantidad eficaz de proteína(s) celulolítica para material celulósico es de aproximadamente 0,5 a aproximadamente 50 mg, preferiblemente de aproximadamente 0,5 a aproximadamente 40 mg, más preferiblemente de aproximadamente 0,5 a aproximadamente 25 mg, más preferiblemente de aproximadamente 0,75 a aproximadamente 20 mg, más preferiblemente de aproximadamente 0,75 a aproximadamente 15 mg, incluso más preferiblemente de aproximadamente 0,5 a aproximadamente 10 mg y de forma más preferida de aproximadamente 2,5 a aproximadamente 10 mg por g de material celulósico.

[0257] En un aspecto preferido, una cantidad eficaz de polipéptido(s) con actividad de mejora celulolítica para proteína(s) celulolítica es de aproximadamente 0,005 a aproximadamente 1,0 g, preferiblemente de aproximadamente 0,01 a aproximadamente 1,0 g, más preferiblemente de aproximadamente 0,15 a aproximadamente 0,75 g, más preferiblemente de aproximadamente 0,15 a aproximadamente 0,5 g, más preferiblemente de aproximadamente 0,1 a aproximadamente 0,5 g, incluso más preferiblemente de aproximadamente 0,1 a aproximadamente 0,5 g y de forma más preferida de aproximadamente 0,05 a aproximadamente 0,2 g por g de proteína(s) celulolítica(s).

Composiciones detergentes

[0258] Los polipéptidos de la presente invención con actividad de mejora celulolítica se pueden añadir y así convertirse en un componente de una composición detergente.

[0259] La composición detergente de la presente invención puede ser formulada, por ejemplo, como una composición de detergente para ropa de lavado a mano o a máquina que incluye una composición aditiva de lavandería adecuada para pretratamiento de tejidos manchados y una composición suavizante añadida de enjuague, o ser formulada como una composición detergente para uso en operaciones de limpieza de superficies duras del hogar en general, o ser formulada para operaciones de lavado de la vajilla a mano o a máquina.

[0260] En un aspecto específico, la presente divulgación proporciona un aditivo detergente comprendiendo un polipéptido de la invención. El aditivo detergente al igual que la composición detergente puede comprender una o más enzimas tal como una proteasa, lipasa, cutinasa, una amilasa, carbohidrasa, celulasa, pectinasa, mananasa, arabinasa, galactanasa, xilanasas, oxidasas, por ejemplo, una lacasa y/o peroxidasa.

[0261] En general, las propiedades de la enzima(s) seleccionada deberían ser compatibles con el detergente seleccionado, (es decir, pH óptimo, compatibilidad con otros ingredientes no enzimáticos y enzimáticos, etc.), y la enzima(s) debería estar presente en cantidades eficaces.

[0262] Celulasas: celulasas adecuadas incluyen aquellas de origen fúngico o bacteriano. Mutantes modificados químicamente o creados genéticamente de proteína se incluyen. Celulasas adecuadas incluyen celulasas de los géneros *Bacillus*, *Pseudomonas*, *Humicola*, *Fusarium*, *Thielavia*, *Acremonium*, por ejemplo, las celulasas fúngicas producidas de *Humicola insolens*, *Myceliophthora thermophila* y *Fusarium oxysporum* descritas en US 4,435,307, US 5,648,263, US 5,691,178, US 5,776,757 y WO 89/09259.

[0263] Celulasas especialmente adecuadas son las celulasas neutras o alcalinas con beneficios de cuidado del color. Ejemplos de tales celulasas son celulasas descritas en las EP 0 495 257, EP 0 531 372, WO 96/11262, WO 96/29397, WO 98/08940. Otros ejemplos son variantes de celulasa tales como las descritas en las WO 94/07998, EP 0 531 315, US 5,457,046, US 5,686,593, US 5,763,254, WO 95/24471, WO 98/12307 y PCT/DK98/00299. [0264] Celulasas disponibles comercialmente incluyen Celluzyme™ y Carezyme™ (Novozymes A/S), Clazinase™ y Puradax HA™ (Genencor International Inc.) y KAC-500(B)™ (Kao Corporation).

[0265] Proteasas: proteasas adecuadas incluyen aquellas de origen animal, vegetal o microbiano. Origen microbiano es preferido. Mutantes químicamente modificados o creados genéticamente de proteína se incluyen. La proteasa puede ser

- una serina proteasa o una metaloproteasa, preferiblemente una proteasa alcalina microbiana o una proteasa de tipo tripsina. Ejemplos de proteasas alcalinas son subtilisinas, especialmente aquellas derivadas de *Bacillus*, por ejemplo, subtilisin Novo, subtilisin Carlsberg, subtilisin 309, subtilisin 147 y subtilisin 168 (descritas en WO 89/06279). Ejemplos de proteasas de tipo tripsina son tripsina (p. ej., de origen bovino o porcino) y la proteasa de *Fusarium* descrita en WO 89/06270 y WO 94/25583.
- [0266] Ejemplos de proteasas útiles son las variantes descritas en WO 92/19729, WO 98/20115, WO 98/20116 y WO 98/34946, especialmente las variantes con sustituciones en una o más de las siguientes posiciones: 27, 36, 57, 76, 87, 97, 101, 104, 120, 123, 167, 170, 194, 206, 218, 222, 224, 235 y 274.
- [0267] Enzimas proteásicas preferidas disponibles comercialmente incluyen Alcalase™, Savinase™, Primase™, Durafase™, Esperase™ y Kannase™ (Novozymes A/S), Maxatase™, Maxacal™, Maxapem™, Properase™, Purafect™, Purafect OxP™, FN2™ y FN3™ (Genencor Internacional Inc.).
- [0268] Lipasas: lipasas adecuadas incluyen aquellas de origen fúngico o bacteriano. Mutantes químicamente modificados o creados genéticamente de proteína se incluyen. Ejemplos de lipasas útiles incluyen lipasas de *Humicola* (sinónimo de *Thermomyces*), por ejemplo, de *H. lanuginosa* (*T. lanuginosus*) como se describe en EP 258 068 y EP 305 216 o de *H. insolens* como se describe en WO 96/13580, una lipasa de *Pseudomonas*, por ejemplo, de *P. alcaligenes* o *P. pseudoalcaligenes* (EP 218 272), *P. cepacia* (EP 331 376), *P. stutzeri* (GB 1,372,034), *P. fluorescens*, *Pseudomonas* sp. cepa SD 705 (WO 95/06720 y WO 96/27002), *P. wisconsinensis* (WO 96/12012), una lipasa de *Bacillus*, por ejemplo, de *B. subtilis* (Dartois *et al.*, 1993, *Biochemica et Biophysica Acta*, 1131: 253-360), *B. stearothermophilus* (JP 64/744992) o *B. pumilus* (WO 91/16422).
- [0269] Otros ejemplos son variantes de lipasa tales como los descritos en WO 92/05249, WO 94/01541, EP 407 225, EP 260 105, WO 95/35381, WO 96/00292, WO 95/30744, WO 94/25578, WO 95/14783, WO 95/22615, WO 97/04079 y WO 97/07202.
- [0270] Enzimas de lipasa preferidas disponibles comercialmente incluyen Lipolase™ y Lipolase Ultra™ (Novozymes A/S).
- [0271] Amilasas: amilasas adecuadas (α y/o β) incluyen aquellas de origen fúngico o bacteriano. Mutantes químicamente modificados o creados genéticamente de proteína se incluyen. Amilasas incluyen, por ejemplo, α -amilasas obtenidas de *Bacillus*, por ejemplo, una cepa especial de *Bacillus licheniformis*, descrita con más detalle en GB 1,296,839.
- [0272] Ejemplos de amilasas útiles son las variantes descritas en WO 94/02597, WO 94/18314, WO 96/23873, y WO 97/43424, especialmente las variantes con sustituciones en una o más de las siguientes posiciones: 15, 23, 105, 106, 124, 128, 133, 154, 156, 181, 188, 190, 197, 202, 208, 209, 243, 264, 304, 305, 391, 408 y 444.
- [0273] Amilasas disponibles comercialmente son Duramyl™, Termamyl™, Fungamyl™ y BAN™ (Novozymes A/S), Rapidase™ y Purastar™ (de Genencor Internacional Inc.).
- [0274] Peroxidasas/oxidásas: peroxidasas/oxidásas adecuadas incluyen aquellas de origen de planta, fúngico o bacteriano. Mutantes químicamente modificados o mutantes creados genéticamente de proteína se incluyen. Ejemplos de peroxidasas útiles incluyen peroxidasas de *Coprinus*, por ejemplo, de *C. cinereus* y variantes de estas como aquellas descritas en WO 93/24618, WO 95/10602 y WO 98/15257.
- [0275] Peroxidasas disponibles comercialmente incluyen Guardzyme™ (Novozymes A/S).
- [0276] La enzima(s) detergente se puede incluir en una composición detergente añadiendo aditivos separados con una o más enzimas, o añadiendo un aditivo combinado comprendiendo todas estas enzimas. Un aditivo detergente de la invención, es decir, un aditivo separado o un aditivo combinado, puede ser formulado, por ejemplo, como un granulado, líquido, composición acuosa, etc. Formulaciones de aditivo detergentes preferidas son granulados, en particular, granulados no polvorientos, líquidos, en particular, líquidos estabilizados, o compuestos acuosos.
- [0277] Granulados no polvorientos pueden ser producidos, por ejemplo, como se describe en US 4,106,991 y 4,661,452 y pueden ser revestidos opcionalmente por métodos conocidos en la técnica. Ejemplos de materiales de recubrimiento cerosos son productos de poli(óxido de etileno) (polietilenglicol, PEG) con pesos molares medios de 1000 a 20000; nonilfenoles etoxilados con de 16 a 50 unidades de óxido de etileno; alcoholes etoxilados grasos en los que el alcohol contiene de 12 a 20 átomos de carbono y en el cual hay de 15 a 80 unidades de óxido de etileno; alcoholes grasos; ácidos grasos; y mono- y di- y triglicéridos de ácidos grasos. Ejemplos de materiales de recubrimiento que forman películas adecuadas para aplicación por técnicas de lecho fluidificado se dan en GB 1483591. Preparaciones enzimáticas líquidas pueden, por ejemplo, ser estabilizadas añadiendo un poliol tal como propilenglicol, un azúcar o alcohol de azúcar, ácido láctico o ácido bórico según métodos establecidos. Enzimas protegidas se pueden preparar según el método descrito en EP 238,216.

[0278] La composición detergente de la invención puede estar en cualquier forma conveniente, por ejemplo, una barra, una pastilla, un polvo, un gránulo, una pasta o un líquido. Un detergente líquido puede ser acuoso, típicamente con hasta 70% de agua y 0-30% de solvente orgánico, o no acuoso.

5 [0279] La composición detergente comprende uno o más tensioactivos, que pueden ser no iónicos incluyendo semipolares y/o aniónicos y/o zwitteriónicos y/o catiónicos. Los tensioactivos están presentes típicamente a un nivel de 0,1 % a 60% en peso.

10 [0280] Cuando se incluye en esto el detergente normalmente contiene de aproximadamente 1% a aproximadamente 40% de un surfactante aniónico tal como alquilbencenosulfonato lineal, alfa-olefinsulfonato, alquil sulfato (sulfato de alcohol graso), etoxisulfato alcohólico, alcanosulfonato secundario, metil éster alfa-sulfo ácido graso, ácido alquil o alquenilsuccínico, o jabón.

15 [0281] Cuando se incluye en esto el detergente normalmente contiene de aproximadamente 0,2% a aproximadamente 40% de un surfactante no iónico tal como alcohol etoxilato, nonilfenol, etoxilato alquilpoliglicósido, óxido de alquildimetilamina, monoetanolamida de ácido graso etoxilado, monoetanolamida de ácido graso, amida de polihidroxi alquil ácido graso, o derivados de N-acil N-alquil glucosamina ("glucamidas").

20 [0282] El detergente puede contener 0-65% de un constructor de detergente o agente complejante tal como zeolita, difosfato, trifosfato, fosfonato, carbonato, citrato, ácido nitrilotriacético, ácido etilendiaminatetraacético, ácido dietilendiaminopentaacético, ácido alquil o alquenilsuccínico, silicatos solubles, o silicatos estratificados (p. ej., SKS-6 de Hoechst).

25 [0283] El detergente puede comprender uno o más polímeros. Ejemplos son carboximetilcelulosa, poli(vinilpirrolidona), poli(etilenglicol), poli(vinil alcohol), poli(vinilpirrolidona-N-óxido), poli(vinilimidazol), policarboxilatos tales como poliacrilatos, copolímeros de ácido maléico/acrílico y copolímeros de ácido de lauril metacrilato/acrílico.

30 [0284] El detergente puede contener un sistema blanqueante que puede comprender una fuente de H₂O₂ tal como perborato o percarbonato que se puede combinar con un activador blanqueante de formación de perácido tal como tetraacetiletilendiamina o nonanoiloxibencenosulfonato. Alternativamente, el sistema blanqueante puede comprender peroxiácidos del tipo, por ejemplo, amida, imida, o sulfona.

35 [0285] La(s) enzima(s) de la composición detergente de la invención puede ser estabilizada usando agentes estabilizantes convencionales, por ejemplo, un poliol tal como propilenglicol o glicerol, un azúcar o alcohol de azúcar, ácido láctico, ácido bórico, o un derivado de ácido bórico, por ejemplo, un éster de borato aromático, o un derivado de ácido fenil borónico tal como ácido 4-formilfenilborónico y la composición se puede formular como se describe en, por ejemplo, WO 92/19709 y WO 92/19708.

40 [0286] El detergente puede también contener otros ingredientes de detergentes convencionales tales como, por ejemplo, acondicionadores de tejido incluyendo arcillas, reforzadores de espuma, supresores de espuma, agentes anticorrosivos, agentes de suspensión de suciedad, agentes de reposición antisuciedad, tintes, bactericidas, blanqueadores ópticos, hidrótrofos, inhibidores de decoloración, o perfumes.

45 [0287] En las composiciones detergentes, cualquier enzima se puede añadir en una cantidad correspondiente a 0,01-100 mg de proteína enzimática por litro de solución de lavado, preferiblemente 0,05-5 mg de proteína enzimática por litro de solución de lavado, en particular 0,1-1 mg de proteína enzimática por litro de solución de lavado.

50 [0288] En las composiciones detergentes, un polipéptido de la presente invención con actividad de mejora celulolítica se puede añadir en una cantidad correspondiente a 0,001-100 mg de proteína, preferiblemente 0,005-50 mg de proteína, más preferiblemente 0,01-25 mg de proteína, incluso más preferiblemente 0,05-10 mg de proteína, de forma más preferida 0,05-5 mg de proteína, e incluso de forma más preferida 0,01-1 mg de proteína por litro de solución de lavado.

55 [0289] Un polipéptido de la invención que tiene actividad de mejora celulolítica puede también ser incorporado en las formulaciones detergentes descritas en WO 97/07202, que se incorpora en la presente como referencia.

Otros usos

60 [0290] En general, tratamiento de cualquier material de pared celular vegetal se puede aumentar por complementación de los polipéptidos de la presente invención con actividad de mejora celulolítica.

Péptido señal

65 [0291] La presente descripción también se refiere a construcciones de ácidos nucleicos que comprenden un gen que codifica una proteína unida operativamente a una secuencia de ácido nucleico que consiste en los nucleótidos 1 a 54 de la SEQ ID N°: 7 que codifica un péptido señal consistente en los aminoácidos 1 a 18 de SEQ ID N°: 8, que permite la secreción de la proteína en un medio de cultivo, en el que el gen es extranjero a la secuencia de ácido nucleico .

[0292] La presente descripción también se refiere a vectores de expresión recombinantes y células huéspedes recombinantes que comprenden dichas construcciones de ácidos nucleicos.

5 [0293] La presente descripción también se refiere a métodos para producir una proteína que comprende (a) cultivar una célula huésped recombinante bajo condiciones adecuadas para la producción de la proteína; y (b) recuperar la proteína.

10 [0294] La proteína puede ser nativa o heteróloga a una célula huésped. El término "proteína" no se entiende en el presente documento para referirse a una longitud específica del producto codificado y, por lo tanto, abarca péptidos, oligopéptidos, y proteínas. El término "proteína" también abarca dos o más polipéptidos combinados para formar el producto codificado. Las proteínas también incluyen polipéptidos híbridos que comprenden una combinación de secuencias de polipéptidos parciales o completas obtenidas a partir de al menos dos proteínas diferentes de las cuales una o más pueden ser heterólogas o nativas a la célula huésped. Las proteínas incluyen además variaciones de ingeniería y alélicas de origen natural de las proteínas mencionadas anteriormente y de las proteínas híbridas.

15 [0295] Preferiblemente, la proteína es una hormona o variante de la misma, enzima, receptor o parte del mismo, anticuerpo o parte del mismo, o reportero. En un aspecto más preferido, la proteína es una oxidoreductasa, transferasa, hidrolasa, liasa, isomerasa o ligasa. En un aspecto aún más preferido, la proteína es una aminopeptidasa, amilasa, carbohidrasa, carboxipeptidasa, catalasa, celulasa, quitinasa, cutinasa, ciclodextrina glicosiltransferasa, desoxirribonucleasa, esterasa, alfa-galactosidasa, beta-galactosidasa, glucoamilasa, alfa-glucosidasa, beta-glucosidasa, invertasa, lacasa, lipasa, manosidasa, mutanasa, oxidasa, enzima pectinolítica, peroxidasa, fitasa, polifenoloxidasas, enzima proteolítica, ribonucleasa, transglutaminasa o xilanasas.

[0296] El gen puede ser obtenido de cualquier fuente procariótica, eucariótica, o de otro tipo.

25 EJEMPLOS

Materiales

30 [0297] Productos químicos usados como tampones y sustratos fueron productos comerciales de al menos grado reactivo.

Cepas

35 [0298] *Thielavia terrestris* NRRL 8126 fue usada como la fuente de los polipéptidos de la familia 61 con actividad de mejora celulolítica. Cepa de *Aspergillus oryzae* JaL250 (WO 99/61651) y *Trichoderma reesei* RutC30 (ATCC 56765; Montencourt and Eveleigh, 1979, *Adv. Chem. Ser. 181: 289-301*) se usaron para la expresión de los polipéptidos de la familia 61 con actividad de mejora celulolítica.

Medios

40 [0299] Medio YEG se compuso por litro de 0,5% de extracto de levadura y 2% de glucosa.

[0300] Medio de dextrosa de patata se compuso por litro de 39 gramos de dextrosa de patata (Difco).

45 [0301] Placas PDA se compusieron por litro de 39 gramos de agar de dextrosa de patata.

[0302] Medio MDU2BP se compuso por litro de 45 g de maltosa, 1 g de $MgSO_4 \cdot 7H_2O$, 1 g de NaCl, 2 g de K_2SO_4 , 12 g de KH_2PO_4 , 7 g de extracto de levadura, 2 g de urea y 0,5 ml de solución de metales traza AMG, pH a 5,0.

50 [0303] Solución de metales traza AMG se compuso por litro de 14,3 g de $ZnSO_4 \cdot 7H_2O$, 2,5 g de $CuSO_4 \cdot 5H_2O$, 0,5 g de $NiCl_2 \cdot 6H_2O$, 13,8 g de $FeSO_4 \cdot 7H_2O$, 8,5 g de $MnSO_4 \cdot H_2O$ y 3 g de ácido cítrico.

[0304] Placas COVE se compusieron por litro de 342,3 g de sacarosa, 20 ml de solución salina COVE, 10 ml de 1 M de acetamida, 10 ml de 1,5 M $CsCl_2$ y 25 g de agar noble.

55 [0305] Solución salina COVE se compuso por litro de 26 g de KCl, 26 g de $MgSO_4 \cdot 7H_2O$, 76 g de KH_2PO_4 y 50 ml de solución de metales traza COVE.

60 [0306] Solución de metales traza COVE se compuso por litro de 0,04 g de $Na_2B_4O_7 \cdot 10H_2O$, 0,4 g de $CuSO_4 \cdot 5H_2O$, 1,2 g de $FeSO_4 \cdot 7H_2O$, 0,7 g de $MnSO_4 \cdot H_2O$, 0,8 g de $Na_2MoO_4 \cdot 2H_2O$ y 10 g de $ZnSO_4 \cdot 7H_2O$.

[0307] Medio CIM se compuso por litro de 20 g de celulosa, 10 g de sólidos maíz, 1,45 g de $(NH_4)_2SO_4$, 2,08 g de KH_2PO_4 , 0,28 g de $CaCl_2$, 0,42 g de $MgSO_4 \cdot 7H_2O$ y 0,42 ml de solución de metales traza, pH a 6,0.

65 [0308] Solución de metales traza se compuso por litro de 41,2 mg de $FeCl_3 \cdot 6H_2O$, 11,6 mg de $ZnSO_4 \cdot 7H_2O$, 5,4 mg de $MnSO_4 \cdot H_2O$, 2,0 mg de $CuSO_4 \cdot 5H_2O$, 0,48 mg de H_3BO_3 y 67,2 mg de ácido cítrico.

[0309] Medio NNCYP se compuso por litro de 5,0 g de NH_4NO_3 , 0,5 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,3 g de CaCl_2 , 2,5 g de ácido cítrico, 1,0 g de bacto-peptona, 5,0 g de extracto de levadura, solución de metales traza COVE y suficiente K_2HPO_4 para conseguir el pH final de aproximadamente 5,4.

[0310] Medio NNCYPmod se compuso por litro de 1,0 g de NaCl , 5,0 g de NH_4NO_3 , 0,2 g de $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0,2 g de CaCl_2 , 2,0 g de ácido cítrico, 1,0 g de bacto-peptona, 5,0 g de extracto de levadura, solución de metales traza COVE y suficiente K_2HPO_4 para conseguir el pH final de aproximadamente 5,4.

[0311] Placas LB se compusieron por litro de 10 g de triptona, 5 g de extracto de levadura, 5 g de cloruro sódico y 15 g de Agar Bacto.

[0312] Medio SOC se compuso por 2% de triptona, 0,5% de extracto de levadura, 10 mM de NaCl , 2,5 mM de KCl , 10 mM de MgCl_2 y 10 mM de MgSO_4 y glucosa esterilizada por filtro a 20 mM añadido después de autoclave.

[0313] Medio refrigerante se compuso por 60% de SOC y 40% de glicerol.

[0314] Medio 2XYT se compuso por litro de 16 g de triptona, 10 g de extracto de levadura, 5 g de NaCl y 15 g de agar Bacto.

[0315] Los ejemplos 1 a 28 ilustran la provisión al igual que características determinadas de polinucleótido de *Thielavia terrestris* NRRL 8126 GH61B y polipéptido (SEC ID n°: 7 y 8, respectivamente). La provisión al igual que características determinadas de secuencias de polinucleótidos GH61 B, GH61C, GH61D y GH61G y polipéptidos (SEC ID n°: 1-6 y 9-10) se incluyen solo para comparación.

Ejemplo 1: Identificación de polipéptidos con actividad de mejora celulolítica de *Thielavia terrestris* NRRL 8126

[0316] Un tapón de agarosa de una placa fresca de *Thielavia terrestris* NRRL 8126 creció en medio NNCYPmod suplementado con 1% de Sigmacell (Sigma Chemical Co., St. Louis, MO) se inoculó en 50 ml de medio NNCYPmod complementado con 1% de glucosa e incubado a 45°C y 200 r.p.m. durante 25 horas. Dos ml de este cultivo se usó para inocular 15 x 100 ml (matraz 500 ml) y 2 x 50 ml (matraz 250 ml) de medio NNCYPmod complementado con 2% de Sigmacell-20 y se incubó a 45°C, 200 r.p.m. durante 4 días. Cultivos agrupados se centrifugaron a 3000 x g durante 10 minutos y el sobrenadante se filtró a través de un prefiltro de fibra de vidrio Nalgene 281-5000 (Nalge Nunc Int'l., Rochester, NY). El filtrado se enfrió a 4°C para almacenamiento.

[0317] **Electroforesis en gel de poli(acrilamida) bidimensional.** Un ml de filtrado se precipitó añadiendo 100 μl de ácido tricloroacético (ATC) saturado a 4°C y se incubó durante 10 minutos en hielo seguido de adición de 9 ml de acetona enfriada con hielo y además incubación en el hielo durante 20 minutos. La solución precipitada se centrifugó a 10000 x g durante 10 minutos a 4°C, el sobrenadante se decantó y el granulado se enjuagó dos veces con acetona enfriada con hielo y se secó al aire. El granulado seco se disolvió en 0,2 ml de tampón de muestra de isoelectroenfoque (IEF) (9,0 M de urea, 3,1 % (p/v) 3-[(3-colamidopropil) dimetil-amonio]-1-propanosulfonato (CHAPS, Pierce Chemical Co. Rockford, IL), 1% (v/v) pH 4-7 anfólitos, 50 mM de ditioneitol (DTT) y 0,005% de bromofenol azul en agua destilada). Solución madre de urea se desionizó usando AG 501-X8 (D), 20- 5-malla, resina de lecho mezclada de BioRad Laboratories (Hercules, CA). La solución desionizada fue almacenada a -20°C. A la mezcla resultante se le permitió solubilizar durante diferentes horas con mezcla suave en un agitador LabQuake™ (Lab Industries, Berkeley, CA). Doscientos μl de cada mezcla de proteína de tampón de muestra IEF se aplicó a una banda de 11 cm IPG (BioRad Laboratories, Hercules, CA) en una bandeja de rehidratación IPG (Amersham Biosciences, Piscataway, NJ). Una parte alícuota de 750 μl de fluido de cobertura de banda seco (Amersham Biosciences, Piscataway, NJ) se estratificó sobre las bandas IPG para prevenir evaporación y permitir rehidratar durante 12 horas mientras aplicación de 30 voltios usando una unidad de isoelectroenfoque IPGPhor (Amersham Biosciences, Piscataway, NJ) a 20°C. La unidad IPGfor se programó para voltaje constante pero con una corriente máxima de 50 μA por banda. Después de 12 horas de rehidratación, las condiciones de isoelectroenfoque fueron de la siguiente manera: 1 hora a 200 voltios, 1 hora a 500 voltios y 1 hora a 1000 voltios. Luego, se aplicó un gradiente de 1000 voltios a 8000 voltios durante 30 minutos e isoelectroenfoque se programó para ejecutarse a 8000 voltios y se completó cuando >30,000 voltio horas se consiguió. Bandas de gel IPG se redujeron y alquilieron antes del segundo análisis de dimensión por primera reducción durante 15 minutos en 100 mg de ditioneitol por 10 ml de tampón de equilibrio de SDS (50 mM de Tris HCl pH 8,8, 6,0 M de urea, 2% (p/v) dodecilsulfato de sodio (SDS), 30% de glicerol y 0,002% (p/v) de bromofenol azul) seguido de 15 minutos de alquilación en 250 mg de iodoacetamida por 10 ml de tampón de equilibrio a oscuras. Las bandas IPG se enjuagaron rápidamente en el tampón en ejecución de SDS-PAGE (Invitrogen/Novex, Carlsbad, CA) y se colocaron en un pocillo 11 cm 8-16% de gel de tris-glicina SDS-PAGE (BioRad Laboratories, Hercules, CA) y se sometieron a electroforesis usando una unidad de electroforesis Criterion (BioRad Laboratories, Hercules, CA) a 50 voltios hasta que la muestra se introdujo en el gel y luego el voltaje se aumentó a 200 voltios y se le permitió ejecutar hasta que el tinte azul de bromofenol alcanzó el fondo del gel.

[0318] **Detección de polipéptido.** Los dos geles dimensional se tiñeron con un SYPRO Orange Protein Stain fluorescente (Molecular Probes, Eugene, OR). Métodos de coloración fluorescentes se optimizaron y adaptaron de

Malone *et al.*, 2001, *Electrophoresis*, 22, 919-932. Geles de SDS-PAGE se fijaron en 40% de etanol, 2% de ácido acético y 0,0005% de SDS en una plataforma oscilante durante 1 hora durante toda la noche. Solución de fijación se retiró y sustituyó con tres pasos de lavado repetidos que consisten en 2% de ácido acético y 0,0005% de SDS durante 30 minutos cada uno. Geles se tiñeron durante 1,5 horas para durante toda la noche a oscuras con 2% de ácido acético, 0,0005% de SDS y 0,02% de SYPRO Orange Protein Stain. Coloración y decoloración se optimizó además para mejorar reproductibilidad y automatización en una unidad de coloración Hoefer Processor Plus (Amersham Biosciences, Piscataway, NJ). Imágenes de los geles de SDS-PAGE fluorescentes manchados se obtuvieron por escaneado en un sistema de imágenes Molecular Dynamics STORM 860 (Amersham Biosciences, Piscataway, NJ) usando fluorescencia azul y tamaños de píxel de 200 μm y una ganancia de tubo fotomultiplicadora de 800 V. Imágenes se vieron y ajustaron usando ImageQuant versión 5.0 (Amersham Biosciences, Piscataway, NJ). Geles se visualizaron además en un transiluminador Dark Reader Blue con filtro de naranja (Clare Chemical Co, Denver, CO). Manchas de gel de proteínas observadas se cortaron usando un punzón Acu-Punch Biopsy Punch de 2 mm (Acuderm Inc., Ft. Lauderdale, FL) y almacenaron en noventa y seis placas de pocillos que se prelavaron con 0,1% de ácido trifluoroacético (TFA) en 60% de acetonitrilo seguido de dos lavados adicionales con agua de calidad de HPLC. Las manchas de gel bidimensionales teñidas se almacenaron en 25-50 μl de agua en las placas prelavadas a -20°C hasta ser digeridas.

[0319] **Digestión en gel de polipéptidos para secuenciación peptídica.** Un robot de manipulación de líquidos MultiPROBE® II (PerkinElmer Life and Analytical Sciences, Boston, MA) se usó para realizar las digestiones en gel. Dos manchas de gel dimensionales con polipéptidos de interés se redujeron con 50 μl de 10 mM de DTT en 100 mM de bicarbonato de amonio pH 8,0 durante 30 minutos. Después de reducción, las piezas de gel se alquilaron con 50 μl de 55 mM de yodoacetamida en 100 mM de tampón de bicarbonato de amonio pH 8,0 durante 20 minutos. Se les permitió a las piezas de gel secas dilatarse en una solución de digestión de tripsina (6 ng/ μl de tripsina de calidad de secuenciación (Promega, Madison, WI) en 50 mM de tampón de bicarbonato de amonio pH 8) durante 30 minutos a temperatura ambiente, seguido de una digestión de 8 horas a 40°C . A cada uno de los pasos de reacción descritos les siguieron lavados numerosos y prelavados con las soluciones apropiadas después del protocolo estándar de fabricación. Cincuenta μl de acetonitrilo se usó para deshidratar el gel entre reacciones y piezas de gel se secaron al aire entre pasos. Péptidos se extrajeron dos veces con 1% de ácido fórmico/2% de acetonitrilo en el agua de calidad de HPLC durante 30 minutos. Soluciones de extracción peptídicas se transfirieron a una placa de tipo PCR bordeada con 96 pocillos (ABGene, Rochester, NY) que se ha enfriado a $10-15^{\circ}\text{C}$ y cubierto con un tapa de placa de 96 pocillos (PerkinElmer Life and Analytical Sciences, Boston, MA) para prevenir evaporación. Las placas se almacenaron además a 4°C hasta que análisis de espectrometría de masas pudiera realizarse.

[0320] **Secuenciación peptídica por espectrometría de masas en serie.** Para secuenciación peptídica por espectrometría de masas en serie, un espectrómetro de masas de tiempo de vuelo cuádruple ortogonal híbrido Q-ToF micro™ (Waters Micromass® MS Technologies, Milford, MA) se usó para análisis LC-MS/MS. El espectrómetro de masas Q-ToF micro™ se equipó con un sistema HPLC capilar y de nanoflujo Ultimate™ (Dionex, Sunnyvale, CA) que se acopló con un micro automuestreador FAMOS (Dionex, Sunnyvale, CA) y un dispositivo de conmutación de columna Switchos II (Dionex, Sunnyvale, CA) para concentrar y desalar muestras. Muestras se cargaron sobre una columna de dispositivo de seguridad (300 μm de ID X 5 cm, C18 PepMap™) (Dionex, Sunnyvale, CA) se equiparon en el bucle de inyección y lavaron con 0,1% de ácido fórmico en agua a 40 $\mu\text{l}/\text{minuto}$ durante 2 minutos usando una bomba Switchos II (Dionex, Sunnyvale, CA). Péptidos se separaron en 75 μm de ID x 15 cm, C18, 3 μm , columna capilar fundida nanoflujo 100 PepMap™ (Dionex, Sunnyvale, CA) a una velocidad de flujo de 175 nl/minuto de un flujo de división de 175 $\mu\text{l}/\text{minuto}$ usando un calibrador NAN-75 (Dionex, Sunnyvale, CA). El gradiente de elución lineal fue 5% a 60% de acetonitrilo en 0,1% de ácido fórmico aplicado sobre un periodo de 45 minutos. El eluyente de columna se monitoreó a 215 nm y se introdujo en el espectrómetro de masas Q-ToF micro™ a través de una fuente de iones de electrospray equipada con la interfaz de nanopulverización. El espectrómetro de masas Q-ToF micro™ fue completamente controlado por microprocesador usando software MassLynx™ versión 3.5 (Waters Micromass® MS Technologies, Milford, MA). Los datos se adquirieron en el modo de escaneado de encuesta y a partir un intervalo de masa de 50 a 2000 m/z con criterio de conmutación para MS a MS/MS para incluir una intensidad iónica mayor de 10,0 cuentas/segundo y estados de carga de +2, +3 y +4. Espectros de análisis de hasta 4 especies de co-elución con un tiempo de escaneado de 1,9 segundos y tiempo de interescaado de 0,1 segundos podrían obtenerse. Un voltaje de cono de 65 voltios se usó típicamente y la energía de colisión se programó para variar según la masa y estado de carga del péptido de elución y en el intervalo de 10 - 60 voltios. Los espectros adquiridos se combinaron, homogenizaron y centraron en una forma automatizada y una lista de valor máximo generado. La lista de valor máximo generada se buscó contra bases de datos seleccionadas usando el software ProteinLynx™ Global Server 1.1 (Waters Micromass® MS Technologies, Milford, MA). Resultados a partir de búsquedas de ProteinLynx™ se evaluaron y proteínas desconocidas se analizaron además evaluando los espectros MS/MS de cada ión de interés y secuencia de novo determinado identificando la serie iónica "y" y "b" y coincidiendo diferencias de masa para el aminoácido apropiado.

[0321] Secuencias peptídicas de *Thielavia terrestris* GH61 B de secuenciación de novo por espectrometría de masas se obtuvieron de diferentes iones con múltiples cargas para aproximadamente una mancha de gel de polipéptido de 40 kDa. Un ión doblemente cargado de péptido tríptico de 878,422 m/z secuencia se determinó para ser [Ile o Leu]-Pro-Ala-Ser-Asn-Ser-Pro-Val-Thr-Asn-Val-Ala- Ser-Asp-Asp-[Ile o Leu]-Arg (SEC ID n°: 11). Un segundo ión doblemente cargado de péptido tríptico de 765,460 se determinó para ser [Ile o Leu]-pro-Glu-Asp-[Ile o Leu]-Glu-pro-Gly-Asp-Tyr-[Ile o Leu]-[Ile o Leu]-Arg (SEC ID n°: 12).

[0322] Veinte μ l de la fracción agrupada con la actividad de mejora celulolítica mayor después del paso de purificación de fenil sefarosa descrito en el ejemplo 24 se mezcló con 20 μ l de tampón de carga 2X SDS-PAGE con 50 mM de DTT y se hirvió durante 4 minutos. Cuarenta μ l se separó por SDS-PAGE usando un gel gradiente 11 cm 8-16% de tris-glicina SDS-PAGE (BioRad Laboratories, Hercules, CA) y tampón en ejecución de tris-glicina (Invitrogen, Carlsbad, CA). El SDS-PAGE se ejecutó bajo condiciones de reducción y el protocolo recomendado del fabricante (BioRad Laboratories, Hercules, CA). El gel se retiró del casete y se enjuagó 3 veces con agua durante al menos 5 minutos cada y se tiñó con Bio-Safe Coomassie Stain (BioRad Laboratories, Hercules, CA) durante 1 hora seguido de decoloración con agua doblemente destilada durante más que 30 minutos. Un enlace proteínico visible a aproximadamente 27 kDa fue digerido en gel con tripsina como se ha descrito anteriormente. Los péptidos recuperados del digerido se sometieron a secuenciación peptídica por espectrometría de masas en serie como se ha descrito anteriormente. Un ión de péptido triptico doblemente cargado de 903,895 m/z se determinó para tener una secuencia de Cys-Pro-Gly-Ser-Phe-Ser-Ser-Cys-Asp-Gly-Ser-Gly-Ala-Gly-Trp-Phe-Lys (SEC ID n°: 13, familia GH61 D). Un ión de péptido triptico triplemente cargado de 415,553 m/z se determinó para ser [Ile o Leu]-Asp-Glu-Ala-Gly-Phe-His-Gly-Asp-Gly-Val-Lys (SEC ID n°: 14, familia GH61D). Otro ión doblemente cargado de péptido triptico de 565,857 m/z se determinó para ser X-X-Ala-Pro-Gly-Asn-Tyr-[Ile or Leu]-[Ile or Leu]-Arg (SEC ID n°: 15, familia GH61C). Estas secuencias peptídicas se usaron para diseño de cebadores para clonación.

Ejemplo 2: Extracción de ADN genómico de *Thielavia terrestris*

[0323] *Thielavia terrestris* NRRL 8126 se cultivó en 25 ml de medio YEG a 37°C y 250 r.p.m. durante 24 horas. Micelios se recogieron luego por filtración a través de Miracloth™ (Calbiochem, La Jolla, CA) y se lavaron una vez con 25 ml de tampón 10 mM de Tris-1 mM de EDTA (TE). El exceso de tampón se drenó de la preparación de micelios, que se congeló posteriormente en nitrógeno líquido. La preparación de micelios congelados se molió a un polvo fino en un molinillo de café eléctrico, y el polvo se añadió a un tubo centrifugador de plástico desechable con 20 ml de tampón TE y 5 ml de 20% (p/v) de dodecilsulfato de sodio (SDS). La mezcla se invirtió suavemente varias veces para asegurar la mezcla y se extrajo dos veces con un volumen igual de alcohol de fenol:cloroformo:isoamil (25:24:1 v/v/v). Acetato de sodio (3 M de solución) se añadió a la muestra extraída a una concentración final de 0,3 M seguido de 2,5 volúmenes de etanol helado para precipitar el ADN. El tubo se centrifugó a 15000 x g durante 30 minutos para granular el ADN. Se le permitió al granulado de ADN secarse por ventilación durante 30 minutos antes de resuspensión en 0,5 ml de tampón TE. Ribonucleasa A sin ADNsa se añadió al granulado de ADN resuspendido a una concentración de 100 μ g por ml y la mezcla se incubó luego a 37°C durante 30 minutos. Proteinasa K (200 μ g/ml) se añadió y el tubo se incubó una hora adicional a 37°C. Finalmente, la muestra se centrifugó durante 15 minutos a 12000 x g, y el sobrenadante se aplicó a un colector Qiaprep 8 (QIAGEN Inc., Valencia, CA). Las columnas se lavaron dos veces con 1 ml de PB (QIAGEN Inc., Valencia, CA) y 1 ml de PE (QIAGEN Inc., Valencia, CA) al vacío. El ADN aislado se eluyó con 100 μ l de TE, se precipitó con etanol, se lavó con 70% de etanol, se secó al vacío, se resuspendió en tampón TE y se almacenó a 4°C.

[0324] Para generar ADN genómico para amplificación PCR, *Thielavia terrestris* se cultivó en 50 ml de medio NNCYP complementado con 1% de glucosa en un matraz de agitación disipado a 42°C y 200 r.p.m. durante 24 horas. Los micelios se recogieron por filtración, se lavaron dos veces en TE (10 mM de Tris-1 mM de EDTA), y se congelaron bajo nitrógeno líquido. Una pieza del tamaño de un guisante de micelios congelados se suspendió en 0,7 ml de 1% de sulfato de dodecilo de litio en TE y se interrumpió por agitación con un volumen igual de 0,1 mm de perlas de zirconio/sílice (Biospec Products, Inc., Bartlesville, OK) durante 45 segundos en un FastPrep FP120 (ThermoSavant, Holbrook, NY). Detrito se retiró por centrifugado a 13000 x g durante 10 minutos y el sobrenadante aclarado se llevó a 2,5 M de acetato amónico y se incubó en hielo durante 20 minutos. Después del periodo de incubación, los ácidos nucleicos se precipitaron por adición de 2 volúmenes de etanol. Después de centrifugado durante 15 minutos en un microcentrifugador a 4°C, el granulado se lavó en 70% de etanol y se secó al aire. El ADN se resuspendió en 120 μ l de 0,1X TE y se incubó con 1 μ l de ribonucleasa pancreática sin ADNsa a 37°C durante 20 minutos. Acetato amónico se añadió a 2,5 M y el ADN se precipitó con 2 volúmenes de etanol. El granulado se lavó en 70% de etanol, se secó al aire y se resuspendió en tampón TE.

Ejemplo 3: Construcción de vector de expresión pAILo2

[0325] Vector de expresión pAILo1 se construyó por modificación de pBANE6 (patente U.S. n° 6,461,837), que comprende un híbrido de los promotores de los genes para alfa-amilasa neutra de *Aspergillus niger* e triosa fosfato isomerasa de *Aspergillus oryzae* (promotor NA2-tpi), secuencia terminadora de amiloglicosidasa de *Aspergillus niger* (terminador AMG) y gen de acetamidasa de *Aspergillus nidulans* (amdS). Todos los pasos de mutagénesis se verificaron por secuenciación usando la química de terminador Big-Dye™ (Applied Biosystems, Inc., Foster City, CA). Modificación de pBANE6 se realizó por primera eliminación de tres sitios de restricción Nco I en las posiciones 2051, 2722 y 3397 pares de bases del marcador de selección amdS por mutagénesis dirigida. Todos los cambios se diseñaron para ser "silenciosos" dejando la secuencia de proteína real de el producto genético amdS sin cambios. Eliminación de estos tres sitios se realizó simultáneamente con un kit GeneEditor™ *in vitro* Site-Directed Mutagenesis (Promega, Madison, WI) según las instrucciones del fabricante usando los siguientes cebadores (el nucleótido subrayado representa la base cambiada):

65 AMDS3NcoMut (2050):
5'-GTGCCCATGATACGCCTCCGG-3' (SEC ID N°: 16)

AMDS2NcoMut (2721):
 5'-GAGTCGATTTCCAAGGCTCCTGACC-3' (SEC ID N°: 17)
 AMDS1 NcoMut (3396):
 5'-GGAGCCATGAAGTGGACCAACGG-3' (SEC ID N°: 18)

[0326] Un plásmido comprendiendo todos los tres cambios de secuencia previstos se sometió luego a mutagénesis dirigida, usando un kit de mutagénesis dirigida QuickChange™ (Stratagene, La Jolla, CA), para eliminar el sitio de restricción Nco I al final del terminador AMG en la posición 1643. Los siguientes cebadores (el nucleótido subrayado representa la base cambiada) se usaron para mutagénesis:

Cebador superior para mutagenizar la secuencia del terminador AMG:
 5'-CACCGTAAAGCCATGCTCTTTCCTTCGTGTAGAAGACCAGACAG-3' (SEC ID n°: 19)
 Cebador inferior para mutagenizar la secuencia del terminador AMG:
 5'-CTGGTCTTCTACACGAAGGAAAGAGCATGGCTTTCACGGTGTCTG-3' (SEC ID n°: 20)

[0327] El último paso en la modificación de pBANE6 fue la adición de un nuevo sitio de restricción Nco I al principio del poliligador usando un equipo de mutagénesis dirigida QuickChange™ y los siguientes cebadores (los nucleótidos subrayados representan las bases cambiadas) para producir pALo1 (Figura 6).

Cebador superior para mutagenizar el promotor NA2-tpi:
 5'-CTATATACACAACCTGGATTTACCATGGGCCCGCGCCGCAGATC-3' (SEC ID n°: 21)
 Cebador inferior para mutagenizar el promotor NA2-tpi:
 5'-GATCTGCGGCCGCGGGCCCATGGTAAATCCAGTTGTGTATATAG-3' (SEC ID n°: 22)

[0328] El gen amdS de pALo1 se cambió con el gen pyrG de *Aspergillus nidulans*. Plásmido pBANE10 (Figura 7) se usó como fuente para el gen pyrG como etiqueta de selección. Análisis de la secuencia de pBANE10 mostro que el marcador pyrG estaba contenido dentro de un fragmento de restricción Nsi I y no contiene sitios de restricción ni Nco I o Pac I. Ya que el amdS es también flanqueado por sitios de restricción Nsi I, la estrategia para cambiar el marcador de selección fue un simple intercambio de fragmentos de restricción Nsi I. ADN plásmido de pALo1 y pBANE10 se digirieron con la enzima de restricción Nsi I y los productos purificados por electroforesis en gel de agarosa. El fragmento Nsi I de pBANE10 con gen thepyrG se ligó al esqueleto de pALo1 para reemplazar el fragmento de ADN Nsi I original con el gen amdS. Clones recombinantes se analizaron por digestión de enzima de restricción para determinar que estos tenían el inserto correcto y también su orientación. Un clon con el gen pyrG transcrito en el sentido contrario a las agujas del reloj se seleccionó. El nuevo plásmido se designó pALo2 (Figura 8).

Ejemplo 4: Amplificación PCR de un fragmento GH61 B de gen de ADN genómico de *Thielavia terrestris* NRRL 8126

[0329] Cebadores se diseñaron basados en secuencias peptídicas obtenidas a través de espectrometría de masas en serie como se describe en el ejemplo 1. Las secuencias peptídicas específicas fueron de la siguiente manera:

[L,I]PASNSPVTNVASDD[L,I]R (SEC ID n°: 23)
 [L,I]PED[L,I]EPGDY[L,I][L,I]R (SEC ID n°: 24)

[0330] Las regiones de las secuencias mostradas en negrita se seleccionaron para diseño del cebador de ADN. Una extensión 5' más larga para la segunda secuencia se creó genéticamente en la suposición de que el siguiente aminoácido en la secuencia fue alanina (basado en conservación de secuencia en homólogos). Los cebadores fueron:

Cebador sentido:
 5'-CCTCCAACCTCCCCGTACNAAYGTNGC-3' (SEC ID n°: 25)
 Cebador antisentido:
 5'-GGCGCGGAGGAGGTARTCNCCNGGYTC-3' (SEC ID n°: 26)

[0331] Amplificación PCR se realizó en un volumen de 50 µl con tampón 1X AmpliTaq (Applied Biosystems, Foster City, CA), 2,5 unidades de polimerasa de ADN AmpliTaq (Applied Biosystems, Foster City, CA), 1 µM de cada cebador sentido y antisentido, y aproximadamente 1 µg de ADN genómico de *Thielavia terrestris* NRRL 8126. Amplificación se realizó en un Robocycler (Stratagene, La Jolla, CA) usando parámetros de ciclo de 3 minutos a 96°C y 3 minutos a 72°C (durante los que ADN polimerasa se añadió), 35 ciclos de 45 segundos a 94°C, 45 segundos a 58°C, y 1 minuto a 72°C, seguido de una extensión final de 7 minutos a 72°C. Los productos reactivos se fraccionaron en un 3% de gel de agarosa usando 40 mM de Tris base-20 mM de acetato de sodio-1 mM de tampón EDTA disodio (TAE) y una banda de aproximadamente 450 pares de bases se cortó, purificó usando un kit de extracción de gel QIAEX II (QIAGEN Inc., Valencia, CA), y se subclonó usando un kit TOPO TA (Invitrogen, Carlsbad, CA). El plásmido de un transformante de *E. coli* se secuenció y se encontró que contenía un inserto de 452 de pares de bases que codifican a una proteína de la familia 61 (GH61 B). Este plásmido se designó pPH27 (Figura 9).

Ejemplo 5: Amplificación de un fragmento de gen GH61C y GH61 D de ADN genómico

[0332] Cebadores se diseñaron basados en secuencias peptídicas obtenidas a través de espectrometría de masas en serie como se describe en el ejemplo 1, con la excepción de que la fuente de los polipéptidos fue una banda de aproximadamente 27 kDa cortada de un gel de SDS-PAGE unidimensional (Novex 4-12% de Bis-Tris, Invitrogen, Carlsbad, CA) después de la purificación parcial a través de Q Sefarosa como se describe en el ejemplo 24. La masa molecular similar de los polipéptidos GH61C y GH61 D dieron lugar a la contaminación cruzada de los polipéptidos en la porción de gel usado para espectrometría de masas y así las secuencias peptídicas determinadas durante el análisis de espectrometría de masas derivaron de ambos polipéptidos. Un cebador se diseñó para la secuencia de dos péptidos adyacentes trípticos posibles con secuencia combinada SGAGWFKIDEAGFHGD (SEC ID n°: 27). Esta resultó ser la secuencia correcta para GH61 D; no obstante, el cebador diseñado estuvo suficientemente cerca de la secuencia GH61C (GDGWFKIDE) para amplificar ese gen también. El cebador antisentido se basó en la secuencia derivada de espectrometría de masas APGNY[I,L][I,L]R (SEC ID n°: 28). Este se refinó y se extendió basado en conservación en un alineamiento de secuencia múltiple para APGNYLIRHEL (SEC ID n°: 29). Esto resultó ser la secuencia correcta para GH61C, pero estuvo suficientemente cerca a la secuencia GH61D (APGNYLVRHEL, SEC ID n°: 30) para amplificar también ese gen. Las secuencias peptídicas específicas usadas para diseñar el cebador fueron de la siguiente manera:

GAGWFKIDE (SEC ID n°: 31)
APGNYLIRHEL (SEC ID n°: 29)

[0333] Los cebadores fueron:

Cebador sentido:
5'-CGGCGCGGGCTGGTTTAARATHGAYGA-3' (SEC ID n°: 32)
Cebador antisentido:
5' AGTTCATGGCGAATCAGATARTTNCNGGNGC-3' (SEC ID n°: 33)

[0334] Amplificación PCR se realizó en un volumen de 50 µl con tampón 1X AmpliTaq, 2,5 unidades de ADN polimerasa AmpliTaq, 1 µM de cada cebador sentido y antisentido, y aproximadamente 1 µg de ADN genómico de *Thielavia terrestris* NRRL 8126. Amplificación se realizó en un Stratagene Robocycler usando parámetros de ciclo de 3 minutos a 96°C y 3 minutos a 72°C (durante los que ADN polimerasa se añadió), 35 ciclos de 45 seg a 94°C, 45 seg a 53, 56, o 59 °C y 1 minuto a 72°C, seguido de una extensión final de 7 minutos a 72°C. Los productos reactivos se fraccionaron en un 3% de gel de agarosa y dos bandas de aproximadamente 150-200 pares de bases se cortaron, se purificaron usando el kit de extracción de gel QIAEX II (QIAGEN Inc., Valencia, CA), y se subclonaron juntos usando el kit Topo TA (Invitrogen, Carlsbad, CA). El plásmido de un transformante de *E. coli* se secuenció y se encontró que contuvo un inserto de 156 pares de bases que codifican una proteína de la familia 61 (GH61C). Este plásmido se designó pPH29 (Figura 10). El plásmido de otro transformante de *E. coli* se secuenció y se encontró que contenía un inserto de 180 pares de bases que codifican otra proteína de la familia 61 (GH61 D). Este plásmido se designó pPH28 (Figura 11).

Ejemplo 6: Construcción de biblioteca de ADN genómico

[0335] Bibliotecas de ADN genómico se construyeron usando el vector de clonación bacteriófago λZipLox (Life Technologies, Gaithersburg, MD) con células de *E. coli* Y1090ZL (Life Technologies, Gaithersburg, MD) como huésped para colocación en placas y purificación de bacteriófago recombinante y *E. coli* DH10Bzip (Life Technologies, Gaithersburg, MD) para escisión de clones pZL1 individuales con el gen GH61 B.

[0336] ADN genómico de *Thielavia terrestris* NRRL 8126 se digirió parcialmente con Tsp 5091 y se dividió por tamaño en 1% de geles de agarosa usando tampón TAE. Fragmentos de ADN migratorio en el intervalo de tamaño 3-7 kb se cortaron y eluyeron del gel usando reactivos Prep-α-Gene (BioRad Laboratories, Hercules, CA). Los fragmentos de ADN eluidos se ligaron con Eco RI dividido y se desforilaron brazos de vector λZipLox (Life Technologies, Gaithersburg, MD), y las mezclas de ligamiento se embalaron usando extractos de embalaje comercial (Stratagene, La Jolla, CA). Las bibliotecas de ADN empaquetadas se colocaron en placas y se amplificaron en células de *E. coli* Y1090ZL. La biblioteca de ADN genómico no amplificada contuvo 3,1 X 10⁶ a Pfu/ml (títulos de antecedentes sin ADN fueron 2.0 X 10⁴ a Pfu/ml).

Ejemplo 7: Identificación de clones de *Thielavia terrestris* GH61 B, GH61 C y GH61 D

[0337] Fragmentos de sonda de *Thielavia terrestris* GH61 B, GH61C y GH61 D se amplificaron a partir de pPH27, pPH29 y pPH28, respectivamente, usando homólogos de cebadores del vector TOPO y ADN polimerasa Herculase (Stratagene, La Jolla, CA), como se muestra debajo.

5'-CTTGGTACCGAGCTCGGATCCACTA-3' (SEC ID n°: 34)
5'-ATAGGGCGAATTGGGCCCTCTAGAT-3' (SEC ID n°: 35)

[0338] Cinquenta picomoles de cada uno de los cebadores se usaron en una reacción PCR con 10 ng de tampón de amplificación Herculase pPH27, pPH28, o pPH29, 1X (Stratagene, La Jolla, CA), 1 µl de 10 mM de mezcla de dATP, dTTP, dGTP y dCTP y 2,5 unidades de ADN polimerasa Herculase en un volumen final de 50 µl. Las condiciones de

amplificación fueron un ciclo a 94°C durante 1 minuto y 20 ciclos cada uno a 94°C durante 30 segundos, 55°C durante 30 segundos y 72°C durante 1 minuto. El bloque de calor luego fue a un ciclo de remojo de 4°C. Los productos reactivos se aislaron en 1,0% de gel de agarosa usando un tampón TAE donde tres <500 pares de bases de bandas de producto se cortaron del gel y se purificaron usando un kit de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA) según las instrucciones del fabricante. Veinticinco ng de cada fragmento se radiomarcó con ³²P usando un kit Prime It II (Stratagene, La Jolla, CA).

[0339] Aproximadamente 90.000 placas de la biblioteca descrita en el ejemplo 6 se seleccionaron por hibridación de placa usando los tres fragmentos de PCR marcados como las sondas. El ADN se reticuló sobre membranas (Hybond N+, Amersham, Arlington Heights, IL) usando un UV Stratalinker (Stratagene, La Jolla, CA). Cada fragmento de gen ³²P-radiomarcado se desnaturalizó añadiendo hidróxido sódico a una concentración final de 0,1 M, y se añadió a una solución de hibridación con 6X SSPE, 7% de SDS a una actividad de aproximadamente 1 x 10⁶ a cpm por ml de solución de hibridación. Cada una de las mezclas se incubó durante toda la noche a 55°C en un baño maría de agitación. Después de incubación, las membranas se lavaron tres veces durante quince minutos en 0,2X SSC con 0,1% de SDS a 65°C. Las membranas se secaron en papel secante durante 15 minutos, se envolvieron en SaranWrap™ y se expusieron a película de rayos X durante toda la noche a 70°C con pantallas intensificantes (Kodak, Rochester, NY).

[0340] Basado en la producción de señales de hibridación fuertes con las sondas GH61 anteriormente descritas, diferentes placas se eligieron para estudio posterior. Las placas se purificaron dos veces en células de *E. coli* Y1090ZL y los genes insertados y el plásmido pZL1 se cortaron posteriormente del vector λZipLox como derivados pZL1 (D'Alessio et al., 1992, Focus® 14:76) usando escisión de in vivo por infección de células de *E. coli* DH10BZL (Life Technologies, Gaithersburg, MD). Las colonias se inocularon en tres ml de LB más 50 µg/ml de medio de ampicilina y crecieron durante toda la noche a 37°C. ADN miniprep se preparó a partir de cada de estos cultivos usando un Biorobot 9600 (QIAGEN Inc., Valencia, CA). Clon pEJG120 (Figura 12) se mostró por secuenciación de ADN para contener el gen genómico en toda su longitud para GH61 B.

[0341] Los genes genómicos en toda su longitud para GH61C y GH61 D se aislaron de la misma manera como se describe para GH61 B.

[0342] *E. coli* pEJG120 con plásmido pEJG120 se depositó con la Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, con una fecha de depósito de 19 de diciembre de 2003.

[0343] Para depósito, el gen Tter61C fue amplificado por PCR de su construcción de expresión pEJG111 (véase ejemplo 13, Figura 13) con los cebadores mostrados debajo.

Cebador directo InFusion:

5'-ACAACCTGGATTACC**ATGCGGTTTCGACGCCTC**-3' (SEC ID n°: 36)

Cebador inverso InFusion:

5'-GTCAGTCACCTCTAGTTACT**AAAACCTCGAAGCC**-3' (SEC ID n°: 37)

Las letras en negrita representan secuencia codificante. La secuencia restante contiene identidad de secuencia en los sitios de inserción de pAILo2.

[0344] Cincuenta picomoles de cada uno de los cebadores de arriba se usaron en una reacción PCR con 50 ng de pEJG111 ADN, 1X Pfx tampón de amplificación (Invitrogen, Carlsbad, CA), 6 µl de 10 mM de mezcla de dATP, dTTP, dGTP y dCTP, 1 µl de 50 mM de MgSO₄, 5 µl de 10X pCRx solución intensificadora (Invitrogen, Carlsbad, CA), y 2,5 unidades de platino ADN polimerasa Pfx (Invitrogen, Carlsbad, CA), en un volumen final de 50 µl. Un Eppendorf Mastercycler 5333 (Eppendorf Scientific, Inc., Westbury, NY) se usó para amplificar el fragmento de ADN y se programó para un ciclo a 98°C durante 2 minutos y 35 ciclos cada uno a 94°C durante 30 segundos, 60°C durante 30 segundos y 68°C durante 1 minuto. Después de los 35, la reacción se incubó a 68°C durante 10 minutos y luego se enfrió a 10°C hasta procesado posterior. Un producto de reacción PCR de aproximadamente 800 pares de bases se aisló en 0,8% de gel de agarosa GTG (Cambrex Bioproducts, East Rutherford, NJ) usando tampón TAE y 0,1 µg de bromuro de etidio por ml. La banda de ADN se visualizó con la ayuda de un Dark Reader™ (Clare Chemical Research, Dolores, CO) para evitar mutaciones inducidas por los UV. Los 800 pares de bases de banda de ADN se cortaron con una hoja de afeitar desechable y se purificaron con un Ultrafree-DA Spin Cup (Millipore, Billerica, MA) según las instrucciones del fabricante.

[0345] La banda de ADN purificada se clonó en un vector pCR4-Blunt TOPO según las instrucciones del fabricante (Invitrogen, Carlsbad, CA). Dos microlitros de la reacción TOPO se transformaron en células de *E. coli* TOP10 según las instrucciones del fabricante (Invitrogen, Carlsbad, CA). Las células transformadas se colocaron en placas sobre placas de agar 2X YT complementadas con 100 µg de ampicilina por ml y se incubaron durante toda la noche a 37°C.

[0346] Ocho colonias se seleccionaron al azar para preparación de ADN plásmido. Cada colonia creció durante toda la noche en 3 ml de LB complementado con 100 µg de ampicilina por ml. De estos cultivos, 10 µl se usó para inocular en el lugar de una placa de agar 2X YT complementada con 100 µg de ampicilina por ml para tener una materia prima

bacteriana de cada colonia. El resto del cultivo se usó luego para preparar ADN plásmido con un QIAGEN BioRobot 9600. Clones se analizaron por digestión de enzima de restricción Eco RI. Ocho clones tuvieron el modelo de digerido de restricción previsto. De estos clones, tres se ordenaron luego con química de terminador Big-Dye™ como se ha descrito anteriormente, usando cebadores directos e inversos M13 estándar. Todos los tres clones mostraron tener la secuencia correcta del gen Tter61 E. Clon #7 se renombró pTter61C (Figura 15) y su materia prima bacteriana se resembró sobre una placa de 2X YT fresco complementada con 100 µg de ampicilina por ml y incubada durante toda la noche a 37°C. De esta placa, células se usaron para inocular dos 1,8 ml de crioviales con aproximadamente 1,5 ml de agarosa LB complementada con 100 µg de ampicilina por ml. Los viales se sellaron con PetriSeal™ (Diversified Biotech, Boston MA) y se depositaron con la Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, como pTter61 G NRRL B-30811, con fecha de depósito 21 de enero de 2005.

[0347] Para depósito, el gen Tter61D fue amplificado por PCR a partir de su construcción de expresión pEJG112 (véase el ejemplo 13, Figura 14) con los siguientes cebadores.

Cebador directo:

5'- CATGCCATGGATGCTTCTCAC-3' (SEC ID n°: 38)

Cebador inverso:

5'- CCTTAATTAATCAGGCGGTGAAGTC-3' (SEC ID n°: 39)

Letras en negrita representan secuencia codificante. La secuencia restante contiene sitios de restricción únicos para futuros experimentos de clonación.

[0348] La reacción PCR se realizó como se ha descrito anteriormente excepto con pEJG112. Un producto de reacción PCR de aproximadamente 1 kb se aisló en 0,8% de gel de GTG-agarosa usando tampón TAE y 0,1 µg de bromuro de etidio por ml. La banda de ADN se visualizó con la ayuda de un Dark Reader™ para evitar mutaciones inducidas por los UV. La banda de ADN 1 kb se cortó con una hoja de afeitar desechable y se purificó con un kit de purificación QIAquick PCR según las instrucciones del fabricante (QIAGEN Inc., Valencia, CA). La banda de ADN purificada se clonó en el vector pCR4-Blunt TOPO según las instrucciones del fabricante y transformantes de células de *E. coli* TOP10 se aislaron como se ha descrito anteriormente.

[0349] Diez colonias se seleccionaron al azar para preparación del ADN plásmido. Cada colonia creció durante toda la noche en 3 ml de LB complementado con 100 µg de ampicilina por ml y usada para preparar ADN plásmido con un QIAGEN BioRobot 9600. Clones se analizaron por digestión de enzima de restricción Eco RI para confirmar la presencia del inserto clonado. Todos los diez clones tuvieron el modelo de digestión de restricción previsto. De estos clones cinco se ordenaron luego con química del terminador de Big-Dye™ como se ha descrito anteriormente, usando cebadores directos e inversos M13. Uno de estos clones mostró que tenía la secuencia correcta del gen Tter61D. Este clon se renombró pTter61D (Figura 16) y se retransformó en células de *E. coli* TOP10 como se ha descrito anteriormente. De una única sembra de colonia, células se usaron para inocular dos crioviales de 1,8 ml con aproximadamente 1,5 ml de agarosa LB complementado con 100 µg/ml de ampicilina por ml. Los viales se sellaron con PetriSeal™ y se depositaron con la Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, como pTter61D NRRL B-30812, con fecha de depósito 21 de enero de 2005.

Ejemplo 8: Construcción de librería de ADNc de etiquetas de secuencia expresadas (EST)

[0350] Dos ml de alícuota de un cultivo líquido de 24 horas (50 ml de medio NNCYPmod complementado con 1% de glucosa en matraz 250 ml, 45°C, 200 r.p.m.) de *Thielavia terrestris* NRRL 8126 se usó para sembrar un matraz 500 ml con 100 ml de medio NNCYPmod complementado con 2% de Sigmacell-20. El cultivo se incubó a 45°C, 200 r.p.m. durante 3 días. Los micelios se cosecharon por filtración a través de una unidad de filtración desechable con un prefiltro de fibra de vidrio (Nalgene, Rochester NY), se lavó dos veces con 10 mM de Tris-HCl-1 mM de EDTA pH 8 (TE), y se congeló rápidamente en nitrógeno líquido.

[0351] El ARN total se aisló usando el siguiente método. Micelios congelados de *Thielavia terrestris* NRRL 8126 se molieron en un molinillo de café eléctrico. El material molido se mezcló 1:1 V/V con 20 ml de Fenazol (Ambion, Inc., Austin, TX) en un tubo Falcon de 50 ml. Una vez los micelios se suspendieron, estos se extrajeron con cloroformo y tres veces con una mezcla de alcohol de fenol-cloroformo-isoamil 25:24:1 v/v/v. De la fase acuosa resultante, el ARN se precipitó añadiendo volumen 1/10 de 3 M de acetato sódico pH 5,2 y 1,25 volúmenes de isopropanol. El precipitado de ARN se recuperó por centrifugado a 12.000 x g durante 30 minutos a 4°C. El granulado final se lavó con etanol 70% frío, se secó al aire y se resuspendió en 500 ml de agua tratada con dietilpirocarbonato (agua DEPC).

[0352] La calidad y cantidad del ARN purificado se evaluó con un Agilent Bioanalyzer 2100 (Agilent Technologies, Inc., Palo Alto, CA). ARNm poliadenilado se aisló de 360 µg de ARN total con la ayuda de un kit Poly(A) Purist Magnetic (Ambion, Inc., Austin, TX) según las instrucciones del fabricante.

[0353] Para crear la genoteca de ADNc, un kit CloneMiner™ (Invitrogen, Carlsbad, CA) se empleó para construir una biblioteca direccional que no requiere el uso de clonación de enzima de restricción, reduciendo así el número de clones quiméricos y sesgo de tamaño.

5 [0354] Para asegurar la síntesis exitosa del ADNc, dos reacciones se realizaron en paralelo con dos concentraciones diferentes de ARNm (2,2 y 4,4 µg de poli(A)⁺ ARNm). Las muestras de ARNm se mezclaron con un cebador Biotin-attB2-Oligo(dt) (kit CloneMiner™, Invitrogen, Carlsbad, CA), 1X primer tampón de cadena (Invitrogen, Carlsbad, CA), 2 µl de 0,1 M de DTT, 10 mM de cada dNTP, y agua a un volumen final de 18 y 16 µl respectivamente.

10 [0355] Las mezclas de reacción se mezclaron cuidadosamente y luego 2 y 4 µl de transcriptasa inversa SuperScript™ (Invitrogen, Carlsbad, CA) se añadieron e incubaron a 45°C durante 60 minutos para sintetizar la primera cadena complementaria. Para segunda síntesis de cadena, para cada reacción de primera cadena se añadió 30 µl de 5X tampón de segunda cadena (Invitrogen, Carlsbad, CA), 3 µl de 10 mM de cada dNTP, 10 unidades de ADN ligasa de E. coli (Invitrogen, Carlsbad, CA), 40 unidades de ADN polimerasa I de E.coli (Invitrogen, Carlsbad, CA) y 2 unidades de ribonucleasa H de E. coli (Invitrogen, Carlsbad, CA) en un volumen total de 150 µl. Las mezclas se incubaron luego a 16°C durante dos horas. Después de la incubación de dos horas, 2 µl de ADN polimerasa T4 (Invitrogen, Carlsbad, CA) se añadieron a cada reacción y se incubaron a 16 °C durante 5 minutos para crear un ADNc de extremo romo. Las reacciones de ADNc se extrajeron con una mezcla de alcohol de fenol-cloroformo-isoamil 25:24:1 v/v/v y se precipitaron en presencia de 20 µg de glicógeno, 120 µl de 5 M de acetato amónico y 660 µl de etanol. Después de centrifugado a 12.000 x g, 4°C durante 30 minutos, los gránulos de ADNc se lavaron con etanol frío 70%, se secaron al vacío durante 2-3 minutos y se resuspendieron en 18 µl de agua DEPC. Por cada muestra de ADNc resuspendida se añadió 10 µl de 5X tampón adaptado (Invitrogen, Carlsbad, CA), 10 µg de adaptador attB1 (Invitrogen, Carlsbad, CA), 7 µl de 0,1 M de DTT y 5 unidades de ADN ligasa T4 (Invitrogen, Carlsbad, CA).

25 5'-TCGTCGGGGACAACCTTTGTACAAAAAAGTTGG-3' (SEC ID n°: 40)
3'-CCCCTGTTGAAACATGTTTTTCAACCp-5' (SEC ID n°: 41)

[0356] Reacciones de ligamiento se incubaron durante toda la noche a 16°C. Exceso de adaptadores se retiró por cromatografía de exclusión de tamaño con 1 ml de resina Sephacryl™ S-500 HR (Amersham Biosciences, Piscataway, NJ). Fracciones de columna se recogieron según las instrucciones del kit CloneMiner™ y las fracciones 3 a 14 se analizaron con un Agilent Bioanalyzer para determinar la fracción en la que los adaptadores attB1 comenzaron elución. El análisis mostró que los adaptadores comenzaron elución fracción alrededor de elución 10 o 11. Para las primeras fracciones de la biblioteca 6 a 11 se agruparon y para las segundas fracciones de la biblioteca 4-11 se agruparon.

35 [0357] Clonación del ADNc se realizó por recombinación de ADN homóloga según el protocolo Gateway System (Invitrogen, Carlsbad, CA) usando BP Clonase™ (Invitrogen, Carlsbad, CA) como la recombinasa. Cada reacción de recombinación de BP clonase™ contuvo aproximadamente 70 ng de attB-flanqueado-ADNc, 250 ng de pDONR™222 (Invitrogen, Carlsbad, CA), 2 µl de 5X tampón de BP Clonase™ (Invitrogen, Carlsbad, CA), 2 µl de TE y 3 µl de BP Clonase™. Reacciones de recombinación se incubaron a 25°C durante toda la noche.

40 [0358] Reacciones de recombinación activadas por calor BP se dividieron luego en 6 partes alícuotas y se sometieron a electroporación en las células electrocompetentes ElectroMax™ DH10B (Invitrogen, Carlsbad, CA) usando un BioRad Gene Pulser II (BioRad, Hercules, CA) con los siguientes parámetros: 2,0 kV, 200 Ω y 25 µF. Las células sometidas a electroporación se resuspendieron en 1 ml de SOC y se incubaron a 37°C durante 60 minutos con agitación constante (200 r.p.m.). Después del periodo de incubación, las células transformadas se agruparon y se mezclaron 1:1 con medio refrigerante. 200 µl de alícuota se retiró para titulación de biblioteca y luego el resto de cada biblioteca se dividió en partes alícuotas en 1,8 ml de crioviales (Wheaton Science Products, Millville, NJ) y se almacenó congelado a -80°C.

50 [0359] Cuatro diluciones en serie de cada biblioteca se prepararon: 1/100, 1/1000, 1/10⁴, 1/10⁵. De cada dilución, 100 µl se colocaron en placas sobre placas LB 150 mm complementados con 50 µg de kanamicina por ml y se incubaron a 37°C durante toda la noche. El número de colonias en cada placa de dilución se contaron y se usaron para calcular el número total de transformantes en cada biblioteca.

55 [0360] La primera biblioteca mostró tener 5,4 millones de clones independientes y la segunda biblioteca mostró tener 9 millones de clones independientes.

Ejemplo 9: Preparación de molde y secuenciación nucleótida de clones de ADNc

60 [0361] Partes alícuotas de ambas bibliotecas se mezclaron y colocaron sobre placas LB 25 x 25 cm complementado con 50 µg de kanamicina por ml. Colonias individuales se dispusieron sobre placas de 96 pocillos con 100 µl de medio LB complementado con 50 µg de kanamicina por ml con la ayuda de un robot Genetix QPix (Genetix Inc., Boston, MA). Cuarenta y cinco placas de 96 pocillos se obtuvieron para un total de 4320 clones individuales. Las placas se incubaron durante toda la noche a 37°C con agitación a 200 r.p.m. Tras la incubación, 100 µl de glicerol 50% estéril se añadió a cada pocillo. Los transformantes se replicaron con la ayuda de una herramienta de 96 pines (Boekel, Feasterville, PA) en placas hondas secundarias de microcultivo de 96 pocillos (Advanced Genetic Technologies Corporation, Gaithersburg, MD) con 1 ml de Magnificent Broth™ (MacConnell Research, San Diego, CA) complementado con 50 µg

de kanamicina por ml en cada pocillo. Las placas de microtitulación primarias se almacenaron congeladas a -80°C. Las placas hondas secundarias se incubaron a 37°C durante toda la noche con agitación vigorosa (300 r.p.m.) en un agitador rotatorio. Para prevenir vertido y contaminación cruzados y para permitir aireación suficiente, cada placa de cultivo secundaria se cubrió con una almohadilla de polipropileno (Advanced Genetic Technologies Corporation, Gaithersburg, MD) y una cobertura de plato de microtitulación plástica. ADN plásmido se preparó con un MWG Robot-Smart 384 (MWG Biotech Inc., High Point, NC) y kits Montage Plasmid Miniprep (Millipore, Billerica, MA).

[0362] Reacciones de secuenciación se realizaron usando química de terminador Big-Dye™ (Giesecke *et al.*, 1992, *Journal of Virology Methods* 38: 47-60) y un cebador de secuenciación directo M13 (-20):

5'-GTAAAACGACGGCCAG-3' (SEC ID n°: 42)

[0363] Las reacciones de secuenciación se realizaron en un formato de 384 pocillos con un Robot-Smart 384 (MWG Biotech Inc., High Point, NC). Eliminación de terminador se realizó con kits Millipore MultiScreen Seq384 Sequencing Clean-up (Millipore, Billerica, MA). Las reacciones contuvieron 6 µl de ADN plásmido y 4 µl de mezcla maestra de secuenciación con 1 µl de 5X tampón de secuenciación (Millipore, Billerica, MA), 1 µl de terminador de Big-Dye™ (Applied Biosystems, Inc., Foster City, CA), 1,6 pmols de cebador directo M13 y 1 µl de agua. Secuenciación del ADN monopaso se realizó con un ABI PRISM Automated DNA Sequencer Model 3700 (Applied Biosystems, Foster city, CA).

20 Ejemplo 10: Análisis de datos de secuencia de ADN de clones de ADNc

[0364] Llamada de bases, asignación de valores de calidad y recorte de vectores se realizaron con la asistencia de software PHRED/PHRAP (University of Washington, Seattle, WA). Análisis de agrupamiento de los EST se realizó con un Parcel Transcript Assembler v. 2.6.2. (Paracel, Inc., Pasadena, CA). Análisis del agrupamiento de EST indicó la presencia de 395 clústeres independientes.

[0365] Análisis de homología secuencial del las secuencias EST ensambladas contra la base de datos PIR se realizó con el programa Blastx (Altschul *et al.*, 1990, *J. Mol. Biol.* 215:403-410) en un clúster de Linux de 32 nodos (Paracel, Inc., Pasadena, CA) usando la matriz BLOSUM 62 (Henikoff, 1992, *Proc. Natl. Acad. Sci. USA* 89: 10915-10919). De los 395 clusters, 246 tuvieron una identidad significativa para genes conocidos en la base de datos de proteína pública y 149 no tuvo ninguna identidad significativa contra esta base de datos. Entre estos 246 clústeres, 13 tuvieron golpes contra homólogos bien caracterizados de genes de glicosil hidrolasa.

35 Ejemplo 11: Identificación de clones de ADNc que codifican dos polipéptidos de la familia 61 con actividad de mejora celulolítica (GH61 E y GH61 G)

[0366] Dos clones de ADNc que codifican polipéptidos de la familia 61 con actividad de mejora celulolítica (GH61 E y GH61 G) se identificaron inicialmente por su identidad para la proteína de la familia 61 a partir de *Volvariella volvacea* (GenPept g49333361). Este análisis indicó que las proteínas fueron 41% y 38%, respectivamente, idénticos a la *V. volvacea* al nivel de proteína sobre un extensión de 289 aminoácidos (867 pares de bases). Después de esta identificación inicial, el clones de EST Tter39E1 y Tter39H8 se recuperaron de sus placas de materia prima original congelada y se sembraron sobre LB complementado con 50 µg de kanamicina por ml. Las placas se incubaron durante toda la noche a 37°C y el día siguiente una única colonia de cada placa se usó para inocular 3 ml de LB complementado con 50 µg de kanamicina por ml. Los cultivos líquidos se incubaron durante toda la noche a 37°C y ADN plásmido se preparó con un QIAGEN BioRobot 9600. ADN plásmido de cada clon de EST se ordenó nuevamente con química de terminador de Big-Dye™ como se ha descrito anteriormente, usando el cebador directo M1 y un cebador poly-T mostrado a continuación a la secuencia del 3' final del clon.

5'-TTTTTTTTTTTTTTTTTTTTTTTTTVN-3' (SEC ID n°: 43)

Donde V=G, A, o C y N= G, A, C, o T

[0367] Análisis de secuencias de clon Tter39E1 indicó que este clon no era de longitud completa y faltaban algunos nucleótidos al 5' final ya que no tiene un codón de inicio "ATG". Para obtener un clon de longitud completa, un par de cebadores se diseñaron amplificar el 5' final de este gen de la agrupación de ADNc original usada para construir la biblioteca.

[0368] El cebador sentido o directo se diseñó para encontrar el adaptador attB1 usado durante la construcción de la genoteca de ADNc y el cebador inverso se diseñó para amplificar de aproximadamente 350 pares de bases del 5' final truncado, como se muestra debajo.

Cebador attB1:
5'-GGGGACAACCTTTGTACAAAAAAGTTGG-3' (SEC ID n°: 44)

Cebador inverso:
5'-AAAGGTAGGATGGTCCTCGTACACCTT-3' (SEC ID n°: 45)

[0369] Cincuenta picomoles de cada uno de los cebadores de arriba se usaron en una reacción PCR con 1 µl del ADNc agrupado, 1X tampón de amplificación Taq (New England BioLabs, Beverly, MA), 6 µl de 10 mM de mezcla de dATP, dTTP, dGTP y dCTP, y 2,5 unidades de ADN polimerasa Taq polimerasa (New England BioLabs, Beverly, MA), en un volumen final de 50 µl. Un Eppendorf Mastercycler 5333 se usó para amplificar el fragmento programado para un ciclo a 98°C durante 2 minutos; 5 ciclos cada a 96°C durante 30 segundos, 55°C durante 30 segundos y 68°C durante 1 minuto; y 35 ciclos cada uno a 96°C durante 30 segundos, 61 °C durante 30 segundos, y 68°C durante 1 minuto. Después de los 35 ciclos, la reacción se incubó a 68°C durante 10 minutos y luego se enfrió a 10°C hasta proceso posterior.

[0370] Un producto de reacción PCR (aproximadamente 400 pares de bases) se aisló en 0,8% de gel de GTG-agarosa usando un tampón TAE y 0,1 µg de bromuro de etidio por ml. La banda de ADN se visualizó con la ayuda de un Dark Reader™ para evitar mutaciones inducidas por UV. Los 400 pares de bases de la banda de ADN se cortaron con una hoja de afeitar desechable y se purificaron con una Ultrafree-DA Spin Cup según las instrucciones del fabricante. La banda de ADN purificada se clonó en un vector pCR2.1 TA-TOPO según las instrucciones del fabricante (Invitrogen, Carlsbad, CA). Dos microlitros de la reacción TA se transformaron en células *E. coli* TOP10 según las instrucciones del fabricante. Las células transformadas se colocaron luego sobre placas de agar 2X YT complementadas con 100 µg/ml de ampicilina y se incubaron durante toda la noche a 37°C. Ocho colonias se seleccionaron aleatoriamente para preparación de ADN plásmido. Cada colonia creció durante toda la noche en 3 ml de medio LB complementado con 100 µg de ampicilina por ml y ADN plásmido se obtuvo a partir de cada uno usando un QIAGEN BioRobot 9600. Los clones se analizaron por digestión de enzima de restricción Eco RI. Todos los ocho clones tuvieron el modelo de digestión de restricción prevista. A partir de estos tres clones se ordenaron luego con química de terminador Big-Dye™ como se ha descrito anteriormente, usando un cebador inverso M13 estándar. Análisis de secuencias indicaron que todos los tres clones tienen un 5' final completo y que al clon EST pTter39E1 original le faltaron solo cinco nucleótidos en el 5' final.

[0371] De la secuencia en toda su longitud de los cebadores PCR InFusion del gen Tter61 R se diseñaron como se muestra debajo para amplificar el gen truncado, para clonación en el vector de expresión pAILo2 y añadir al mismo tiempo los 5 nucleótidos que faltan en su 5' final.

Cebador directo en fusión:
 5'-ACTGGATTACC**ATGCTCGCAAACGGTGCCATCGTCT**-3' (SEC ID n°: 46)
 Cebador inverso en fusión:
 5'- TCACCTCTAGTTAATTAAT**CAGCAGCTGAAGACGGCCG**-3' (SEC ID n°: 47)

Las letras en negrita representan secuencia codificante. Las letras en negrita y subrayadas representan los cinco nucleótidos que faltan del 5' final del gen como se ha descrito anteriormente. La secuencia restante contiene identidad de secuencia en los sitios de inserción de pAILo2.

[0372] La reacción PCR se llevó a cabo como se describe en el ejemplo 7 excepto con pTter39E1. Un producto de reacción PCR de aproximadamente 700 pares de bases se aisló en un 0,8% de gel de GTG-agarosa usando tampón TAE y 0,1 µg de bromuro de etidio por ml. La banda de ADN se visualizó con la ayuda de un Dark Reader™ para evitar mutaciones inducidas por los UV. La banda de ADN de 700 pares de bases se cortó con una hoja de afeitar desechable y se purificó con una Ultrafree-DA Spin Cup según las instrucciones del fabricante.

[0373] La banda de ADN purificada se clonó en el vector pCR4-Blunt TOPO según las instrucciones del fabricante y transformantes de células *E. coli* TOP10 se aislaron como se describe en el ejemplo 7. Ocho colonias se seleccionaron aleatoriamente para preparación del ADN plásmido. Cada colonia creció durante toda la noche en 3 ml de LB complementado con 100 µg de ampicilina por ml y se usó para preparar ADN plásmido con un QIAGEN BioRobot 9600. Los clones se analizaron por digestión de enzima de restricción Pac I/Pst I. Siete de ocho clones tuvieron el patrón de digestión de restricción previsto. De estos siete clones, tres se secuenciaron luego con química de terminador Big-Dye™ como se ha descrito anteriormente, usando cebadores directos e inversos M13 estándar. Todos los tres clones mostraron tener la secuencia correcta del gen Tter61 E. Clon #2 se renombró pTter61E (Figura 17) y se retransformó en células de *E. coli* TOP10 como se ha descrito anteriormente. De una única colonia sembrada, células se usaron para inocular dos crioviales de 1,8 ml con aproximadamente 1,5 ml de agarosa LB complementada con 100 µg/ml de ampicilina por ml. Los viales se sellaron con PetriSeal™ y se depositaron con la Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, como pTter61E NRRL B-30814, con fecha de depósito 21 de enero 21 de 2005.

[0374] Una vez la identidad del clon Tter39H8 se confirmó, el plásmido se renombró pTter61G (Figura 18). 0,5 µl de alícuota de ADN plásmido de este clon se transfirió en un frasco de células de *E. coli* TOP10, se mezcló suavemente y se incubó en hielo durante 10 minutos. Las células se sometieron a choque térmico luego a 42°C durante 30 segundos y se incubaron nuevamente en hielo durante 2 minutos. Las células se resuspendieron en 250 µl de SOC y se incubaron a 37°C durante 60 minutos con agitación constante (200 r.p.m.). Después del periodo de incubación, dos partes alícuotas de 30 µl se colocaron sobre placas LB complementadas con 50 µg de kanamicina por ml y se incubaron durante toda la noche a 37°C. El día siguiente una única colonia se seleccionó y sembró sobre una placa 2X YT fresca complementada con 50 µg de kanamicina por ml y se incubó durante toda la noche a 37°C. De esta placa, células se usaron para inocular dos crioviales de 1,8 ml con aproximadamente 1,5 ml de agarosa LB complementada con 50 µg de kanamicina

por ml. Los viales se sellaron con PetriSeal™ y se depositaron con la colección de Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, como pTter61G NRRL B-30811, con fecha de depósito 21 enero de 2005.

5 **Ejemplo 12: Caracterización de las secuencias genómicas de *Thielavia terrestris* que codifican polipéptidos de la familia GH61 B, GH61C y GH61 D con actividad de mejora celulolítica y secuencias de ADNc que codifican polipéptidos de la familia GH61G y GH61E con actividad de mejora celulolítica.**

10 [0375] Secuenciación del ADN del clon genómico (clon 15) de *Thielavia terrestris* GH61 B y superposición múltiple de clones genómicos GH61C y GH61 D clones genómicos se realizó con un Applied Biosystems Model 3700 Automated DNA Sequencer usando química de terminador BigDye™ version 3.1 y química de dGTP (Applied Biosystems, Inc., Foster City, CA) y estrategia de "primer walking". Datos de secuencia de nucleótidos se escrutinaron para calidad y todas las secuencias se compararon entre sí con asistencia de software PHRED/PHRAP (University of Washington, Seattle, WA).

15 [0376] Modelos de gen para las secuencias de ADN genómico de *Thielavia terrestris* GH61 B, GH61C y GH61 D se construyeron basadas en similitud con genes homólogos de *Diplodia gossypina*, *Trichophaea saccata* y *Pseudopezizomyces nigrella*.

20 [0377] La secuencia de nucleótidos (SEC ID n°: 1) y secuencia de aminoácidos deducida (SEC ID n°: 2) del gen de *Thielavia terrestris* GH61 se muestra en las figuras 1 A y 1 B. La secuencia codificante es 1104 pares de bases que incluye el codón de terminación y se interrumpe por intrones de 60 y 63 pares de bases. La proteína codificada predicha es 326 aminoácidos. La región codificante es 65,8% G+C. Usando el programa SignalP (SignalP), un péptido señal de 19 residuos se predijo. La proteína predicha madura contiene 307 aminoácidos con una masa molecular de 31,3 kDa.

25 [0378] Análisis de la secuencia de aminoácidos deducida del gen GH61B con el programa Interproscan (Zdobnov and Apweiler, 2001, Bioinformatics 17: 847-848) mostró que el gen GH61 B contuvo la firma de secuencia del dominio de enlace de celulosa fúngica. Esta firma de secuencia conocida como el patrón Prosite PS00562 (Sigrist et al., 2002, Brief Bioinform. 3: 265-274) se encontró de aproximadamente los residuos 271 a 307 del polipéptido maduro.

30 [0379] Un alineamiento comparativo de secuencias de aminoácidos se determinó usando el algoritmo de Smith-Waterman (Waterman et al., 1976, Adv. Math. 20: 367) con penalización abierta de espacio de 11, penalización de extensión del espacio de 1 y la matriz BLOSUM62. El alineamiento mostró que la secuencia de aminoácidos deducida del gen de *Thielavia terrestris* que codifica un polipéptido de la familia GH61 B con actividad de mejora celulolítica compartió 70% y 64% de identidad (incluyendo espacios) con las secuencias de aminoácidos deducidas de dos proteínas de la familia 61 de *Neurospora crassa* (número de acceso EAA26873 y EAA36262).

35 [0380] La secuencia de nucleótidos (SEC ID n°: 3) y la secuencia de aminoácidos deducida (SEC ID n°: 4) del gen de *Thielavia terrestris* GH61C se muestran en la figura 2. La secuencia codificante es 778 pares de bases que incluye el codón de terminación y se interrumpe por un intrón de 58 pares de bases. La proteína codificada predicha es 240 aminoácidos. La región codificante es 66,2% G+C. Usando el programa SignalP, un péptido señal de 17 residuos se predijo. La proteína predicha madura contiene 223 aminoácidos con una masa molecular de 24,1 kDa.

45 [0381] Un alineamiento comparativo de secuencias de aminoácidos se determinó usando el algoritmo de Smith-Waterman (Waterman et al., 1976, Adv. Math. 20: 367) con penalización abierta de espacio de 11, penalización de extensión del espacio de 1 y la matriz BLOSUM62. El alineamiento mostró que la secuencia de aminoácidos deducida del gen *Thielavia terrestris* que codifica un polipéptido de la familia GH61C con actividad de mejora celulolítica compartió 76% y 72% de identidad con las secuencias de aminoácidos deducidas de dos proteínas de la familia 61 de *Neurospora crassa* (número de acceso Q9P3R7 y Q7RV41, respectivamente).

50 [0382] La secuencia de nucleótidos (SEC ID n°: 5) y la secuencia de aminoácidos deducida (SEC ID n°: 6) del gen de *Thielavia terrestris* GH61 D se muestran en la figura 3. La secuencia codificante es 913 pares de bases que incluye el codón de terminación y se interrumpe por intrones de 70 y 66 pares de bases. La proteína codificada predicha es 258 aminoácidos. La región codificante es 63,1% G+C. Usando el programa SignalP, un péptido señal de 19 residuos se predijo. La proteína predicha madura contiene 239 aminoácidos con una masa molecular de 25,7 kDa.

55 [0383] Un alineamiento comparativo de secuencias de aminoácidos se determinó usando el algoritmo de Smith-Waterman (Waterman et al., 1976, Adv. Math. 20: 367) con penalización abierta de espacio de 11, penalización de extensión del espacio de 1 y la matriz BLOSUM62. El alineamiento mostró que la secuencia de aminoácidos deducida del gen de *Thielavia terrestris* que codifica un polipéptido de la familia GH61 D con actividad de mejora celulolítica compartió 53% y 53% de identidad con las secuencias de aminoácidos deducidas de dos proteínas de la familia 61 de *Neurospora crassa* (número de acceso Q9P3R7 y Q7RV41, respectivamente).

60 [0384] La secuencia de ADNc (SEC ID n°: 7) y la secuencia de aminoácidos deducida (SEC ID n°: 8) del gen de *Thielavia terrestris* GH61 E se muestran en la figura 4. La secuencia codificante es 681 pares de bases que incluye el codón de terminación. La proteína codificada predicha es 226 aminoácidos. El % de contenido G+C dl clon de ADNc

GH61E es 65,6% y de la región codificante de proteína madura (nucleótidos 55 a 681 de la SEC ID n°: 7) es 65,4%. Usando el programa SignalP, un péptido señal de 18 residuos se predijo. La proteína predicha madura contiene 208 aminoácidos con una masa molecular de 22,3 kDa.

5 [0385] Un alineamiento comparativo de secuencias de aminoácidos se determinó usando el algoritmo de Smith-Waterman (Waterman et al., 1976, Adv. Math. 20: 367) con penalización abierta de espacio de 11, penalización de extensión del espacio de 1 y la matriz BLOSUM62. El alineamiento mostró que la secuencia de aminoácidos deducida del gen de *Thielavia terrestris* que codifica un polipéptido de la familia GH61 E con actividad de mejora celulolítica compartió 73% y 53% de identidad con las secuencias de aminoácidos deducidas de una proteína de la familia 61 de *Neurospora crassa* (número de acceso Q873G1) y otro de *Magnaporthe grisea* (número de acceso EAA54517), respectivamente.

15 [0386] La secuencia de nucleótidos (SEC ID n°: 9) y la secuencia de aminoácidos deducida (SEC ID n°: 10) del gen de *Thielavia terrestris* GH61 G se muestran en la figura 5. La secuencia codificante es 915 pares de bases que incluye el codón de terminación. La proteína codificada predicha es 304 aminoácidos. El % de contenido G+C del clon de ADNc GH61 G es 64,5% y de la región codificante de proteína madura (nucleótidos 58 a 912 de la SEC ID n°: 9) es 64,4%. Usando el programa SignalP, un péptido señal de 19 residuos se predijo. La proteína predicha madura contiene 285 aminoácidos con una masa molecular de 30,0 kDa.

20 [0387] Análisis de la secuencia de aminoácidos deducida del gen GH61 G con el programa Interproscan mostró que el gen GH61 G contuvo la firma de secuencia del dominio de enlace de celulosa fúngica. Esta firma de secuencia estuvo presente de aproximadamente los residuos 271 a 304 del polipéptido maduro.

25 [0388] Un alineamiento comparativo de secuencias de aminoácidos se determinó usando el algoritmo de Smith-Waterman (Waterman et al., 1976, Adv. Math. 20: 367) con penalización abierta de espacio de 11, penalización de extensión del espacio de 1 y la matriz BLOSUM62. El alineamiento mostró que la secuencia de aminoácidos deducida del gen *Thielavia terrestris* que codifica un polipéptido de la familia GH61 G con actividad de mejora celulolítica compartió 61 % y 61 % de identidad con las secuencias de aminoácidos deducidas de una proteína de la familia 61 de *Neurospora crassa* (número de acceso Q7SCJ5) y otro de *Gibberella zeae* (número de acceso EAA72972), respectivamente.

30 [0389] Un alineamiento comparativo de secuencias de la familia 61 de *Thielavia terrestris* NRRL 8126 se determinó usando el método MAFFT NW con refinamiento iterativo y parámetros predeterminados (Kato et al., 2002, Nucleic Acids Research 30: 3059). Una matriz de identidad se calculó usando software LASERGENE™ MEGALIGN™ (DNASTAR, Inc., Madison, WI). Los resultados del alineamiento se muestran en la tabla 1.

Tabla 1. Alineamiento de secuencias polipeptídicas de *Thielavia terrestris* GH61

Porcentaje de identidad

		1	2	3	4	5	
Divergencia	1		59,6	35,2	33,2	31,2	1 <i>Thielavia terrestris</i> GH61C
	2	59,6		32,2	27,5	26,7	2 <i>Thielavia terrestris</i> GH61D
	3	135,0	138,4		43,3	41,7	3 <i>Thielavia terrestris</i> GH61G
	4	142,6	174,6	129,6		42,1	4 <i>Thielavia terrestris</i> GH61E
	5	134,4	157,0	116,4	98,8		5 <i>Thielavia terrestris</i> GH61B
		1	2	3	4	5	

Ejemplo 13: Construcción de un vector de expresión de *Aspergillus oryzae* para los genes de la familia de *Thielavia terrestris* GH61 B, GH61C, Gh61 D, GH61 E y GH61 G

55 [0390] Dos cebadores oligonucleótidos sintéticos mostrados debajo se diseñaron para amplificar por PCR el gen de *Thielavia terrestris* de la familia GH61 B a partir del clon genómico. Un kit InFusion Cloning (BD Biosciences, Palo Alto, CA) se usó para clonar el fragmento directamente en el vector de expresión, pAILo2, sin necesidad de restricción de digestión y ligamiento.

60 Cebador directo:
 5'-ACTGGATTTACCAT**GAAGTCGTTCA**CCATTG-3' (SEC ID n°: 48)
 Cebador inverso:
 5'-AGTCACCTCTAGTTAGAGGC**ACTGCGAGTAG**-3' (SEC ID n°: 49)

65 Las letras en negrita representan la secuencia codificante. La secuencia restante es homóloga a los sitios de inserción de pAILo2.

[0391] Cinquenta picomoles de cada uno de los cebadores arriba se usaron en una reacción PCR con 100 ng de clon genómico 15 de *Thielavia terrestris* ADN (preparado como se describe en el ejemplo 2), tampón de amplificación 1X Pfx, 1,5 µl de 10 mM de mezcla de dATP, dTTP, dGTP y dCTP, 2,5 unidades de platino Pfx ADN polimerasa, 1 µl de 50 mM de MgSO₄ y 5 µl de solución intensificadora 10X pCRx en un volumen final de 50 µl. Las condiciones de amplificación fueron un ciclo a 94°C durante 2 minutos; y 30 ciclos cada uno a 94°C durante 15 segundos, 55°C durante 30 segundos y 68°C durante 3 minutos. El bloque de calor luego fue a un 4°C ciclo de remojo.

[0392] Los productos reactivos se aislaron en un 1,0% de gel de agarosa usando tampón TAE donde una banda de producto 3 kb se cortó del gel y se purificó usando un kit QIAquick Gel Extraction según las instrucciones del fabricante.

[0393] El fragmento se clonó luego en el vector de expresión pAILo2 usando un kit de InFusion Cloning. El vector se digirió con endonucleasas de restricción Nco I y Pac I (usando condiciones específicas por el fabricante). El fragmento se purificó por electroforesis en gel y purificación de gel QIAquick. El fragmento de gen y el vector digerido se unieron en una reacción que dio lugar al plásmido de expresión pEJG108 (Figura 19) en el que transcripción del gen de la familia GH61 B fue bajo el control del promotor NA2-tpi. La reacción de ligamiento (20 µl) se compuso por el tampón de infusión 1X (BD Biosciences, Palo Alto, CA), 1X BSA (BD Biosciences, Palo Alto, CA), 1 µl de enzima InFusion (diluida 1:10) (BD Biosciences, Palo Alto, CA), 100 ng de pAILo2 digerido con Nco I y Pac I y 100 ng de producto purificado PCR de *Thielavia terrestris* GH61B. La reacción se incubó a temperatura ambiente durante 30 minutos. Un µl de la reacción se usó para transformar células de *E. coli* XL10 Solopac Gold (Stratagene, La Jolla, CA). Un transformante de *E. coli* con pEJG108 (gen GH61 B) se detectó por digestión de enzima de restricción y ADN plásmido se preparó usando un QIAGEN BioRobot 9600.

[0394] Los genes de *Thielavia terrestris* de la familia GH61 C y GH61 D se generaron de la misma manera como se ha descrito anteriormente usando los siguientes cebadores.

Para Tter61 C:

Cebador directo InFusion:

5'-ACA**ACTGGATT**TACCAT**GC**GGTT**CGACGCCTC**-3' (SEC ID n°: 50)

Cebador inverso InFusion:

5'-GTCAGTCACCTCTAGTT**ACTAAA**ACT**CGAAGCC**-3' (SEC ID n°: 51)

Para Tter61D

Cebador directo InFusion:

5'-ACTGGATT**ACCATGCTTCTCACATCAG**-3' (SEC ID n°: 52)

Cebador inverso InFusion:

5'-AGTCACCTCTAGTT**ATCAGGCGGTGAAGTC**-3' (SEC ID n°: 53)

Las letras en **negrita** representan la secuencia codificante. La secuencia restante es homóloga a los sitios de inserción de pAILo2.

[0395] Transformantes de *E. coli* con las construcciones de expresiones correctas recombinantes pEJG111 (gen GH61C) y pEJG112 (gen GH61 D) se identificaron por análisis de digestión de enzima de restricción y ADN plásmido se preparó usando un QIAGEN BioRobot 9600.

[0396] Dos cebadores oligonucleótidos sintéticos, mostrados debajo, se diseñaron para amplificar por PCR el marco de lectura abierto en toda su longitud de EST pTter61 G de *Thielavia terrestris* que codifica el gen de la familia GH61 G. Un kit In-Fusion Cloning (BD Biosciences, Palo Alto, CA) se usó para clonar el fragmento directamente en pAILo2.

Cebador directo InFusion:

5'- ACTGGATT**ACCATGAAGGGACTTTTCAGTGC**-3' (SEC ID n°: 54)

Cebador inverso InFusion:

5'- TCACCTCTAGTTAATTAAT**TACAAGCACTGCGAGTAGT**-3' (SEC ID n°: 55)

Las letras en **negrita** representan la secuencia codificante. La secuencia restante contiene identidad de secuencia comparado con los sitios de inserción de pAILo2.

[0397] Cinquenta picomoles de cada uno de los cebadores de arriba se usaron en una reacción PCR con 50 ng de ADN pTter61G, tampón de amplificación 1X Pfx, 6 µl de 10 mM de mezcla de dATP, dTTP, dGTP y dCTP, 2,5 unidades de platino Pfx ADN polimerasa, 1 µl de 50 mM de MgSO₄ y 5 µl de solución intensificadora 10X pCRx en un volumen final de 50 µl. Un Eppendorf Mastercycler 5333 se usó para amplificar el fragmento programado para un ciclo a 98°C durante 2 minutos; y 35 ciclos cada a 94°C durante 30 segundos, 65°C durante 30 segundos y 68°C durante 1.5 minutos. Después de los 35 ciclos, la reacción se incubó a 68°C durante 10 minutos y luego se enfrió a 10°C hasta proceso posterior. Un producto de reacción PCR 1 kb se aisló en un 0,8% gel de GTG-agarosa usando tampón TAE y 0,1 µg de bromuro de etidio por ml. La banda de ADN se visualizó con la ayuda de un Dark Reader™ para evitar mutaciones inducidas por los UV. La banda de ADN 1 kb se cortó con una hoja de afeitar desechable y se purificó con un Ultrafree-DA Spin Cup según las instrucciones del fabricante.

[0398] El vector pAILo2 fue linealizado por digestión con Nco I y Pac I. El fragmento se purificó por electroforesis en gel y ultrafiltración como se ha descrito anteriormente. Clonación del fragmento PCR purificado en el vector pAILo2 linealizado y purificado se realizó con un kit InFusion Cloning. La reacción (20 µl) contuvo de tampón InFusion terrestreis, 1X BSA, 1 µl de enzima InFusion (diluida 1:10), 100 ng de pAILo2 digerido con Nco I y Pac I y 50 ng del producto PCR purificado de *Thielavia terrestris* GH61 G. La reacción se incubó a temperatura ambiente durante 30 minutos. Dos microlitros de la reacción se usaron para transformar células de *E. coli* XL10 SoloPac® à Gold (Stratagene, La Jolla, CA) según las instrucciones del fabricante. Después del periodo de recuperación, dos partes alícuotas de 100 µl de la reacción de transformación se colocaron sobre placas 150 mm 2X YT complementadas con 100 µg de ampicilina por ml. Las placas se incubaron durante toda la noche a 37°C. Seis clones putativos recombinantes se seleccionaron de manera aleatoria de las placas de selección y ADN plásmido se obtuvo a partir de cada uno usando un QIAGEN BioRobot 9600. Los clones se analizaron por digestión de enzima de restricción Pst I. Cinco de seis clones tuvieron el patrón de digestión de restricción previsto, dos clones se secuenciaron luego para confirmar que no había ninguna de las mutaciones en el inserto clonado. Clon #1 se seleccionó y se designó pAILo24 (Figura 20).

[0399] El gen Tter61 E truncado se amplificó por PCR con cebadores In-Fusion específicos diseñados para añadir los 5 nucleótidos faltantes en su 5' final como se describe en el ejemplo 11. La banda de gel purificada se clono luego en pAILo2 como se ha descrito anteriormente. Dos microlitros de la reacción TOPO se transformaron en células de *E. coli* TOP10 según las instrucciones del fabricante. Las células transformadas se colocaron sobre placas de agar 2XYT complementadas con 100 µg/ml de ampicilina y se incubaron durante toda la noche a 37°C. Ocho colonias se seleccionaron de manera aleatoria y ADN plásmido se preparó como se ha descrito anteriormente. Los clones se analizaom por digestión de enzima de restricción Sac I. Siete de ocho clones tuvieron el patrón de digestión de restricción previsto, cuatro clones se secuenciaron luego para confirmar que no había ninguna de las mutaciones en el inserto clonado. Clon #5 se seleccionó y se designó pAILo28 (Figura 21).

Ejemplo 14: Expresión de genes de *Thielavia terrestris* que codifican polipéptidos de la Familia gH61B, GH61C, GH61D, GH61E y GH61G con actividad de mejora celulolítica en *Aspergillus oryzae* JaL250

[0400] Protoplastos de *Aspergillus oryzae* JaL250 se prepararon según el método de Christensen *et al.*, 1988, *Bio/Technology* 6: 1419-1422. Cinco µg de pEJG108 (al igual que pAILo2 como vector de control) se usó para transformar *Aspergillus oryzae* JaL250.

[0401] La transformación de *Aspergillus oryzae* JaL2500 con pEJG108 (gen GH61 B) liberó aproximadamente 100 transformantes. Diez transformantes se aislaron en placas PDA individuales.

[0402] Placas PDA confluyentes de todos los transformantes se lavaron con 5 ml de 0,01% de Tween 80 y se inocularon separadamente en 25 ml de medio MDU2BP en frascos de agitación de vidrio 125 ml y se incubaron a 34°C, 200 r.p.m. Cinco días tras la incubación, 5 µl de sobrenadante de cada cultivo se analizó usando 8-16% de geles de SDS-PAGE de tris-glicina (Invitrogen, Carlsbad, CA) según las instrucciones del fabricante. Perfiles de SDS-PAGE de los cultivos mostraron que 9 de los 10 transformantes tuvieron una nueva banda mayor a aproximadamente 42 kDa.

[0403] Una placa confluyente de transformante 10 (crecida en PDA) se lavó con 10 ml de 0,01% de Tween 20 y se inoculó en un Fernbach de 2 litros con 500 ml de medio MDU2BP para generar caldo para caracterización de la enzima. El matraz se recogió en el día 5 y se filtró usando una membrana GP Express plus 0,22 µm GP (Millipore, Bedford, MA).

[0404] Plásmidos pEJG111 (gen GH61 C) y pEJG112 (gen GH61 D) se expresaron en *Aspergillus oryzae* JaL250 usando el mismo protocolo anteriormente descrito.

[0405] Protoplastos de *Aspergillus oryzae* Jal250 (WO 99/61651) se prepararon según el método de Christensen *et al.*, 1988, *supra*. Cinco microgramos de pAILo24 y pAILo28 (así como pAILo2 como un vector de control) se usaron para transformar protoplastos de *Aspergillus oryzae* Jal250.

[0406] Las transformaciones de *Aspergillus oryzae* Jal250 liberaron aproximadamente 50 transformantes por construcción de expresión. Ocho transformantes se aislaron de cada transformación a placas PDA individuales y se incubaron durante cinco días a 34°C.

[0407] Placas de espora confluyentes se lavaron con 3 ml de 0,01% de Tween 80 y la suspensión de esporas se usó para inocular 25 ml de medio MDU2BP en frascos 125 ml de agitación de vidrio. Cultivos transformantes se incubaron a 34°C con agitación constante a 200 r.p.m. El día cinco después de inoculación, los cultivos se centrifugaron a 6000 x g y sus sobrenadantes se recogieron. Cinco micro-litros de cada sobrenadante se mezclaron con un volumen igual de tampón de carga 2X (10% de β-mercaptoetanol) y se cargaron sobre un gel SDS-PAGE 1,5 mm de tris-glicina 8 %-16 % y se tiñeron con Simply Blue SafeStain (Invitrogen, Carlsbad, CA). Los perfiles de SDS-PAGE de los caldos de cultivo mostraron que siete de ocho transformantes pAILo24 tienen un nuevo enlace proteínico de aproximadamente 45 kDa. Esta nueva proteína se ejecuta en los geles de SDS-poliacrilamida como una mancha de grasa más que una banda bien definida, sugiriendo la presencia de formas múltiples quizá debido a modificaciones postraduccionales como glicosilación. Este tipo de modificaciones explicaría la diferencia entre el peso molecular deducido del *T. terrestris*

GH61G, 33,7 kDa, y el peso aparente molecular de la nueva proteína presente en estos transformantes. Clon #3 se seleccionó para estudios posteriores. En el caso de los transformantes pAilo28 6 de ocho transformantes pAilo28 tienen un nuevo enlace proteínico de aproximadamente 25 kDa muy cerca del peso calculado para la proteína madura de 22,5 kDa.

5

Ejemplo 15: Construcción de pMJ09

[0408] Vector pMJ04 se construyó por PCR amplificando el terminador de gen de celobiohidrolasa 1 (*cbh1*) de *Trichoderma reesei* Cel7A de ADN genómico de *Trichoderma reesei* RutC30 usando 993429 (antisentido) y 993428 (sentido) mostrados debajo. El cebador antisentido se creó genéticamente para tener un sitio Pac I en el 5'-final y un sitio Spe en el 5'-final del cebador sentido.

10

Cebador 993429 (antisentido):

5'-AACGTTAATTAAGGAATCGTTTTGTGTTT-3' (SEC ID n°: 56)

15

Cebador 993428 (sentido):

5'-AGTACTAGTAGCTCCGTGGCGAAAGCCTG-3' (SEC ID n°: 57)

[0409] Las reacciones de amplificación (50 µl) se compusieron de tampón de reacción 1X TermoPol (New England Biolabs, Beverly, MA), 0,3 mM de dNTPs, 100 ng de ADN genómico de *Trichoderma reesei* RutC30 (que fue aislado usando un kit DNeasy Plant Maxi, QIAGEN Inc., Valencia, CA), 0,3 µM de cebador 993429, 0,3 µM de cebador 993428 y 2 unidades de polimerasa Vent (New England BioLabs, Beverly, MA). Las reacciones se incubaron en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 30 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 55°C y 30 segundos a 72°C (15 minutos extensión final).

20

[0410] Los productos reactivos se aislaron en un gel de agarosa 1,0% usando tampón TAE donde una banda de producto de 229 pares de bases se cortó del gel y se purificó usando un kit QIAGEN QIAquick Gel Extraction según las instrucciones del fabricante.

25

[0411] El fragmento de PCR resultante se digirió con Pac I y Spe I y se ligó en pAilo1 digerido con las mismas enzimas de restricción que usan un kit Rapid Ligation (Roche, Indianapolis, IN), para generar pMJ04 (Figura 22).

30

[0412] Vector pMJ06 se construyó por PCR amplificando el promotor de gen de celobiohidrolasa 1 de *Trichoderma reesei* Cel7A de ADN genómico de *Trichoderma reesei* RutC30 usando los cebadores 993696 (antisentido) y 993695 (sentido) mostrados debajo. El cebador antisentido se creó genéticamente para tener un sitio Sal I en el 5'-final del cebador sentido y un sitio Nco I en el 5'-final del cebador antisentido.

35

Cebador 993695 (sentido):

5'-ACTAGTCGACCGAATGTAGGATTGTT-3' (SEC ID n°: 58)

40

Cebador 993696 (antisentido):

5'-TGACCATGGTGCGCAGTCC-3' (SEC ID n°: 59)

[0413] Las reacciones de amplificación (50 µl) se compusieron de tampón de reacción 1X TermoPol, 0,3 mM de dNTPs, 100 ng de ADN genómico de *Trichoderma reesei* RutC30 (que fue preparado usando un QIAGEN DNeasy Plant Maxi Kit), 0,3 µM de cebador 993696, 0,3 µM de cebador 993695 y 2 unidades de polimerasa Vent. Las reacciones se incubaron en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 30 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 55°C y 60 segundos a 72°C (15 minutos extensión final).

45

[0414] Los productos reactivos se aislaron en un gel de agarosa 1,0% usando tampón TAE donde una banda de producto de 988 pares de bases se cortó del gel y se purificó usando un kit QIAGEN QIAquick Gel Extraction según las instrucciones del fabricante.

50

[0415] El fragmento de PCR resultante se digirió con Nco I y sal I y se ligó en pMJ04 digerido con las mismas enzimas de restricción, usando un kit Rapid Ligation, para generar pMJ06 (Figura 23).

55

[0416] Vector de expresión pMJ09 se construyó por PCR amplificando el terminador del gen de celobiohidrolasa (*cbh1*) de *Trichoderma reesei* Cel7A de ADN genómico de *Trichoderma reesei* RutC30 usando los cebadores 993843 (antisentido) y 993844 (sentido) mostrados debajo. El cebador antisentido se creó genéticamente para tener un sitio Pac I y un Spe I en el 5'-final y un sitio Pvu I en el 5'-final del cebador de sentido.

60

Cebador 993844 (sentido):

5'-CGATCGTCTCCCTATGGGTCATTACC-3' (SEC ID n°: 60)

Cebador 993843 (antisentido):

5'-ACTAGTTAATTAAGCTCCGTGGCGAAAG-3' (SEC ID n°: 61)

[0417] Las reacciones de amplificación (50 µl) se compusieron de tampón de reacción 1 X TermoPol, 0,3 mM de dNTPs, 100 ng de ADN genómico de *Trichoderma reesei* RutC30 (que se extrajo usando un kit QIAGEN DNeasy Plant Maxi)

65

0,3 µM de cebador 993844, 0,3 µM de cebador 993843 y 2 unidades de polimerasa Vent. Las reacciones se incubaron en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 30 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 55°C y 60 segundos a 72°C (15 minutos extensión final).

5 [0418] Los productos reactivos se aislaron en un gel de agarosa 1,0% usando tampón TAE donde una banda de producto de 473 pares de bases se cortó del gel y se purificó usando un kit QIAGEN QIAquick Gel Extraction según las instrucciones del fabricante.

10 [0419] El fragmento de PCR resultante se digirió con Pvu I y Spe I y se ligó en pMJ06 digerido con Pac I y Spe I usando un kit de Rapid Ligation para generar pMJ09 (Figura 24).

Ejemplo 16: Construcción de vector de expresión pSMaI155

15 [0420] Basado en el gen htp de *E. coli* encontrado en el plásmido pPHTI (Cummings *et al.*, 1999, *Current Genetics* 36: 371-82) los siguientes cebadores se diseñaron para amplificación:

Directo (993863):

5'-GGGttcgaaTTCATTTAAACGGCT-3' (SEC ID n°: 62) Bst BI

Inverso (993864):

20 5'-GGGagcgctCAATATTCATCTCTC-3' (SEC ID n°: 63) Eco 47III

25 [0421] El sistema de ADN polimerasa Vent® (New England BioLabs, Inc., Beverly, MA) se utilizó para amplificar el gen de interés bajo las siguientes condiciones de termociclado gradientes que usan un Robocycler Gradient 40 (Stratagene, La Jolla, CA): un ciclo a 95°C durante 2 minutos; 25 ciclos cada uno a 95°C durante 1 minuto, 51 °C - 65°C durante 1 minuto, y 72°C durante 2 minutos; y 1 ciclo a 72°C durante 7 minutos. El fragmento 1,8 kb resultante se amplificó exitosamente a temperatura de reconocimiento 51 °C y gel purificado usando un kit QIAquick Gel Extraction. El fragmento de gel purificado se clonó luego en pCR- BluntII-TOPO (Invitrogen, Carlsbad, CA) creando pEmY10 (Figura 25).

30 [0422] Plásmido pMJ09 se digirió con Nsi I para eliminar el gen marcador amdS. El digerido se fraccionó en un gel de agarosa 0,7% usando tampón TAE y una banda cortada 3,9 kb y purificada usando un kit QIAquick Gel Extraction según el protocolo sugerido del fabricante. El vector fue defosforilado usando fosfatasa alcalina de camarón (Roche Diagnostics Corporation, Indianapolis, IN) según el protocolo del fabricante. El gen de resistencia a higromicina se retiró del plásmido pEmY10 por digestión con Nsi I y un fragmento 2,0 kb fragmento como arriba. Este fragmento se ligó con el fragmento 3,9 kb de pMJ09. La mezcla de ligamiento se transformó en células SURE de *E. coli* (Stratagene, La Jolla, CA), y ocho colonias resultantes se seleccionaron para ligamiento correcto por análisis de restricción y secuenciación del ADN. Uno de los clones verificados se designó pSMaI149. El sitio Nco I en el gen de resistencia a higromicina fue quitado de pSMaI149 usando el kit Quickchange Site-Directed Mutagenesis (Stratagene, La Jolla, CA) según el protocolo del fabricante. Los cebadores de mutagénesis/amplificación tuvieron las siguientes secuencias, donde la G subrayado representa el sitio de mutación.

5'-GGTCGCGGAGGCGATGGATGCGATCG-3' (SEC ID n°: 64)

5'-CGATCGCATCCATCGCCTCCGCGACC-3' (SEC ID n°: 65)

45 [0423] Ocho colonias seleccionadas de forma aleatoria se cribaron por digestión con Nco I, y una de estas, se designó clon 7, que mostró pérdida del sitio Nco I fue además verificado por secuenciación de ADN y designado pSMaI155 (Figura 26).

50 **Ejemplo 17: Expresión de genes de *Thielavia terrestris* que codifican polipéptidos de la familia GH61 B, GH61C, GH61 D, GH61 E y GH61 G con actividad de mejora celulolítica en *Trichoderma reesei***

[0424] Dos cebadores oligonucleótidos sintéticos mostrados debajo se diseñaron para amplificar por PCR un gen de *Thielavia terrestris* que codifica una familia putativa GH61 B del clon genómico. Un kit de InFusion Cloning se usó para clonar el fragmento directamente en el vector de expresión, pSMaI155, sin necesidad de restringir digeridos y ligamiento.

Cebador directo:

5'-cgcgactgcgacc**ATGAAGTCGTTCA**CCATTG-3' (SEC ID n°: 66)

Cebador inverso:

60 5'-tcgccacggagctta**GAGGCACTGCGAGTAG**-3' (SEC ID n°: 67)

Las letras en negrita representan la secuencia codificante. La secuencia restante es homóloga a los sitios de inserción del clon 7 pSMaI155.

65 [0425] Cincuenta picomoles de cada uno de los cebadores de arriba se usaron en una reacción PCR con 10 ng de ADN 15 de clon genómico de *Thielavia terrestris*, tampón de amplificación 1X Pfx, 1,5 µl de 10 mM de mezcla de dATP,

dTTP, dGTP y dCTP, 2,5 unidades de platino Pfx ADN polimerasa, 1 µl de 50 mM de MgSO₄ y 5 µl de solución intensificadora 10X pCRx en un volumen final de 50 µl. Las condiciones de amplificación fueron un ciclo a 94°C durante 2 minutos; y 30 ciclos cada uno a 94°C durante 15 segundos, 55°C durante 30 segundos y 68°C durante 3 minutos. El bloque de calor luego fue a un 4°C ciclo de remojo.

5

[0426] Los productos reactivos se aislaron en un del de agarosa 1,0% usando tampón TAE donde una banda de producto 3 kb se cortó del gel y se purificó usando un kit QIAquick Gel Extraction según las instrucciones del fabricante.

10

[0427] El fragmento se clonó luego clonado en el vector de expresión pSMA1155 usando un kit InFusion Cloning. El vector se digirió con endonucleasas de restricción Nco I y Pac I. El fragmento se purificó por electroforesis en gel y gel de purificación QIAquick. El fragmento de gen y el vector digerido se unieron en una reacción resultante en el plásmido de expresión pEJG110 (Figura 27) en el que transcripción del gen de la familia GH61 B fue bajo el control del promotor y terminador CBHI de *Trichoderma reesei*. La reacción de ligamiento (20 µl) se compuso por tampón de infusión 1X, 1X BSA (BD Biosciences, Palo Alto, CA), 1 µl de enzima InFusion (diluida 1:10), 100 ng de pSma155 digerido con Nco I y Pac I, y 100 ng del producto PCR purificado de *Thielavia terrestris* GH61 B. La reacción se incubó a temperatura ambiente durante 30 minutos. Un µl de la reacción se usó para transformar células Solopac Gold de *E. coli* XL10 (Stratagene, La Jolla, CA). Un transformante de *E. coli* con el plásmido pEJG110 se detectó por digestión de enzima de restricción y ADN plásmido se preparó usando un QIAGEN BioRobot 9600.

15

20

[0428] Secuenciación del ADN de pEJG110 se realizó con un Applied Biosystems Model 3700 Automated DNA Sequencer usando terminador version 3.0 Big-Dye™ (Applied Biosystems, Inc., Foster City, CA) y estrategia de "walking primer". Datos de secuencia de nucleótidos se escrutinaron para calidad y todas las secuencias se compararon entre sí con asistencia de software PHRED/PHRAP (University of Washington, Seattle, WA).

25

[0429] Para la expresión en *T. reesei* el gen Tter61 E se amplificó por PCR de su construcción de expresión pAILo28 con los siguientes cebadores.

Cebador directo:

5'-GCCCATGGACCATGCTCGCAAAC-3' (SEC ID n°: 68)

30

Cebador inverso:

5'-CCTCTAGTTAATTAATCAGCAGC-3' (SEC ID n°: 69)

Asimismo, el gen Tter61 G se amplificó por PCR de su construcción de expresión pAILo24 con los siguientes cebadores.

35

Cebador directo:

5'-GCCCATGGACCATGAAGGGACTT-3' (SEC ID n°: 70)

Cebador inverso:

5'-CCTCTAGTTAATTAATTACAAGC-3' (SEC ID n°: 71)

40

Ambos experimentos se llevaron a cabo de la siguiente manera:

45

[0430] Cincuenta picomoles de cada uno de los cebadores de arriba se usaron en una reacción PCR con 50 ng de ADN pAILo28, 1 µl de 10 mM de mezcla de dATP, dTTP, dGTP y dCTP, 5 µl de tampón de ADN polimerasa I 10X AmpliTaq® (Perkin-Elmer/Applied Biosystems, Inc., Foster City, CA) y 5 unidades de ADN polimerasa AmpliTaq® (Perkin-Elmer/Applied Biosystems, Inc., Foster City, CA), en un volumen final de 50 µl. Un Eppendorf Mastercycler 5333 se usó para amplificar el fragmento de ADN y se programó para un ciclo a 95°C durante 3 minutos; y 25 ciclos cada uno a 95°C durante 45 segundos, 55°C durante 60 segundos y 72°C durante 1 minuto 30 segundos. Después de los 25 ciclos, la reacción se incubó a 72°C durante 10 minutos y luego se enfrió a 4°C hasta proceso posterior. Un producto de reacción PCR de aproximadamente 681 bases se aisló en un gel de GTG-agarosa 0,8% usando tampón TAE y 0,1 µg de bromuro de etidio por ml. La banda de ADN se visualizó con la ayuda de un Dark Reader™ para evitar mutaciones inducidas por los UV. La banda ADN apropiada se cortó con una hoja de afeitar desechable y se purificó con un kit QIAquick PCR Purification según las instrucciones del fabricante. La banda ADN purificada se clonó en el vector pCR2.1-TOPO según las instrucciones del fabricante (Invitrogen, Carlsbad, CA). Dos microlitros de la TOPO-reacción se transformaron en células de *E. coli* TOP10 según las instrucciones del fabricante. Las células transformadas se colocaron sobre placas de agar 2X YT complementadas con 100 µg/ml de ampicilina y se incubaron durante toda la noche a 37°C. Ocho colonias se seleccionaron de manera aleatoria para preparación de ADN plásmido. Cada colonia creció durante toda la noche en 3 ml de medio LB complementado con 100 µg de ampicilina por ml y se usó para preparar ADN plásmido con un QIAGEN BioRobot 9600. El ADN plásmido se analizó por mapeo de restricción para identificar clones positivos para inserción GH61 E o GH61 G usando endonucleasas de restricción Pac I y Nco I. Una vez que un clon se validó que hubo inserción exitosa del gen GH61 E o GH61 G, el clon se secuenció para fidelidad usando BigDye Terminator Version 3 y se analizó usando ABI 3700 DNA Analyzer (Foster City, CA) según las instrucciones del fabricante.

60

[0431] El clon de ADN de *E. coli* TOPO con la secuencia GH61 E o GH61 G confirmada se digirió con Pac I y Nco I y el producto de reacción se resolvió luego en un gel de agarosa 0,8%. La banda ADN apropiada se cortó con una hoja de afeitar desechable y se purificó con un kit QIAquick PCR Purification según las instrucciones del fabricante.

5 [0432] Plásmido pSMai155 se digirió de la misma manera con Pac I y Nco I para crear extremos compatibles con los fragmentos GH61. El producto de digestión se resolvió en un gel de agarosa 0,8% y el plásmido lineal se cortó y purificó usando kit QIAquick PCR Purification según las instrucciones del fabricante.

10 [0433] El fragmento de gen Pac I/Nco I GH61 E o GH61 G se ligó en el plásmido pSMai155 digerido PacI/NcoI usando el kit Rapid DNA Ligation (Roche, Indianapolis, IN) siguiendo las instrucciones del fabricante. Este ligamiento se usó luego para transformar células SURE de *E. coli* siguiendo las instrucciones del fabricante. Colonias se seleccionaron, cultivaron y plásmido se preparó como se ha descrito anteriormente. El ADN plásmido se analizó por mapeo de restricción para identificar clones positivos para inserción GH61 E o GH61 G usando Pac I y Nco I. Uno de los clones que tuvo el patrón de restricción correcto para el gen GH61 E se designó pCW095 y uno que tuvo el patrón correcto para el gen GH61 G se designó pCW096.

15 [0434] Un vector de expresión, pCW076, se creó para utilizar el promotor CBHII de *Trichoderma reesei* para expresión genética. Vector de expresión pSMai155 se digirió con endonucleasas de restricción Sal I y Nco I para eliminar el fragmento promotor CBHI. Construcción de expresión pEJG114 (Figura 28) se digirió también con Sall y NcoI para aislar el promotor CBHII para ligamiento en el plásmido pSMai155 digerido Sal I/Nco I.

20 [0435] El promotor Sal I/Nco I CBHII se ligó en el Sall/NcoI pSMai155 digerido usando el kit Rapid DNA Ligation siguiendo las instrucciones del fabricante. Este ligamiento se usó luego para transformar células SURE de *E. coli* siguiendo las instrucciones del fabricante. Colonias se seleccionaron, clonaron y plásmido se preparó como se ha descrito anteriormente. El ADN plásmido se analizó por digestión de enzima de restricción para identificar clones positivos para inserción de promotor CBHII usando Sall y Nco I. Estos clones se secuenciaron luego usando cebador TrCBH2PrSeqF1, mostrado debajo, para confirmar la presencia del promotor CBHII. Uno de los clones que tuvo la secuencia correcta se designó pCW076 (Figura 29).

25
30 TrCBH2PrSeqF1:
5'-GGATGAAGCTCATTAGCCG-3' (SEC ID n°: 72)

[0436] Para la expresión en *Trichoderma reesei*, el gen Tter61 D gen se aisló de pTter61 D (como se describe en el ejemplo 7 usando endonucleasas de restricción Pac I y Nco I). Plásmido pCW076 se digirió de la misma manera con Pac I y Nco I para crear extremos compatibles con el fragmento GH61 D. Los productos de digestión se resolvieron en gel de agarosa 0,8% usando tampón TAE y los fragmentos lineales se cortaron y purificaron usando un kit QIAquick PCR Purification según las instrucciones del fabricante.

35 [0437] El fragmento del gen GH61D Pac I/Nco I se ligó en Pac I/Nco I digerido pCW076 usando el kit Rapid DNA Ligation siguiendo las instrucciones del fabricante. Un microlitro de la reacción de ligamiento se usó luego para transformar células SURE de *E. coli* siguiendo las instrucciones del fabricante. Colonias se seleccionaron, cultivaron y plásmido se preparó como se ha descrito anteriormente. El ADN plásmido se analizó por digestión de enzima de restricción para identificar clones positivos para inserción GH61 D usando Pac I y Nco I. Uno de los clones que tuvo el patrón de restricción correcto se designó pCW078 (Figura 30).

40 [0438] Protoplastos de *Trichoderma reesei* RutC30 se prepararon según el método de Christensen *et al.*, 1988, *supra*. Cinco µg de pEJG110, pCW095, pCW096, o pCW078 se usaron para transformar *Trichoderma reesei* RutC30. Cada transformación individual liberó aproximadamente 100 transformantes. Sesenta transformantes se aislaron, de cada transformación, a placas de uridina Cove/10 mM individuales y se incubaron a 28°C durante siete días.

45 [0439] Cepas transformantes se analizaron en lotes de ocho a la vez. Esporas se recogieron de ocho placas individuales con un asa estéril, se inocularon separadamente en 25 ml de medio CIM en frascos de agitación de vidrio 125 ml y se incubaron a 28°C, 250 r.p.m. Cinco días después de inoculación, 0,5 µl de sobrenadante de cada cultivo se analizó usando geles Tris SDS-PAGE 7,5% (BioRad, Hercules, CA) según las instrucciones del fabricante. Perfiles de SDS-PAGE de los cultivos mostraron que de 20% a 50% de los transformantes produjeron un nuevo enlace proteínico correspondiente al gen recién introducido. Un transformante apropiado se seleccionó de cada construcción de expresión para estudios posteriores.

50 **Ejemplo 18: Identificación de una familia de una beta-glucosidasa del gen de la familia GH3A en la secuencia genómica de *Aspergillus fumigatus***

[0440] Una búsqueda BLAST (Altschul *et al.*, 1997, *Nucleic Acids Res.* 25: 3389-3402) de la secuencia del genoma parcial de *Aspergillus fumigatus* (The Institute for Genomic Research, Rockville, MD) se llevó a cabo usando como consulta una secuencia de proteína de beta-glucosidasa de *Aspergillus aculeatus* (N° de acceso P48825). Diferentes genes se identificaron como homólogos de la familia putativa GH3A basados en un alto grado de similitud a la secuencia

de consulta en el nivel de aminoácido. Una región genómica de aproximadamente 3000 pares de bases con más de 70% de identidad a la secuencia de consulta en el nivel de aminoácido se eligió para estudio posterior.

Ejemplo 19: Extracción de ADN genómico de *Aspergillus fumigatus*

[0441] *Aspergillus fumigatus* creció en 250 ml de medio de dextrosa de patata en un matraz de agitación disipado a 37°C y 240 r.p.m. Micelios se recogieron por filtración, se lavaron dos veces en TE (10 mM de Tris-1 mM EDTA) y se congelaron bajo nitrógeno líquido. Micelios congelados se molieron, por mortero y punzón, a un polvo fino, que se resuspendió en pH 8,0 tampón con 10 mM de Tris, 100 mM de EDTA, 1% Tritón X-100, 0,5 M de guanidina-HCl y 200 mM de NaCl. Ribonucleasa pancreática sin ADNsa se añadió a una concentración de 20 mg/litro y el lisado se incubó a 37°C durante 30 minutos. Detrito celular se retiró por centrifugado y ADN se usó usando una columna Maxi 500 (QIAGEN Inc., Valencia, CA). Las columnas se equilibraron en 10 ml de QBT se lavaron con 30 ml de QC y se eluyeron con 15 ml de QF (todos los tampones de QIAGEN Inc., Valencia, CA). ADN se precipitó en isopropanol, se lavo en etanol 70% y se recuperó por centrifugado. El ADN se resuspendió en tampón TE.

Ejemplo 20: Clonación del gen de beta-glucosidasa de la familia GH3A y construcción de un vector de expresión de *Aspergillus oryzae*

[0442] Dos cebadores oligonucleótidos sintéticos mostrados debajo se diseñaron para amplificar por PCR el gen de *Aspergillus fumigatus* que codifica la beta-glucosidasa de la familia GH3A del ADN genómico preparado en el ejemplo 19. Un kit InFusion Cloning se usó para clonar el fragmento directamente en el vector de expresión, pAILo2, sin necesidad de restricción de digeridos y ligamiento.

Cebador directo:

5'-ACTGGATTTACCATGAGATTCGGTTGGCTCG-3' (SEC ID nº: 73)

Cebador inverso:

5'-AGTCACCTCTAGTTACTAGTAGACACGGGGC-3' (SEC ID nº: 74)

Las letras en negrita representan la secuencia codificante. La secuencia restante es homóloga a los sitios de inserción de pAILo2.

[0443] Cincuenta picomoles de cada uno de los cebadores de arriba se usaron en una reacción PCR con 100 ng de ADN genómico de *Aspergillus fumigatus*, tampón de amplificación 1X Pfx, 1,5 µl de 10 mM de mezcla de dATP, dTTP, dGTP y dCTP, 2,5 unidades de platino Pfx ADN polimerasa, 1 µl de 50 mM de MgSO₄ y 2,5 µl de solución intensificadora 10X pCRx en un volumen final de 50 µl. Las condiciones de amplificación fueron un ciclo a 94°C durante 2 minutos; y 30 ciclos cada uno a 94°C durante 15 segundos, 55°C durante 30 segundos y 68°C durante 3 minutos. El bloque de calor luego fue a un 4°C ciclo de remojo.

[0444] Los productos reactivos se aislaron en un gel de agarosa 1,0% usando tampón TAE donde una banda de productor 3 kb se cortó del gel y se purificó usando un kit QIAquick Gel Extraction según las instrucciones del fabricante.

[0445] El fragmento se clonó luego en el vector de expresión pAILo2 usando un kit InFusion Cloning. El vector se digirió con endonucleasas de restricción Nco I y Pac I (usando condiciones especificadas por el fabricante). El fragmento se purificó por electroforesis en gel y purificación de gel QIAquick. El fragmento de gen y el vector digerido se ligaron juntos en una reacción resultante en el plásmido de expresión pEJG97 (Figura 31) en la transcripción del gen de beta-glucosidasa de la familia GH3A fue bajo el control del promotor NA2-tpi. La reacción de ligamiento (20 µl) se compuso por tampón InFusion 1X, 1X BSA (BD Biosciences, Palo Alto, CA), 1 µl de enzima InFusion (diluida 1:10), 150 ng de pAILo2 digerido con Nco I y Pac I, y 50 ng de la beta-glucosidasa de *Aspergillus fumigatus* purificada producto PCR. La reacción se incubó a temperatura ambiente durante 30 minutos. Un µl de la reacción se usó para transformar células XL10 Solopac Gold de *E. coli* (Stratagene, La Jolla, CA). Un transformante de *E. coli* con el plásmido pEJG97 se detectó por digestión de enzima de restricción y ADN plásmido se preparó usando un QIAGEN BioRobot 9600.

Ejemplo 21: Caracterización de la secuencia genómica de *Aspergillus fumigatus* que codifica una beta-glucosidasa de la familia GH3A

[0446] Secuenciación del ADN del gen de beta-glucosidasa de *Aspergillus fumigatus* de pEJG97 se realizó con un Perkin-Elmer Applied Biosystems Model 377 XL Automated DNA Sequencer (Perkin-Elmer/Applied Biosystems, Inc., Foster City, CA) usando química de terminador Dye (Giesecke et al., 1992, Journal of Virology Methods 38: 47-60) y estrategia "primer walking". Datos de secuencia de nucleótidos se escrutinaron para calidad y todas las secuencias se compararon entre sí con asistencia de software PHRED/PHRAP (University of Washington, Seattle, WA).

[0447] Un modelo de gen para la secuencia de *Aspergillus fumigatus* se construyó basado en similitud a genes homólogos de *Aspergillus aculeatus*, *Aspergillus niger* y *Aspergillus kawachii*. La secuencia de nucleótidos (SEC ID nº: 75) y secuencia de aminoácidos deducida (SEC ID nº: 76) se muestran en las Figuras 32A y 32B. El fragmento genómico codifica un polipéptido de 863 aminoácidos, interrumpido por 8 intrones de 62, 55, 58, 63, 58, 58, 63 y 51 pares de bases. El %G+C contenido del gen es 54,3%. Usando el programa de software SignalP (Nielsen et al., 1997,

supra), un péptido señal de 19 residuos se predijo. La proteína predicha madura contiene 844 aminoácidos con una masa molecular de 91,7 kDa.

5 [0448] Un alineamiento comparativo de secuencias de beta-glucosidasa se determinó usando el método de Clustal W (Higgins, 1989, *CABIOS* 5: 151-153) usando el software LASERGENE™ MEGALIGN™ (DNASTAR, Inc., Madison, WI) con una tabla de identidad y los siguientes parámetros de alineamiento múltiples: penalización del espacio de 10 y penalización de longitud del espacio de 10. Parámetros de alineamiento de pareja fueron Ktuple=1, penalización de espacio espacio=3, ventanas=5 y diagonales=5. El alineamiento mostró que la secuencia de aminoácidos deducida del gen de beta-glucosidasa de *Aspergillus fumigatus* gen compartió 78%, 76% y 76% de identidad a las secuencias de aminoácidos deducidas de las beta-glucosidasas de *Aspergillus aculeatus* (número de acceso P48825), *Aspergillus niger* (000089) y *Aspergillus kawachii* (P87076).

Ejemplo 22: Expresión del gen de beta-glucosidasa de la familia GH3A de *Aspergillus fumigatus* en *Aspergillus oryzae* JaL250

15 [0449] Protoplastos de *Aspergillus oryzae* JaL250 se prepararon según el método de Christensen *et al.*, 1988, *supra*. Cinco µg de pEJG97 (al igual que pAlLo2 como un vector de control) se usaron para transformar *Aspergillus oryzae* JaL250.

20 [0450] La transformación de *Aspergillus oryzae* JaL250 con pEJG97 liberó aproximadamente 100 transformantes. Diez transformantes se aislaron a placas PDA individuales.

25 [0451] Placas confluyentes PDA de cinco de los diez transformantes se lavaron con 5 ml de Tween 80 0,01% y se inocularon separadamente en 25 ml de medio MDU2BP en frascos de agitación de vidrio 125 ml y se incubaron a 34°C, 200 r.p.m. Cinco días tras la incubación, 0,5 µl de sobrenadante de cada cultivo se analizó usando 8-16% de geles Tris-Glicina SDS-PAGE (Invitrogen, Carlsbad, CA) según las instrucciones del fabricante. Perfiles de SDS-PAGE de los cultivos mostraron que uno de los transformantes (designado transformante 1) tuvo una banda principal de aproximadamente 130 kDa.

30 [0452] Una placa confluyente de transformante 1 (crecida en PDA) se lavó con 10 ml de Tween 20 0,01 % y se inoculó en un Fernbach 2 litros con 400 ml de medio MDU2BP para generar caldo para caracterización de la enzima. El matraz se recogió el día 5 y se filtró usando una membrana 0,22 µm GP Express plus (Millipore, Bedford, MA).

35 **Ejemplo 23: Expresión del gen de beta-glucosidasa de la familia GH3A de *Aspergillus fumigatus* en gen de beta-glucosidasa de *Trichoderma reesei***

[0453] Dos cebadores oligonucleótidos sintéticos mostrados debajo se diseñaron para amplificar por PCR el gen de beta-glucosidasa de *Aspergillus fumigatus* de pEJG97 descrito en el ejemplo 20. Un kit InFusion Cloning se usó para clonar el fragmento directamente en el vector de expresión, pMJ09, sin necesidad de restringir digeridos y ligamiento.

40 Cebador directo:
5'-GGACTGCGCACCATGAGATTCGGTTGGCTC-3' (SEC ID n°: 77)
Cebador inverso:
5'-TCGCCACGGAGCTTACTAGTAGACACGGGG-3' (SEC ID n°: 78)

45 Las letras en negrita representan la secuencia codificante. La secuencia restante es homóloga a los sitios de inserción de pMJ09.

50 [0454] Cinquenta picomoles de cada uno de los cebadores de arriba se usaron en una reacción PCR (50 µl) con 100 ng de pEJG97 ADN, tampón de amplificación 1X Pfx, 1,5 µl de 10 mM de mezcla de dATP, dTTP, dGTP y dCTP, 2,5 unidades de PlatinumPfx ADN polimerasa, 1 µl de 50 mM de MgSO4 y 2,5 µl de solución intensificadora 10X pCRx. Las condiciones de amplificación fueron un ciclo a 94°C durante 2 minutos; y 30 ciclos cada uno a 94°C durante 15 segundos, 55°C durante 30 segundos y 68°C durante 3 minutos. El bloque de calor luego fue a un 4°C ciclo de remojo.

55 [0455] Los productos reactivos se aislaron en un gel de agarosa 1,0% usando tampón TAE donde una banda de producto 3 kb se cortó del gel y se purificó usando un kit QIAGEN QIAquick Gel Extraction según las instrucciones del fabricante.

60 [0456] El fragmento 3 kb purificado se clonó luego en pMJ09 usando un kit InFusion Cloning. El vector se digirió con Nco I y Pac I. El fragmento se purificó por electroforesis en gel y purificación de gel QIAquick, como se ha descrito anteriormente. El fragmento de gen y el vector digerido se ligaron juntos en una reacción resultante en el plásmido de expresión pEJG107 (Figura 33) en la que transcripción del gen de beta-glucosidasa de la familia GH3A fue bajo el control del promotor cbhl de *Trichoderma reesei*. El ligamiento (50 µl) se compuso por tampón InFusion 1X, 1X BSA, 1 µl de enzima InFusion (diluida 1:10), 100 ng de pMJ09 digerido con Nco I y Pac I, y 100 ng del producto PCR purificado de la beta-glucosidasa de *Aspergillus fumigatus*. La reacción se incubó a temperatura ambiente durante 30 minutos. Un µl de la reacción se usó para transformar células DURE electrocompetentes de *E. coli* (Stratagene, La Jolla, CA). Un

transformante de *E.coli* con el plásmido pEJG107 se detectó por digestión de enzima de restricción y ADN plásmido se preparó usando un QIAGEN BioRobot 9600.

5 [0457] Secuenciación del ADN del gen de beta-glucosidasa de *Aspergillus fumigatus* de pEJG107 se realizó con un PPerkin-Elmer Applied Biosystems Model 377 XL Automated DNA Sequencer (Perkin-Elmer/Applied Biosystems, Inc., Foster City, CA) usando química de terminador Dye (Giesecke *et al.*, 1992, *Journal of Virology Methods* 38: 47-60) y estrategia "primer walking". Datos de secuencia de nucleótidos se escrutinaron para calidad y todas las secuencias se compararon entre sí con asistencia de software PHRED/PHRAP (University of Washington, Seattle, WA).

10 [0458] Protoplastos de *Trichoderma reesei* RutC30 se prepararon según el método de Christensen *et al.*, 1988, *supra*. Cinco µg de pEJG107 se usaron para transformar *Trichoderma reesei* RutC30. La transformación liberó aproximadamente 100 transformantes. Sesenta transformantes se aislaron en placas de uridina Cove/10 mM y se incubaron a 28°C.

15 [0459] Esporas se recogieron de placas de 8 de los 60 transformantes deslizando tres veces con un asa estéril y se inocularon separadamente en 25 ml de medio CIM en frascos de agitación de vidrio 125 ml y se incubaron a 28°C, 250 r.p.m. Cinco días tras la incubación, 0,5 µl de sobrenadante de cada cultivo se analizó usando geles Tris SDS-PAGE 7,5% (Biorad, Hercules, CA) según las instrucciones del fabricante. Perfiles de SDS-PAGE de los cultivos mostraron que uno de los transformantes (designado transformante 1) tuvo una nueva banda principal de aproximadamente 130 kDa.

20 **Ejemplo 24: Purificación y caracterización de los polipéptidos de *Thielavia terrestris* con actividad de mejora celulolítica**

25 [0460] Cultivos en frascos de agitación de *Thielavia terrestris* NRRL 8126 crecieron en 100 ml de medio NNCYPmod en un matraz de agitación disipado 500 ml a 44°C, 175 r.p.m. Los matraces se inocularon con un tapón (aproximadamente 2 cm cuadrado) de una placa de agar fresca del mismo medio y se les permitió crecer durante aproximadamente 3-5 días antes de cosecha. Muestras de caldo crudas se obtuvieron por cultivos de centrifugado durante aproximadamente 10 minutos a 9500 x g para eliminar materia celular y cualquier fuente de carbono residual. El sobrenadante resultante se centrifugó luego otra vez como se ha descrito anteriormente. El sobrenadante se usó como la materia prima para análisis bioquímico y cuando no estuvo en uso se almacenó a 4°C.

30 [0461] Cultivos fermentadores de *Thielavia terrestris* crecieron por cultivo de *Thielavia terrestris* en medio NNCYP con 52 g de celulosa por litro, que se dosificó durante fermentación (sin fuente de carbono adicional) a 42°C y se mantuvieron a pH 5,0. A la fermentación se le permitió ejecutarse hasta que la celulosa dosificada se había agotado (típicamente aproximadamente 40 horas) en este tiempo el caldo se recogió y centrifugó para eliminar micelios como se ha descrito anteriormente.

35 [0462] Antes de realizar experimentos de purificación de proteína, preparaciones a pequeña escala de caldo de *Thielavia terrestris* se prepararon por concentración de muestras de sobrenadante aclaradas usando Centricon Plus 20 (Millipore, Bedford, MA) filtrando los dispositivos en un centrifugador rotor de cubo oscilante (Sorvall, RC3B Plus). Aproximadamente 3 ml de cada concentrado se cargó sobre una columna de desalación 10DG Econo PAC (BioRad, Hercules, CA) equilibrada con 50 mM de acetato sódico pH 5,0, y muestras se eluyeron usando 4 ml de 50 mM de acetato sódico pH 5,0. Para preparaciones a gran escala de caldo de *Thielavia terrestris* (aproximadamente 0,5 litro a 10 litros), el protocolo descrito abajo se utilizó. Muestras de caldo crudas se aclararon de detrito celular por centrifugado de cultivos durante aproximadamente 20 minutos a 9500 x g. Muestras de caldo esclarecidas se filtraron luego (membrana GP Express, polietersulfona, 0,22 µm, Millipore, Bedford, MA), tampón cambiado con 50 mM de acetato de sodio pH 5,0 (Pall Filtron, North Borough, MA, 10 kDa de membrana de polietersulfona, aproximadamente 10-20 psi), y concentrado usando un dispositivo de ultrafiltración Amicon (Millipore, Bedford, MA, 10 kDa de membrana, 40 psi, 4°C). Concentraciones de proteína se determinaron usando un kit BCA Protein Assay (Pierce, Rockford, IL) en el que albúmina de suero bovino se usó como estándar de proteína. Partes alícuotas del procedimiento de desalación se examinaron típicamente en geles SDS-PAGE 8-16% G (Invitrogen, Carlsbad, CA; 200 V durante 1 hora) en los que estándares de peso moleculares de precisión (BioRad, Hercules, CA) se incluyeron. Geles se tiñeron para proteína usando Biosafe Coomassie Stain (BioRad, Hercules, CA) y se destiñeron usando H₂O desionizada.

40 [0463] Los sobrenadantes de *Thielavia terrestris* se cargaron sobre una columna Q Sepharose Big Bead (Amersham Pharmacia, Uppsala Sweden) equilibrada con 20 mM de Tris-HCl pH 8,2. Flujo a través de material se recogió y se almacenó a 4°C hasta uso. Antes de elución, la columna se lavó con cinco volúmenes de columna de tampón de inicio. Material unido se eluyó con un gradiente lineal de 0 a 0,40 M de NaCl (15 volúmenes de columna; 10 ml fracciones) en 20 mM de tampón Tris-HCl pH 8,2. Basado en el perfil UV (A280nm) fracciones individuales representativas de picos de proteína resueltos se agruparon para caracterización y tampón se cambió (Ultrafree Biomax 10 kDa NMWL membrane, Millipore, Bedford, MA) usando 50 mM de acetato sódico pH 5,0.

45 [0464] Fracciones Q Sepharose Big Bead de *Thielavia terrestris* designadas se fraccionaron luego además usando una columna Phenyl Superose (HR 16/10 Pharmacia Biotech, Uppsala Sweden) usando el protocolo descrito abajo. Sulfato amónico sólido se añadió a la muestra para fraccionarse para dar una concentración final de 1,7 M de (NH₄)₂SO₄. La muestra se centrifugó y se filtró luego (10 minutos, 5,000 x g, RT; 0,22 µm de membrana GP Express, Millipore,

Bedford, MA) para eliminar materia granulosa antes de carga sobre una columna Phenyl Superose equilibrada con 1,7 M de $(\text{NH}_4)_2 \text{SO}_4$ en 20 mM de Tris-HCl pH 8,2. Elución de materia unida se consiguió usando un gradiente lineal decreciente (15 volúmenes de columna usando H_2O desionizada) de 1,7 M de $(\text{NH}_4)_2 \text{SO}_4$ en 20 mM de Tris-HCl pH 8,2. Fracciones se agruparon basadas en absorbancia A280nm para enriquecer áreas de valor máximo resueltas. Agrupaciones se intercambiaron en tampón en 50 mM de acetato sódico pH 5,0 usando un dispositivo de ultrafiltración Amicon (membrana de polietersulfona, 5 kDa de NMWL, Bedford, MA). Concentraciones de proteína se determinaron usando kit BCA Protein Assay y muestras se analizaron en geles de Tris-glicina SDS-PAGE 8-16% como se describe en el ejemplo 1.

[0465] Restos de maíz fueron pretratados en el U.S. Department of Energy National Renewable Energy Laboratory (NREL) usando ácido sulfúrico diluido. Las siguientes condiciones se usaron para el pretratamiento: 1,4 % en peso de ácido sulfúrico a 165°C y 107 psi durante 8 minutos. Según NREL, los sólidos insolubles en agua en los restos de maíz pretratados (PCS) contuvieron 56,5% de celulosa, 4,6% de hemicelulosa y 28,4% de lignina. Celulosa y hemicelulosa se determinaron por una hidrólisis de ácido sulfúrico de dos etapas con análisis posterior de azúcares por cromatografía líquida de alto rendimiento usando NREL Standard Analytical Procedure #002. Lignina se determinó gravimétricamente después de hidrolizar la celulosa y fracciones de hemicelulosa con ácido sulfúrico usando NREL Standard Analytical Procedure #003. Antes de la hidrólisis enzimática, los PCS se lavaron con un volumen grande de agua DDI en un filtro de vidrio; el peso en seco del lavado de agua PCS resultó ser 24,54%. PCS molidos (peso en seco 32,35%) se obtuvieron a partir del lavado de agua PCS por fresado en un desfibrador de café y lavado posterior con agua desionizada en un filtro Millipore 22 μm (6P Express Membrane, Stericup, Millipore, Bedford, MA).

[0466] Hidrólisis de PCS se llevó a cabo usando microtubos Immunoware 1,1 ml (Pierce, Rockford, IL) usando un volumen de reacción total de 1,0 ml. En este protocolo, hidrólisis de PCS (10 mg/ml en 50 mM de tampón acetato sódico pH 5,0) se realizó usando cargas de proteína diferentes (expresado como mg de enzima por gramo de PCS) de un caldo de *Thielavia terrestris* o muestra Celluclast® 1,5L (Novozymes A/S, Bagsværd, Denmark) en presencia de 3% de beta-glucosidasa de *Aspergillus oryzae* (producida de manera recombinante en *Aspergillus oryzae* según la WO 02/095014) o 3% de beta-glucosidasa de *Aspergillus fumigatus* (producida de manera recombinante en *Aspergillus oryzae* según el Ejemplo 22) de carga de proteína celulosa. Selección de fracciones de proteína de *Thielavia terrestris* para capacidad de hidrolización PCS se realizaron a 50°C (baños maría Isotemp 102S o TS Autoflow CO₂ Jacketed Incubator). Típicamente, las reacciones se ejecutaron en cuadruplicados y partes alícuotas tomadas durante el curso de la hidrólisis. Reacciones de hidrólisis PCS se detuvieron mediante la mezcla de una parte alícuota 20 μl de cada hidrolizado con 180 μl de 0,11 M de NaOH (reactivo de parada). Diluciones apropiadas en serie se generaron para cada muestra y el contenido de azúcar reductor se determinó usando un ensayo de ácido para-hidroxibenzoico hidrácido (PHBAH, Sigma, St. Louis, MO) adaptado a un formato de microplaca de 96 pocillos como se describe debajo. Brevemente, una alícuota de 90 μl de una muestra apropiadamente diluida se colocó en una microplaca de fondo cónico de 96 pocillos. Reacciones se iniciaron añadiendo 60 μl de PHBAH 1,5% (p/v) en 2% de NaOH en cada pocillo. Las placas se calentaron descubiertas a 95°C durante 10 minutos. A las placas se les permitió enfriarse a temperatura ambiente (RT) y 50 μl de H_2O destilada se añadió a cada pocillo. Una alícuota de 100 μl de cada pocillo se transfirió a una placa de fondo plano de 96 pocillos y la absorbancia a A410nm se midió usando un SpectraMax Microplate Reader (Molecular Devices, Sunnyvale, CA). Estándares de glucosa (0,1-0,0125 mg/ml diluido con hidróxido sódico 0,4%) se usaron para preparar una curva estándar para traducir los valores A410nm obtenidos en los equivalentes de glucosa. Los equivalentes resultantes se usaron para calcular el porcentaje de conversión de celulosa PCS para cada reacción.

[0467] El grado de conversión de celulosa a azúcar reductor (conversión; %) se calculó usando la siguiente ecuación:

$$\text{Conversión (\%)} = \text{RS}_{(\text{mg/ml})} * 100 * 162 / (\text{Celulosa}_{(\text{mg/ml})} * 180) = \\ = \text{RS}_{(\text{mg/ml})} * 100 / (\text{Celulosa}_{(\text{mg/ml})} * 1,111)$$

En esta ecuación, RS es la concentración de azúcar reductor en la solución medido en los equivalentes de glucosa (mg/ml), y el factor 1,111 refleja el aumento de peso en la conversión de celulosa a glucosa.

[0468] En un intento por reconstruir la actividad de hidrólisis PCS de caldo de *Thielavia terrestris* original, depósitos seleccionados de una columna Q Sepharose Big Bead (Grupos A a F) se mezclaron en proporciones iguales y su actividad de hidrólisis PCS se examinó. La mezcla que se aproxima más estrechamente a la actividad PCS del material de inicio original (caldo de *Thielavia terrestris*, 5 mg de carga enzimática = 77% conversión de celulosa PCS) fue una mezcla que consiste en grupos A+B+C+E (5 mg de carga enzimática = 83% conversión de celulosa PCS).

[0469] Un FPLC ejecutado (AKTA Design, Amersham Pharmacia Biotech, Uppsala Sweden) en el que un grupo de Q Sepharose Big Bead de *Thielavia terrestris* (flujo a través de material) se fraccionó en una columna Phenyl Superose generando múltiples fracciones que se agruparon posteriormente y se intercambiaron en tampón como se describe previamente. Basado en las concentraciones de proteína para cada una de las muestras agrupadas (kit BCA Protein Assay) aproximadamente 43% de la proteína cargada se recuperó de este paso de fraccionamiento.

[0470] Para seleccionar muestras de *Thielavia terrestris* que podrían mejorar rendimiento Celluclast® 1,5L, reacciones de hidrólisis PCS (1,0 ml protocolo, 10 g de PCS por litro, 50°C, complementado por adición de 3% de carga total de beta-glucosidasa de *Aspergillus oryzae*) se realizaron en las que 2,5 mg de carga enzimática de Celluclast® 1,5L se

mezcló con 2,5 mg de carga enzimática de cada muestra agrupada (5 mg de proteína de carga enzimática por reacción). Reacciones de control Celluclast® 1,5L que consisten en 10 mg de carga enzimática, 5 mg de carga enzimática y 2,5 mg carga enzimática dio valores de conversión de celulosa PCS de 86%, 75% y 53%, respectivamente. Análisis de la hidrólisis resultante de PCS mostró que una muestra agrupada cuando se añadió a Celluclast® 1,5L en una proporción de proteína de 50:50 y carga total de 5 mg de enzima superó los 5 mg de carga enzimática de Celluclast® 1,5L de reacción de control (79% vs. 75%). Esta muestra cuando se analizó en un gel de gradiente de acrilamida 8-16% mostró comprender enlace proteínico principal a aproximadamente 42 kDa.

Ejemplo 25: Efecto de *Thielavia terrestris* GH61B en la hidrólisis de restos de maíz pretratados por *Trichoderma reesei* que expresa beta-glucosidasa de *Aspergillus fumigatus*

[0471] *Thielavia terrestris* GH61 B (producido de manera recombinante en *Aspergillus oryzae* como se describe en el ejemplo 12) se desaló y cambió a 50 mM de acetato sódico pH 5,0 usando un filtro centrífugo Centricon Plus-20 con una membrana Biomax-5 (5000 NMWL; Millipore, Bedford, MA) antes de experimentos de hidrólisis. Caldo de fermentación de *Trichoderma reesei* sin células que expresa beta-glucosidasa de *Aspergillus fumigatus* (véase ejemplo 23) se usó en experimentos de hidrólisis sin desalación o intercambio de tampón.

[0472] La concentración de proteína en las muestras enzimáticas se determinó por el ensayo de microplaca BCA como se describe en las instrucciones para kit BCA Protein Assay Reagent (Pierce Chemical Co., Rockford, IL). Caldo de fermentación de *Trichoderma reesei* que expresa beta-glucosidasa de *Aspergillus fumigatus* (véase ejemplo 23) se desaló adicionalmente por paso a través de columnas Bio Spin 6 (Bio-Rad Laboratories, Hercules, CA) según las instrucciones del fabricante antes de medición de la concentración de proteína.

[0473] Diluciones enzimáticas se prepararon frescas antes de cada experimento a partir de soluciones de enzima de materia prima, que fueron almacenadas a -20°C.

[0474] Azúcares reductores (RS) se determinaron usando un ensayo de ácido p-hidroxibenzoico hidrácido (PHBAH) (Lever, M., 1972, *A new reaction for colorimetric determination of carbohydrates*, *Anal. Biochem.*, 47: 273-279), que se modificó y adaptó a un formato de microplaca 96 pocillos.

[0475] Una alícuota de 90- μ l de la muestra diluida se colocó en cada pocillo de un microplaca de fondo cónico de 96 pocillos (Corning Inc., Acton, MA, Costar., policarbonato claro). El ensayo se empezó añadiendo 60 μ l de PHBAH 1,25% en hidróxido sódico 2% para cada pocillo. La placa descubierta se calentó en un bloque de calefacción a medida durante 10 minutos a 95°C. Después de que la microplaca se enfrió a temperatura ambiente, 35 μ l de agua desionizada se añadió a cada pocillo. Una alícuota de 100- μ l se retiró de cada pocillo y se transfirió a una placa de fondo plano de 96 pocillos (Corning Inc., Acton, MA, Costar., poliestireno de unión medio). La absorbancia a 410 nm (A410) se midió usando un lector de microplaca SpectraMAX (Molecular Devices, Sunnyvale, CA). El valor A410 se tradujo en los equivalentes de glucosa usando una curva estándar.

[0476] La curva estándar se obtuvo con seis estándares de glucosa (0,005, 0,010, 0,025, 0,050, 0,075 y 0,100 mg/ml), que se trataron de forma similar a las muestras. Estándares de glucosa se prepararon por dilución de 10 mg/ml de solución de glucosa de materia prima con agua desionizada.

[0477] El grado de conversión de celulosa a azúcar reductor (conversión; %) se calculó usando la misma ecuación descrita en el ejemplo 24.

[0478] Hidrólisis de PCS molido (1% p/v en una base de peso en seco) por un caldo de fermentación de *Trichoderma reesei* sin células que expresa beta-glucosidasa de *Aspergillus fumigatus* (véase ejemplo 23) a 2,5, 5, 10 y 20 mg proteína por g de PCS se llevó a cabo en placas 96 Deepwell (1,2 ml, Brinkmann, Westbury, NY) cubierto con Deepwell Mats 96 (Brinkmann, Westbury, NY). Todas las reacciones con el volumen inicial de 1 ml se ejecutaron en 50 mM de acetato sódico pH 5,0 con agitación intermitente a 50°C, 55°C y 60°C.

[0479] Reacciones con el caldo de fermentación de *Trichoderma reesei* en 5 mg por g de PCS se complementaron con *Thielavia terrestris* GH61 B (producido de manera recombinante en *Aspergillus oryzae* como se describe en este caso) a 1 mg por g de PCS (20% de carga de proteína celulosa), y los resultados se compararon con reacciones no complementadas.

[0480] Partes alícuotas de 20 μ l se retiraron de reacciones de hidrólisis PCS en puntos de tiempo específicos que usan una pipeta de 8 canales, y se añadieron a 180 μ l de mezcla alcalina (102 mM de Na₂CO₃ y 58 mM de NaHCO₃) en placa de filtración de 96 pocillos MultiScreen HV (Millipore, Bedford, MA) para terminar la reacción. Las muestras se filtraron al vacío en otra microplana de fondo plana para eliminar el residuo PCS. Después de dilución apropiada, los filtrados se analizaron para RS usando el ensayo PHBAH como se describe en el ejemplo 24.

[0481] Los resultados como se muestra en las figuras 34 y 35 indicaron que caldo de fermentación de *Trichoderma reesei* que expresa beta-glucosidasa de *Aspergillus fumigatus* (véase ejemplo 23) con la proteína de *Thielavia terrestris* GH61 B mejoró el rendimiento de hidrólisis de 94 horas por 16-30% en comparación con caldo no complementado y

requisitos de carga enzimática se redujeron. A 50°C y 55°C, la misma conversión de celulosa en 94 horas podría conseguirse usando caldo de fermentación complementado en carga de proteína total de 6 mg por g de caldo de fermentación no complementado y PCS a 10 mg por g de PCS.

5 **Ejemplo 26: Hidrólisis de restos de maíz pretratados (PCS) con *Thielavia terrestris* GH61 B producido de manera recombinante en *Aspergillus oryzae***

10 [0482] Transformantes de *Aspergillus oryzae* que expresan cultivos de *Thielavia terrestris* GH61 B crecieron en matraz de agitación (SF) de la siguiente manera. Esporas se recogieron con 5 mililitros de una solución acuosa de Tween 80 0,01% y uno o lavados más con MDU2BP para maximizar el número de esporas recogidas. La suspensión de esporas se usó luego para inocular 500 mililitros de medio MDU2BP en un matraz Fernbach 2 litros. Cultivos transformantes se incubaron a 34°C con agitación constante (200 r.p.m.). El día cinco después de inoculación, los caldos de cultivo se recogieron por filtración en una unidad filtro de nilón 500 mililitro, 75 mm con un tamaño de poro de 0,45 µm. Una muestra de 5 µl del caldo se analizó por SDS-PAGE. El caldo con rGH61 B contuvo una banda principal de aproximadamente 42.000 PM.

20 [0483] Los caldos se almacenaron a -20°C antes de concentración por ultrafiltración de presión en células agitadas (Amicon, membrana PM10 con 10kD MWCO) a aproximadamente concentración 15X, seguido por desalado/intercambio de tampón en el tampón de citrato (50 mM, pH 5,0) usando columnas Econo-Pac 10DG (BioRad). Después de ensayo de la concentración de proteína por un kit BCA Protein Assay, las materias primas enzimáticas se almacenaron a -20°C. Las proteínas no se purificaron posteriormente, y materias primas se añadieron a mezclas de reacción basadas en proteína medida.

25 [0484] Hidrólisis de PCS se llevó a cabo usando microtubos 1,1 ml Immunoware o 1 ml bloques deep-well (VWR) usando un volumen de reacción total de 1,0 ml. Hidrólisis de PCS (10 mg/ml en 50 mM de tampón de acetato sódico pH 5,0) se realizó usando cargas de proteína diferentes (expresado como mg enzimático por gramo PCS) de proteína de *Thielavia terrestris* GH61B recombinante con o sin Celluclast® 1,5L añadido en presencia de beta-glucosidasa de *Aspergillus fumigatus* 3% (BG, 3% de carga de proteína celulasa). Incubación de mezclas de hidrólisis se realizó a 50°C (baños maría Isotemp 102S o TS Autoflow CO₂ Jacketed Incubator con humidificación). Típicamente, reacciones se ejecutaron en triplicados y partes alícuotas tomadas durante el curso de hidrólisis. Datos proporcionados fue la media de triplicados. Reacciones de hidrólisis PCS se detuvieron mediante la mezcla una alícuota de 20 µl de cada hidrolizado con 180 µl de tampón de parada PCS (0,102 M de Na₂CO₃, 0,058 M de NaHCO₃, pH ~10). Muestras se evaluaron inmediatamente o se almacenaron congeladas a -20°C. El contenido de azúcar reductor determinado usando un ensayo de ácido p-hidroxibenzoico hidrácido (PHBAH) se adaptó a un formato de microplaca de 96 pocillos. Brevemente, una alícuota de 100 µl de una muestra apropiadamente diluida se colocó en una microplaca de fondo cónico de 96 pocillos. Reacciones se iniciaron añadiendo 50 µl de 1,5% (p/v) PHBAH en NaOH 2% a cada pocillo. Placas se calentaron descubiertas a 95°C durante 10 minutos. A las placas se les permitió refrescar a RT y 50ul de H₂O se añadió a cada pocillo. Una alícuota de 100 µl de cada pocillo se transfirió a una placa de fondo plano de 96 pocillos y la absorbancia a A_{410nm} se midió usando un lector de microplaca SpectraMax (Molecular Devices). Estándares de glucosa (0,1-0,0125 mg/ml diluido con hidróxido sódico 0,4%) se usaron para preparar una curva estándar para traducir los valores A410 obtenidos en los equivalentes de glucosa (basado en el rendimiento previsto de la conversión teórica completa de celulosa a glucosa). Los equivalentes resultantes se usaron para calcular el porcentaje de conversión de celulosa PCS para cada reacción. Los resultados se representan como equivalentes de glucosa (porcentaje conversión) vs. carga de proteína total (mg enzima / g PCS).

45 [0485] Para examinar la capacidad de la proteína de *Thielavia terrestris* GH61 B para mejorar la actividad de Celluclast® 1,5L, la hidrólisis dependiente de dosis de PCS se evaluó para mezclas de 2,5 mg de Celluclast® 1,5L más 0,078 a 1,25 mg de la proteína GH61 B por g de PCS usando el micro protocolo 1,0 ml PCS. Para fines de comparación, cargas de proteína totales similares se realizaron también para Celluclast® 1,5L sin adición de la proteína GH61 B. Todas las muestras tuvieron beta-glucosidasa de *Aspergillus fumigatus* adicional a 3% de carga de celulasa total.

50 [0486] Los resultados mostrados en la tabla 2 demostraron que la adición de varias cantidades de proteína de *Thielavia terrestris* GH61 B a 2,5 mg/g de Celluclast® PCS 1,5L con una carga enzimática resultante total de 2,578 a 3,75 mg/g de conversión de glucosa aumentada PCS de arriba que se obtuvo por Celluclast® 1,5L sin la proteína GH61 B a 2,5 o 3,75 mg/g de niveles de carga enzimática PCS. Adición de tan solo 0,078 mg de la proteína GH61 B a 2,5 mg de Celluclast® 1,5L mejoró hidrólisis de PCS sobre la cantidad medida para 3,75 mg de Celluclast® 1,5L sin la proteína GH61 B. La adición de la proteína GH61 B a soluciones con Celluclast® 1,5L aumentó, por lo tanto, el rendimiento de azúcares reductores (predominantemente glucosa y celobiosa residual) sobre hidrólisis de restos de maíz pretratados con ácido (PCS). El aumento de rendimiento fue superior al conseguido por adición de cantidades iguales o inferiores de Celluclast® 1,5L, dando como resultado requisitos de carga enzimática reducidos y eficiencia de enzima aumentada.

65 Tabla 2. Valores de Celluclast® y de GH61 B están en mg enzima/g de PCS y glucosa está en porcentaje de rendimiento teórico. Todas las mezclas contienen 3% p/p de beta-glucosidasa añadida para mejorar conversión de celobiosa a glucosa. Conversión de glucosa se muestra como porcentaje de glucosa teórica disponible por hidrólisis de PCS.

Celluclast®	GH61 B OD	Enzima total (mg/g PCS)	Conversión de glucosa (%)
2,5	0	2,5	59,3
2,5	0,078	2,578	65,4
2,5	0,16	2,66	65,3
2,5	0,31	2,81	66,9
2,5	0,63	3,13	67,8
2,5	1,25	3,75	66,8
0	1,25	1,25	2,6
3,75	0	3,75	62,7
6	0	6	70,9
8	0	8	73,7

Ejemplo 27: Hidrólisis de restos de maíz pretratados (PCS) con polipéptidos GH61 B, GH61 E y GH61 G producidos de manera recombinante en *Aspergillus oryzae*

5 [0487] Tres polipéptidos de la familia 61, incluyendo *Thielavia terrestris* GH61 B, GH61 E y GH61 G se evaluaron en la misma serie de reacciones de hidrólisis PCS como se describe en los ejemplos 24 y 25. Todos los polipéptidos de la familia 61 se expresaron en *Aspergillus oryzae* como se describe en el ejemplo 14 y los caldos se concentraron usando una célula agitada de Amicon equipada con una membrana PM10, 10 kDa de corte (Millipore, Billerica, MA) y se desalaron usando una columna Econo-Pac 10DG (BioRad Laboratories, Hercules, CA). Después de ensayo de la concentración de proteína por BCA (Pierce, BSA usado como estándar), estas materias primas enzimáticas se almacenaron a -20°C. Las proteínas no se purificaron posteriormente y materias primas se añadieron a mezclas de reacción basadas en proteína medida.

15 [0488] Hidrólisis de PCS (10 mg/ml en 50 mM de tampón de acetato sódico pH 5,0) se llevó a cabo usando microtubos 1,1 ml Immunoware (Pierce, Rockford, IL) con un volumen de reacción total de 1,0 ml. Los polipéptidos de la familia 61 se evaluaron para su capacidad de mejorar la capacidad hidrolítica de una preparación con celulasa derivada de fermentación de *Trichoderma reesei* una beta-glucosidasa de *Aspergillus oryzae* (WO 02/095014), de ahora en adelante llamada Tr/AoBG y obtenida de Novozymes A/S, Bagsværd, Denmark. Hidrólisis de PCS se realizó usando 2,5 mg de Tr/AoBG por gramo de PCS, complementado con 0,2 mg de polipéptido GH61 por gramo de PCS. La hidrólisis de PCS se realizó a 50°C (incubadora de doble pared TS Autoflow CO₂). Reacciones se ejecutaron en duplicados y partes alícuotas tomadas durante el curso de la hidrólisis. Reacciones de hidrólisis de PCS se detuvieron mediante la mezcla de una alícuota 20 µl de cada hidrolizado con 180 µl de 0,11 M de NaOH (reactivo de parada). Diluciones en serie apropiadas se generaron para cada muestra y el contenido de azúcar reductor se determinó usando un ensayo de ácido para-hidroxibenzoico hidrácido (PHBAH, Sigma, St. Louis, MO) adaptado a un formato de microplaca de 96 pocillos como se describe abajo. Brevemente, una alícuota de 90 µl de una muestra apropiadamente diluida se colocó en una microplaca de fondo cónico de 96 pocillos. Reacciones se iniciaron añadiendo 60 µl de 1,5% (p/v) PHBAH en NaOH 2% a cada pocillo. Placas se calentaron descubiertas a 95°C durante 10 minutos. A las placas se les permitió refrescar a temperatura ambiente (RT) y 50 µl de H₂O destilada se añadió a cada pocillo. Una alícuota de 100 µl de cada pocillo se transfirió a una placa de fondo plano de 96 pocillos y la absorbancia a A410nm se midió usando un lector de microplaca SpectraMax (Molecular Devices, Sunnyvale, CA). Estándares de glucosa (0,1-0,0125 mg/ml diluido con hidróxido sódico 0,4%) se usaron para preparar una curva estándar para traducir los valores A410nm obtenidos en los equivalentes de glucosa. Los equivalentes resultantes se usaron para calcular el porcentaje de conversión de celulosa PCS para cada reacción.

35 [0489] El grado de conversión de celulosa a azúcar reductor (conversión; %) se calculó usando la misma ecuación descrita en el ejemplo 24.

40 [0490] Conversión de celulosa por Tr/AoBG sola (2,5 mg/g PCS) o complementada con cada uno de los tres polipéptidos GH61 (0,2 mg/g PCS) se resumen en la tabla 3 y en la Figura 36.

Tabla 3. Conversión de celulosa por Tr/AoBG sola o Tr/AoBG complementada con polipéptido GH61 a 115 h, 50°C, pH 5,0.

Prueba #	Nombre	Carga, mg/ g PCS	Conversión a 115 h, %
1	Tr/AoBG	2,5	74,6
2	Tr/AoBG + T.t.GH61 B	2,5+0,2	82,0

3	Tr/AoBG + T.t.GH61 E	2,5+0,2	82,0
4	Tr/AoBG+ T.t.GH61 G	2,5+0,2	82,2
5	Tr/AoBG	3,5	82,4

[0491] La tabla 3 y la Figura 36 muestran que todos los tres polipéptidos GH61 mejoraron la actividad de Tr/AoBG en PCS. Complementación de 0,2 mg de T.t. GH61 B, T.t. GH61 E, o T.t. GH61 G a 2,5 mg de Tr/AoBG liberó un nivel de conversión cerca o superior que la de 3,5 mg de Tr/AoBG, indicando actividad Tr/AoBG en PCS fue impulsada por los tres polipéptidos GH61.

[0492] *Thielavia terrestris* GH61 C se evaluó en un experimento separado usando 2,5 mg/g PCS de Tr/AoBG y 0,125 mg/g PCS de T.t. GH61C con incubación durante 120 horas. Conversión fue 63,1% con Tr/AoBG solo y 66,2% con adición T.t. GH61 C.

Ejemplo 28: Hidrólisis de restos de maíz pretratados (PCS) con polipéptido GH61 D producido de manera recombinante en *Aspergillus oryzae*

[0493] Polipéptido de *Thielavia terrestris* GH61 D se expresó en *Aspergillus oryzae* como se describe en el ejemplo 14 y los caldos se desalaron y se concentraron como en el ejemplo 27. Hidrólisis de PCS se llevo a cabo usando bloques deep-well 1 ml (VWR) usando un volumen de reacción total de 1,0 ml. Hidrólisis de PCS (10 mg/ml en 50 mM de tampón de acetato sódico pH 5,0) se realizó usando cargas de proteína diferentes (expresado como mg enzimático por gramo PCS) de una proteína de *Thielavia terrestris* recombinante con o sin Celluclast® 1,5L añadido en presencia de β-glucosidasa de *Aspergillus fumigatus* 3% (BG, 3% de carga de proteína celulasa). Incubación de mezclas de hidrólisis se realizó a 50°C (incubadora de doble TS Autoflow CO₂ con humidificación). Típicamente, reacciones se ejecutaron en triplicados y partes alícuotas tomadas durante el curso de la hidrólisis. Los datos proporcionados son la media de triplicados. Reacciones de hidrólisis de PCS se detuvieron mediante la mezcla de una alícuota de 20 µl de cada hidrolizado con 180 µl de tampón de parada PCS (0,102 M de Na₂CO₃, 0,058 M de NaHCO₃, pH ~10). Muestras se evaluaron inmediatamente o se almacenaron congeladas a -20°C. El contenido de azúcar reductor se determinó usando un ensayo de ácido p-hidroxibenzoico hidrácida (PHBAH) adaptado a un formato de microplaca de 96 pocillos. Brevemente, una alícuota 100 µl de una muestra apropiadamente diluida se colocó en una microplaca de fondo cónico de 96 pocillos. Reacciones se iniciaron añadiendo 50 µl de 1,5% (p/v) PHBAH en NaOH 2% a cada pocillo. Placas se calentaron descubiertas a 95°C durante 10 minutos. A las placas se les permitió refrescar a RT y 50 µl de H₂O añadido a cada pocillo. Una alícuota de 100 µl de cada pocillo se transfirió a una placa de fondo plano de 96 pocillos y la absorbancia a A410nm se midió usando un lector de microplaca SpectraMax (Molecular Devices). Estándares de glucosa (0,1-0,0125 mg/ml diluido con hidróxido sódico 0,4%) se usaron para preparar una curva estándar para traducir los valores A410 obtenidos en los equivalentes de glucosa (basado en el rendimiento previsto de la conversión teórica completa de celulosa a glucosa). Los equivalentes resultantes se usaron para calcular el porcentaje de conversión de celulosa PCS para cada reacción. Los resultados se representaron como equivalentes de glucosa (porcentaje conversión) vs. carga de proteína total (mg enzima/g PCS). Para examinar la capacidad de GH61 D para mejorar la actividad de Celluclast® 1,5L, la hidrólisis dependiente de dosis de PCS se evaluó para mezclas de 2,5 mg de Celluclast® más 0,1 a 0,8 mg GH61 D/g PCS usando el micro protocolo 1,0 ml PCS. Para fines de comparación, cargas de proteína totales similares se realizaron también para Celluclast® 1,5L sin adición de GH61 D. Todas las muestras tuvieron β-glucosidasa de *Aspergillus fumigatus* adicional a 3% de carga de celulasa total.

[0494] Adición de varias cantidades de GH61 D a 2,5 mg/g de Celluclast® PCS con una carga enzimática resultante total de 2,6 a 3,3 mg/g PCS aumenta conversión de glucosa de arriba de la obtenida por Celluclast® sin GH61 D a 3,4 mg/g PCS niveles de carga de enzima (tabla 4).

Tabla 4. Conversión de celulosa por Celluclast® solo o Celluclast® complementado con GH61 D

Celluclast®	GH61 D	Enzima total (mg/g PCS)	Conversión de glucosa (%)
2,5	0	2,5	66,1
2,5	0,1	2,6	73,2
2,5	0,2	2,7	75,4
2,5	0,4	2,9	76,5
2,5	0,8	3,3	78,5
2,7	0	2,7	67,3
2,9	0	2,9	68,3

3,4	0	3,4	71,7
4,0	0	4,0	73,6

Depósito de material biológico

5 [0495] El siguiente material biológico se ha depositado según las condiciones del tratado de Budapest con la Agricultural Research Service Patent Culture Collection, Northern Regional Research Center, 1815 University Street, Peoria, Illinois, 61604, y se le ha dado los siguientes números de registro:

	Depósito	Número de acceso	Fecha de depósito
10	Cepa de <i>E. coli</i> pEJG120	NRRL B-30699	19 diciembre, 2003
	Cepa de <i>E. coli</i> pTter61 C	NRRL B-30823	21 enero, 2005
	Cepa de <i>E. coli</i> pTter61 D	NRRL B-30812	21 enero, 2005
	Cepa de <i>E. coli</i> pTter61 e	NRRL B-30814	21 enero, 2005
	Cepa de <i>E. coli</i> pTter61 G	NRRL B-30811	21 enero, 2005

15 [0496] Las cepas se han depositado bajo condiciones que aseguran que acceso a los cultivos estará disponible durante la pendency de esta solicitud de patente para uno determinado por el Comisionado de Patentes y Marcas para tener derecho a ellas bajo 37 C.F.R. §1, 4 y 35 U.S.C. §122. Los depósitos representan cultivos substancialmente puros de las cepas depositadas. Los depósitos están disponibles según sea necesario por leyes de patentes extranjeras en países donde equivalentes de la presente solicitud, o su descendiente se solicitan. No obstante, se debe entender que
 20 la disponibilidad de un depósito no constituye una licencia para practicar la presente invención en deroga de los derechos de patentes concedidas por acción gubernamental.

Listado de secuencias

25 [0497]

<110> NOVOZYMES BIOTECH, INC.

<120> Polipéptidos con actividad de mejora celulolítica y ácidos nucleicos que los codifican

30 <130> 10587.204-WO

<150> 60/540,661

35 <151> 2004-01-30

<160> 78

<170> Versión de patentIn 3.2

40 <210> 1

<211> 1846

<212> ADN

<213> Thielavia terrestris

45 <400> 1

ES 2 468 366 T3

aattgaagga gggagtggcg gagtggccac caagtcaggc ggctgtcaac taaccaagga 60
tgggaacagt tcggctcgcc ttgcccgagg gcagcgttcc ctgatgggga cgaaccatgg 120
gactggggtc agctgctgta taaaagttca aatcgatgat ctctcagatg gcgctgctgg 180
ggtgttctgc gcttttccat cctcgcaacc tggtatccca ctagtccagc gttcggcacc 240
atgaagtcgt tcaccattgc cgccttggca gccctatggg cccaggaggc cgccgcccac 300
gcgaccttc aggacctctg gattgatgga gtcgactacg gctcgcaatg tgtccgcctc 360
ccggcgcca actccccgt caccaatggt gcgtccgacg atatccgatg caatgtcggc 420
acctcgaggc ccaccgtcaa gtgcccggtc aaggccggct ccacggtcac gatcgagatg 480
caccaggttc gcacgcctct ctgcgtaggc cccccagcta ctatatggca ctaacacgac 540
ctccagcaac ctggcgaccg gtcttgcgcc aacgaggcta tcggcggcga ccactacggc 600
cccgtaatgg tgtacatgtc caaggtcgat gacgcggtga cagccgacgg ttcategggc 660
tggttcaagg tgttccagga cagctgggcc aagaaccctg cgggttcgac gggcgacgac 720
gactactggg gcaccaagga cctcaactcg tgctgcggca agatgaacgt caagatcccc 780
gaagacatcg agccgggcca ctacctgctc cgcgccgagg ttatcgcgct gcacgtggcc 840
gccagctcgg gcggcgcgca gttctacatg tcctgctacc agctgaccgt gacgggctcc 900
ggcagcgcca ccccctcgac cgtgaatttc ccgggcgcct actcggccag cgaccggggc 960
atcctgatca acatccacgc gcccatgtcg acctacgtcg tcccgggccc gaccgtgtac 1020
gcgggcggct cgaccaagtc ggctggcagc tcctgctccg gctgcgaggc gacctgcacg 1080
gttggttccg gccccagcgc gacactgacg cagcccacct ccaccgacgac cgcgacctcc 1140

ES 2 468 366 T3

gccctggcg gcggcggctc cggctgcacg gcggccaagt accagcagtg cggcggcacc 1200
 ggctacactg ggtgcaccac ctgcgctgta agttccctcg tgatatgcag cggaacaccg 1260
 tctggactgt tttgctaact cgcgctcgtag tccgggtcta cctgcagcgc cgtctcgcct 1320
 ccgtactact cgcagtgcct ctaagccggg agcgccttgct cagcgggctg ctgtgaagga 1380
 gctccatgtc cccatgccgc catggccgga gtaccgggct gagcgcccaa ttcttgata 1440
 tagttgagtt ttcccaatca tgaatacata tgcactctgca tggactggtg cgtcgtcagt 1500
 ctacatcctt tgctccactg aactgtgaga ccccatgtca tccggaccat tcgatcggtg 1560
 ctgcctctac catctcgggt gatgggtctg ggcttgagag tcaactggcac gtcctcggcg 1620
 gtaatgaaat gtggaggaaa gtgtgagctg tctgacgcac tcggcgctga tgagacgttg 1680
 agcgcggccc aactggtgt tctgtaagcc agcacacaaa agaatactcc aggatggccc 1740
 atagcggcaa atatacagta tcagggatgc aaaaagtgca aaagtaaggg gctcaatcgg 1800
 ggatcgaacc cgagacctcg cacatgactt atttcaagtc aggggt 1846

<210> 2

5 <211> 326

<212> PRT

<213> Thielavia terrestris

<400> 2

10

ES 2 468 366 T3

Met Lys Ser Phe Thr Ile Ala Ala Leu Ala Ala Leu Trp Ala Gln Glu
 1 5 10 15

Ala Ala Ala His Ala Thr Phe Gln Asp Leu Trp Ile Asp Gly Val Asp
 20 25 30

Tyr Gly Ser Gln Cys Val Arg Leu Pro Ala Ser Asn Ser Pro Val Thr
 35 40 45

Asn Val Ala Ser Asp Asp Ile Arg Cys Asn Val Gly Thr Ser Arg Pro
 50 55 60

Thr Val Lys Cys Pro Val Lys Ala Gly Ser Thr Val Thr Ile Glu Met
 65 70 75 80

His Gln Gln Pro Gly Asp Arg Ser Cys Ala Asn Glu Ala Ile Gly Gly
 85 90 95

Asp His Tyr Gly Pro Val Met Val Tyr Met Ser Lys Val Asp Asp Ala
 100 105 110

ES 2 468 366 T3

Val Thr Ala Asp Gly Ser Ser Gly Trp Phe Lys Val Phe Gln Asp Ser
 115 120 125

Trp Ala Lys Asn Pro Ser Gly Ser Thr Gly Asp Asp Asp Tyr Trp Gly
 130 135 140

Thr Lys Asp Leu Asn Ser Cys Cys Gly Lys Met Asn Val Lys Ile Pro
 145 150 155 160

Glu Asp Ile Glu Pro Gly Asp Tyr Leu Leu Arg Ala Glu Val Ile Ala
 165 170 175

Leu His Val Ala Ala Ser Ser Gly Gly Ala Gln Phe Tyr Met Ser Cys
 180 185 190

Tyr Gln Leu Thr Val Thr Gly Ser Gly Ser Ala Thr Pro Ser Thr Val
 195 200 205

Asn Phe Pro Gly Ala Tyr Ser Ala Ser Asp Pro Gly Ile Leu Ile Asn
 210 215 220

Ile His Ala Pro Met Ser Thr Tyr Val Val Pro Gly Pro Thr Val Tyr
 225 230 235 240

Ala Gly Gly Ser Thr Lys Ser Ala Gly Ser Ser Cys Ser Gly Cys Glu
 245 250 255

Ala Thr Cys Thr Val Gly Ser Gly Pro Ser Ala Thr Leu Thr Gln Pro
 260 265 270

Thr Ser Thr Ala Thr Ala Thr Ser Ala Pro Gly Gly Gly Gly Ser Gly
 275 280 285

Cys Thr Ala Ala Lys Tyr Gln Gln Cys Gly Gly Thr Gly Tyr Thr Gly
 290 295 300

Cys Thr Thr Cys Ala Ser Gly Ser Thr Cys Ser Ala Val Ser Pro Pro
 305 310 315 320

Tyr Tyr Ser Gln Cys Leu
 325

- <210> 3
- 5 <211> 880
- <212> ADN
- <213> Thielavia terrestris

ES 2 468 366 T3

<400> 3

accccgggat cactgcccct aggaaccagc acacctcggg ccaatcatgc ggttcgacgc 60
 cctctccgcc ctcgctcttg cgccgcttgt ggctggccac ggcgccgtga ccagctacat 120
 catcggcggc aaaacctatc ccggctacga gggcttctcg cctgcctcga gcccgccgac 180
 gatccagtac cagtggcccc actacaaccc gaccctgagc gtgaccgacc cgaagatgcy 240
 ctgcaacggc ggcacctcgg cagagctcag cgcgcccgtc caggccggcg agaacgtgac 300
 ggccgtctgg aagcagtgga cccaccagca aggccccgtc atggtctgga tgttcaagtg 360
 ccccggcgac ttctcgtcgt gccacggcga cggcaagggc tggttcaaga tcgaccagct 420
 gggcctgtgg ggcaacaacc tcaactcgaa caactggggc accgcgatcg tctacaagac 480
 cctccagtgg agcaaccgga tccccaaaga cctcgcgccg ggcaactacc tcatccgcca 540
 cgagctgctc gccctgcacc aggccaacac gccgcagttc tacgccgagt gcgcccagct 600
 ggtcgtctcc ggcagcggct ccgcctctgc cccgtccgac tacctctaca gcatccccgt 660
 ctacgcgccc cagaacgacc ccggcatcac cgtgagtggg cttccgttcc gcggcgagct 720
 ctgtggaat cttgctgacg atgggctagg ttgacatcta caacggcggg cttacctcct 780
 acaccccgcc cggcggcccc gtctgggtctg gcttcgagtt ttaggcgcat tgagtcgggg 840
 gctacgaggg gaaggcatct gttcgcgatga gcgtgggtac 880

5 <210> 4
 <211> 478
 <212> PRT
 <213> Thielavia terrestris

10 <400> 4

Met Arg Phe Asp Ala Leu Ser Ala Leu Ala Leu Ala Pro Leu Val Ala
 1 5 10 15
 Gly His Gly Ala Val Thr Ser Tyr Ile Ile Gly Gly Lys Thr Tyr Pro
 20 25 30
 Gly Tyr Glu Gly Phe Ser Pro Ala Ser Ser Pro Pro Thr Ile Gln Tyr
 35 40 45
 Gln Trp Pro Asp Tyr Asn Pro Thr Leu Ser Val Thr Asp Pro Lys Met
 50 55 60
 Arg Cys Asn Gly Gly Thr Ser Ala Glu Leu Ser Ala Pro Val Gln Ala
 65 70 75 80

15

ES 2 468 366 T3

Gly Glu Asn Val Thr Ala Val Trp Lys Gln Trp Thr His Gln Gln Gly
85 90 95

Pro Val Met Val Trp Met Phe Lys Cys Pro Gly Asp Phe Ser Ser Ser
100 105 110

His Gly Asp Gly Lys Gly Trp Phe Lys Ile Asp Gln Leu Gly Leu Trp
115 120 125

Gly Asn Asn Leu Asn Ser Asn Asn Trp Gly Thr Ala Ile Val Tyr Lys
130 135 140

Thr Leu Gln Trp Ser Asn Pro Ile Pro Lys Asn Leu Ala Pro Gly Asn
145 150 155 160

Tyr Leu Ile Arg His Glu Leu Leu Ala Leu His Gln Ala Asn Thr Pro
165 170 175

Gln Phe Tyr Ala Glu Cys Ala Gln Leu Val Val Ser Gly Ser Gly Ser
180 185 190

Ala Leu Pro Pro Ser Asp Tyr Leu Tyr Ser Ile Pro Val Tyr Ala Pro
195 200 205

Gln Asn Asp Pro Gly Ile Thr Val Asp Ile Tyr Asn Gly Gly Leu Thr
210 215 220

Ser Tyr Thr Pro Pro Gly Gly Pro Val Trp Ser Gly Phe Glu Phe Met
225 230 235 240

Arg Phe Asp Ala Leu Ser Ala Leu Ala Leu Ala Pro Leu Val Ala Gly
245 250 255

His Gly Ala Val Thr Ser Tyr Ile Ile Gly Gly Lys Thr Tyr Pro Gly
260 265 270

Tyr Glu Gly Phe Ser Pro Ala Ser Ser Pro Pro Thr Ile Gln Tyr Gln
275 280 285

Trp Pro Asp Tyr Asn Pro Thr Leu Ser Val Thr Asp Pro Lys Met Arg
290 295 300

Cys Asn Gly Gly Thr Ser Ala Glu Leu Ser Ala Pro Val Gln Ala Gly
305 310 315 320

ES 2 468 366 T3

Glu Asn Val Thr Ala Val Trp Lys Gln Trp Thr His Gln Gln Gly Pro
 325 330 335

Val Met Val Trp Met Phe Lys Cys Pro Gly Asp Phe Ser Ser Ser His
 340 345 350

Gly Asp Gly Lys Gly Trp Phe Lys Ile Asp Gln Leu Gly Leu Trp Gly
 355 360 365

Asn Asn Leu Asn Ser Asn Asn Trp Gly Thr Ala Ile Val Tyr Lys Thr
 370 375 380

Leu Gln Trp Ser Asn Pro Ile Pro Lys Asn Leu Ala Pro Gly Asn Tyr
 385 390 395 400

Leu Ile Arg His Glu Leu Leu Ala Leu His Gln Ala Asn Thr Pro Gln
 405 410 415

Phe Tyr Ala Glu Cys Ala Gln Leu Val Val Ser Gly Ser Gly Ser Ala
 420 425 430

Leu Pro Pro Ser Asp Tyr Leu Tyr Ser Ile Pro Val Tyr Ala Pro Gln
 435 440 445

Asn Asp Pro Gly Ile Thr Val Asp Ile Tyr Asn Gly Gly Leu Thr Ser
 450 455 460

Tyr Thr Pro Pro Gly Gly Pro Val Trp Ser Gly Phe Glu Phe
 465 470 475

<210> 5

5 <211> 1000

<212> ADN

<213> Thielavia terrestris

<400> 5

10

ES 2 468 366 T3

```

ctcctggtcc tgggccaccg cttggtgect gcactattgg tagagttggt ctattgctag      60
agttggccat gcttctcaca tcagtcctcg gctcggctgc cctgcttget agcggcgctg      120
cggcacacgg cgccgtgacc agctacatca tcgccggcaa gaattacccg gggtaggtag      180
ctgattattg agggcgcatt caaggttcat accggtgtgc atggctgaca accggtggc      240
agataccaag gcttttctcc tgcgaactcg ccgaacgtca tccaatggca atggcatgac      300
tacaaccccc tcttgtcgtg cagcgactcg aagcttcgct gcaacggcgg cacgtcggcc      360
accctgaacg ccacggccgc accgggagac accatcaccg ccatctgggc gcagtggacg      420
cacagccagg gccccatcct ggtgtggatg tacaagtgcc cgggctcctt cagctcctgt      480
gacggctccg gcgctggctg gttcaagatc gacgaggccg gcttccacgg cgacggcgtc      540
aaggtcttcc tcgacaccga gaaccgctcc ggctgggaca tcgccaagct cgtcggcggc      600
aacaagcagt ggagcagcaa ggtccccgag ggcctcgccc ccggcaacta cctcgtccgc      660
cacgagttga tcgccctgca ccaggccaac aaccgcagt tctaccgga gtgcgccag      720
gtcgtcatca ccggctccgg caccgcgag ccggatgcct catacaaggc ggctatcccc      780
ggctactgca accagaatga cccgaacatc aaggtgagat ccaggcgtaa tgcagtctac      840
tgctggaaag aaagtgtcc aagctaaacc gcgctccagg tgcccatcaa cgaccactcc      900
atccctcaga cctacaagat tcccggcct cccgtcttca agggcaccgc cagcaagaag      960
gccccgggact tcaccgctg aagttgttga atcgatggag      1000

```

5

<210> 6
 <211> 516
 <212> PRT
 <213> Thielavia terrestris

10

<400> 6

ES 2 468 366 T3

Met	Leu	Leu	Thr	Ser	Val	Leu	Gly	Ser	Ala	Ala	Leu	Leu	Ala	Ser	Gly
1				5					10					15	
Ala	Ala	Ala	His	Gly	Ala	Val	Thr	Ser	Tyr	Ile	Ile	Ala	Gly	Lys	Asn
			20					25					30		
Tyr	Pro	Gly	Tyr	Gln	Gly	Phe	Ser	Pro	Ala	Asn	Ser	Pro	Asn	Val	Ile
		35					40						45		
Gln	Trp	Gln	Trp	His	Asp	Tyr	Asn	Pro	Val	Leu	Ser	Cys	Ser	Asp	Ser
	50					55						60			
Lys	Leu	Arg	Cys	Asn	Gly	Gly	Thr	Ser	Ala	Thr	Leu	Asn	Ala	Thr	Ala
65					70					75					80
Ala	Pro	Gly	Asp	Thr	Ile	Thr	Ala	Ile	Trp	Ala	Gln	Trp	Thr	His	Ser
				85					90					95	
Gln	Gly	Pro	Ile	Leu	Val	Trp	Met	Tyr	Lys	Cys	Pro	Gly	Ser	Phe	Ser
			100					105					110		
Ser	Cys	Asp	Gly	Ser	Gly	Ala	Gly	Trp	Phe	Lys	Ile	Asp	Glu	Ala	Gly
		115					120					125			

ES 2 468 366 T3

Phe His Gly Asp Gly Val Lys Val Phe Leu Asp Thr Glu Asn Pro Ser
 130 135 140

Gly Trp Asp Ile Ala Lys Leu Val Gly Gly Asn Lys Gln Trp Ser Ser
 145 150 155 160

Lys Val Pro Glu Gly Leu Ala Pro Gly Asn Tyr Leu Val Arg His Glu
 165 170 175

Leu Ile Ala Leu His Gln Ala Asn Asn Pro Gln Phe Tyr Pro Glu Cys
 180 185 190

Ala Gln Val Val Ile Thr Gly Ser Gly Thr Ala Gln Pro Asp Ala Ser
 195 200 205

Tyr Lys Ala Ala Ile Pro Gly Tyr Cys Asn Gln Asn Asp Pro Asn Ile
 210 215 220

Lys Val Pro Ile Asn Asp His Ser Ile Pro Gln Thr Tyr Lys Ile Pro
 225 230 235 240

Gly Pro Pro Val Phe Lys Gly Thr Ala Ser Lys Lys Ala Arg Asp Phe
 245 250 255

Thr Ala Met Leu Leu Thr Ser Val Leu Gly Ser Ala Ala Leu Leu Ala
 260 265 270

Ser Gly Ala Ala Ala His Gly Ala Val Thr Ser Tyr Ile Ile Ala Gly
 275 280 285

Lys Asn Tyr Pro Gly Tyr Gln Gly Phe Ser Pro Ala Asn Ser Pro Asn
 290 295 300

Val Ile Gln Trp Gln Trp His Asp Tyr Asn Pro Val Leu Ser Cys Ser
 305 310 315 320

Asp Ser Lys Leu Arg Cys Asn Gly Gly Thr Ser Ala Thr Leu Asn Ala
 325 330 335

Thr Ala Ala Pro Gly Asp Thr Ile Thr Ala Ile Trp Ala Gln Trp Thr
 340 345 350

His Ser Gln Gly Pro Ile Leu Val Trp Met Tyr Lys Cys Pro Gly Ser
 355 360 365

ES 2 468 366 T3

Phe Ser Ser Cys Asp Gly Ser Gly Ala Gly Trp Phe Lys Ile Asp Glu
 370 375 380

Ala Gly Phe His Gly Asp Gly Val Lys Val Phe Leu Asp Thr Glu Asn
 385 390 395 400

Pro Ser Gly Trp Asp Ile Ala Lys Leu Val Gly Gly Asn Lys Gln Trp
 405 410 415

Ser Ser Lys Val Pro Glu Gly Leu Ala Pro Gly Asn Tyr Leu Val Arg
 420 425 430

His Glu Leu Ile Ala Leu His Gln Ala Asn Asn Pro Gln Phe Tyr Pro
 435 440 445

Glu Cys Ala Gln Val Val Ile Thr Gly Ser Gly Thr Ala Gln Pro Asp
 450 455 460

Ala Ser Tyr Lys Ala Ala Ile Pro Gly Tyr Cys Asn Gln Asn Asp Pro
 465 470 475 480

Asn Ile Lys Val Pro Ile Asn Asp His Ser Ile Pro Gln Thr Tyr Lys
 485 490 495

Ile Pro Gly Pro Pro Val Phe Lys Gly Thr Ala Ser Lys Lys Ala Arg
 500 505 510

Asp Phe Thr Ala
 515

<210> 7

<211> 681

5 <212> ADN

<213> Thielavia terrestris

<400> 7

atgctcgcaa acggtgccat cgtcttctcg gccgccgccc tcggcgtcag tggccactac 60

acctggccac gggttaacga cggcgccgac tggcaacagg tccgtaaggc ggacaactgg 120

caggacaacg gctacgtcgg ggatgtcagc tcgccacaga tccgctgttt ccaggcgacc 180

ccgtccccgg ccccatccgt cctcaacacc acggccggct cgaccgtgac ctactgggcc 240

aaccccgacg tctaccaccc cgggcctgtg cagttttaca tggcccgcgt gcccgatggc 300

gaggacatca actcgtggaa cggcgacggc gccgtgtggt tcaaggtgta cgaggaccat 360

10 cctacctttg gcgctcagct cacatggccc agcacgggca agagctcgtt cgcggttccc 420

ES 2 468 366 T3

atccccccgt gcatcaagtc cggctactac ctctccggg cggagcaaat cggcctgcac 480
 gtcgccaga gcgtaggcgg agcgcagttc tacatctcat gcgccagct cagcgtcacc 540
 ggcgccggca gcaccgagcc gccgaacaag gtggccttcc ccggcgctta cagtgcgacg 600
 gaccgggca ttctgatcaa catctactac cctgttccca cgtcctacca gaaccccggc 660
 ccggccgtct tcagctgctg a 681

<210> 8

<211> 452

5 <212> PRT

<213> Thielavia terrestris

<400> 8

Met Leu Ala Asn Gly Ala Ile Val Phe Leu Ala Ala Ala Leu Gly Val
 1 5 10 15
 Ser Gly His Tyr Thr Trp Pro Arg Val Asn Asp Gly Ala Asp Trp Gln
 20 25 30
 Gln Val Arg Lys Ala Asp Asn Trp Gln Asp Asn Gly Tyr Val Gly Asp
 35 40 45
 Val Thr Ser Pro Gln Ile Arg Cys Phe Gln Ala Thr Pro Ser Pro Ala
 50 55 60
 Pro Ser Val Leu Asn Thr Thr Ala Gly Ser Thr Val Thr Tyr Trp Ala
 65 70 75 80
 Asn Pro Asp Val Tyr His Pro Gly Pro Val Gln Phe Tyr Met Ala Arg
 85 90 95
 Val Pro Asp Gly Glu Asp Ile Asn Ser Trp Asn Gly Asp Gly Ala Val
 100 105 110
 Trp Phe Lys Val Tyr Glu Asp His Pro Thr Phe Gly Ala Gln Leu Thr
 115 120 125
 Trp Pro Ser Thr Gly Lys Ser Ser Phe Ala Val Pro Ile Pro Pro Cys
 130 135 140
 Ile Lys Ser Gly Tyr Tyr Leu Leu Arg Ala Glu Gln Ile Gly Leu His
 145 150 155 160
 Val Ala Gln Ser Val Gly Gly Ala Gln Phe Tyr Ile Ser Cys Ala Gln
 165 170 175

10

ES 2 468 366 T3

Leu Ser Val Thr Gly Gly Gly Ser Thr Glu Pro Pro Asn Lys Val Ala
180 185 190

Phe Pro Gly Ala Tyr Ser Ala Thr Asp Pro Gly Ile Leu Ile Asn Ile
195 200 205

Tyr Tyr Pro Val Pro Thr Ser Tyr Gln Asn Pro Gly Pro Ala Val Phe
210 215 220

Ser Cys Met Leu Ala Asn Gly Ala Ile Val Phe Leu Ala Ala Ala Leu
225 230 235 240

Gly Val Ser Gly His Tyr Thr Trp Pro Arg Val Asn Asp Gly Ala Asp
245 250 255

Trp Gln Gln Val Arg Lys Ala Asp Asn Trp Gln Asp Asn Gly Tyr Val
260 265 270

Gly Asp Val Thr Ser Pro Gln Ile Arg Cys Phe Gln Ala Thr Pro Ser
275 280 285

Pro Ala Pro Ser Val Leu Asn Thr Thr Ala Gly Ser Thr Val Thr Tyr
290 295 300

Trp Ala Asn Pro Asp Val Tyr His Pro Gly Pro Val Gln Phe Tyr Met
305 310 315 320

Ala Arg Val Pro Asp Gly Glu Asp Ile Asn Ser Trp Asn Gly Asp Gly
325 330 335

Ala Val Trp Phe Lys Val Tyr Glu Asp His Pro Thr Phe Gly Ala Gln
340 345 350

Leu Thr Trp Pro Ser Thr Gly Lys Ser Ser Phe Ala Val Pro Ile Pro
355 360 365

Pro Cys Ile Lys Ser Gly Tyr Tyr Leu Leu Arg Ala Glu Gln Ile Gly
370 375 380

Leu His Val Ala Gln Ser Val Gly Gly Ala Gln Phe Tyr Ile Ser Cys
385 390 395 400

Ala Gln Leu Ser Val Thr Gly Gly Gly Ser Thr Glu Pro Pro Asn Lys
405 410 415

ES 2 468 366 T3

Val Ala Phe Pro Gly Ala Tyr Ser Ala Thr Asp Pro Gly Ile Leu Ile
 420 425 430

Asn Ile Tyr Tyr Pro Val Pro Thr Ser Tyr Gln Asn Pro Gly Pro Ala
 435 440 445

Val Phe Ser Cys
 450

5 <210> 9
 <211> 960
 <212> ADN
 <213> Thielavia terrestris

<400> 9

atgaagggac	ttttcagtgc	cgccgccctc	tccctggccg	tcggccagge	ttcgcccacat	60
tacatcttcc	agcaactctc	catcaacggg	aaccagtttc	cggtgtacca	atatattcgc	120
aagaacacca	attataacag	tcccgttacc	gatctcacgt	ccgacgatct	tcggtgcaat	180
gtcggcgccc	agggtgctgg	gacagacacc	gtcacggtga	aggccggcga	ccagttcacc	240
ttcacccttg	acaccctgt	ttaccaccag	gggcccacat	ccatctacat	gtccaaggcc	300
ccgggcgcgg	cgtcagacta	cgatggcagc	ggcggctggt	tcaagatcaa	ggactggggc	360
ccgactttca	acgccgacgg	cacggccacc	tgggacatgg	ccggctcata	cacctacaac	420
atcccgaact	gcattcccga	cggcgactat	ctgctccgca	tccagtcgct	ggccatccac	480
aaccctggc	cggcgggcat	cccgcagttc	tacatctcct	gcgcccagat	caccgtgacc	540
ggcggcggca	acggcaacc	tggcccgcag	gccctcatcc	ccggcgcctt	caaggacacc	600
gaccgggct	acacggtgaa	catctacacg	aacttcaca	actacacggt	tcccggcccg	660
gaggtcttca	gctgcaacgg	cggcggctcg	aaccgcgcc	cgccggtgag	tagcagcacg	720
cccgcgacca	cgacgctggt	cacgtcgacg	cgcaaccagt	cctccacgtc	ctccgcctcg	780
acgccggcct	cgaccggcgg	ctgcaccgtc	gccaaagtgg	gccagtgagg	cggcaacggg	840
tacaccggct	gcacgacctg	cgccggccgg	tccacctgca	gcaagcagaa	cgactactac	900
tcgcagtgct	tgtaagggag	gccgcaaagc	atgaggtggt	tgaagaggag	gagaggggct	960

10 <210> 10
 <211> 608
 <212> PRT
 15 <213> Thielavia terrestris

<400> 10

ES 2 468 366 T3

Met Lys Gly Leu Phe Ser Ala Ala Ala Leu Ser Leu Ala Val Gly Gln
 1 5 10 15

Ala Ser Ala His Tyr Ile Phe Gln Gln Leu Ser Ile Asn Gly Asn Gln
 20 25 30

Phe Pro Val Tyr Gln Tyr Ile Arg Lys Asn Thr Asn Tyr Asn Ser Pro
 35 40 45

Val Thr Asp Leu Thr Ser Asp Asp Leu Arg Cys Asn Val Gly Ala Gln
 50 55 60

Gly Ala Gly Thr Asp Thr Val Thr Val Lys Ala Gly Asp Gln Phe Thr
 65 70 75 80

Phe Thr Leu Asp Thr Pro Val Tyr His Gln Gly Pro Ile Ser Ile Tyr
 85 90 95

Met Ser Lys Ala Pro Gly Ala Ala Ser Asp Tyr Asp Gly Ser Gly Gly
 100 105 110

Trp Phe Lys Ile Lys Asp Trp Gly Pro Thr Phe Asn Ala Asp Gly Thr
 115 120 125

Ala Thr Trp Asp Met Ala Gly Ser Tyr Thr Tyr Asn Ile Pro Thr Cys
 130 135 140

Ile Pro Asp Gly Asp Tyr Leu Leu Arg Ile Gln Ser Leu Ala Ile His
 145 150 155 160

Asn Pro Trp Pro Ala Gly Ile Pro Gln Phe Tyr Ile Ser Cys Ala Gln
 165 170 175

Ile Thr Val Thr Gly Gly Gly Asn Gly Asn Pro Gly Pro Thr Ala Leu
 180 185 190

Ile Pro Gly Ala Phe Lys Asp Thr Asp Pro Gly Tyr Thr Val Asn Ile
 195 200 205

Tyr Thr Asn Phe His Asn Tyr Thr Val Pro Gly Pro Glu Val Phe Ser
 210 215 220

Cys Asn Gly Gly Gly Ser Asn Pro Pro Pro Pro Val Ser Ser Ser Thr
 225 230 235 240

ES 2 468 366 T3

Pro Ala Thr Thr Thr Leu Val Thr Ser Thr Arg Thr Thr Ser Ser Thr
245 250 255

Ser Ser Ala Ser Thr Pro Ala Ser Thr Gly Gly Cys Thr Val Ala Lys
260 265 270

Trp Gly Gln Cys Gly Gly Asn Gly Tyr Thr Gly Cys Thr Thr Cys Ala
275 280 285

Ala Gly Ser Thr Cys Ser Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu
290 295 300

Met Lys Gly Leu Phe Ser Ala Ala Ala Leu Ser Leu Ala Val Gly Gln
305 310 315 320

Ala Ser Ala His Tyr Ile Phe Gln Gln Leu Ser Ile Asn Gly Asn Gln
325 330 335

Phe Pro Val Tyr Gln Tyr Ile Arg Lys Asn Thr Asn Tyr Asn Ser Pro
340 345 350

Val Thr Asp Leu Thr Ser Asp Asp Leu Arg Cys Asn Val Gly Ala Gln
355 360 365

Gly Ala Gly Thr Asp Thr Val Thr Val Lys Ala Gly Asp Gln Phe Thr
370 375 380

Phe Thr Leu Asp Thr Pro Val Tyr His Gln Gly Pro Ile Ser Ile Tyr
385 390 395 400

Met Ser Lys Ala Pro Gly Ala Ala Ser Asp Tyr Asp Gly Ser Gly Gly
405 410 415

Trp Phe Lys Ile Lys Asp Trp Gly Pro Thr Phe Asn Ala Asp Gly Thr
420 425 430

Ala Thr Trp Asp Met Ala Gly Ser Tyr Thr Tyr Asn Ile Pro Thr Cys
435 440 445

Ile Pro Asp Gly Asp Tyr Leu Leu Arg Ile Gln Ser Leu Ala Ile His
450 455 460

Asn Pro Trp Pro Ala Gly Ile Pro Gln Phe Tyr Ile Ser Cys Ala Gln
465 470 475 480

ES 2 468 366 T3

Ile Thr Val Thr Gly Gly Gly Asn Gly Asn Pro Gly Pro Thr Ala Leu
 485 490 495

Ile Pro Gly Ala Phe Lys Asp Thr Asp Pro Gly Tyr Thr Val Asn Ile
 500 505 510

Tyr Thr Asn Phe His Asn Tyr Thr Val Pro Gly Pro Glu Val Phe Ser
 515 520 525

Cys Asn Gly Gly Gly Ser Asn Pro Pro Pro Pro Val Ser Ser Ser Thr
 530 535 540

Pro Ala Thr Thr Thr Leu Val Thr Ser Thr Arg Thr Thr Ser Ser Thr
 545 550 555 560

Ser Ser Ala Ser Thr Pro Ala Ser Thr Gly Gly Cys Thr Val Ala Lys
 565 570 575

Trp Gly Gln Cys Gly Gly Asn Gly Tyr Thr Gly Cys Thr Thr Cys Ala
 580 585 590

Ala Gly Ser Thr Cys Ser Lys Gln Asn Asp Tyr Tyr Ser Gln Cys Leu
 595 600 605

<210> 11
 <211> 17
 5 <212> PRT
 <213> Thielavia terrestris

<220>
 <221> MISC_FEATURE
 10 <222> (1).. (1)
 <223> Xaa= Ile o Leu

<220>
 <221> MISC_FEATURE
 15 <222> (16).. (16)
 <223> Xaa= Ile o Leu

<400> 11

Xaa Pro Ala Ser Asn Ser Pro Val Thr Asn Val Ala Ser Asp Asp Xaa
 1 5 10 15

20 Arg

<210> 12
 <211> 13
 <212> PRT
 25 <213> Thielavia terrestris

<220>
 <221> MISC_FEATURE
 30 <222> (1).. (1)
 <223> Xaa= Ile o Leu

<220>
 <221> MISC_FEATURE
 <222> (5).. (5)
 <223> Xaa= Ile o Leu
 5

<220>
 <221> MISC_FEATURE
 <222> (11)..(12)
 <223> Xaa= Ile o Leu
 10

<400> 12
 Xaa Pro Glu Asp Xaa Glu Pro Gly Asp Tyr Xaa Xaa Arg
 1 5 10

<210> 13
 <211> 17
 <212> PRT
 <213> Thielavia terrestris
 15

<400> 13
 Cys Pro Gly Ser Phe Ser Ser Cys Asp Gly Ser Gly Ala Gly Trp Phe
 1 5 10 15

Lys

<210> 14
 <211> 12
 <212> PRT
 <213> Thielavia terrestris
 25

<220>
 <221> MISC_FEATURE
 <222> (1) .. (1)
 <223> Xaa= Ile o Leu
 30

<400> 14
 Xaa Asp Glu Ala Gly Phe His Gly Asp Gly Val Lys
 1 5 10

<210> 15
 <211> 10
 <212> PRT
 <213> Thielavia terrestris
 35

<220>
 <221> MISC_FEATURE
 <222> (1)..(2)
 <223> Xaa= cualquier aminoácido
 40

<220>
 <221> MISC_FEATURE
 <222> (8)..(9)
 <223> Xaa= Ile o Leu
 45

<400> 15
 Xaa Xaa Ala Pro Gly Asn Tyr Xaa Xaa Arg
 1 5 10

<210> 16
 <211> 22
 <212> ADN
 55

<213> Aspergillus nidulans
 <400> 16
 5 gtgcccatg atacgcctcc gg 22
 <210> 17
 <211> 26
 <212> ADN
 <213> Aspergillus nidulans
 10 <400> 17
 gagtcgtatt tccaaggctc ctgacc 26
 <210> 18
 <211> 24
 <212> ADN
 <213> Aspergillus nidulans
 15 <400> 18
 20 ggaggccatg aagtgacca acgg 24
 <210> 19
 <211> 45
 <212> ADN
 25 <213> Aspergillus niger
 <400> 19
 caccgtgaaa gccatgctct ttcctctgtg tagaagacca gacag 45
 30 <210> 20
 <211> 45
 <212> ADN
 <213> Aspergillus niger
 35 <400> 20
 ctgtcttct acacgaagga aagagcatgg ctttcacggt gtctg 45
 <210> 21
 <211> 44
 <212> ADN
 40 <213> Aspergillus niger
 <400> 21
 45 ctatatacac aactggatt accatgggcc cgcggccgca gatc 44
 <210> 22
 <211> 44
 <212> ADN
 <213> Aspergillus niger
 50 <400> 22
 gatctgcggc cgcgggcca tggtaaacc agttgtgtat atag 44
 <210> 23
 <211> 17
 <212> PRT
 <213> Thielavia terrestris
 55 <220>
 <221> MISC_FEATURE
 <222> (1) .. (1)
 <223> Xaa= Ile o Leu
 60 <220>
 <221> MISC_FEATURE
 <222> (16)..(16)
 65

ES 2 468 366 T3

<223> Xaa= Ile o Leu
<400> 23

Xaa Pro Ala Ser Asn Ser Pro Val Thr Asn Val Ala Ser Asp Asp Xaa
1 5 10 15

Arg

5
<210> 24
<211> 13
<212> PRT
<213> Thielavia terrestris

10
<220>
<221> MISC_FEATURE
<222> (1)..(1)
<223> Xaa= Ile o Leu

15
<220>
<221> MISC_FEATURE
<222> (5)..(5)
<223> Xaa= Ile o Leu

20
<220>
<221> MISC_FEATURE
<222> (11)..(12)
<223> Xaa= Ile o Leu

25
<400> 24

Xaa Pro Glu Asp Xaa Glu Pro Gly Asp Tyr Xaa Xaa Arg
1 5 10

30
<210> 25
<211> 28
<212> ADN
<213> Thielavia terrestris

35
<220>
<221> MISC_FEATURE
<222> (20)..(20)
<223> N = A, C, G, o T

40
<220>
<221> MISC_FEATURE
<222> (26)..(26)
<223> N = A, C, G, o T

45
<400> 25
cctccaactc ccccgtcacn aaygtngc 28

50
<210> 26
<211> 27
<212> ADN
<213> Thielavia terrestris

55
<220>
<221> MISC_FEATURE
<222> (19)..(19)
<223> N=A,C,G, o T

60
<220>
<221> MISC_FEATURE
<222> (22)..(22)

ES 2 468 366 T3

<223> N=A, C, G, o T

<400> 26
 5 ggcgcgagg aggtartcnc cnggitc 27

<210> 27
 <211> 16
 <212> PRT
 <213> Thielavia terrestris

10 <400> 27

Ser Gly Ala Gly Trp Phe Lys Ile Asp Glu Ala Gly Phe His Gly Asp
 1 5 10 15

15 <210> 28
 <211> 8
 <212> PRT
 <213> Thielavia terrestris

20 <220>
 <221> MISC_FEATURE
 <222> (6) .. (7)
 <223> Xaa= Ile o Leu

25 <400> 28

Ala Pro Gly Asn Tyr Xaa Xaa Arg
 1 5

30 <210> 29
 <211> 11
 <212> PRT
 <213> Thielavia terrestris

35 <400> 29

Ala Pro Gly Asn Tyr Leu Ile Arg His Glu Leu
 1 5 10

<210> 30
 <211> 11
 <212> PRT

40 <213> Thielavia terrestris

<400> 30

Ala Pro Gly Asn Tyr Leu Val Arg His Glu Leu
 1 5 10

45 <210> 31
 <211> 9
 <212> PRT
 <213> Thielavia terrestris

50 <400> 31

Gly Ala Gly Trp Phe Lys Ile Asp Glu
 1 5

55 <210> 32
 <211> 27
 <212> ADN
 <213> Thielavia terrestris

<220>
 <221> MISC_FEATURE
 <222> (19) .. (19)
 <223> R= A o G
 5

<220>
 <221> MISC_FEATURE
 <222> (22) .. (22)
 <223> H= A, C, o T
 10

<220>
 <221> MISC_FEATURE
 <222> (25) .. (25)
 <223> Y= C o T
 15

<400> 32
 cggcgcgggc tggttaara tgayga 27

<210> 33
 <211> 32
 <212> ADN
 <213> Thielavia terrestris

<220>
 <221> misc_feature
 <222> (21) .. (21)
 <223> R= A o G

<220>
 <221> misc_feature
 <222> (24) .. (24)
 <223> N= A, C, G, o T

<220>
 <221> misc_feature
 <222> (27) .. (27)
 <223> N= A, C, G, o T

<220>
 <221> misc_feature
 <222> (30) .. (30)
 <223> N= A, C, G, o T

<400> 33
 agttcatggc gaatcagata rttncnggn gc 32

<210> 34
 <211> 25
 <212> ADN
 <213> Thielavia terrestris

<400> 34
 cftggtaccg agctcggatc cacta 25

<210> 35
 <211> 25
 <212> ADN
 <213> Thielavia terrestris

<400> 35
 atagggcgaa ttggccctc tagat 25

<210> 36
 <211> 32
 <212> ADN
 <213> Thielavia terrestris

65

ES 2 468 366 T3

<400> 36
 acaactggat ttaccatgcg gttcgacgcc tc 32

5 <210> 37
 <211> 33
 <212> ADN
 <213> Thielavia terrestris

10 <400> 37
 gtcagtcacc tctagtact aaaactcgaa gcc 33

15 <210> 38
 <211> 21
 <212> ADN
 <213> Thielavia terrestris

20 <400> 38
 catgccatgg atgcttctca c 21

25 <210> 39
 <211> 25
 <212> ADN
 <213> Thielavia terrestris

<400> 39
 ccttaattaa tcaggcgggtg aagtc 25

30 <210> 40
 <211> 32
 <212> ADN
 <213> Thielavia terrestris

35 <400> 40
 tcgtcgggga caacttgta caaaaaagtt gg 32

40 <210> 41
 <211> 27
 <212> ADN
 <213> Thielavia terrestris

<400> 41
 cccctgtga aacatgtttt ttaacc 27

45 <210> 42
 <211> 16
 <212> ADN
 <213> Thielavia terrestris

50 <400> 42
 gtaaacgac ggccag 16

55 <210> 43
 <211> 25
 <212> ADN
 <213> Escherichia coli

60 <220>
 <221> misc_feature
 <222> (24) .. (27)
 <223> V= A, C, o G

65 <220>
 <221> misc_feature
 <222> (25) .. (25)
 <223> N=A, C, G, o T

ES 2 468 366 T3

<400> 43
 tttttttt tttttttt tttvn 25

5 <210> 44
 <211> 27
 <212> ADN
 <213> Escherichia coli

10 <400> 44
 ggggacaact ttgtacaaaa aagtgg 27

 <210> 45
 <211> 27
 <212> ADN
 <213> Escherichia coli

15

 <400> 45
 aaaggtagga tggctctgt acacctt 27

20

 <210> 46
 <211> 36
 <212> ADN
 <213> Thielavia terrestris

25

 <400> 46
 actggattac catgctcgca aacggtgcca tcgtct 36

30

 <210> 47
 <211> 38
 <212> ADN
 <213> Thielavia terrestris

35

 <400> 47
 tcacctctag ttaattaatc agcagctgaa gacggccg 38

 <210> 48
 <211> 31
 <212> ADN
 <213> Thielavia terrestris

40

 <400> 48
 actggattta ccatgaagtc gttcaccatt g 31

45

 <210> 49
 <211> 31
 <212> ADN
 <213> Thielavia terrestris

50

 <400> 49
 agtcacctct agttagaggc actgcgagta g 31

 <210> 50
 <211> 32
 <212> ADN
 <213> Thielavia terrestris

55

 <400> 50
 acaactggat ttaccatgcg gttcgagcc tc 32

60

 <210> 51
 <211> 33
 <212> ADN
 <213> Thielavia terrestris

65

 <400> 51

	gtcagtcacc tctagttact aaaactcgaa gcc	33
5	<210> 52 <211> 27 <212> ADN <213> Thielavia terrestris	
10	<400> 52 actggattac catgcttctc acatcag	27
15	<210> 53 <211> 30 <212> ADN <213> Thielavia terrestris	
20	<400> 53 agtcacctct agttatcagg cgggaagtc	30
25	<210> 54 <211> 31 <212> ADN <213> Thielavia terrestris	
30	<400> 54 actggattac catgaaggga ctttcagtg c	31
35	<210> 55 <211> 31 <212> ADN <213> Thielavia terrestris	
40	<400> 55 agtcacctct agttagaggc actgcgagta g	31
45	<210> 56 <211> 29 <212> ADN <213> Trichoderma reesei	
50	<400> 56 aacgtaatt aaggaatcgt tttgtgtt	29
55	<210> 57 <211> 29 <212> ADN <213> Trichoderma reesei	
60	<400> 57 agtactagta gctccgtggc gaaagcctg	29
65	<210> 58 <211> 26 <212> ADN <213> Trichoderma reesei	
70	<400> 58 actagtcgac cgaatgtagg attgtt	26
75	<210> 59 <211> 19 <212> ADN <213> Trichoderma reesei	
80	<400> 59 tgaccatggt ggcagtc	19

ES 2 468 366 T3

	<210> 60	
	<211> 26	
	<212> ADN	
	<213> <i>Trichoderma reesei</i>	
5		
	<400> 60	
	cgatcgtctc cctatgggtc attacc	26
10		
	<210> 61	
	<211> 28	
	<212> ADN	
	<213> <i>Trichoderma reesei</i>	
15		
	<400> 61	
	actagttaat taagctccgt ggcgaaag	28
20		
	<210> 62	
	<211> 24	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
25		
	<400> 62	
	gggttcgaat tcatttaaac ggct	24
30		
	<210> 63	
	<211> 24	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
35		
	<400> 63	
	gggagcgtc aatattcatc tctc	24
40		
	<210> 64	
	<211> 26	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
45		
	<400> 64	
	ggtcgcggag gcgatggatg cgatcg	26
50		
	<210> 65	
	<211> 26	
	<212> ADN	
	<213> <i>Escherichia coli</i>	
55		
	<400> 65	
	cgatcgcattc catcgctcc gcgacc	26
60		
	<210> 66	
	<211> 34	
	<212> ADN	
	<213> <i>Thielavia terrestris</i>	
65		
	<400> 66	
	cgcggactgc gcacatgaa gtcgtcacc attg	34
70		
	<210> 67	
	<211> 31	
	<212> ADN	
	<213> <i>Thielavia terrestris</i>	
75		
	<400> 67	
	tcgccacgga gcttagaggc actgcgagta g	31
80		
	<210> 68	
	<211> 23	

ES 2 468 366 T3

	<212> ADN		
	<213> Thielavia terrestris		
	<400> 68		
5	gcccatggac catgctcgca aac	23	
	<210> 69		
	<211> 23		
	<212> ADN		
10	<213> Thielavia terrestris		
	<400> 69		
	cctctagtta attaatacagc agc	23	
	<210> 70		
	<211> 23		
	<212> ADN		
	<213> Thielavia terrestris		
15			
	<400> 70		
20	gcccatggac catgaagga ctt	23	
	<210> 71		
	<211> 23		
	<212> ADN		
	<213> Thielavia terrestris		
25			
	<400> 71		
	attaattaca de cctctagtta agc	23	
30			
	<210> 72		
	<211> 19		
	<212> ADN		
	<213> Trichoderma reesei		
35			
	<400> 72		
	ggatgaagct cattagccg	19	
	<210> 73		
	<211> 31		
	<212> ADN		
	<213> Aspergillus fumigatus		
40			
	<400> 73		
45	actggattta ccatgagatt cggttggctc g	31	
	<210> 74		
	<211> 31		
	<212> ADN		
	<213> Aspergillus fumigatus		
50			
	<400> 74		
	agtcacctct agttactagt agacacgggg c	31	
	<210> 75		
	<211> 3060		
	<212> ADN		
	<213> Aspergillus fumigatus		
55			
	<400> 75		
60			

ES 2 468 366 T3

atgagattcg gttggctcga ggtggccgct ctgacggccg cttctgtagc caatgccag 60
 gtttgtgatg ctttcccgtc attgtttcgg atatagttga caatagtcac ggaaataatc 120
 aggaattggc tttctctcca ccattctacc cttcgccttg ggctgatggc cagggagagt 180
 gggcagatgc ccatcgacgc gccgtcgaga tcgtttctca gatgacactg gcggagaagg 240
 ttaaccttac aacgggtact ggggtgggtg cgactttttt gttgacagtg agctttcttc 300
 actgaccatc tacacagatg ggaaatggac cgatgcgtcg gtcaaaccgg cagcgttccc 360
 aggtaagctt gcaattctgc aacaacgtgc aagtgtagt gctaaaacgc ggtggtgcag 420
 acttgggtatc aactggggtc tttgtggcca ggattcccct ttgggtatcc gtttctgtga 480
 gctatacccg cggagtcttt cagtccttgt attatgtgct gatgattgtc tctgtatagc 540
 tgacctcaac tccgccttcc ctgctggtac taatgtcgcc gcgacatggg acaagacact 600
 cgcctacctt cgtggcaagg ccatgggtga ggaattcaac gacaagggcg tggacatttt 660
 gctggggcct gctgctggtc ctctcggcaa ataccggac ggcggcagaa tctgggaagg 720
 cttctctcct gatccggttc tcactgggtg acttttcgcc gaaactatca agggatcca 780
 agacgcgggt gtgattgcta ctgccaagca ttacattctg aatgaacagg agcatttccg 840
 acaggttggc gagggcccagg gatatggtta caacatcacg gagacgatca gctccaacgt 900
 ggatgacaag accatgcacg agttgtacct ttgggtgagta gttgacactg caaatgagga 960
 ccttgattga tttgactgac ctggaatgca ggccctttgc agatgctgtg cgcggtaaga 1020
 ttttccgtag acttgacctc gcgacgaaga aatcgctgac gaaccatcgt agctggcggt 1080
 ggcgctgtca tgtgttctca caatcaaate aacaacagct acggttgtca aaacagtcaa 1140
 actctcaaca agctcctcaa ggctgagctg ggcttccaag gcttcgcatc gatgactgg 1200
 agcgcctacc acagcgggtg cggcgtgcc ctcgctgggt tggatatgtc gatgcctgga 1260
 gacatttctc tcgacgacgg actctccttc tggggcacga acctaactgt cagtgttctt 1320

ES 2 468 366 T3

aacggcaccg ttccagcctg gcgtgtcgat gacatggctg ttcgtatcat gaccgcgtac 1380
tacaagggtg gtcgtgaccg tcttcgtatt ccccctaact tcagctcctg gaccocgggat 1440
gagtacggct gggagcattc tgctgtctcc gagggagcct ggaccaaggt gaacgacttc 1500
gtcaatgtgc agcgcagtca ctctcagatc atccgtgaga ttggtgccgc tagtacagtg 1560
ctcttgaaga acacgggtgc tcttcctttg accggcaagg aggttaaagt ggggtgttctc 1620
gggtgaagacg ctggttccaa cccgtggggg gctaacggct gccccgaccg cggctgtgat 1680
aacggcactc ttgctatggc ctggggtagt ggtactgcc aactccctta ccttgtcacc 1740
cccgagcagg ctatccagcg agaggctatc agcaacggcg gcaatgtctt tgctgtgact 1800
gataacgggg ctctcagcca gatggcagat gttgcatctc aatccagggt agtgccggct 1860
cttagaaaaa gaacgttctc tgaatgaagt tttttaacca ttgcgaacag cgtgtctttg 1920
gtgtttgtca acgccgactc tggagagggg ttcatcagtg tcgacggcaa cgaggggtgac 1980
cgcaaaaatc tcaactctgtg gaagaacggc gaggccgtca ttgacactgt tgtcagccac 2040
tgcaacaaca cgattgtggg tattcacagt gttgggcccg tcttgatcga ccggtggtat 2100
gataacccca acgtcactgc catcatctgg gccggcttgc ccggtcagga gagtggcaac 2160
tccctggtcg acgtgctcta tggccgcgct aaccccagcg ccaagacccc gttcacctgg 2220
ggcaagactc gggagtetta cggggctccc ttgctcaccg agcctaacia tggcaatggg 2280
gctccccagg atgatttcaa cgagggcgct ttcattgact accgtcactt tgacaagcgc 2340
aatgagacce ccatttatga gtttgccat ggcttgagct acaccacett tggttactct 2400
caccttcggg ttcaggccct caatagttcg agttcggcat atgtcccgc tagcggagag 2460
accaagcctg cgccaaccta tgggtgagatc ggtagtgcc cggactacct gtatcccag 2520
ggctcaaaa gaattaccaa gtttatttac ccttggtcga actcgaccga cctcgaggat 2580
tcttctgacg acccgaacta cggctgggag gactcggagt acattcccga aggcgctagg 2640
gatgggtctc ctcaaccctt cctgaaggct ggcggcgctc ctggtggtaa ccctaccctt 2700
tatcaggatc ttgttagggg gtcggccacc ataaccaaca ctggtaacgt cgccggttat 2760
gaagtccctc aattgggtgag tgaccegc atgttccttgcg ttgcaatttg gctaactcgc 2820
ttctagtatg tttcactggg cggaccgaac gagcctcggg tcgttctgcg caagttcgac 2880
cgaatcttcc tggctcctgg ggagcaaaaag gtttgacca cgactcttaa ccgtcgtgat 2940
ctcgccaatt gggatgtgga ggctcaggac tgggtcatca caaagtaccc caagaaagtg 3000
cacgtcggca gctcctcgcg taagctgcct ctgagagcgc ctctgccccg tgtctactag 3060

<210> 76
<211> 863
<212> PRT
<213> *Aspergillus fumigatus*

5

<400> 76

ES 2 468 366 T3

Met Arg Phe Gly Trp Leu Glu Val Ala Ala Leu Thr Ala Ala Ser Val
 1 5 10 15

Ala Asn Ala Gln Glu Leu Ala Phe Ser Pro Pro Phe Tyr Pro Ser Pro
 20 25 30

Trp Ala Asp Gly Gln Gly Glu Trp Ala Asp Ala His Arg Arg Ala Val
 35 40 45

Glu Ile Val Ser Gln Met Thr Leu Ala Glu Lys Val Asn Leu Thr Thr
 50 55 60

Gly Thr Gly Trp Glu Met Asp Arg Cys Val Gly Gln Thr Gly Ser Val
 65 70 75 80

Pro Arg Leu Gly Ile Asn Trp Gly Leu Cys Gly Gln Asp Ser Pro Leu
 85 90 95

Gly Ile Arg Phe Ser Asp Leu Asn Ser Ala Phe Pro Ala Gly Thr Asn
 100 105 110

Val Ala Ala Thr Trp Asp Lys Thr Leu Ala Tyr Leu Arg Gly Lys Ala
 115 120 125

Met Gly Glu Glu Phe Asn Asp Lys Gly Val Asp Ile Leu Leu Gly Pro
 130 135 140

Ala Ala Gly Pro Leu Gly Lys Tyr Pro Asp Gly Gly Arg Ile Trp Glu
 145 150 155 160

Gly Phe Ser Pro Asp Pro Val Leu Thr Gly Val Leu Phe Ala Glu Thr
 165 170 175

Ile Lys Gly Ile Gln Asp Ala Gly Val Ile Ala Thr Ala Lys His Tyr
 180 185 190

Ile Leu Asn Glu Gln Glu His Phe Arg Gln Val Gly Glu Ala Gln Gly
 195 200 205

Tyr Gly Tyr Asn Ile Thr Glu Thr Ile Ser Ser Asn Val Asp Asp Lys

ES 2 468 366 T3

210				215				220							
Thr 225	Met	His	Glu	Leu	Tyr 230	Leu	Trp	Pro	Phe	Ala 235	Asp	Ala	Val	Arg	Ala 240
Gly	Val	Gly	Ala	Val 245	Met	Cys	Ser	Tyr	Asn 250	Gln	Ile	Asn	Asn	Ser	Tyr 255
Gly	Cys	Gln	Asn 260	Ser	Gln	Thr	Leu	Asn 265	Lys	Leu	Leu	Lys	Ala 270	Glu	Leu
Gly	Phe	Gln 275	Gly	Phe	Val	Met	Ser 280	Asp	Trp	Ser	Ala	His 285	His	Ser	Gly
Val 290	Gly	Ala	Ala	Leu	Ala	Gly 295	Leu	Asp	Met	Ser	Met 300	Pro	Gly	Asp	Ile
Ser 305	Phe	Asp	Asp	Gly	Leu 310	Ser	Phe	Trp	Gly	Thr 315	Asn	Leu	Thr	Val	Ser 320
Val	Leu	Asn	Gly	Thr 325	Val	Pro	Ala	Trp	Arg 330	Val	Asp	Asp	Met	Ala	Val 335
Arg	Ile	Met	Thr 340	Ala	Tyr	Tyr	Lys	Val 345	Gly	Arg	Asp	Arg	Leu	Arg	Ile 350
Pro	Pro	Asn 355	Phe	Ser	Ser	Trp	Thr 360	Arg	Asp	Glu	Tyr	Gly 365	Trp	Glu	His
Ser 370	Ala	Val	Ser	Glu	Gly	Ala 375	Trp	Thr	Lys	Val	Asn 380	Asp	Phe	Val	Asn
Val 385	Gln	Arg	Ser	His	Ser 390	Gln	Ile	Ile	Arg	Glu 395	Ile	Gly	Ala	Ala	Ser 400
Thr	Val	Leu	Leu	Lys 405	Asn	Thr	Gly	Ala	Leu 410	Pro	Leu	Thr	Gly	Lys	Glu 415
Val	Lys	Val	Gly 420	Val	Leu	Gly	Glu	Asp 425	Ala	Gly	Ser	Asn	Pro 430	Trp	Gly
Ala	Asn 435	Gly	Cys	Pro	Asp	Arg	Gly 440	Cys	Asp	Asn	Gly	Thr 445	Leu	Ala	Met
Ala	Trp	Gly	Ser	Gly	Thr	Ala	Asn	Phe	Pro	Tyr	Leu	Val	Thr	Pro	Glu

ES 2 468 366 T3

450	455	460																	
Gln	Ala	Ile	Gln	Arg	Glu	Val	Ile	Ser	Asn	Gly	Gly	Asn	Val	Phe	Ala	465	470	475	480
Val	Thr	Asp	Asn	Gly	Ala	Leu	Ser	Gln	Met	Ala	Asp	Val	Ala	Ser	Gln	485	490	495	
Ser	Ser	Val	Ser	Leu	Val	Phe	Val	Asn	Ala	Asp	Ser	Gly	Glu	Gly	Phe	500	505	510	
Ile	Ser	Val	Asp	Gly	Asn	Glu	Gly	Asp	Arg	Lys	Asn	Leu	Thr	Leu	Trp	515	520	525	
Lys	Asn	Gly	Glu	Ala	Val	Ile	Asp	Thr	Val	Val	Ser	His	Cys	Asn	Asn	530	535	540	
Thr	Ile	Val	Val	Ile	His	Ser	Val	Gly	Pro	Val	Leu	Ile	Asp	Arg	Trp	545	550	555	560
Tyr	Asp	Asn	Pro	Asn	Val	Thr	Ala	Ile	Ile	Trp	Ala	Gly	Leu	Pro	Gly	565	570	575	
Gln	Glu	Ser	Gly	Asn	Ser	Leu	Val	Asp	Val	Leu	Tyr	Gly	Arg	Val	Asn	580	585	590	
Pro	Ser	Ala	Lys	Thr	Pro	Phe	Thr	Trp	Gly	Lys	Thr	Arg	Glu	Ser	Tyr	595	600	605	
Gly	Ala	Pro	Leu	Leu	Thr	Glu	Pro	Asn	Asn	Gly	Asn	Gly	Ala	Pro	Gln	610	615	620	
Asp	Asp	Phe	Asn	Glu	Gly	Val	Phe	Ile	Asp	Tyr	Arg	His	Phe	Asp	Lys	625	630	635	640
Arg	Asn	Glu	Thr	Pro	Ile	Tyr	Glu	Phe	Gly	His	Gly	Leu	Ser	Tyr	Thr	645	650	655	
Thr	Phe	Gly	Tyr	Ser	His	Leu	Arg	Val	Gln	Ala	Leu	Asn	Ser	Ser	Ser	660	665	670	
Ser	Ala	Tyr	Val	Pro	Thr	Ser	Gly	Glu	Thr	Lys	Pro	Ala	Pro	Thr	Tyr	675	680	685	
Gly	Glu	Ile	Gly	Ser	Ala	Ala	Asp	Tyr	Leu	Tyr	Pro	Glu	Gly	Leu	Lys				

690		695		700											
Arg	Ile	Thr	Lys	Phe	Ile	Tyr	Pro	Trp	Leu	Asn	Ser	Thr	Asp	Leu	Glu
705					710					715					720
Asp	Ser	Ser	Asp	Asp	Pro	Asn	Tyr	Gly	Trp	Glu	Asp	Ser	Glu	Tyr	Ile
				725					730					735	
Pro	Glu	Gly	Ala	Arg	Asp	Gly	Ser	Pro	Gln	Pro	Leu	Leu	Lys	Ala	Gly
			740					745					750		
Gly	Ala	Pro	Gly	Gly	Asn	Pro	Thr	Leu	Tyr	Gln	Asp	Leu	Val	Arg	Val
		755					760					765			
Ser	Ala	Thr	Ile	Thr	Asn	Thr	Gly	Asn	Val	Ala	Gly	Tyr	Glu	Val	Pro
	770					775					780				
Gln	Leu	Tyr	Val	Ser	Leu	Gly	Gly	Pro	Asn	Glu	Pro	Arg	Val	Val	Leu
785					790					795					800
Arg	Lys	Phe	Asp	Arg	Ile	Phe	Leu	Ala	Pro	Gly	Glu	Gln	Lys	Val	Trp
				805					810					815	
Thr	Thr	Thr	Leu	Asn	Arg	Arg	Asp	Leu	Ala	Asn	Trp	Asp	Val	Glu	Ala
			820					825					830		
Gln	Asp	Trp	Val	Ile	Thr	Lys	Tyr	Pro	Lys	Lys	Val	His	Val	Gly	Ser
		835					840					845			
Ser	Ser	Arg	Lys	Leu	Pro	Leu	Arg	Ala	Pro	Leu	Pro	Arg	Val	Tyr	
	850					855					860				

5 <210> 77
 <211> 30
 <212> ADN
 <213> Aspergillus fumigatus

10 <400> 77
 ggactgcgca ccatgagatt cggttggctc 30

ES 2 468 366 T3

<210> 78

<211> 30

<212> ADN

5 <213> *Aspergillus fumigatus*

<400> 78

tcgccacgga gctfactagt agacacgggg 30

10

REIVINDICACIONES

- 5 1. Polipéptido aislado con actividad de mejora celulolítica y una secuencia de aminoácidos que tiene al menos 85%, incluso más preferiblemente al menos 90%, de la forma más preferible al menos 95%, e incluso de la forma más preferible al menos 97% de identidad con los aminoácidos 19 a 226 de la SEC ID n°: 8.
2. Polipéptido según la reivindicación 1, que se codifica por el polinucleótido contenido en el plásmido pTter61 E que está contenido en *E. coli* NRRL B-30814.
- 10 3. Polinucleótido aislado comprendiendo una secuencia de nucleótidos que codifica el polipéptido según cualquiera de las reivindicaciones 1-2.
4. Construcción de ácidos nucleicos comprendiendo el polinucleótido según la reivindicación 3 unido operativamente a una o más secuencias de control que dirigen la producción del polipéptido en un huésped de expresión.
- 15 5. Vector de expresión recombinante comprendiendo la construcción de ácidos nucleicos según la reivindicación 4.
6. Célula huésped recombinante comprendiendo la construcción de ácidos nucleicos según la reivindicación 4.
- 20 7. Método para la producción del polipéptido según cualquiera de las reivindicaciones 1-2, comprendiendo (a) cultivo de una célula huésped comprendiendo una construcción de ácidos nucleicos comprendiendo una secuencia de nucleótidos que codifica el polipéptido bajo condiciones propicias para la producción del polipéptido; y (b) recuperación del polipéptido.
- 25 8. Composición detergente comprendiendo un polipéptido con actividad de mejora celulolítica según las reivindicaciones 1-2, una actividad celulolítica, y un tensioactivo
9. Método para degradación o conversión de un material celulósico, comprendiendo: tratamiento del material celulósico con una cantidad eficaz de una proteína celulolítica en presencia de una cantidad eficaz del polipéptido con actividad de mejora celulolítica según cualquiera de las reivindicaciones 1-2, donde la presencia del polipéptido con actividad de mejora celulolítica aumenta la degradación del material celulósico en comparación con la ausencia del polipéptido que tiene actividad de mejora celulolítica
- 30 10. Método según la reivindicación 9, donde una o más enzimas celulolíticas se seleccionan del grupo que consiste en una celulasa, endoglucanasa, celobiohidrolasa y beta-glucosidasa.
- 35 11. Método para la producción de una sustancia orgánica, comprendiendo:
- 40 (a) sacarificación de un material celulósico con una cantidad eficaz de una proteína celulolítica en presencia de una cantidad eficaz del polipéptido que tiene actividad de mejora celulolítica según cualquiera de las reivindicaciones 1-2, donde la presencia del polipéptido que tiene actividad de mejora celulolítica aumenta la degradación de material celulósico en comparación con la ausencia del polipéptido que tiene actividad de mejora celulolítica,
- 45 (b) fermentación del material celulósico sacarificado del paso (a) con uno o más microorganismos fermentadores; y
- (c) recuperación de la sustancia orgánica de la fermentación.
12. Método según la reivindicación 11, donde el material celulósico es restos de maíz.
- 50 13. Método según las reivindicaciones 11-12, donde una o más enzimas celulolíticas se seleccionan del grupo que consiste en una celulasa, endoglucanasa, celobiohidrolasa y beta-glucosidasa.
14. Método según cualquiera de las reivindicaciones 11-13, donde la sustancia orgánica es un alcohol, ácido orgánico, cetona, aminoácido o gas.
- 55 15. Método según la reivindicación 14, donde el alcohol es arabinitol, butanol, etanol, glicerol, metanol, 1,3-propanodiol, sorbitol, o xilitol.

AATTGAAGGAGGAGTGGGGAGTGCCACCACCAAGTCAAGCGGCTGTCAACTAACCAAGGATGGGAACAGTTCCGGCTCGCCTTGCCCCGAGG 90
GCAGCGTTCCCTGATGGGACGAACCATGGACTGGGGTCAGTGTGTATAAAAAGTTCAAATCGATGATCTCTCAGATGGCGCTGTCTGG 180
GGTGTCTGGCCTTTCCATCCTCGCAACCTGGTATCCCACTAGTCCAGCGTTTCGGCACCAATGAAGTCGTTCCACCATGCCCCCTTGCCA 270
M K S F T I A A L A
GCCCTATGGGGCCAGGCGGCCCCCGGACCTTCCAGGACCTCTGGATGATGGAGTCCGACTACGGCTCGCAATGTGTCCGCTC 360
A L W A Q E A A A H A T F Q D L W I D G V D Y G S Q C V R L
CCGGCTCCAACCTCCCCTCACCAATGTTGGCTCCGACGATATCCGATGCAATGTCCGACCTCGAGGCCACCGTCAAGTCCCGCTC 450
P A S N S P V T N V A S D D I R C N V G T S R P T V K C P V
AAGCCGGCTCCACGGTCCAGATCGAGATGCACCCAGGTTCCGACCCCTCTCTCGGTAGGCCCCCAGCTACTATATGGCACATAACACGAC 540
K A G S T V T I E M H Q
CTCCAGCAACCTGGCGACCGGTCTTGGGCCAACGAGGCTATCGGGGGGACCACTACGGCCCCCGTAATGGTGATACATGTCCAAGTTCGAT 630
Q P G D R S C A N E A I G G D H Y G P V M V Y M S K V D
GACGGGTGACCGGACGGTTCATCGGGCTGGTTCAAAGTGTCCAGGACAGTGGGCCAAGAACCCCGTTCGGGTTCGACGGCGGACGAC 720
D A V T A D G S S G W F K V F Q D S W A K N P S G S T G D D
GACTACTGGGGCACCAGGACCTCAACTCGTGTGGGCAAGATGAACGTCAGATCCCCGAAGACATCGAGCCGGGCGACTACCTGCTC 810
D Y W G T K D L N S C C G K M N V K I P E D I E P G D Y L L
CGGCGGAGGTTATCGCGTGCACGTGGCCCGCCAGCTCGGGGGGGCCAGTTCATCATGTCCCTGCTACCAGCTGACCCGTGACGGGCTCC 900
R A E V I A L H V A A S S G G A Q F Y M S C Y Q L T V T G S
GGCAGGCCACCCCTCGACCGTGAATTTCCCGGGCCCTACTCGGCCAGCGACCCGGGCATCTGTGATCAACATCCAGCGCCCATGTCCG 990
G S A T P S T V N F P G A Y S A S D P G I L I N I H A P M S
ACCTACGTCGTCGCGCCCGGACCGTGTACGGGGGCGCTCGACCAAGTGGCTGGCAGCTCCTGCTCCGGTCCGAGCGGACCTGCACG 1080
T Y V V P G P T V Y A G G S T K S A G S S C S G C E A T C T
GTTGGTTCCGGCCCCAGCGGACACTGACGACGACCCACCTCCACCGGACCCCGGACCTCCGGCCCCCTGGGGGGGGGCTCCGGCTGCACG 1170
V G S G P S A T L T Q P T S T A T A T S A P G G G S G C T
GCGGCCAAGTACCAGCAGTGGCGGACCGGCTACACTGGGTGCACCCACTGGCTGTAAGTTCCTCCCTCGTGATATGCAGCGGGAACACCG 1260
A A K Y Q Q C G G T G Y T G C T T C A
TCTGGACTGTTTGTACTCGCGTCGTTAGTCCGGGTCTACCTGCAGCGCCGCTCTCGCCCTCCGTACTACTCGCAGTGCCTTAAGCCGGG 1350
S G S T C S A V S P P Y Y S Q C L .

Fig. 1A

AGCGCTTGCTCAGCGGGCTGCTGTGAAGGAGCTCCCATGTCCCCATGCCCGGAGTACCGGGCTGAGCGCCCAATTCTTGTATA 1440
TAGTTGAGTTTTCCCAATCATGAATACATAATGCATATGCATGGACTGTTGCCGTGTCAGTCTACATCCTTTGCTCCACTGAACITGTGAGA 1530
CCCCATGTCATCCGGACCAATTCGATCCGGTCTCGCTCTACCATCTCGGTTGATGGGCTTGAGAGTCACTGGCACGTCCTCGGCCG 1620
GTAATGAAATGTGGAGGAAAGTGTGAGCTGTCTGACGCCACTCGGGCTGATGAGACGTTGAGCGGCCCCACACTGGTGTCTGTAAAGCC 1710
AGCACACAAAAGAAATACCTCCAGGATGGCCCATFAGCCCAATATACAGTATCAGGGATGCAAAAAGTGCAAAAGTAAGGGCTCAATCGG 1800
GGATCGAACCCGAGACCTCGCACATGACTTATTTCAAGTCAGGGGT

Fig. 1B

ACCCCGGGATCACTGCCCTAGGAACCAGCACACCTCGGTCCAATCATGCGGTTGACGCCCCTCTCCGCCCTCGCTCTTG
80
M R F D A L S A L A L
CGCCGCTTGTGGCTGGCCACGGCGCCGTGACCAGCTACATCATCGGCGGCAAAACCTATCCCGGCTACGAGGGCTTCTCG
160
A P L V A G H G A V T S Y I I G G K T Y P G Y E G F S
CTGCCTCGAGCCCGCCGACGATCCAGTACCAGTGGCCCGACTACAACCCGACCTGAGCGTGACCACCCGAAGATGCG
240
P A S S P P T I Q Y Q W P D Y N P T L S V T D P K M R
CTGCAACGGCGGCACCTCGGCAGAGCTCAGCGCGCCCGTCCAGGCCGGCGAGAACGTGACGGCCGTCTGGAAGCAGTGA
320
C N G G T S A E L S A P V Q A G E N V T A V W K Q W
CCACCAGCAAGGCCCGTCATGGTCTGGATGTTCAAGTGCCCCGGCGACTTCTCGTCGTGCCACGGCGACGGCAAGGGC
400
T H Q Q G P V M V W M F K C P G D F S S S H G D G K G
TGGTTC AAGATCGACCAGCTGGGCCTGTGGGGCAACAACCTCAACTCGAACAAC TGGGGCACCGCGATCGTCTACAAGAC
480
W F K I D Q L G L W G N N L N S N N W G T A I V Y K T
CCTCAGTGGAGCAACCCGATCCCCAAGAACCTCGCGCCGGGCAACTACCTCATCCGCCACGAGCTGCTCGCCCTGCACC
560
L Q W S N P I P K N L A P G N Y L I R H E L L A L H
AGGCCAACACGCCGAGTCTACGCCGAGTGCGCCAGCTGGTCTCTCCGGCAGCGGCTCCGCCCTGCCCCGTCGGAC
640
Q A N T P Q F Y A E C A Q L V V S G S G S A L P P S D
TACCTCTACAGCATCCCCGTCTACGCGCCCCAGAACGCCCGGCATCACCGTGAGTGGGCTTCCGTTCCGCGGGCAGCT
720
Y L Y S I P V Y A P Q N D P G I T
CTGTGAAATCTTGCTGACGATGGGCTAGGTTGACATCTACAACGGCGGGCTTACCTCCTACACCCCGCCGGCGGGCCC
800
V D I Y N G G L T S Y T P P G G P
GTCTGGTCTGGCTTCGAGTTTTAGGCGCATTGAGTCGGGGGTACGAGGGGAAGGCATCTGTTCCGATGAGCGTGGGTAC
880
V W S G F E F .

MRFDALSALALAPLVAGHGA VTSYI IGGKTYPGYEGFSPASSPPTIQYQWPDYNPTLSVTDPKMRCNGGTS AELSAPVQA
GENVTAVWKQWTHQQGPVMVWMFKCPGDFSSSHGDGKGFKIDQLGLWGNNLNSNNWGTAIVYKTLQWSNPIPKNLAPGN
YLIRHELLALHQANTPQFYAECAQLVVS GSGSALPPSDYLYSIPVYAPQNDPGITVDIYNGGLTSYTPPGPVWSGF EF

Fig. 2

CTCCTGTTCTGGGCCACCGCTTGTTCCTGCACTATTGGTAGAGTTGGTCTATTGCTAGAGTTGGCCATGCTTCTCACA 80
M L L T
TCAGTCTCGGCTCGGCTGCCCTGCTTGCTAGCGGCGCTGCGGCACACGGCGCCGTGACCAGCTACATCATCGCCGGCAA
160
S V L G S A A L L A S G A A A H G A V T S Y I I A G K
GAATTACCGGGTGGGTAGCTGATTATTGAGGGCGCATTCAAGGTTCAACCAGGTTGTCATGGCTGACAACCGGCTGGC
240
N Y P G
AGATACCAAGGCTTTTCTCCTGCGAACTCGCCGAACGTCATCCAATGGCAATGGCATGACTACAACCCCGTCTTGTGCGTG
320
Y Q G F S P A N S P N V I Q W Q W H D Y N P V L S C
CAGCGACTCGAAGCTTCGCTGCAACGGCGGCACGTCGGCCACCCTGAACGCCACGGCCGACCGGGCGACACCATCACCG
400
S D S K L R C N G G T S A T L N A T A A P G D T I T
CCATCTGGGCGCAGTGGACGCACAGCCAGGGCCCATCCTGGTGTGGATGTACAAGTGCCCGGGCTCCTTCAGCTCCTGT
480
A I W A Q W T H S Q G P I L V W M Y K C P G S F S S C
GACGGCTCCGGCGCTGGCTGGTTCAAGATCGACGAGGCCGGCTTCCACGGCGACGGCGTCAAGGTCTTCTCGACACCGA
560
D G S G A G W F K I D E A G F H G D G V K V F L D T E
GAACCCGTCCGGTGGGACATCGCCAAGCTCGTCGGCGGCAACAAGCAGTGGAGCAGCAAGGTCCCCGAGGGCCTCGCCC
640
N P S G W D I A K L V G G N K Q W S S K V P E G L A
CCGGCAACTACCTCGTCCGCCACGAGTTGATCGCCCTGCACCAGGCCAACACCCGAGTTCTACCCGGAGTGCGCCCAG
720
P G N Y L V R H E L I A L H Q A N N P Q F Y P E C A Q
GTCGTATCACCGGCTCCGGCACC GCGCAGCCGATGCCTCATAAAGCGGCTATCCCCGGCTACTGCAACCAGAATGA
800
V V I T G S G T A Q P D A S Y K A A I P G Y C N Q N D
CCCGAACATCAAGGTGAGATCCAGGCGTAATGCAGTCTACTGCTGGAAGAAAGTGGTCCAAGCTAAACCGGCTCCAGG
880
P N I K
TGCCCATCAACGACCACTCCATCCCTCAGACCTACAAGATTCCCAGCCCTCCCGTCTTCAAGGGCACCGCCAGCAAGAAG
960
V P I N D H S I P Q T Y K I P G P P V F K G T A S K K
GCCCGGACTTCACCGCCTGAAGTTGTTGAATCGATGGAG 1000
A R D F T A .

MLLTSVLGSAALLASGAAAHGAVTSYIIAGKNYPGYQGFSPANSPNVIQWQWHDYNPVLS CSDSKLRCNGGTSATLNATA
APGDTITAIWAQWTHSQGPILVWMYKCPGSFSSCDGSGAGWFKIDEAGFHGDGVKVF LD TENPSGWDIAKLVGGNKQWSS
KVPEGLAPGNLVRHELIALHQANNPQFYPECAQVVI TGS GTAQPDASYKAAIPGYCNQNDPNIKVPI NDHSIPQTYKIP
GPPVFKGTASKKARDFTA

Fig. 3

ATGCTCGCAAACGGTGCCATCGTCTTCTGCGCCGCCCTCGGCGTCAGTGGCCACTACACCTGGCCACGGGTTAACGA
 80
M L A N G A I V F L A A A L G V S G H Y T W P R V N D
 CGGCGCCGACTGGCAACAGGTCCGTAAGGCGGACAACGGCTACGTCGGGGATGTCACGTCGCCACAGA
 160
 G A D W Q Q V R K A D N W Q D N G Y V G D V T S P Q
 TCCGCTGTTTCCAGGCGACCCCGTCCCCGGCCCCATCCGTCTCAACACCACGGCCGGCTCGACCGTGACCTACTGGGCC
 240
 I R C F Q A T P S P A P S V L N T T A G S T V T Y W A
 AACCCGACGCTACCACCCCGGCCTGTGCAGTTTTACATGGCCCGCTGCCGATGGCGAGGACATCAACTCGTGGAA
 320
 N P D V Y H P G P V Q F Y M A R V P D G E D I N S W N
 CGGCGACGGCGCCGTGTGGTTCAAGGTGTACGAGGACCATCTACCTTTGGCGCTCAGCTCACATGGCCCAGCACGGGCA
 400
 G D G A V W F K V Y E D H P T F G A Q L T W P S T G
 AGAGCTCGTTCGCGGTTCCCATCCCCCGTGCATCAAGTCCGGCTACTACCTCCTCCGGGCGGAGCAAATCGGCCTGCAC
 480
 K S S F A V P I P P C I K S G Y Y L L R A E Q I G L H
 GTCGCCAGAGCGTAGGCGGAGCGCAGTTCTACATCTCATGCGCCAGCTCAGCGTCACCGGCGGCGGCGAGCACCGAGCC
 560
 V A Q S V G G A Q F Y I S C A Q L S V T G G G S T E P
 GCCGAACAAGGTGGCCTTCCCCGGCGCTTACAGTGCAGCGGACCCGGGCATTCTGATCAACATCTACTACCCTGTTCCCA
 640
 P N K V A F P G A Y S A T D P G I L I N I Y Y P V P
 CGTCTACCAGAACCCCGGCCCGGCGTCTTCAGCTGCTGA 681
 T S Y Q N P G P A V F S C .

MLANGAIVFLAALGVSGHYTWPRVNDGADWQQVRKADNWQDNGYVGDVTSPOIRCFQATPSPAPSVLNTTAGSTVTYWA
NPDVYHGPVQFYMARVPDGEDINSWNGDGAVWFKVYEDHPTFGAQLTWPSTGKSSFAVPIPPCIKSGYLLRAEQIGLH
VAQSVGGAQFYISCAQLSVTGGGSTEPPNKVAFPGAYSATDPGILINIYYPVPTS YQNPGPAVFSC

Fig. 4

ATGAAGGGACTTTTCAGTGCCGCCGCCCTCTCCCTGGCCGTCCGCCAGGCTTCGGCCCATTACATCTTCCAGCAACTCTC
 80
M K G L F S A A A L S L A V G Q A S A H Y I F Q Q L S
 CATCAACGGGAACCAGTTTCCGGTGTACCAATATATTCCGAAGAACCAATTATAACAGTCCCCTTACCGATCTCACGT
 160
 I N G N Q F P V Y Q Y I R K N T N Y N S P V T D L T
 CCGACGATCTTCGGTGAATGTCCGGCGCCAGGGTGTCTGGGACAGACACCGTCACGGTGAAGGCCGGCGACCAGTTCACC
 240
 S D D L R C N V G A Q G A G T D T V T V K A G D Q F T
 TTCACCCTTGACACCCCTGTTTACCACCAGGGGCCATCTCCATCTACATGTCCAAGGCCCGGGCGGGCGTCAGACTA
 320
 F T L D T P V Y H Q G P I S I Y M S K A P G A A S D Y
 CGATGGCAGCGGGCTGGTTCAAGATCAAGGACTGGGGCCCGACTTTCAACGCCGACGGCACGGCCACCTGGGACATGG
 400
 D G S G G W F K I K D W G P T F N A D G T A T W D M
 CCGGCTCATAACCTACAACATCCCAGCTGCATTCCCAGCGGGACTATCTGCTCCGCATCCAGTCGCTGGCCATCCAC
 480
 A G S Y T Y N I P T C I P D G D Y L L R I Q S L A I H
 AACCCCTGGCCGGGGCATCCCGCAGTTCTACATCTCCTGCGCCAGATCACCGTGACGGCGGGCGCAACGGCAACCC
 560
 N P W P A G I P Q F Y I S C A Q I T V T G G G N G N P
 TGGCCCGACGGCCCTCATCCCGCGCCTTCAAGGACACCGACCCGGGCTACACGGTGAACATCTACACGAACCTCCACA
 640
 G P T A L I P G A F K D T D P G Y T V N I Y T N F H
 ACTACACGGTCCCAGCCCGGAGGTCTTCAGCTGCAACGGCGGGCTCGAACCCGCCCCCGCGGTGAGTAGCAGCAG
 720
 N Y T V P G P E V F S C N G G G S N P P P P V S S S T
 CCCGACACGACGCTGGTCACGTGACGCGCACCACGTCTCCACGTCTCCGCTCGACGCGGGCTCGACCGGGCGG
 800
 P A T T T L V T S T R T T S S T S S A S T P A S T G G
 CTGACCGTCCGAAGTGGGGCCAGTGCGGGGCAACGGGTACACCGGCTGCACGACCTGCGCGGGCCGGTCCACCTGCA
 880
 C T V A K W G Q C G G N G Y T G C T T C A A G S T C
 GCAAGCAGAACGACTACTACTCGCAGTGTGTAAGGGAGGCCGAAAGCATGAGGTGTTGAAGAGGAGAGGGGGTC
 960
 S K Q N D Y Y S Q C L .

MKGLFSAALSSLAVGQASAHYIFQQLSINGNQFPVYQYIRKNTNYSNVTDLTSDDLRCNVGAQGAGTDTVTVKAGDQFT
 FTLDTPVYHQGPISIIYMSKAPGAASDYDGGGWFKIKDWGPTFNADGTATWDMAGSYTYNIPTCIPDGDYLLRIQSLAIH
 NPWPAGIPQFYISCAQITVTGGNGNPGPTALIPGAFKDTDPGYTVNIYTNFHNITVPGPPEVFSNNGGSPPPPVSSST
 PATTTLVTSRRTTSSTSSASTPASTGGCTVAKWGCNGYTGCTTCAAGSTCSKQNDYYSQCL

Fig. 5

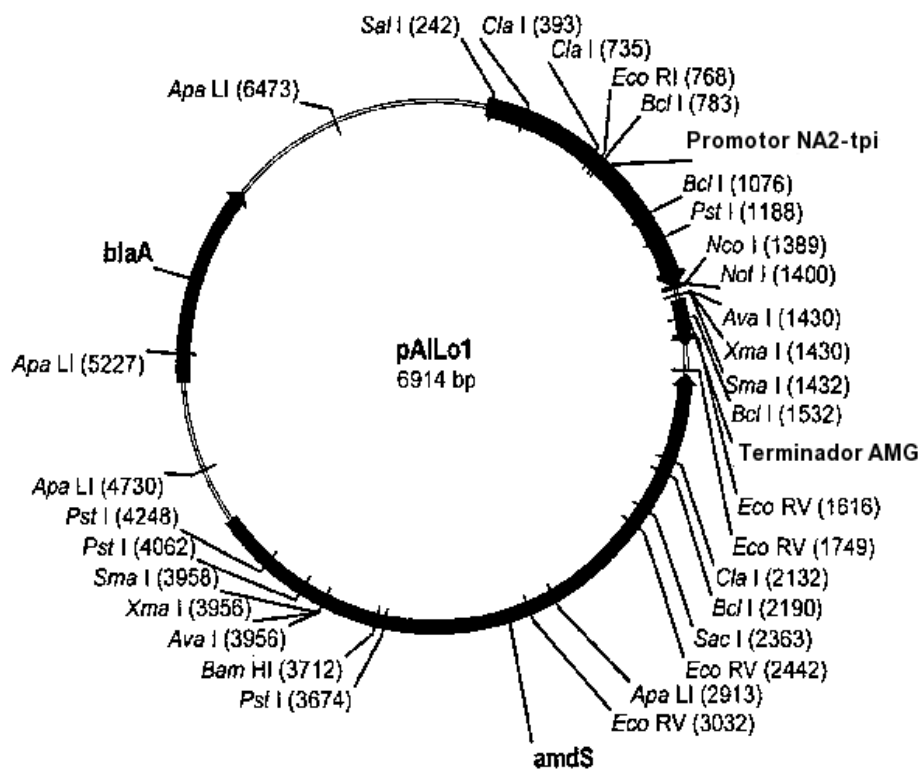


Fig. 6

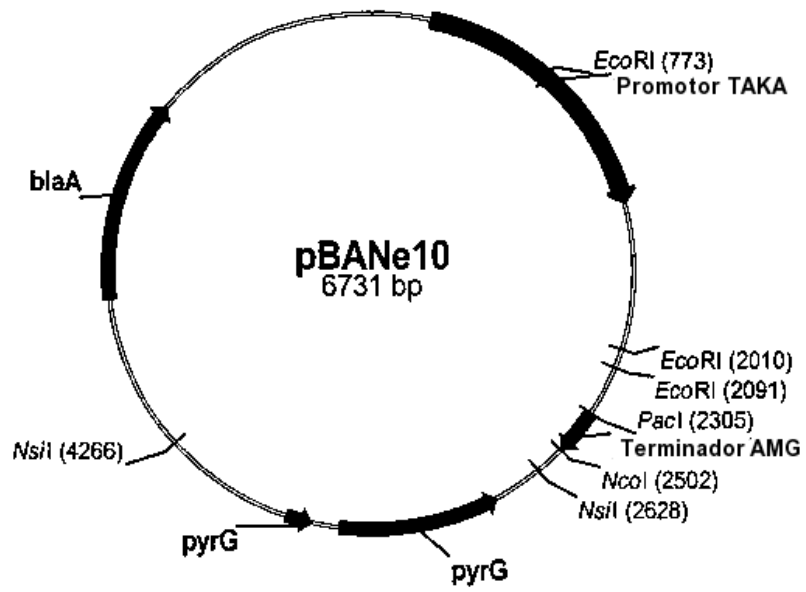


Fig. 7

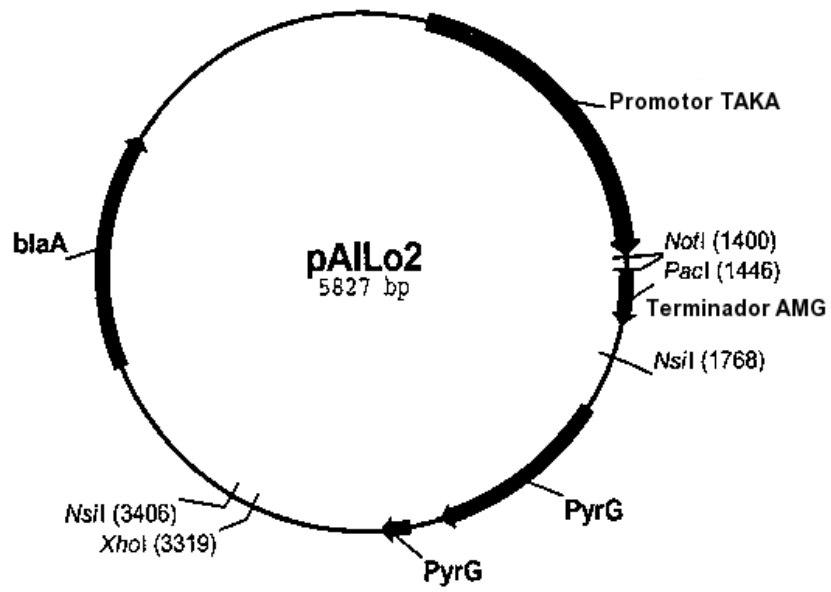


Fig. 8

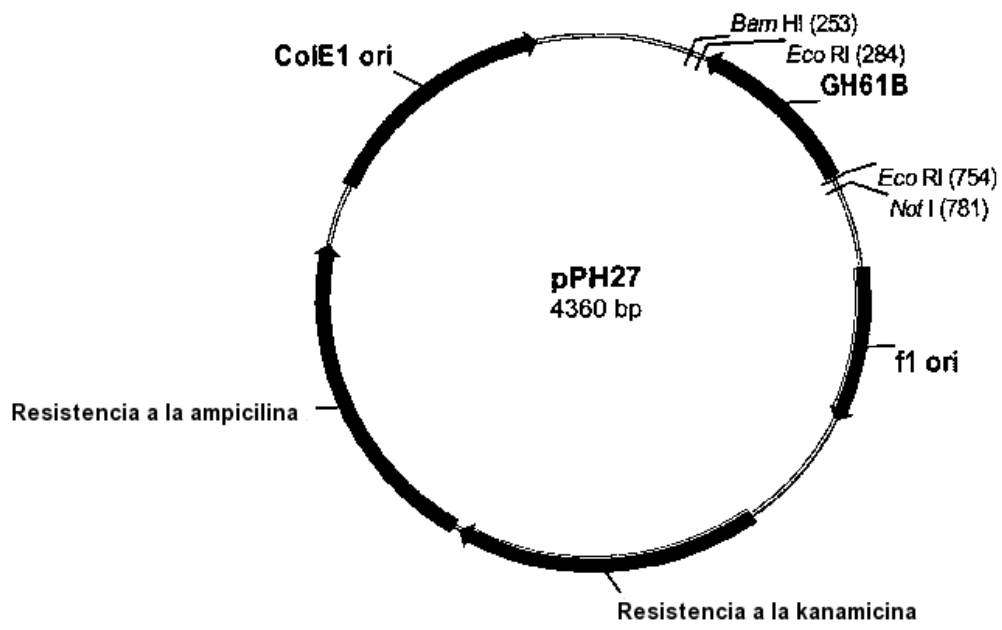


Fig. 9

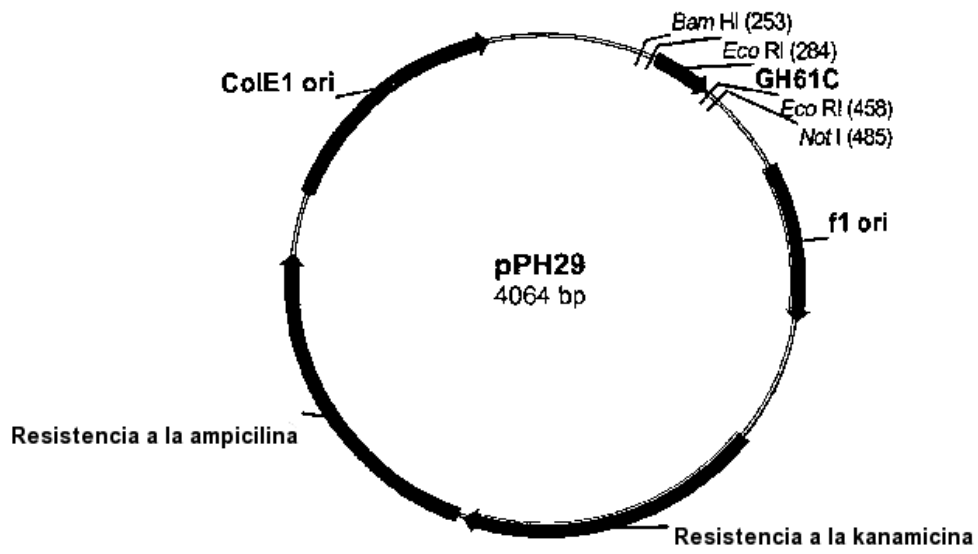


Fig. 10

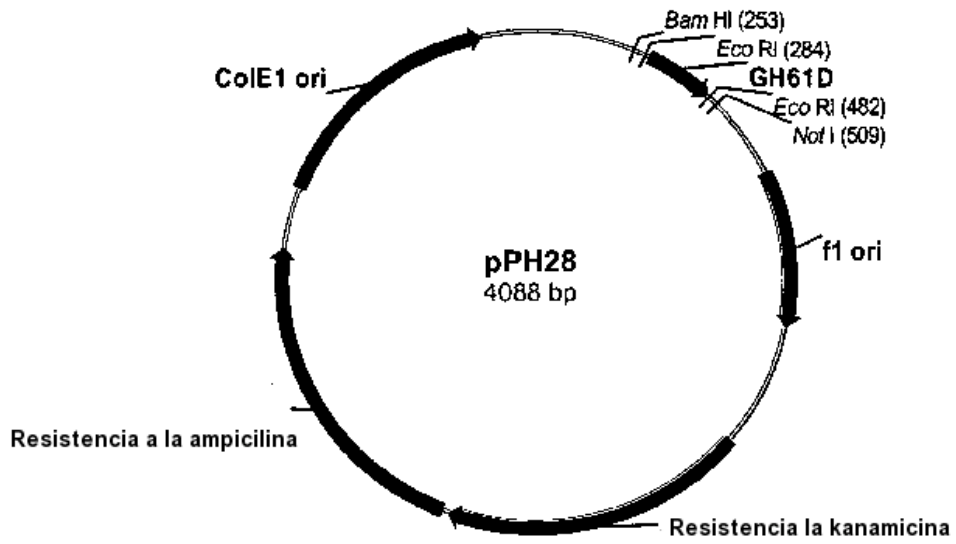


Fig. 11

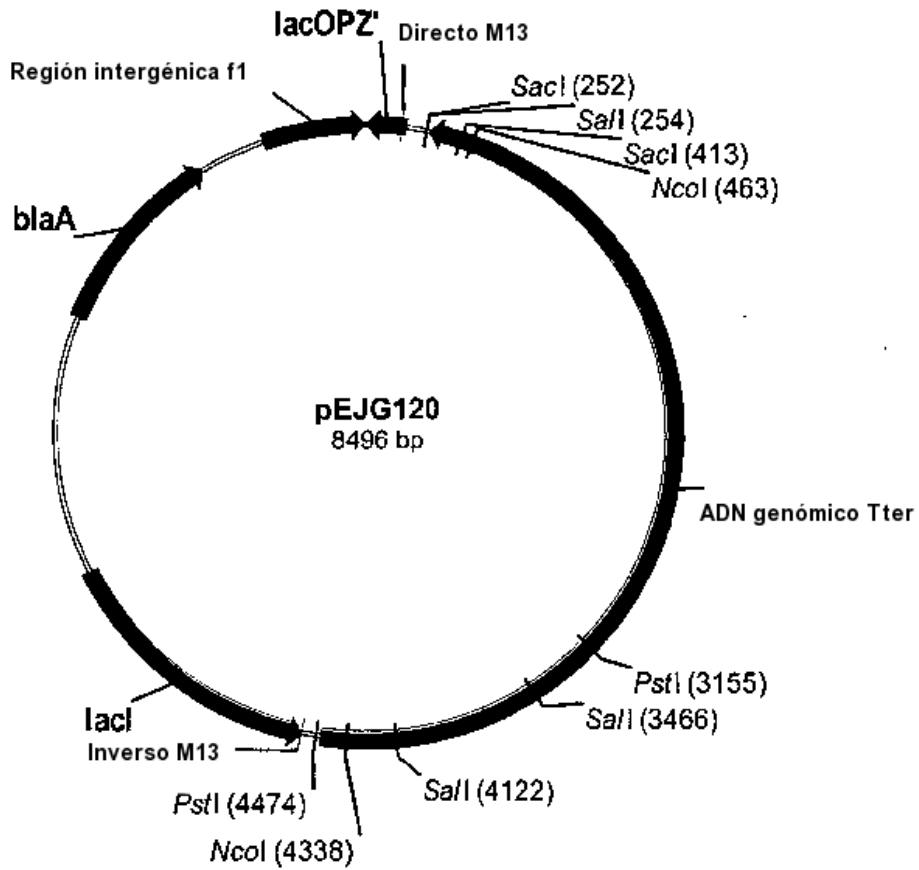


Fig. 12

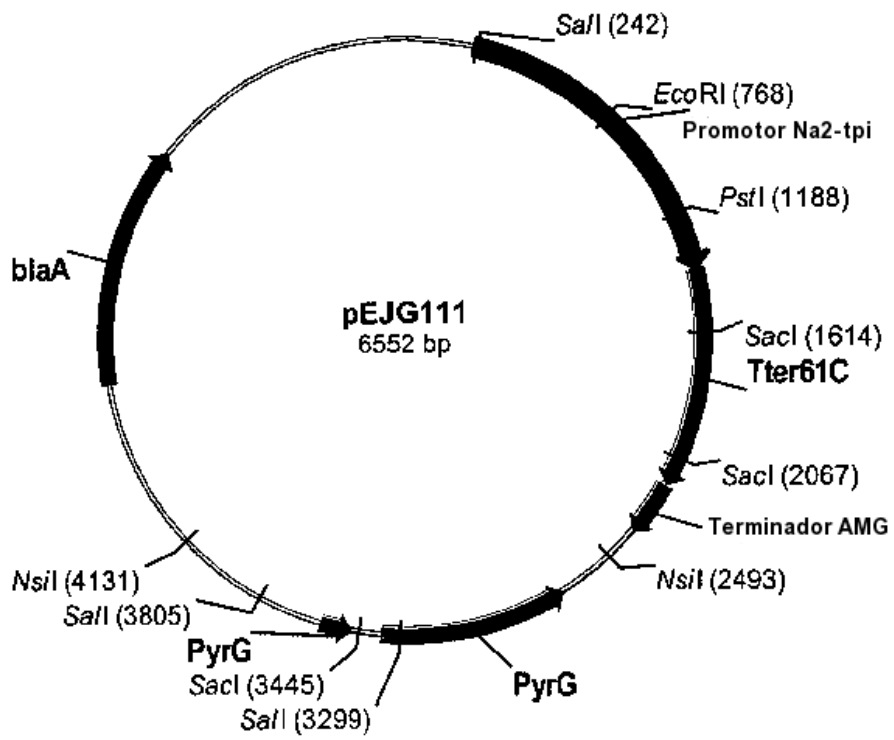


Fig. 13

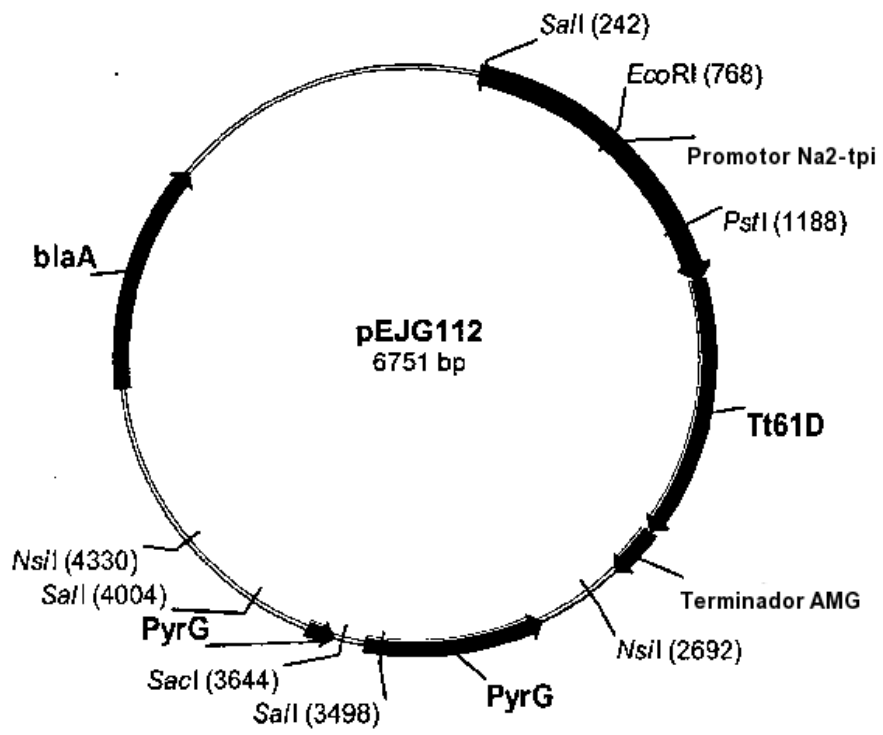


Fig. 14

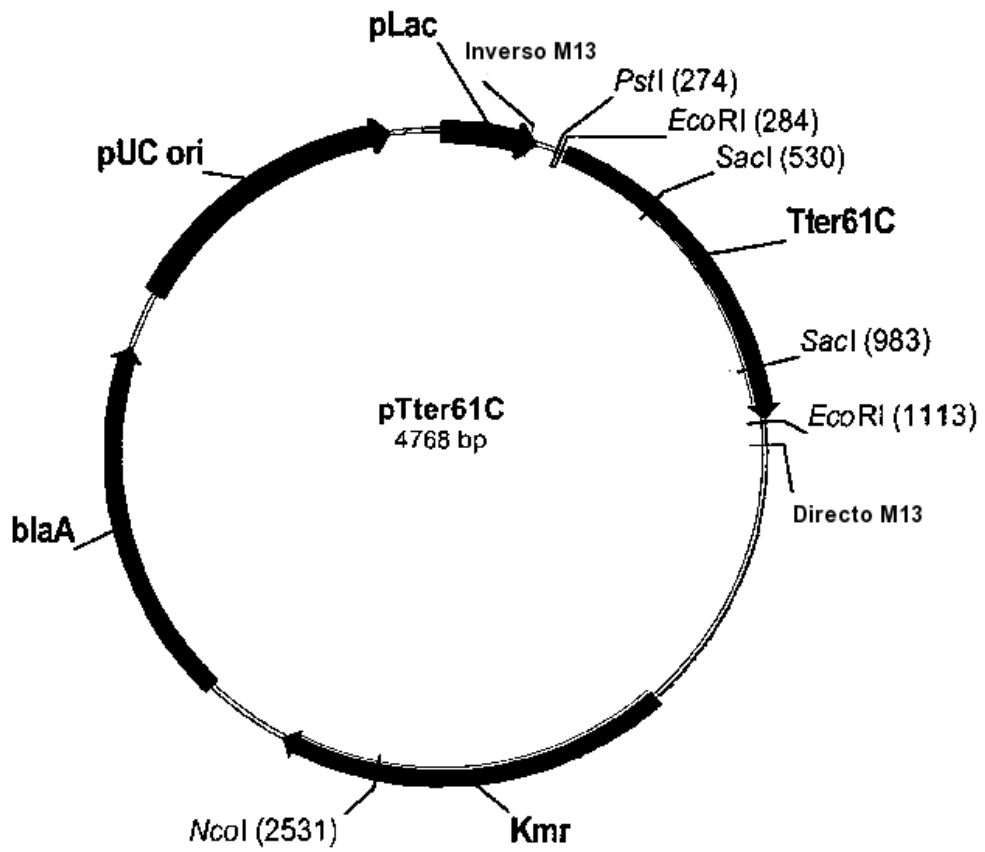


Fig. 15

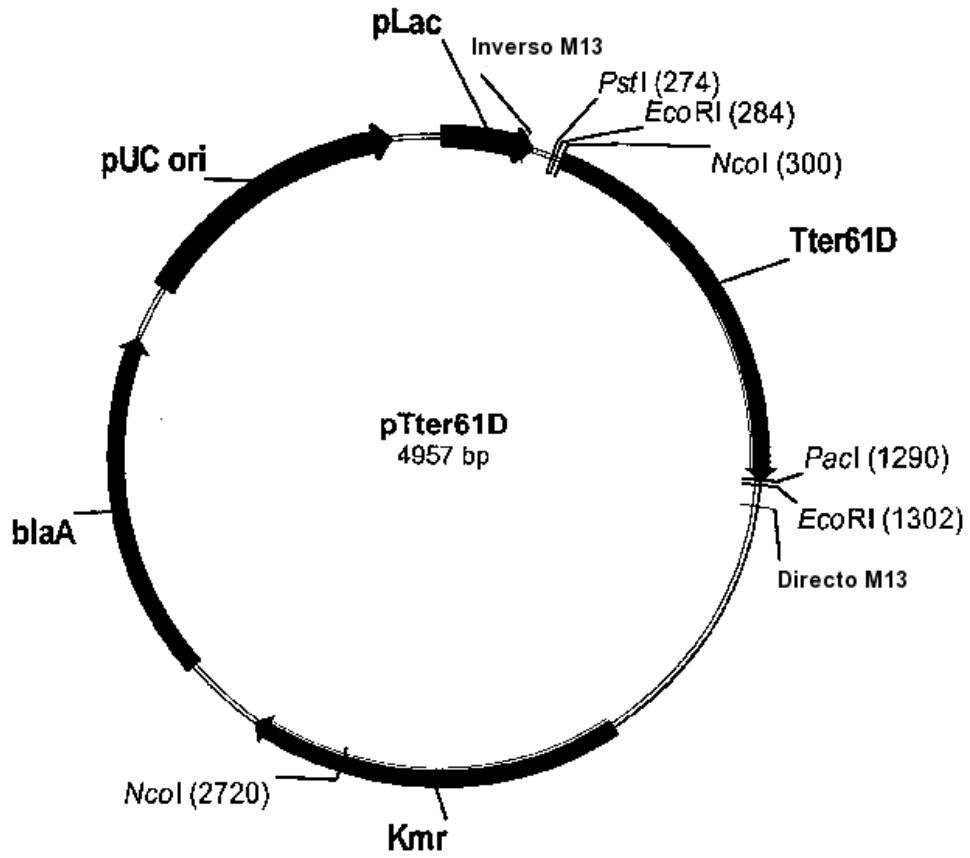


Fig. 16

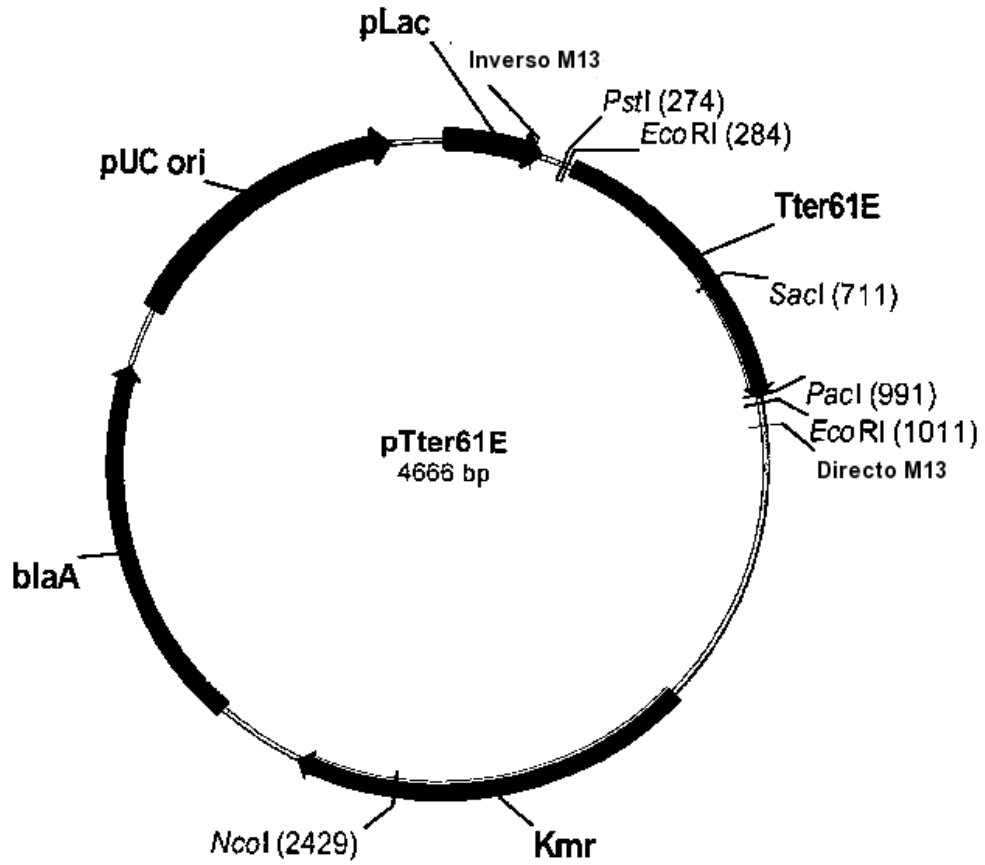


Fig. 17

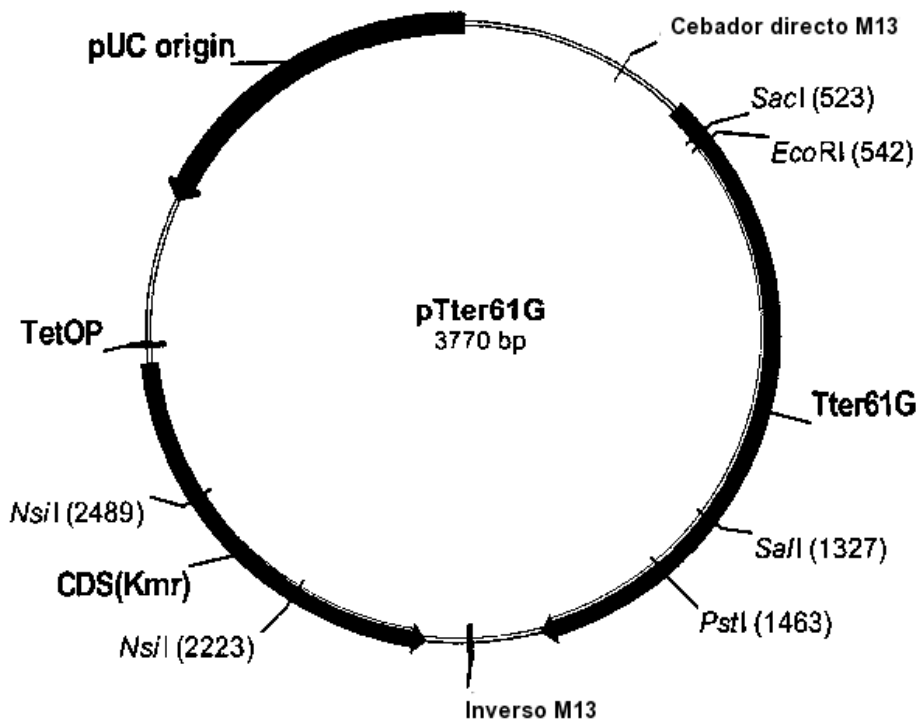


Fig. 18

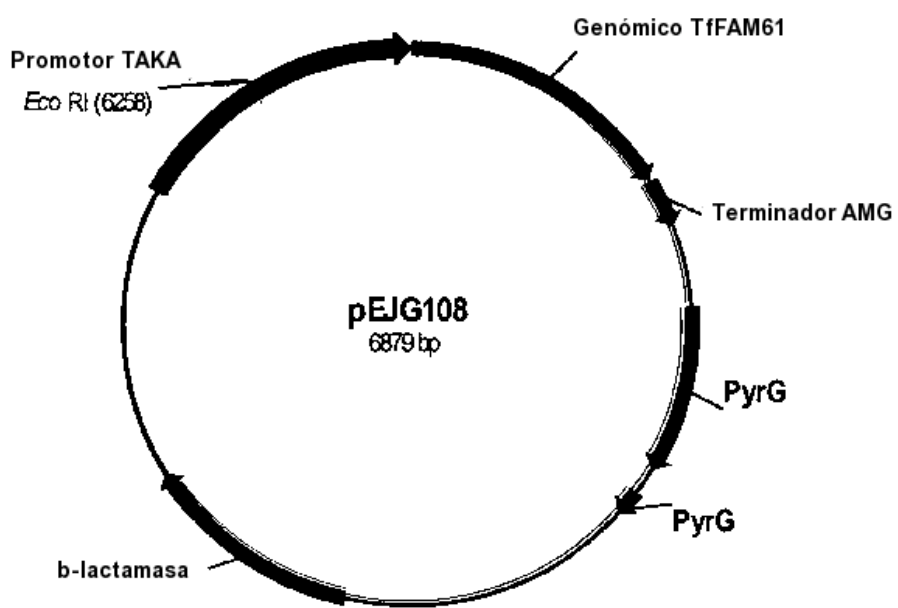


Fig 19

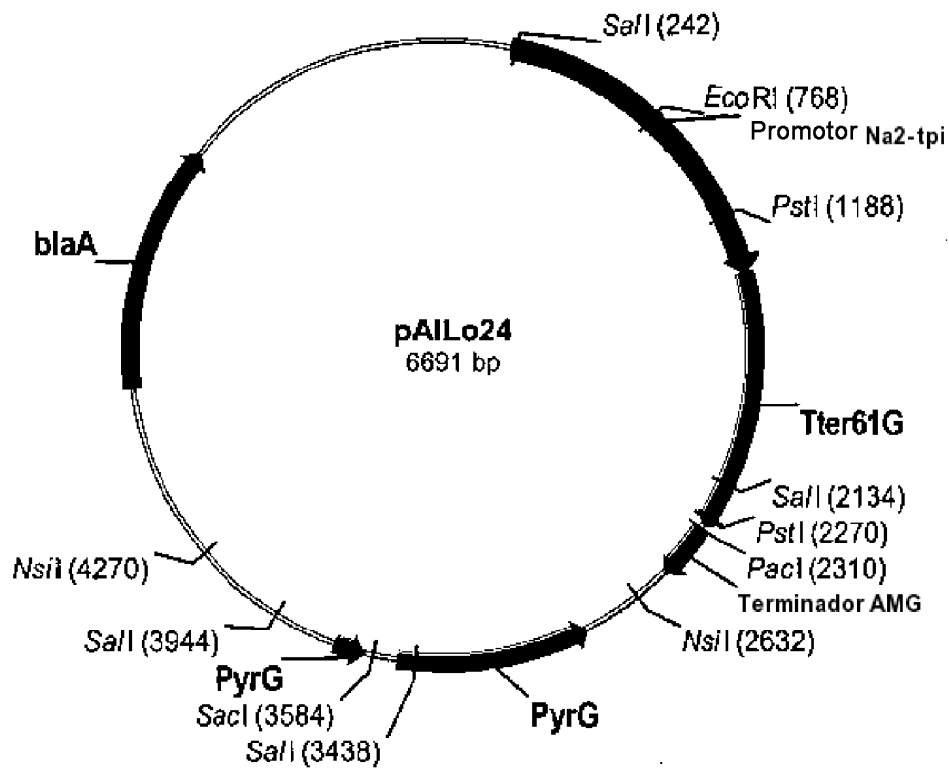


Fig. 20

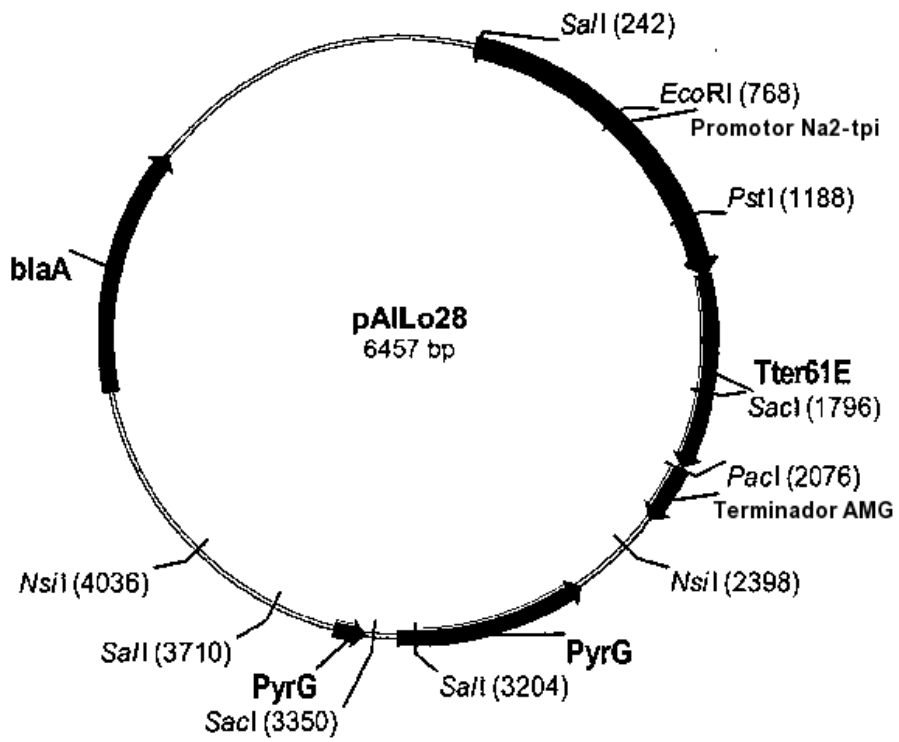


Fig. 21

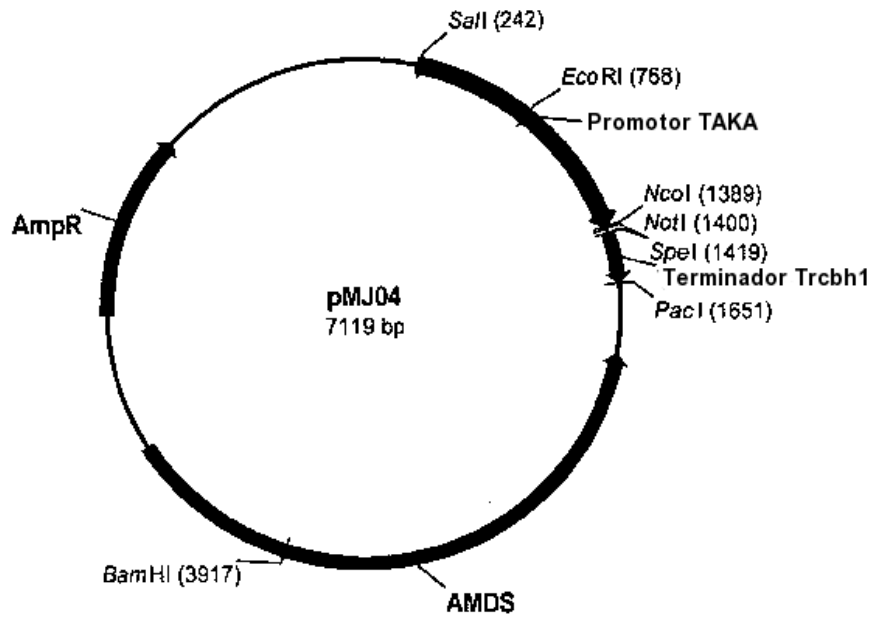


Fig. 22

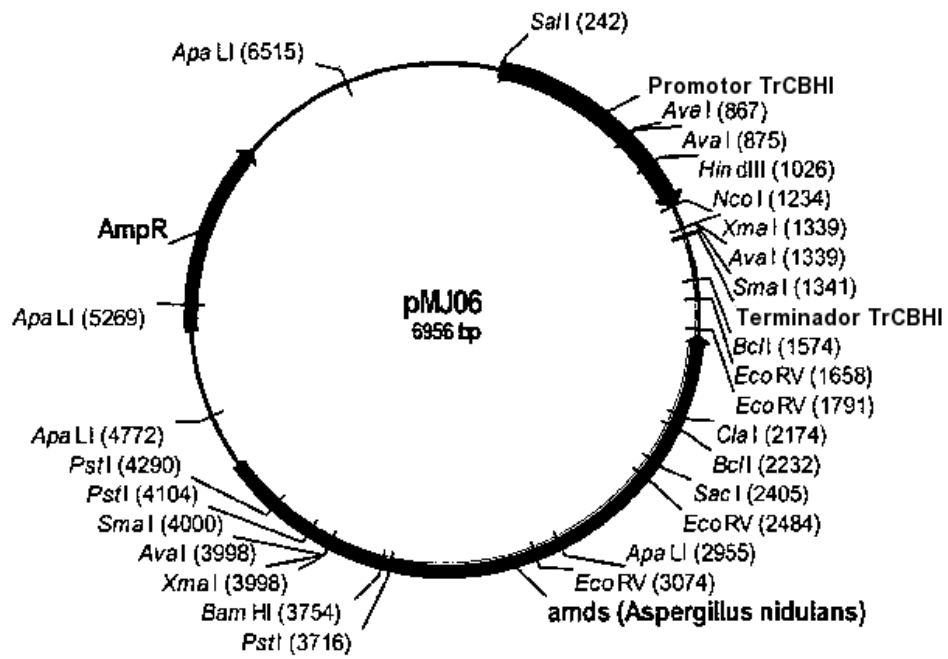


Fig. 23

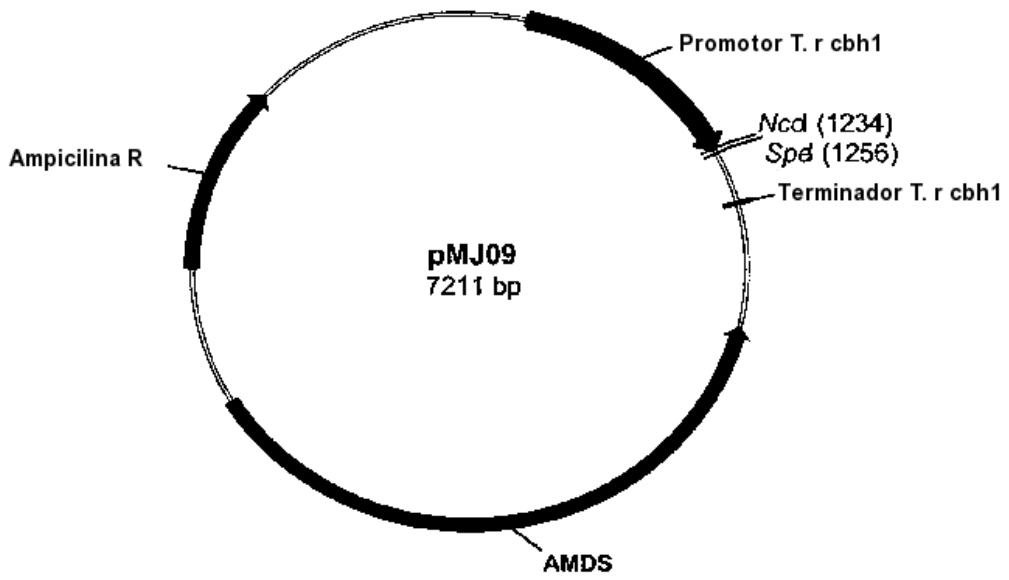


Fig. 24

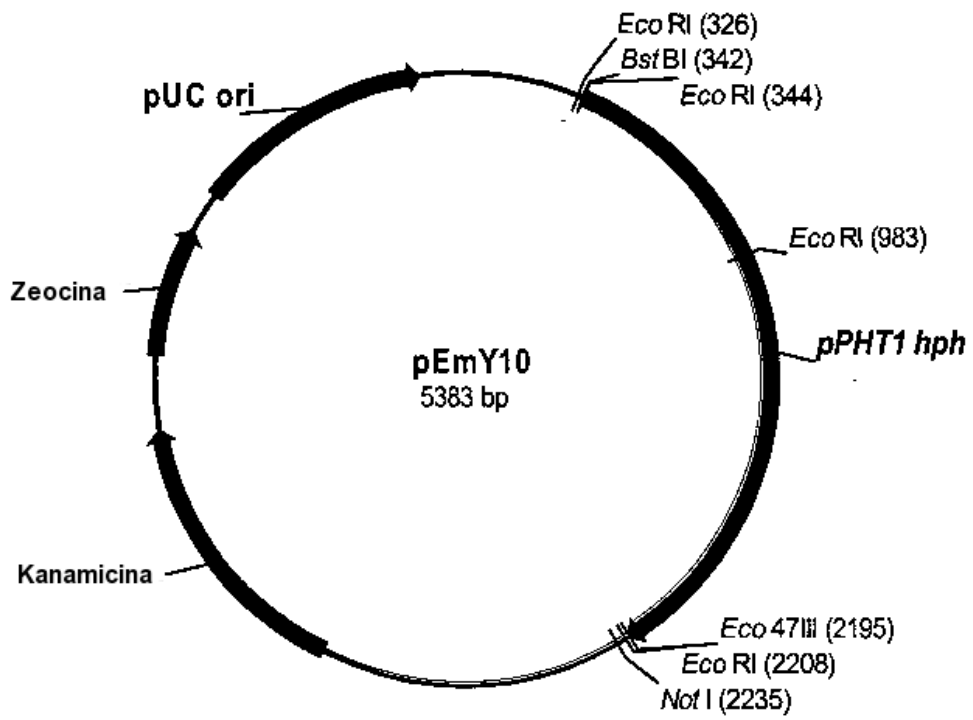


Fig. 25

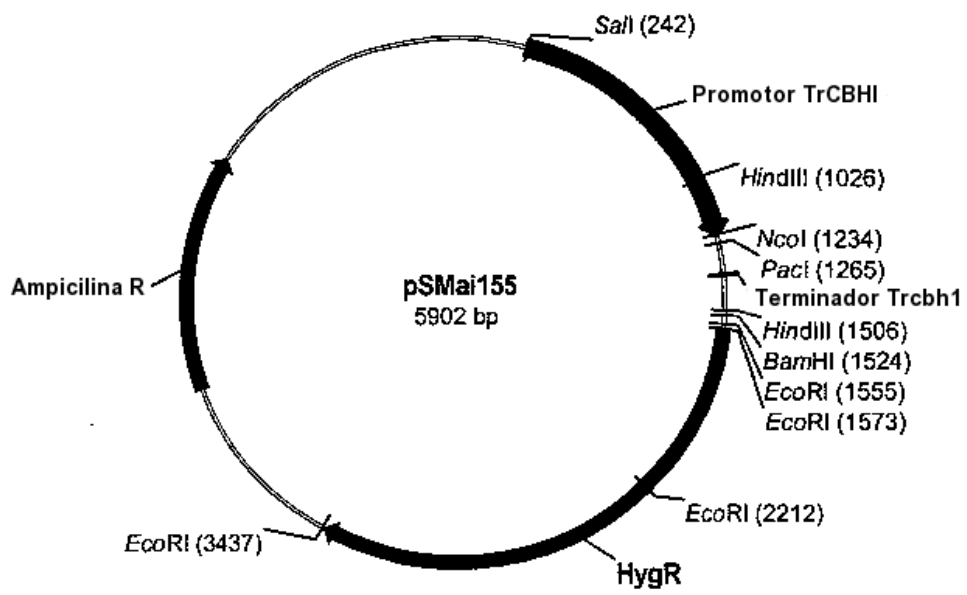


Fig. 26

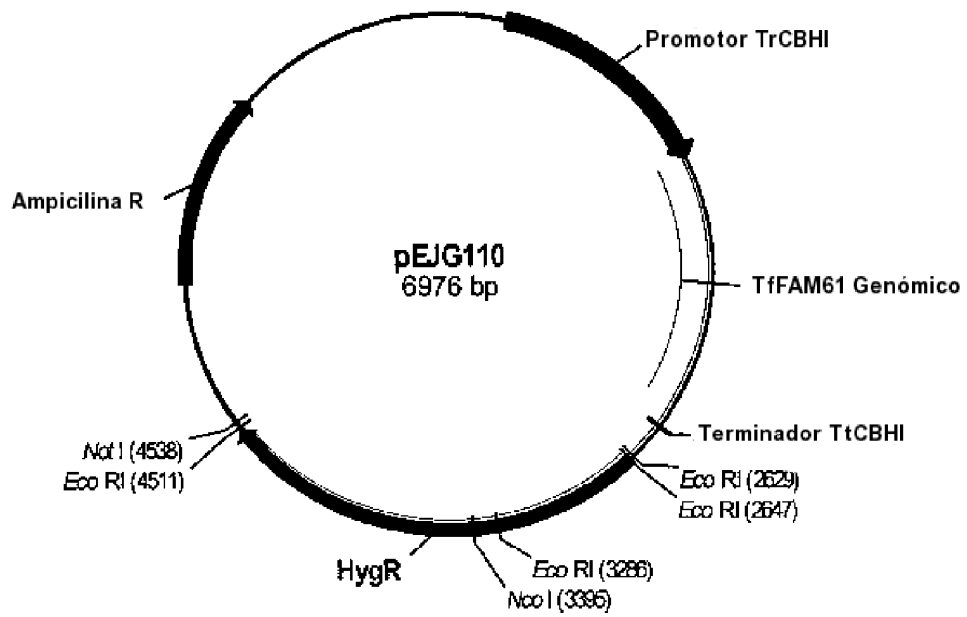


Fig. 27

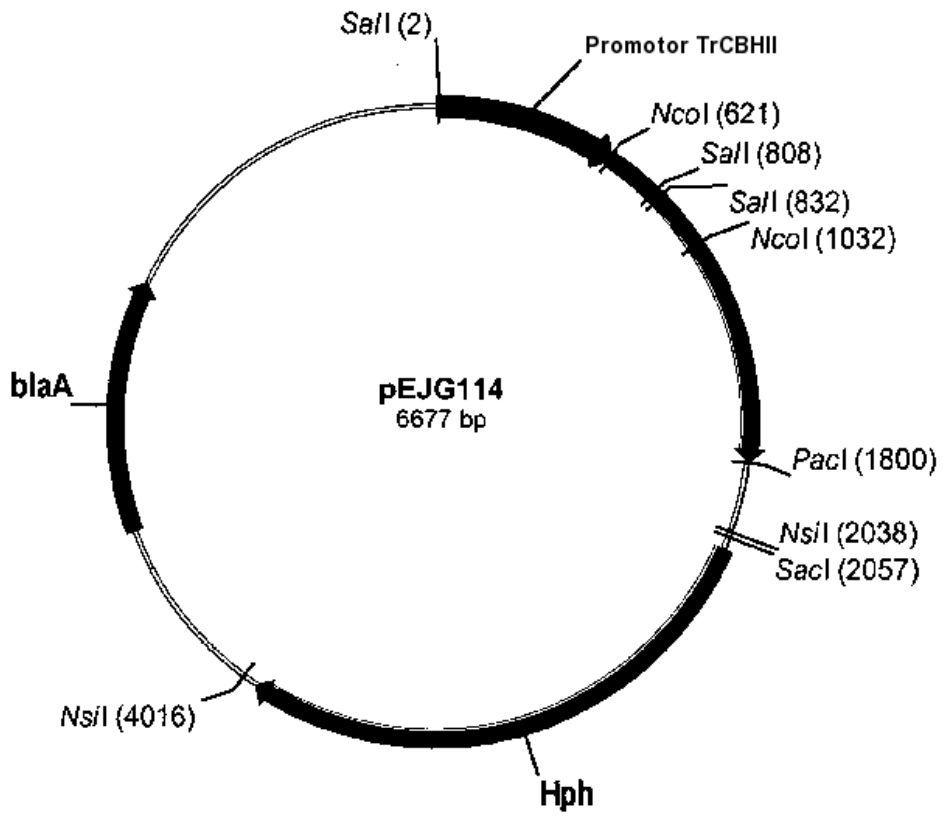


Fig. 28

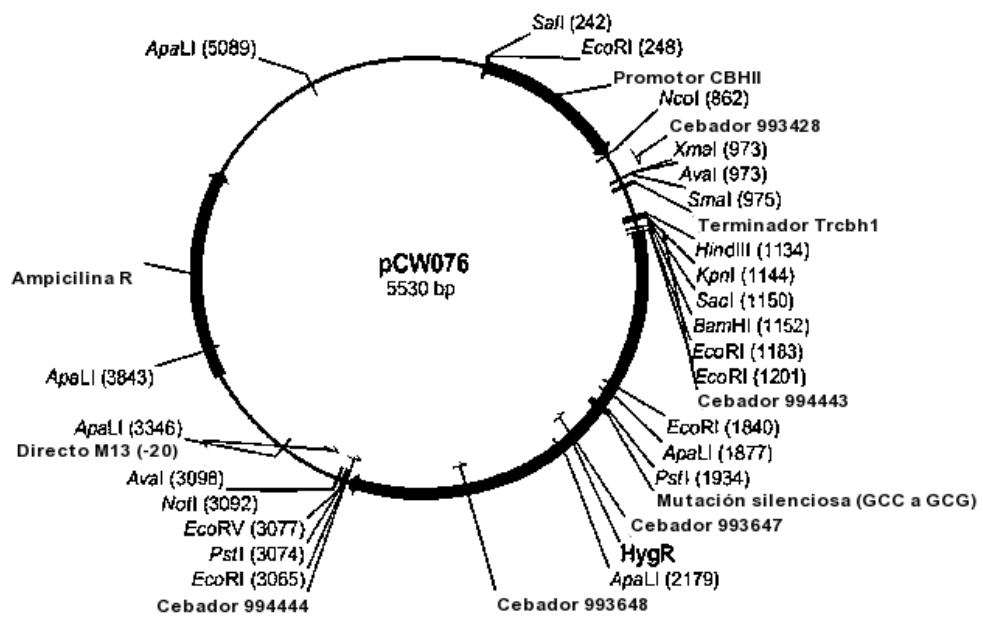


Fig. 29

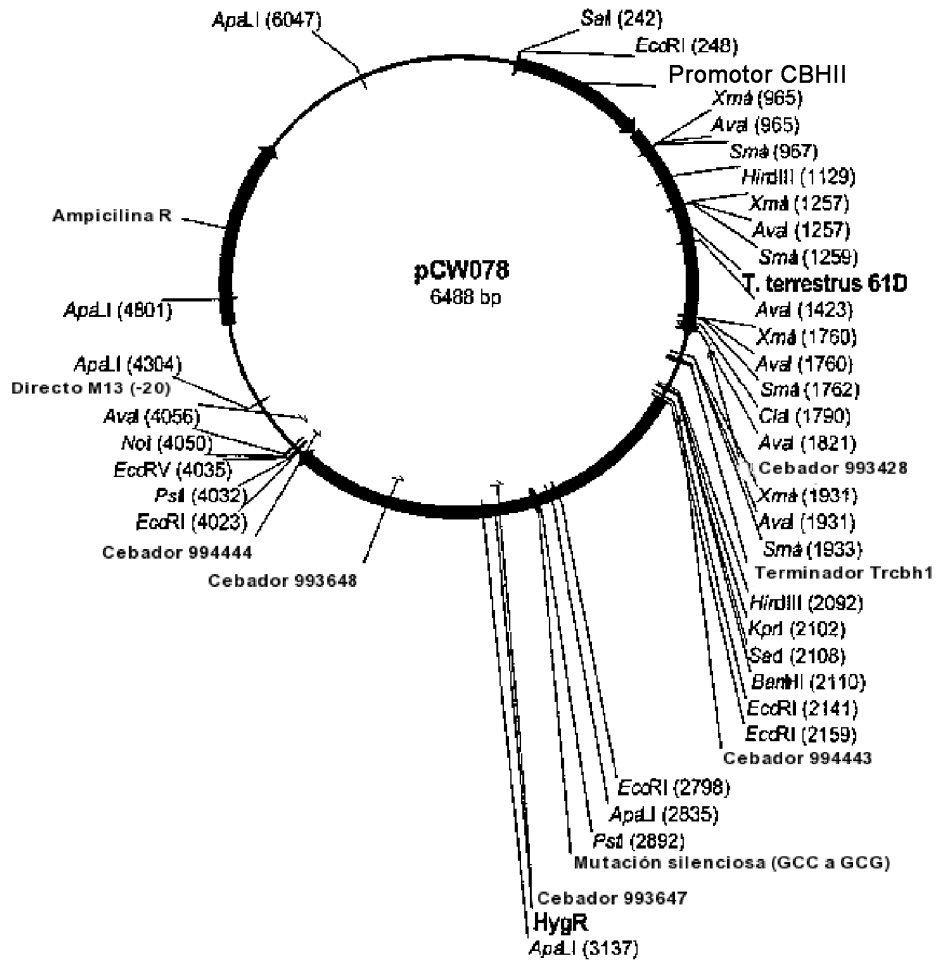


Fig. 30

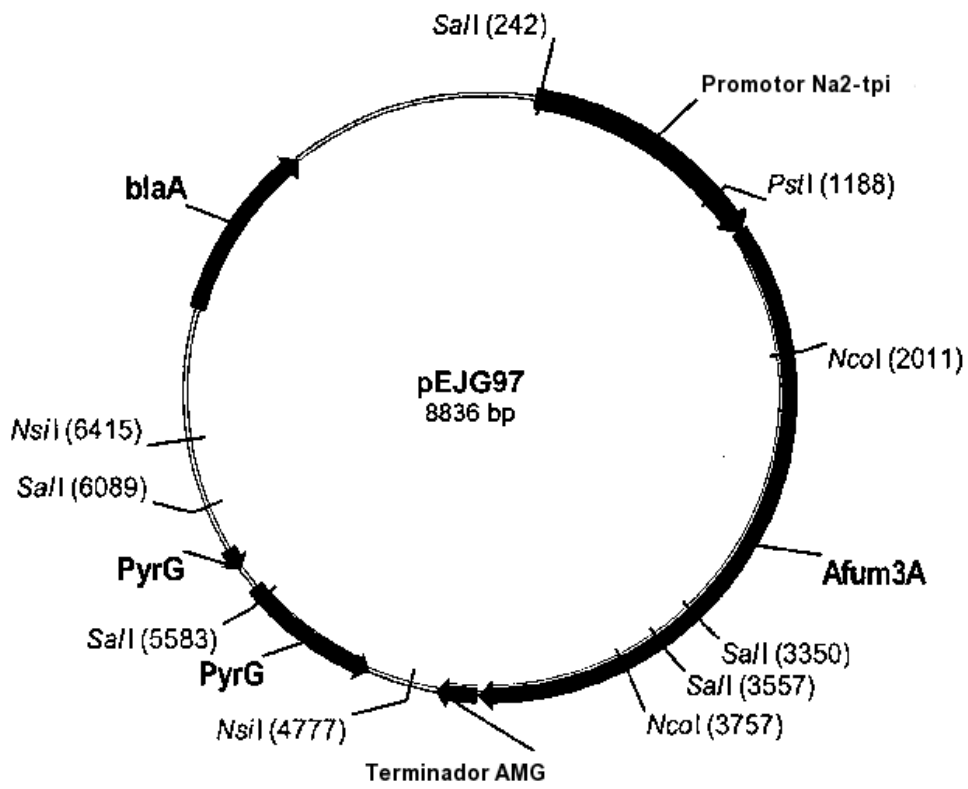


Fig. 31

ATGAGATTCCGGTTGGCTCGAGGTGGCCCGCTCTGACGGCCCGCTTCTGTAGCCAAATGCCCCAGGTTTGTGATGCTTTCCCGTCAATTGTTTCGGATATAGTTGA 100
M R F G W L E V A A L T A A S V A N A Q
 CAATAGTCATGGAATAATCAGGAATTGGCTTTCTCCACCATTCTACCCCTTCCGCTTGGCTGTATGGCCAGGGAGAGTGGCCAGATGCCCATCGACGC 200
 E L A F S P P F Y P S P W A D G Q G E W A D A H R R
 GCCGTCGAGATCGTTTCTCAGATGACACTGGCGGAGAAGTTAACTTACAACGGGTACTGGGTGGGTCCGACTTTTTTGTGACAGTGAAGCTTTCTTC 300
 A V E I V S Q M T L A E K V N L T T G T G
 ACTGACCATCTACAGATGGGAAATGGACCGATCGCTGGTCAAAACCGGCAGCTTCCCAGGTAAGCTTGCAAATCTGCAAAACCGTGCAAGTGTAGTT 400
 W E M D R C V G Q T G S V P R
 GCTAAAAACGGGTGGCAGACTTGGTATCAACTGGGCTTTTGGCCAGGATTTCCCTTTTGGGTATCCGTTTCTGTAGCTATATACCCCGGGAGTCTTT 500
 L G I N W G L C G Q D S P L G I R F
 CAGTCCTTGTAATTATGCTGATGATTGTCTCTGTATAGCTGACCTCAACTCCGCTTCCCTGCTGGTACTAATGTCCGCGGACATGGGACAAAGACACT 600
 S D L N S A F P A G T N V A A T W D K T L
 CGCCTACCTTCGTGGCAAGGCCATGGGTGAGGAATTCACACGACAAGGGCGTGGACATTTTGGCTGGGCTGTCTGGTCCCTCGGCAAAATACCCGGAC 700
 A Y L R G K A M G E F N D K G V D I L L G P A A G P L G K Y P D
 GCGGCAGAAATCGGAAAGCTTCTCCTGATCCGGTTCTCAGTGGTACTTTTCCGCCGAAACTATCAAGGGTATCCAAGACGCGGGTGTGATTGCTA 800
 G G R I W E G F S P D P V L T G V L F A E T I K G I Q D A G V I A
 CTGCCAAGCATTACATTTCTGAATGAACAGGAGCATTTCCGACAGTTTGGCGAGSCTTGGTTACAAACATCACGGAGACGATCAGCTCCAACGT 900
 T A K H Y I L N E Q E H F R Q V G E A Q G Y G Y N I T E T I S S N V
 GGATGACAAGACCATGCACCGAGTTGTACCTTTGGTGTAGTTGACACTGCAAAATGAGGACCTTGATTTGATTTGACTGACCTGGAATGCAAGCCCTTTGC 1000
 D D K T M H E L Y L W
 AGATGCTGTGCGGGTAAGATTTTCCGTAGACTTGACCTCGCGACGAAATAATCGCTGACGAACCATCGTAGCTGGCGTTGGCGTGTGCATGTTCCCTA 1100
 D A V R A G V G A V M C S Y
 CAATCAAAATCAACACAGCTACGGTTGTCAAAAACAGTCAAACTCAACAAGCTCCTCAAGSCTGAGCTGGGCTTCCAAGGCTTCGTCATGAGTACTGG 1200
 N Q I N N S Y G C Q N S Q T L N K L L K A E L G F Q G F V M S D W
 AGCGTCAACACAGCGGTGTGGCGCTGCCCTCGCTGGTGGATATGTCGATGCCCTGGAGACATTTCCCTTCGACGACGACTCCTCCTTCGGGGCACGA 1300
 S A H H S G V G A A L A G L D M S M P G D I S F D D G L S F W G T
 ACCTAACTGTCAGTGTCTTAACGGCACCGTTCCAGCTGGCGTTCGATGACATGGCTGTTCGATCATGACCGGCTACTACAAGGTTGCTCGTGACCG 1400
 N L T V S V L N G T V P A W R V D D M A V R I M T A Y Y K V G R D R
 TCTTCGTATTCCTTAACTTCAGCTCCTGGACCCCGGATGAGTACGGCTGGAGCATTCTGTCTCCGAGGGAGCTGGACCAAGGTGAACGACTTC 1500
 L R I P P N F S S W T R D E Y G W E H S A V S E G A W T K V N D F
 GTCAATGTGACGCGCAGTCACTCTCAGATCATCCGTGAGATTGGTCCCGCTAGTACAGTCTTTGAAGAACACGGGTGCTTCCCTTTGACCCGCAAGG 1600
 V N V Q R S H S Q I I R E I G A A S T V L L K N T G A L P L T G K

Fig. 32A

AGGTTAAAGTGGGTGTTCTCGGTGAAGACGCGTGGTTCACAAACCCGTGGGTGCTAACGGCTGCCCGACCGGGCTGGATAAACGGCACCTTTGGCTATGGC 1700
 EVKVLGVLGLEDAGSNGSNPWGANGNPCPDRCGRCDNGLAMACTGGGTAGTGGTACTCCCTTACCTTGTACCCCGGAGCAGGCTATCCAGCGAGAGGTTCATCAGCAACGGCGGCAATGTCTTTGTGTGACT 1800
 WGSGTANFPYLVTPEQAIQRERVISNNGGNVFAVT
 GATAACGGGGCTCTCAGCCAGATGTCAGATCTCAATCCAGGTGAGTGGGGCTCTAGAAAAAAGAACCTTCTCTGAATGAAGTTTTTAACCA 1900
 DNGALSQMAADVASSQS
 TTGGAAACAGCGTCTTTGGTGTTCACACGGCCGACTCTGGAGAGGGTTTCATCAGTGTCCAGCGCAACGAGGGTGACCGCAAAAATCTCACCTGTG 2000
 VSLVVFVNAADSGEGFISVDGNRGRKRLTLW
 GAAGAACGGCGAGGGCGTCAATGACACTGTTGTACGCCACTGCAACAACACGATTTGGTGTATTCACAGTGTGGGCCCGCTTGTGATCGACCGGTGGTAT 2100
 KNGEAVIDTVVVSHCNNTIVVISHSVGPVLIIDRWY
 GATAACCCCAACGTCACTGCCATCATCTGGGCCGGCTTGGCCGGTTCAGSAGAGTGGCAACTCCCTGGTCGACGTGCTCTATGGCCGGCTCAACCCAGCG 2200
 DNPNAVTAIIWAGLPLPGQESGNSLVDVLYGRVNPSC
 CCAAGACCCCGTTACCTGGGCAAGACTCGGGAGTCTTACGGGGCTCCCTTGTCCACCGAGCCTAACAATGGCAATGGTGTCTCCCGAGGATGATTCAA 2300
 AKTFPTWGTKTRESYGAPELLLTPNNGNGAPQDDFN
 CGAGGGCTTTCATGACTACCGTCACTTTGACAAGCGCAATGAGACCCCAATTAATGAGTTTGGCCATGGCTTGAGCTACACCCACCTTTGGTTACTCT 2400
 EGVFI D Y R H F D K R N E T P I Y E F G H G L S Y T T F G Y S
 CACTTCGGGTTTCAGGCCCTCAATAGTTCGAGTTCGGCATATGTCCCGACTAGCGGAGAGACCAAGCCCTGCCCAACCTATGGTGAGATCGGTAGTGCCG 2500
 HLRVQLANSSSAYVPTSSGETKPAPTYGEISA
 CCGACTACCTGTATCCCGAGGGTCAAAAAGATTACCAAGTTTATACCCTTGGCTCAACTCGACCGACCTCGAGGATTCCTTCTGACGACCCGAACTA 2600
 ADYLYPEGLKRI TKFIYPLWLNS TDL E D S S D D P N Y
 CGGCTGGGAGGACTCGGAGTACATTCGCCGAAGCGCTAGGGATGGGTCTCCCTCAACCCCTCCTGAAGGCTGGCGGGCTCCCTGGTGGTAACCCCTACCCCTT 2700
 GWE D S E Y I P E G A R D G S P Q P L L K A G G A P G G N P T L
 TATCAGGATCTTGTAGGGTTCGGCCACCATAACCAACACTGGTAACGTCGGCCGGTTATGAAGTCCCTCAATTTGGTGAGTGACCCCGCATGTTCCCTTGGC 2800
 YQDLVRSATITNTGNVA G Y E V P Q L
 TTGCAATTTGGCTAACCTCGTCTAGTATGTTTCACTGGGCGGACCGAACGAGCCCTCGGGTCTGGTCTGCGCAAGTTCGACCCGAATCTTCCTGGCTCCTGG 2900
 YVSLGGPNERPRLV V L R K F D R I F L A P G
 GGAGCAAAAAGGTTTGGACCACGACTCTTAACCGTCTGATCTCGCCAAATGGGATGTGGAGGCTCAGGACTGGGTTCATCAAAAAGTACCCCAAGAAAGTG 3000
 EQKVTTLNRRDL ANWDVEAQDWVITKYPKKV
 CAGTCGGCAGCTCCTCGGTAAGTGCCTCTGAGAGCGCCCTCTGCCCGGTGCTACTAG 3060
 H V G S S R K L P L R A P L P R V Y .

Fig. 32B

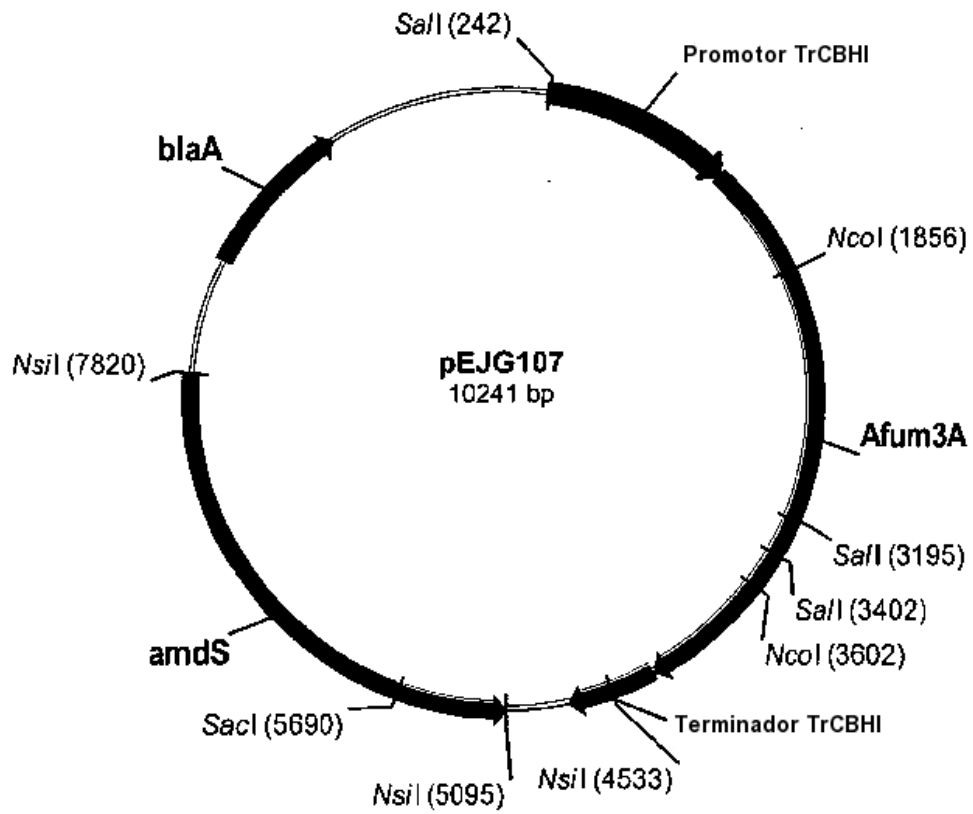


Fig. 33

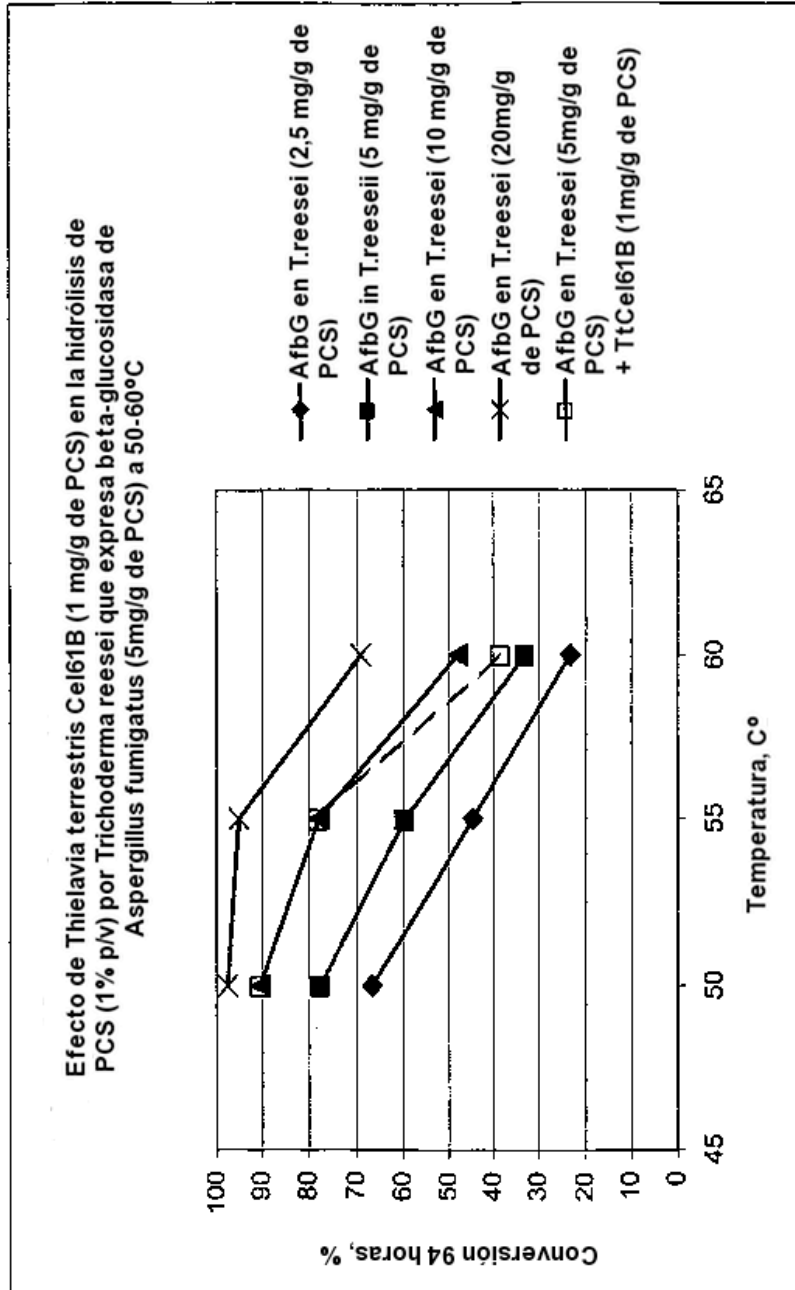


Fig. 34

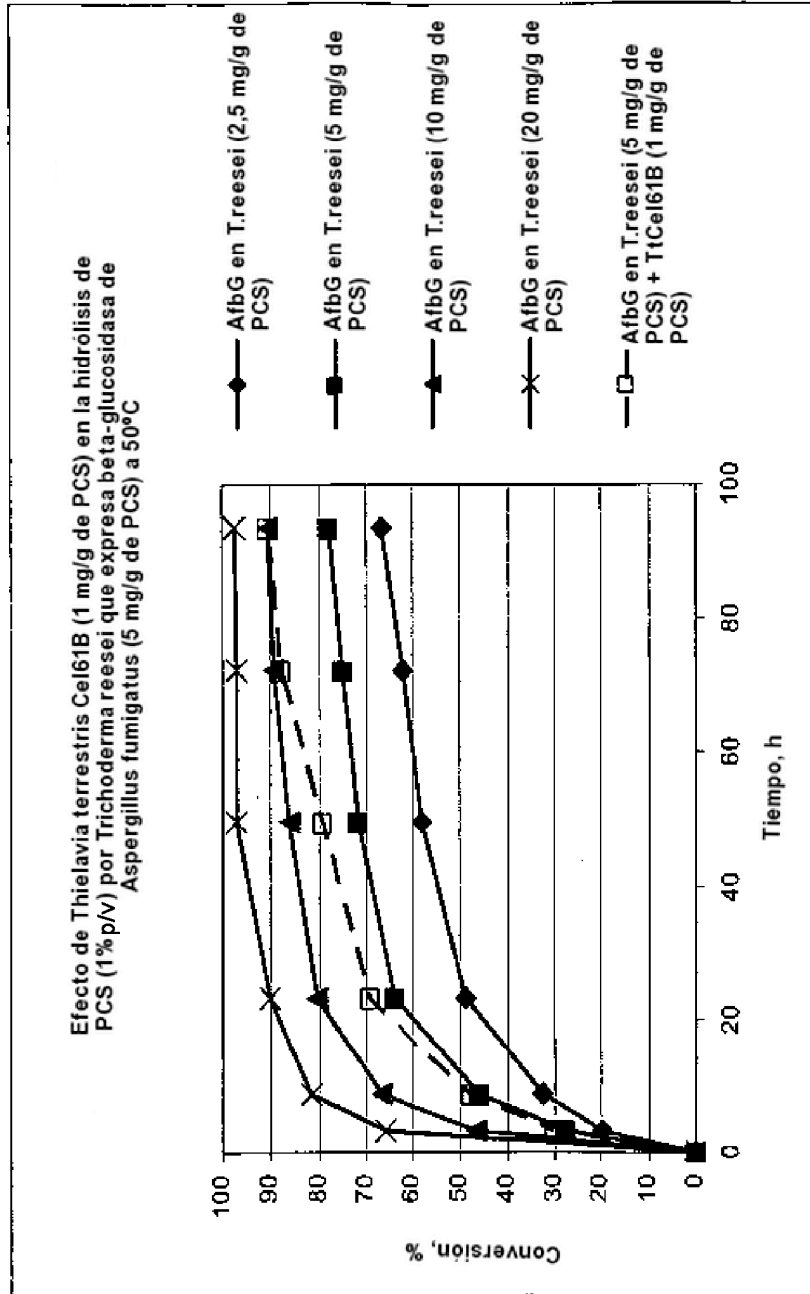


Fig. 35

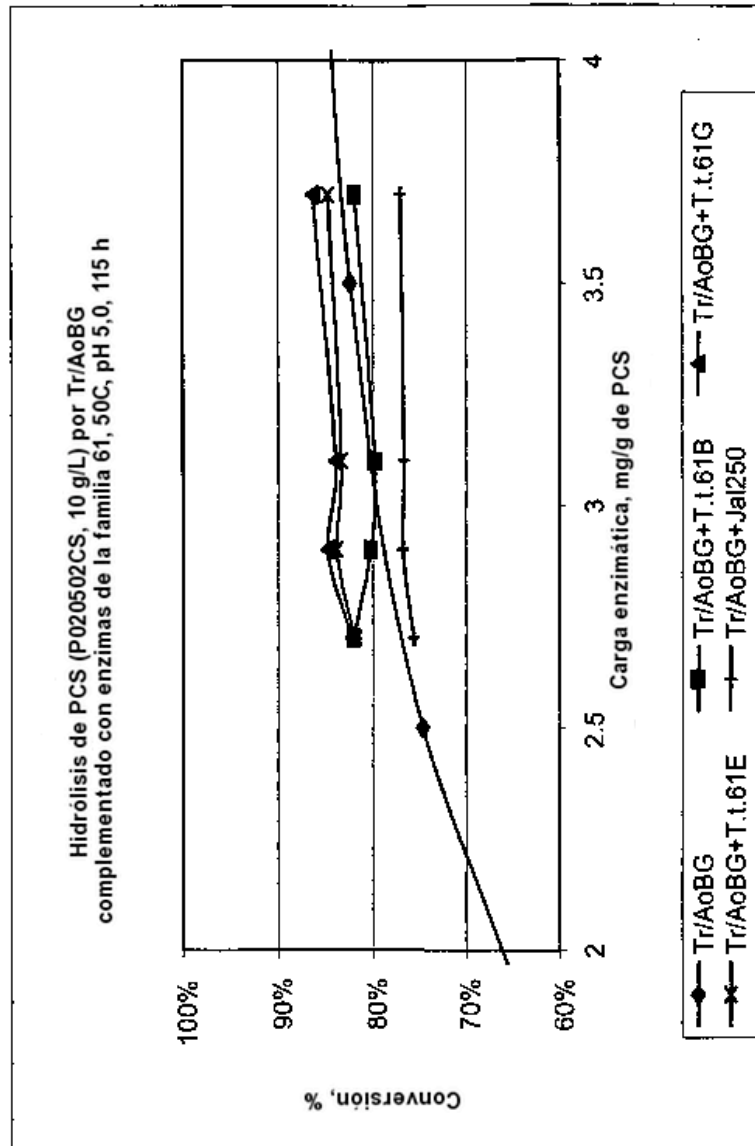


Fig. 36