

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 393 058**

51 Int. Cl.:

C12N 9/00 (2006.01)

C12N 9/24 (2006.01)

C12N 9/42 (2006.01)

C07K 14/38 (2006.01)

C12N 1/12 (2006.01)

C12N 15/00 (2006.01)

C12P 21/06 (2006.01)

C07H 21/04 (2006.01)

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **10165162 .8**

96 Fecha de presentación: **30.04.2004**

97 Número de publicación de la solicitud: **2228440**

97 Fecha de publicación de la solicitud: **15.09.2010**

54 Título: **Variantes de beta-glucosidasa**

30 Prioridad:

02.05.2003 US 467767 P
09.12.2003 US 528342 P

45 Fecha de publicación de la mención BOPI:

18.12.2012

45 Fecha de la publicación del folleto de la patente:

18.12.2012

73 Titular/es:

NOVOZYMES INC. (100.0%)
1445 DREW AVENUE
DAVIS, CA 95616, US

72 Inventor/es:

LAMSA, MICHAEL;
FIDANTSEF, ANA y
GORRE-CLANCY, BRIAN

74 Agente/Representante:

TOMAS GIL, Tesifonte Enrique

ES 2 393 058 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

Variantes de beta-glucosidasa

Antecedentes de la invención

5

Campo de la invención

[0001] La presente invención se realizó con apoyo del Gobierno bajo el Subcontrato NREL N° ZCO-30017-02, contrato principal DE-AC36-98GO10337 otorgado por el Departamento de Energía. El Gobierno tiene algunos derechos en esta invención.

10

Antecedentes de la invención

Campo de la invención

15

[0002] La presente invención se refiere a variantes de beta-glucosidasas que tienen una o más propiedades mejoradas en relación con su enzima progenitora, ácidos nucleicos que codifican las variantes, métodos para producir las variantes y métodos para usar las variantes.

20

Descripción de las técnicas relacionadas

[0003] La celulosa es un polímero de la glucosa de azúcar simple enlazada covalentemente por enlaces beta 1,4. Muchos microorganismos producen enzimas que hidrolizan glucanos beta enlazados. Estas enzimas incluyen endoglucanasas, celobiohidrolasas y beta-glucosidasas. Las endoglucanasas digieren el polímero de celulosa en lugares al azar, abriéndolo al ataque por parte de celobiohidrolasas. Las celobiohidrolasas secuencialmente liberan moléculas de celobiosa de las extremidades del polímero de celulosa. La celobiosa es un dímero de glucosa de enlace beta 1,4 hidrosoluble. Las beta-glucosidasas hidrolizan celobiosa a glucosa.

25

[0004] La conversión de materias primas celulósicas en etanol tiene las ventajas de la disponibilidad inmediata de grandes cantidades de materia prima, la conveniencia de evitar quemar o verter en terrenos los materiales y la limpieza del combustible de etanol. La madera, los residuos agrícolas, los brotes herbáceos y los desperdicios sólidos municipales han sido considerados materias primas para la producción de etanol. Estos materiales principalmente consisten en celulosa, hemicelulosa y lignina. Una vez que la celulosa se convierte en glucosa, la glucosa es fácilmente fermentada por levadura en etanol. Ya que la glucosa es fácilmente fermentada a etanol por una variedad de levaduras mientras que la celobiosa no lo es, cualquier celobiosa restante al final de la hidrólisis representa una pérdida de rendimiento de etanol. De manera más importante, la celobiosa es un inhibidor potente de endoglucanasas y celobiohidrolasas. La acumulación de celobiosa durante la hidrólisis es extremadamente indeseable para la producción de etanol.

30

[0005] La acumulación de celobiosa ha sido un problema importante en la hidrólisis enzimática porque los microorganismos productores de celulasa producen poca beta-glucosidasa. La baja cantidad de beta-glucosidasa da como resultado una escasez de la capacidad de hidrolizar la celobiosa a glucosa. Se han utilizado diferentes enfoques para aumentar la cantidad de beta-glucosidasa en la conversión de celulosa a glucosa.

35

[0006] Un enfoque es producir beta-glucosidasa usando microorganismos que producen poca celulasa y añadir la beta-glucosidasa exógenamente a endoglucanasa y celobiohidrolasa para mejorar la hidrólisis. No obstante; las cantidades requeridas son demasiado costosas para una biomasa comercial para la operación de etanol.

45

[0007] Un segundo enfoque es llevar a cabo hidrólisis de celulosa simultáneamente con fermentación de la glucosa por levadura. Este proceso es conocido como sacarificación y fermentación simultánea (SSF). En un sistema SSF, la fermentación de la glucosa la retira de la solución. No obstante, los sistemas SSF no son aún comercialmente viables debido a que la temperatura operativa para la levadura de 28°C es demasiado baja para las condiciones de 50°C requeridas.

50

[0008] Un tercer enfoque para superar la escasez de beta-glucosidasa es sobreexpresar la beta-glucosidasa en un huésped, aumentando así el rendimiento de la beta-glucosidasa.

55

[0009] Sería una ventaja en la técnica proporcionar variantes de beta-glucosidasa con propiedades mejoradas para convertir materiales celulósicos a monosacáridos, disacáridos y polisacáridos. Las propiedades mejoradas incluyen actividad alterada dependiente de perfiles de temperatura, termoestabilidad, actividad de pH, estabilidad del pH, especificidad de sustrato, especificidad de producto y estabilidad química. WO02/095014 A2 describe una beta-glucosidasa de *A. Oryzae*. También se describen métodos para obtener beta-glucosidasas y métodos de producción de variantes por una o varias sustituciones de aminoácidos.

60

[0010] Es un objeto de la presente invención proporcionar variantes de beta-glucosidasas con propiedades mejoradas en comparación con sus enzimas progenitoras.

65

Resumen de la invención

5 [0011] La presente invención se refiere a variantes aisladas de una beta-glucosidasa progenitora, comprendiendo una sustitución en la posición correspondiente a la posición 285 de los aminoácidos 20 a 861 de SEC ID No: 2 o correspondiente a la posición 285 de los aminoácidos 20 a 863 de SEC ID No: 70, donde la variante tiene una actividad de beta-glucosidasa y comprende las sustituciones Q202R + H285Q + D722G (o D724G); las sustituciones Q202R + H285Q; o las sustituciones H285Q + D722G (o D724G) comparado con la beta-glucosidasa progenitora.

10 [0012] La presente invención también se refiere a secuencias de nucleótidos aisladas que codifican las beta-glucosidasas variantes que tienen actividad de beta-glucosidasa y a constructos de ácidos nucleicos, vectores y células huéspedes que comprenden las secuencias de nucleótidos.

15 [0013] La presente invención también se refiere a métodos para producir una variante de una beta-glucosidasa progenitora que tiene actividad de beta-glucosidasa en una célula huésped.

[0014] La presente invención también se refiere a plantas que codifican beta-glucosidasa variantes que tienen actividad de beta-glucosidasa.

20 [0015] La presente invención además se refiere a métodos para degradar o convertir biomasa con contenido de celulosa y hemicelulosa usando las variantes de beta-glucosidasa de la invención.

Breve descripción de las figuras

25 [0016]
La Figura 1 muestra un mapa de restricción de pSATE101.
La Figura 2 muestra un mapa de restricción de pSATE111.
La Figura 3 muestra un mapa de restricción de pMJ04.
La Figura 4 muestra un mapa de restricción de pCaHj527.
30 La Figura 5 muestra un mapa de restricción de pMT2188.
La Figura 6 muestra un mapa de restricción de pCaHj568.
La Figura 7 muestra un mapa de restricción de pMJ05.
La Figura 8 muestra un mapa de restricción de pSMai130.
La Figura 9 muestra la secuencia de ADNc de un gen de beta-glucosidasa de *Aspergillus oryzae* (SEC ID n°: 1) y la secuencia de aminoácidos deducida del mismo (SEC ID n°: 2).
35 La Figura 10 muestra 63 bp de la secuencia de señal de endoglucanasa V de *Humicola insolens* putativa (codón de inicio ATG a Ala 21, SEC ID n°: 29).
La Figura 11 muestra un mapa de restricción de pSMai135.
La Figura 12 muestra un mapa de restricción de pALFd1.
La Figura 13 muestra un mapa de restricción de pALo1.
40 La Figura 14 muestra un mapa de restricción de pBANE10.
La Figura 15 muestra un mapa de restricción de pALo2.
La Figura 16 muestra un mapa de restricción de pALFd3BG41.
La Figura 17 muestra un mapa de restricción de pALFd3BG48.
La Figura 18 muestra una determinación de termoestabilidad de variantes de beta-glucosidasa de *Aspergillus oryzae*
45 BG41 y BG48.
La Figura 19 muestra los efectos de la termoestabilidad de las mutaciones G161S y H285Q individualmente y combinadas.
La Figura 20 muestra un mapa de restricción de pEJG97.
La Figura 21 muestra la secuencia de ADN genómico y la secuencia de aminoácidos deducida de una beta-glucosidasa de *Aspergillus fumigatus* (SEC ID n°: 69 y 70, respectivamente). El péptido señal predicho está subrayado y los intrones predichos están en cursiva.
50 La Figura 22 muestra un mapa de restricción de pCR4Blunt-TOPOAfcDNA5'.
La Figura 23 muestra un mapa de restricción de pCR4Blunt-TOPOAfcDNA3'.
La Figura 24 muestra un mapa de restricción de pCR4Blunt-TOPOAfcDNA.
55 La Figura 25 muestra un mapa de restricción de pALFd7.
La Figura 26 muestra un mapa de restricción de pALFd6.
La Figura 27 muestra un mapa de restricción de pEJG97AfumFAM3AG142S.
La Figura 28 muestra un mapa de restricción de pALFd7G142S.
La Figura 29 muestra un mapa de restricción de pEJG97AfumFAM3AH266Q.
60

Descripción detallada de la invención

65 [0017] La presente invención se refiere a variantes aisladas de una beta-glucosidasa progenitora, comprendiendo una sustitución en la posición correspondiente a la posición 285 de los aminoácidos 20 a 861 de SEC ID No: 2 o correspondiente a la posición 285 de los aminoácidos 20 a 863 de SEC ID No: 70, donde la variante tiene una actividad

de beta-glucosidasa y comprende las sustituciones Q202R + H285Q + D722G (o D724G); las sustituciones Q202R + H285Q; o las sustituciones H285Q + D722G (o D724G) comparado con la beta-glucosidasa progenitora.

Definiciones

- 5 [0018] El término "beta-glucosidasa" es definido en la presente como una beta-D-glucósido glucohidrolasa (E.C. 3.2.1.21) que cataliza la hidrólisis de residuos de beta-D-glucosa no reductores terminales con la liberación de beta-D-glucosa. Para objetivos de la presente invención, la actividad de beta-glucosidasa se determina según el procedimiento básico descrito por Venturi et al., 2002, J. Basic Microbial. 42: 55-66, excepto varias temperaturas y pH 5 empleadas en la presente. Una unidad de actividad de beta-glucosidasa es definida como 1.0 µmol de beta-D-glucosa producida por minuto a 25°C, pH 5.
- 10 [0019] El término "variante" es definido en la presente como una beta-glucosidasa que comprende una o más alteraciones, tales como sustituciones, inserciones, deleciones, y/o truncamientos de uno o más residuos de aminoácidos específicos en una o más posiciones específicas en el polipéptido.
- 15 [0020] El término beta-glucosidasa de "tipo salvaje" denota una beta-glucosidasa expresada por un microorganismo de origen natural, tal como una levadura u hongo filamentoso encontrado en la naturaleza.
- 20 [0021] El término beta-glucosidasa "progenitora" según se utiliza en este caso significa una beta-glucosidasa a la cual se producen modificaciones, por ejemplo, sustitución/es, inserción/es, deleción/es y/o truncamiento/s, para producir las variantes enzimáticas de la presente invención. Este término también se refiere al polipéptido con el cual una variante es comparada y alineada. El progenitor puede ser un polipéptido de origen natural (tipo salvaje), o incluso puede ser una variante del mismo, preparado por cualquiera de los medios adecuados. Por ejemplo, la proteína progenitora puede ser una variante de un polipéptido de origen natural que ha sido modificada o alterada en la secuencia de aminoácidos. Un progenitor puede también ser una variante alélica que es cualquiera de dos o más formas alternativas de un gen que ocupa la misma localización cromosómica. Una variante alélica de un polipéptido es un polipéptido codificado por la variante alélica correspondiente de un gen.
- 25 [0022] El término "redistribución" significa recombinación de secuencia/s de nucleótidos entre dos o más secuencias de nucleótidos homólogos que dan como resultado secuencias de nucleótidos recombinadas (es decir, secuencias de nucleótidos que han sido sometidas a un ciclo de bajado) con un número de nucleótidos cambiado, en comparación con las secuencias de nucleótidos iniciales.
- 30 [0023] El término, "biblioteca aleatoria", "biblioteca de variantes" o "biblioteca" es definido en la presente como una biblioteca de polipéptidos variantes. La diversidad en la biblioteca de variantes puede ser generada por medio de mutagénesis de los genes que codifican las variantes en el nivel de triplete del ADN, de manera que los codones individuales son diversificados, por ejemplo, usando cebadores de secuencias parcialmente aleatorizadas en una reacción de PCR. Se han descrito diferentes técnicas, por las que puede crearse una biblioteca combinatoria diversa diversificando diferentes posiciones de nucleótido en un gen y recombinándolas, por ejemplo, donde estas posiciones están demasiado alejadas para ser cubiertas por un único cebador oligonucleótido (adicionado o dopado). Estas técnicas incluyen el uso de recombinación *in vivo* de los segmentos de gen individualmente diversificados como se describe en WO 97/07205 en la página 3, líneas 8 a 29. También incluyen el uso de técnicas de bajado de ADN para crear una biblioteca de genes de longitud total, caracterizada por el hecho de que diferentes segmentos de gen son combinados, y caracterizada por el hecho de que cada segmento puede ser diversificado, por ejemplo, por mutagénesis adicionada (Stemmer, 1994; Nature 370: 389-391; US 5,811,238; US 5,605,793 y US 5,830,721). Puede usarse un gen que codifica una proteína "de estructura principal" (polipéptido progenitor de tipo salvaje) como un polinucleótido patrón, y combinarlo con uno o más oligonucleótidos mono o bicatenarios como se describe en WO 98/41623 y WO 98/41622. Los oligonucleótidos monocatenarios pueden ser parcialmente aleatorizados durante la síntesis. Los oligonucleótidos bicatenarios pueden ser productos de PCR que incorporan diversidad en una zona específica. En ambos casos, uno puede diluir la diversidad con segmentos correspondientes que codifican la secuencia de la proteína de estructura principal para limitar el número medio de cambios que se introducen.
- 35 [0024] El término "recombinación" es definido en la presente como un proceso caracterizado por el hecho de que los ácidos nucleicos se asocian entre sí en regiones de homología, conduciendo a intercambio de ADN intercatenario entre esas secuencias. Para objetivos de la presente invención, la recombinación homóloga se determina según los procedimientos resumidos por Paques y Haber, 1999, Microbiology and Molecular Biology Reviews 63: 349-404. La "recombinación homóloga" es definida en la presente como la recombinación en la cual no ocurre ningún cambio en las secuencias de nucleótidos en las regiones de homología relativas a las secuencias de nucleótidos de entrada. Para la recombinación homóloga perfecta, las regiones deberían contener un número suficiente de ácidos nucleicos, tales como 100 a 1.500 bp, preferiblemente 400 a 1.500 bp, y de forma más preferible 800 a 1.500 bp, que son altamente homólogos a la secuencia de ácidos nucleicos correspondientes para mejorar la probabilidad de recombinación homóloga. La recombinación puede también ocurrir por recombinación no homóloga. La "recombinación no homóloga" es definida en la presente como recombinación caracterizada por el hecho de que cualquier modo de reparación del ADN que incorpora intercambio de cadenas da como resultado una secuencia de nucleótidos diferente de cualquiera de las secuencias recombinantes.
- 40
- 45
- 50
- 55
- 60
- 65

5 [0025] El término "propiedad mejorada" es definido en la presente como una característica asociada con una variante que se mejora en comparación con la beta-glucosidasa progenitora. Tales propiedades mejoradas incluyen, de modo enunciativo y no limitativo, perfil de actividad alterada dependiente de la temperatura, termoestabilidad, actividad de pH, estabilidad de pH, especificidad de sustrato, especificidad de producto y estabilidad química.

10 [0026] El término "actividad térmica mejorada" es definido en la presente como una enzima variante que muestra una alteración del perfil de actividad dependiente de la temperatura de una variante de beta-glucosidasa a una temperatura específica relativa al perfil de actividad dependiente de la temperatura de la beta-glucosidasa progenitora. El valor de la actividad térmica proporciona una medida de la eficiencia enzimática en la realización de catálisis de una reacción de hidrólisis sobre una gama de temperaturas. Una beta-glucosidasa tiene una gama de temperatura específica caracterizada por el hecho de que la proteína es estable y retiene su actividad enzimática, pero se vuelve menos estable y así menos activa con temperatura en aumento. Además, el nivel inicial de una reacción catalizada por una beta-glucosidasa se puede acelerar por un aumento en la temperatura que se mide determinando la actividad térmica de una variante. Una variante más termoactiva llevará a un aumento en el índice de hidrólisis disminuyendo el tiempo requerido y/o disminuyendo la concentración enzimática requerida para la hidrólisis. De forma alternativa, una variante con una actividad térmica reducida catalizará una reacción de hidrólisis a una temperatura inferior a la temperatura óptima de la enzima progenitora definida por el perfil de actividad dependiente de la temperatura del progenitor.

20 [0027] El término "termoestabilidad mejorada" es definido en la presente como una enzima variante que muestra retención de actividad enzimática después de un periodo de incubación a temperatura elevada relativa a la enzima progenitora. Tal variante puede o no mostrar un perfil de actividad térmica alterada relativa al progenitor, por ejemplo, puede tener una capacidad mejorada para replegarse después de la incubación a temperatura elevada relativa al progenitor.

25 [0028] En una forma de realización preferida, la actividad térmica de la beta-glucosidasa variante es al menos 1, 5 veces, preferiblemente al menos 2 veces, más preferiblemente al menos 5 veces, de forma más preferible al menos 7 veces, e incluso de forma más preferible al menos 20 veces más activa térmicamente que la enzima progenitora cuando la actividad residual es comparada usando el metilumbeliferil-beta-D-glucopiranosido como sustrato a 60°C y pH 5 durante 30 15 horas.

[0029] El término "especificidad de producto mejorada" es definido en la presente como una enzima variante que muestra un perfil de producto alterado relativo al progenitor en el cual el perfil de producto alterado mejora el rendimiento de la variante en una aplicación relativa al progenitor. El término "perfil del producto" es definido en la 35 presente como la composición química de los productos reactivos producida por hidrólisis enzimática.

[0030] El término "estabilidad química mejorada" es definido en la presente como una enzima variante que muestra retención de actividad enzimática después de un periodo de incubación en presencia de un producto químico o productos químicos, de origen natural o bien sintético, que reducen la actividad enzimática de la enzima progenitora. La 40 estabilidad química mejorada puede también suponer variantes más capaces de catalizar una reacción en presencia de tales productos químicos.

Convenciones para la designación de las variantes

45 [0031] En la presente invención, se emplea una numeración específica de posiciones de residuos de aminoácido en las variantes de beta-glucosidasa. Por ejemplo, alineando las secuencias de aminoácidos de beta-glucosidasas conocidas, es posible designar un número de posición de aminoácido a cualquier residuo de aminoácido en cualquier enzima de beta-glucosidasa.

50 [0032] Usando el sistema de numeración originado de la secuencia de aminoácidos de la beta-glucosidasa descrito en SEC ID n°: 2 o SEC ID n°: 70, alineado con la secuencia de aminoácidos de una cantidad de otras beta-glucosidasas, es posible indicar la posición de un residuo de aminoácido en una beta-glucosidasa en regiones de homología estructural.

55 [0033] Pueden hacerse alineamientos múltiples de secuencias proteínicas, por ejemplo, usando "ClustalW" (Thompson, J.D., Higgins, D.G. y Gibson, T.J., 1994, CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice, Nucleic Acids Research 22: 4673-4680). Pueden realizarse alineamientos múltiples de secuencias de ADN usando la alineación de proteína como un molde, sustituyendo los aminoácidos con el codón correspondiente de la secuencia de ADN.

60 [0034] Los algoritmos de comparación de secuencias por parejas de uso común son adecuados para detectar similitudes entre secuencias proteínicas que no se han bifurcado más allá del punto de aproximadamente 20-30% de identidad de secuencia (Doolittle, 1992, Protein Sci. 1: 191-200; Brenner et al., 1998, Proc. Natl. Acad. Sci. USA 95, 6073-6078). No obstante, las proteínas realmente homólogas con el mismo pliegue y función biológica similar frecuentemente se han bifurcado al punto en el cual la comparación basada en la secuencia tradicional no detecta su 65 relación (Lindahl y Elofsson, 2000, J. Mol. Biol. 295: 613-615). Puede lograrse una sensibilidad superior en la búsqueda

5 basada en secuencia usando programas de búsqueda que utilizan representaciones probabilísticas de familias de proteína (perfiles) para buscar bases de datos. Por ejemplo, el programa PSI-BLAST genera perfiles a través de un proceso de búsqueda de base de datos iterativo y es capaz de detectar homólogos remotos (Atschul et al., 1997, Nucleic Acids Res. 25: 3389-3402). Puede lograrse una sensibilidad aún mayor si la familia o superfamilia para la proteína de interés tiene uno o más representantes en las bases de datos de estructura de proteína. Programas tales como GenTHREADER (Jones 1999, J. Mol. Biol. 287: 797-815; McGuffin y Jones, 2003, Bioinformatics 19: 874-881) utilizan información de una variedad de fuentes (PSI- BLAST, predicción de estructura secundaria, perfiles de alineación estructural y potenciales de disolución) como entrada a una red neuronal que predice el pliegue estructural para una secuencia dudosa. De forma similar, el método de Gough et al., 2000, J. Mol. Biol. 313: 903 919, puede utilizarse para alinear una secuencia de estructura desconocida con los modelos de superfamilia presentes en la base de datos SCOP. Estos alineamientos pueden ser usados sucesivamente para generar modelos de homología para la proteína de interés, y tales modelos pueden evaluarse en cuanto a su exactitud usando una variedad de herramientas desarrolladas para tal fin.

15 [0035] Para las proteínas de estructura conocida, están disponibles diferentes herramientas y recursos para recuperar y generar alineamientos estructurales. Por ejemplo, las superfamilias de SCOP de proteínas han sido estructuralmente alineadas, y esos alineamientos son accesibles y pueden descargarse. Estos alineamientos pueden utilizarse para predecir los residuos de aminoácidos correspondientes estructuralmente y funcionalmente en proteínas en la misma superfamilia estructural. Esta información, con información derivada de modelado por homología y búsquedas de perfiles, puede utilizarse para predecir cuáles residuos mutar cuando se mueven las mutaciones de interés de una proteína a un homólogo remoto o cercano.

25 [0036] Al describir las diferentes variantes de beta-glucosidasa de la presente invención, la nomenclatura descrita a continuación se adapta para facilidad de referencia. En cualquier caso, se utilizan las abreviaturas de aminoácidos aceptadas de IUPAC de una letra o tres letras.

30 [0037] Sustituciones. Para una sustitución de aminoácido, se usa la siguiente nomenclatura: [aminoácido original, posición, aminoácido sustituido]. Por consiguiente, la sustitución de treonina con alanina en la posición 226 es designada como "Thr226Ala" o "T226A". Las mutaciones múltiples se separan por marcas de adición ("+"), por ejemplo, "Gly205Arg + Ser411Phe" o "G205R + S411F", representando mutaciones en las posiciones 205 y 411 sustituyendo la glicina (G) con arginina (R) y la serina (S) con fenilalanina (F), respectivamente.

35 [0038] Delecciones. Para una delección de aminoácido, se usa la siguiente nomenclatura: [aminoácido original, posición*]. Por consiguiente, la delección de glicina en la posición 195 es designada como "Gly195*" o "G195*". Las delecciones múltiples se separan por marcas de adición ("+"), por ejemplo, "Gly195* + Ser411*" o "G195* + S411*".

40 [0039] Inserciones. Para una inserción de aminoácido, se usa la siguiente nomenclatura: [Aminoácido original, posición, aminoácido original, aminoácido nuevo insertado]. Por consiguiente la inserción de lisina después de la glicina en la posición 195 es designada "Gly195GlyLys" o "G195GK". Múltiples inserciones de aminoácidos son designadas [aminoácido original, posición, aminoácido original, aminoácido nuevo insertado n°1, aminoácido nuevo insertado n°2 etc.] Por ejemplo, la inserción de lisina y alanina después glicina en la posición 195 es indicada como "Gly195GlyLysAla" o "G195GKA".

45 [0040] En tales casos el/los residuo/s de aminoácido insertado/s son numerados por la adición de letras minúsculas al número de la posición del residuo de aminoácido que precede a el/los residuo/s de aminoácido insertado/s. En el ejemplo de arriba las secuencias serían entonces:

Progenitor:	Variante:
195	195 195a 195b
G	G - K - A

50 [0041] Indicaciones de degeneración. Para las indicaciones de degeneración donde un residuo de aminoácido idéntico al residuo de aminoácido existente está insertado, surge la degeneración en la nomenclatura. Por ejemplo, una glicina insertada después de la glicina en el ejemplo de arriba sería indicada por "G195GG". Suponiendo que una alanina estuviera presente en la posición 194, el mismo cambio real podría también ser indicado como "A194AG".

	Progenitor:	Variante:
Numeración I:	194 195	194 195 195a
Secuencia:	A - G	A - G - G
Numeración II:		194 194a 195

55 Tales ejemplos serán evidentes para el experto en la materia, y la indicación "G195GG" y las indicaciones correspondientes para este tipo de inserción están así destinadas a comprender tales indicaciones equivalentes de degeneración.

[0042] Si los segmentos de la secuencia de aminoácidos se repiten en el polipéptido progenitor y/o en la variante, surgen indicaciones de degeneración equivalentes, también cuando alteraciones diferentes de las inserciones son catalogadas tales como deleciones y/o sustituciones. Por ejemplo, la deleción de dos aminoácidos "AG" consecutivos en la secuencia "AGAG" desde la posición 194-97 puede ser escrita como "A194*+G195*" o "A196*+G197*":

5

	Progenitor:	Variante:
Numeración I:	194 195 196 197	194 195
Secuencia:	A - G - A - G	A - G
Numeración II:		196 197

10

[0043] Modificaciones múltiples. Las variantes que comprenden modificaciones múltiples se separan por marcas de adición ("+"), por ejemplo, "Arg170Tyr+Gly195Glu" o "R170Y+G195E" que representan modificaciones en las posiciones en 170 y 195 substituyendo ácido glutámico y tirosina por arginina y glicina, respectivamente. De ese modo, "Tyr167Gly,Ala,Ser,Thr + Arg170Gly,Ala,Ser,Thr" designa las siguientes variantes:

15

"Tyr167Gly+Arg170Gly", "Tyr167Gly+Arg170Ala", "Tyr167Gly+Arg170Ser", "Tyr167Gly+Arg170Thr", "Tyr167Ala+Arg170Gly", "Tyr167Ala+Arg170Ala", "Tyr167Ala+Arg170Ser", "Tyr167Ala+Arg170Thr", "Tyr167Ser+Arg170Gly", "Tyr167Ser+Arg170Ala", "Tyr167Ser+Arg170Ser", "Tyr167Ser+Arg170Thr", "Tyr167Thr+Arg170Gly", "Tyr167Thr+Arg170Ala", "Tyr167Thr+Arg170Ser" y "Tyr167Thr+Arg170Thr".

20

[0044] Esta nomenclatura es particularmente pertinente para las modificaciones que implican substitución, inserción o deleción de residuos de aminoácidos que tienen propiedades comunes específicas. Tales modificaciones son denominadas modificación/es conservadora/s de aminoácidos. Ejemplos de modificaciones conservadoras están en el 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 pequeños aminoácidos (glicina, alanina, serina, treonina y metionina). Las modificaciones de aminoácidos, que generalmente no alteran la actividad específica son conocidas en la técnica y son descritas, 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 al igual que a la inversa (Taylor, 1986, *Journal of Theoretical Biology* 119: 205-218; <http://www.compbio.dundee.ac.uk/papers/amas/amas3d.html>).

25

Beta-glucosidasas progenitoras

30

[0045] En la presente invención, la beta-glucosidasa progenitora es preferiblemente (a) un polipéptido que comprende una secuencia de aminoácidos que tiene al menos 70% de identidad con aminoácidos 20 a 861 de la SEC ID n°: 2 o aminoácidos 20 a 863 de la SEC ID n°: 70 o (b) un polipéptido codificado por una secuencia de nucleótidos que se hibrida bajo condiciones de astringencia al menos bajas con los nucleótidos 58 a 2583 de SEC ID n°: 1 o nucleótidos 58 a 2589 de SEC ID n°: 71, o sus hebras complementarias.

35

40

[0046] En un primer aspecto, la beta-glucosidasa progenitora comprende una secuencia de aminoácidos que tiene un grado de identidad de aminoácidos 20 a 861 de SEC ID n°: 2 o aminoácidos 20 a 863 de la SEC ID n°: 70 (es decir, los polipéptidos maduros) de al menos 70% aproximadamente, preferiblemente al menos 75% aproximadamente, más preferiblemente al menos 80% aproximadamente, más preferiblemente al menos 85% aproximadamente, incluso más preferiblemente al menos 90% aproximadamente, de forma más preferible al menos 95% aproximadamente, e incluso de forma más preferible al menos 97% aproximadamente, que tienen actividad de beta-glucosidasa (en adelante, "polipéptidos homólogos"). En una forma de realización preferida, los polipéptidos homólogos tienen una secuencia de aminoácidos que difiere por cinco aminoácidos, preferiblemente por cuatro aminoácidos, más preferiblemente por tres aminoácidos, incluso más preferiblemente por dos aminoácidos y de forma más preferible por un aminoácido de los aminoácidos 20 a 861 de la SEC ID n°: 2 o aminoácidos 20 a 863 de la SEC ID n°: 70. Para objetivos de la presente invención, el grado de identidad entre dos secuencias de aminoácidos se determina por el método de Clustal (Higgins, 1989, *CABIOS* 5: 151-153) usando el Software de MEGALIGN™ de LASERGENE™ (DNASTAR, Inc., Madison, WI) con una tabla de identidad y los siguientes parámetros de alineación múltiple: Penalización por espacio de 10 y penalización por longitud de espacio de 10. Los parámetros de alineación por parejas fueron Ktuple=1, penalizaciones por espacio =3, ventanas =5 y diagonales=5.

45

50

55

[0047] Las beta-glucosidasas progenitoras sustancialmente homólogas pueden tener una o más sustituciones, deleciones o adiciones de aminoácidos. Estos cambios son preferiblemente de una naturaleza menor, es decir, sustituciones de aminoácidos conservadores como se ha descrito anteriormente y otras sustituciones que no afectan significativamente al pliegue tridimensional o a la actividad de la proteína o polipéptido; deleciones pequeñas, típicamente de uno a aproximadamente 30 aminoácidos; y pequeñas extensiones del amino o carboxilo-terminal, tales como un residuo de metionina aminoterminal, un pequeño péptido de enlace de hasta 20-25 residuos aproximadamente, o una pequeña extensión que facilita la purificación (una marca de afinidad), tal como un tracto de polihistidina, o proteína A (Nilsson et al., 1985, *EMBO J.* 4: 1075; Nilsson et al., 1991, *Methods Enzymol.* 198: 3. Véase, también, en general, Ford et al., 1991, *Protein Expression and Purification* 2: 95-107.

60

[0048] Aunque los cambios anteriormente descritos son preferiblemente de una naturaleza menor, tales cambios también pueden ser de una naturaleza considerable tal como fusión de polipéptidos más grandes de hasta 300 aminoácidos o más tanto como extensiones amino-terminales como carboxilo-terminales.

[0049] Además de los 20 aminoácidos estándares, los aminoácidos no estándares (como 4-hidroxiprolina, 6-N-metil lisina, ácido 2-aminoisobutírico, isovalina y alfa-metil serina) pueden sustituirse por residuos de aminoácidos de una beta-glucosidasa de tipo salvaje. Un número limitado de aminoácidos no conservadores, aminoácidos que no son codificados por el código genético, y los aminoácidos no naturales se pueden sustituir por residuos de aminoácidos. "Los aminoácidos no naturales" han sido modificados después de la síntesis de proteína, y/o tiene una estructura química en sus cadena/s lateral/es diferente de los aminoácidos estándares. Los aminoácidos no naturales pueden ser químicamente sintetizados y preferiblemente están disponibles comercialmente, e incluyen ácido pipercolico, ácido carboxílico de tiazolidina, dehidroprolina, 3- y 4-metilprolina y 3,3-dimetilprolina.

[0050] Preferiblemente, el polipéptido progenitor comprende los aminoácidos 20 a 861 de SEC ID nº: 2 o aminoácidos 20 a 863 de la SEC ID nº: 70 o una variante alélica de la misma; o un fragmento de la misma que tiene actividad de beta-glucosidasa. En otra forma de realización preferida, el polipéptido progenitor comprende los aminoácidos 20 a 861 de SEC ID nº: 2 o aminoácidos 20 a 863 de la SEC ID nº: 70. En otra forma de realización preferida, el polipéptido progenitor consiste en los aminoácidos 20 a 861 de SEC ID nº: 2 o aminoácidos 20 a 863 de la SEC ID nº: 70 o una variante alélica de la misma; o un fragmento de la misma que tiene actividad de beta-glucosidasa. En otra forma de realización preferida, el polipéptido progenitor consiste en los aminoácidos 20 a 861 de SEC ID nº: 2 o aminoácidos 20 a 863 de la SEC ID nº: 70. En otra forma de realización preferida, el polipéptido progenitor es codificado por la secuencia de nucleótidos contenida en *Escherichia coli* DSM 14240, caracterizado por el hecho de que la secuencia de ácidos nucleicos codifica un polipéptido que tiene actividad de beta-glucosidasa. En otra forma de realización preferida, el polipéptido progenitor es codificado por la región de codificación de polipéptido maduro contenida en *Escherichia coli* DSM 14240. En otra forma de realización preferida, el polipéptido progenitor es codificado por la secuencia de nucleótidos contenida en el plásmido pEJG113 que está contenido en *Escherichia coli* NRRL B-30695, caracterizado por el hecho de que la secuencia de ácidos nucleicos codifica un polipéptido que tiene actividad de beta-glucosidasa. En otra forma de realización preferida, el polipéptido progenitor es codificado por la región de codificación de polipéptido maduro contenida en el plásmido pEJG113 que está contenido en *Escherichia coli* NRRL B-30695.

[0051] Un fragmento de SEC ID nº: 2 o SEC ID nº: 70 es un polipéptido que tiene uno o más aminoácidos delecionados del término amino y/o carboxilo de esta secuencia de aminoácidos. Preferiblemente, un fragmento contiene al menos 770 residuos de aminoácidos, más preferiblemente, al menos 800 residuos de aminoácidos, y, de forma más preferible, al menos 830 residuos de aminoácidos.

[0052] En un segundo aspecto, la beta-glucosidasa progenitora es codificada por una secuencia de nucleótidos que hibridiza bajo condiciones de astringencia baja, 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 una sonda de nucleótidos que se hibrida bajo las mismas condiciones con (i) los nucleótidos 58 a 2583 de SEC ID nº: 1 o los nucleótidos 58 a 2589 de SEC ID nº: 71, (ii) la secuencia genómica de nucleótidos que comprende los nucleótidos 58 a 2583 de SEC ID nº: 1 o los nucleótidos 58 a 2589 de SEC ID nº: 71, (iii) una subsecuencia de (i) o (ii), o (iv) una cadena complementaria de (i); (ii), o (iii) (J. Sambrook, E.F. Fritsch, y T. Maniatus, 1989, Molecular Cloning, A Laboratory Manual, 2ª edición, Cold Spring Harbor, New York). La subsecuencia de la SEC ID nº: 1 puede ser 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 beta-glucosidasa.

[0053] Una subsecuencia de SEC ID nº: 1 o SEC ID nº: 71 u homólogo de la misma, es una secuencia de nucleótidos donde uno o más nucleótidos han sido delecionados del extremo 5' y/o 3'. Preferiblemente, una subsecuencia contiene al menos 2310 nucleótidos, más preferiblemente al menos 2400 nucleótidos, y, de forma más preferible, al menos 2490 nucleótidos.

[0054] Los polipéptidos progenitores también pueden ser variantes alélicas de los polipéptidos que tienen actividad de beta-glucosidasa. Una variante alélica denota cualquiera de dos o más formas alternativas de un gen que ocupa la misma localización cromosómica. La variación alélica surge naturalmente a través de la mutación, y puede suponer polimorfismo dentro de las poblaciones. Las mutaciones de gen pueden ser silenciosas (ningún cambio en el polipéptido codificado) o pueden codificar polipéptidos que tienen secuencias de aminoácidos alteradas. Una variante alélica de un polipéptido es un polipéptido codificado por una variante alélica de un gen.

[0055] La secuencia de nucleótidos de SEC ID nº: 1 o una subsecuencia de la misma, al igual que la secuencia de aminoácidos de SEC ID nº: 2 o SEC ID nº: 70, o un fragmento de la misma, se puede utilizar para diseñar sondas de nucleótido para identificar y clonar polipéptidos progenitores de codificación de ADN que tienen actividad de beta-glucosidasa 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 ADN genómico o ADNc del género o la especie de interés, siguiendo procedimientos estándares de hibridación de Southern, para identificar y aislar el gen correspondiente. Tales sondas

pueden ser considerablemente más cortas que la secuencia entera, pero deberían ser de al menos 15, preferiblemente al menos 25, y más preferiblemente al menos 35 nucleótidos de longitud. También pueden utilizarse sondas más largas. Pueden utilizarse tanto la sonda de ADN como la de ARN. Las sondas típicamente son marcadas para detectar el gen correspondiente (por ejemplo, con ^{32}P , ^3H , ^{35}S , biotina o avidina).

5

[0056] Una biblioteca genómica de ADN o ADNc obtenida a partir de esos otros organismos se pueden seleccionar para ADN que se hibrida con las sondas anteriormente descritas y que codifica un polipéptido progenitor que tiene actividad de beta-glucosidasa. El ADN genómico u otro ADN de esos otros organismos se pueden separar por agarosa o electroforesis en gel de poliacrilamida, u otras técnicas de separación. El ADN de las bibliotecas o el ADN separado se puede transferir y ser inmovilizado en nitrocelulosa u otro material portador adecuado. Para identificar un clon o ADN que sea homólogo a SEC ID n°: 1 o SEC ID n°: 71, o una subsecuencia de la misma, el material portador se usa en una hibridación de Southern. Para objetivos de la presente invención, la hibridación indica que la secuencia de nucleótidos hibridiza a una sonda de nucleótidos marcada correspondiente a la secuencia de nucleótidos mostrada en SEC ID n°: 1 o SEC ID n°: 71, su cadena complementaria, o una subsecuencia de la misma, bajo condiciones de astringencia de baja a muy alta. Las moléculas a las cuales la sonda marcada con radioactividad se hibrida pueden ser detectadas usando, por ejemplo, película radiográfica.

[0057] En una forma de realización preferida, la sonda de nucleótidos es una secuencia de nucleótidos que codifica el polipéptido de SEC ID n°: 2 o SEC ID n°: 70, o una subsecuencia de la misma. En otra forma de realización preferida, la sonda de nucleótidos es SEC ID n°: 1 o SEC ID n°: 71. En otra forma de realización preferida, la sonda de nucleótidos son los nucleótidos 58 a 2583 de SEC ID n°: 1 o los nucleótidos 58 a 2589 de SEC ID n°: 71. En otra forma de realización preferida, la sonda de nucleótidos es la secuencia de ácidos nucleicos contenida en *Escherichia coli* DSM 14240, caracterizada por el hecho de que la secuencia de ácidos nucleicos codifica un polipéptido que tiene actividad de beta-glucosidasa. En otra forma de realización preferida, la sonda de nucleótidos es la región de codificación de polipéptido maduro contenida en *Escherichia coli* DSM 14240. En otra forma de realización preferida, la sonda de nucleótidos es la secuencia de ácidos nucleicos contenida en el plásmido pEJG113 que está contenido en *Escherichia coli* NRRL B-30695, caracterizado por el hecho de que la secuencia de ácidos nucleicos codifica un polipéptido que tiene actividad de beta-glucosidasa. En otra forma de realización preferida, la sonda de nucleótidos es la región de codificación de polipéptido maduro contenida en el plásmido pEJG113 que está contenido en *Escherichia coli* NRRL B-30695.

[0058] Para las sondas largas de al menos 100 nucleótidos de longitud, las condiciones de astringencia baja a muy alta son definidas como prehibridación e hibridación a 42°C en 5X SSPE, SDS al 0,3%, 200 µg/ml de ADN de esperma de salmón cortado y desnaturalizado, y bien formamida al 25% para astringencias bajas, formamida al 35% para astringencias medias y medias-altas, o formamida al 50% para astringencias altas y muy altas, siguiendo procedimientos estándares de hibridación de Southern.

[0059] Para las sondas largas de al menos 100 nucleótidos de longitud, el material portador es finalmente lavado tres veces cada uno durante 15 minutos usando 2X SSC, SDS al 0,2% 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), de forma más preferible al menos a 65°C (astringencia alta), e incluso de forma más preferible al menos a 70°C (astringencia muy alta).

[0060] Para las sondas cortas que son de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, las condiciones de astringencia son definidas como prehibridación, hibridación y lavado post-hibridación a aproximadamente 5°C hasta 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, NP-40 al 0,5%, 1X Solución de Denhardt, 1 mM de pirofosfato de sodio, 1 mM de fosfato monobásico de sodio, 0,1 mM de ATP y 0,2 mg de ARN de levadura por ml siguiendo procedimientos estándares de hibridación de Southern.

[0061] Para las sondas cortas que son de aproximadamente 15 nucleótidos a aproximadamente 70 nucleótidos de longitud, el material portador se lava una vez en 6X SSC más SDS al 0,1% durante 15 minutos y dos veces cada una durante 15 minutos usando 6X SSC a 5°C a 10°C por debajo de la T_m calculada.

[0062] La beta-glucosidasa progenitora se puede obtener a partir de microorganismos de cualquier género. Para objetivos de la presente invención, el término "obtenido a partir de" según se utiliza en este caso en relación con una fuente dada significa que la beta-glucosidasa progenitora codificada por una secuencia de nucleótidos es producida por la fuente o por una célula en la cual la secuencia de nucleótidos de la fuente ha sido insertada. En una forma de realización preferida, la beta-glucosidasa progenitora es segregada extracelularmente.

[0063] La beta-glucosidasa progenitora puede ser una beta-glucosidasa fúngica. En una forma de realización más preferida, la beta-glucosidasa fúngica es una beta-glucosidasa de levadura tal como una beta-glucosidasa *Candida*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces* o *Yarrowia*. En otra forma de realización más preferida, la beta-glucosidasa fúngica es una beta-glucosidasa filamentosa fúngica tal como una beta-glucosidasa de *Acremonium*, *Agaricus*, *Alternaria*, *Aspergillus*, *Botryosphaeria*, *Ceriporiopsis*, *Chaetomidium*, *Claviceps*, *Cochliobolus*, *Coprinopsis*,

65

Coptotermes, Corynascus, Cryphonectria, Diplodia, Exidia, Fusarium, Gibberella, Holomastigotoides, Humicola, Irpex, Lentinula, Leptosphaeria, Magnaporthe, Melanocarpus, Meripilus, Myceliophthora, Neurospora, Penicillium, Phanerochaete, Poitrasia, Pseudoplectania, Pseudotriconympha, Rhizomucor, Scytalidium, Talaromyces, Thermoascus, Thielavia, Trichoderma, Trichophaea, Verticillium, Volvariella o Xylaria.

5

[0064] En una forma de realización más preferida, la beta-glucosidasa progenitora es una beta-glucosidasa de *Saccharomyces carlsbergensis, Saccharomyces cerevisiae, Saccharomyces diastaticus, Saccharomyces douglasii, Saccharomyces kluyveri, Saccharomyces norbensis* o *Saccharomyces oviformis*.

10

[0065] En otra forma de realización más preferida, la beta-glucosidasa progenitora es una beta-glucosidasa de *Acremonium cellulolyticus, Aspergillus aculeatus, Aspergillus awamori, Aspergillus fumigatus, Aspergillus foetidus, Aspergillus fumigatus, Aspergillus japonicus, Aspergillus nidulans, Aspergillus niger, Aspergillus oryzae, 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 sarcochrourm, Fusarium solani, Fusarium sporotrichioides, Fusarium sulphureum, Fusarium torulosum, Fusarium trichothecioides, Fusarium venenatum, Humicola grisea, Humicola insolens, Humicola lanuginosa, Irpex lacteus, Mucor miehei, Myceliophthora thermophila, Neurospora crassa, Penicillium funiculosum, Penicillium purpurogenum, Phanerochaete chrysosporium, Schizophyllum commune, Sclerotium rolfsii, Sporotrichum cellulophilum, Talaromyces emersonii, Thielavia terrestris, Trichoderma harzianum, Trichoderma koningii, Trichoderma longibrachiatum, Trichoderma reesei* o *Trichoderma viride*.

15

20

[0066] En una forma de realización aún más preferida, la beta-glucosidasa progenitora es una beta-glucosidasa de *Aspergillus oryzae*, y más preferiblemente la beta-glucosidasa de *Aspergillus oryzae* de SEC ID n°: 2 o el polipéptido maduro de la misma. En otra forma de realización preferida, la beta-glucosidasa progenitora es codificada por la secuencia de nucleótidos contenida en *E. coli* DSM 14240, en la que la secuencia de nucleótidos codifica un polipéptido que tiene actividad de beta-glucosidasa. En otra forma de realización, la más preferida, la beta-glucosidasa progenitora es codificada por la región de codificación del polipéptido maduro contenida en *E. coli* DSM 14240.

25

30

[0067] En una forma de realización más preferida, la beta-glucosidasa progenitora es una beta-glucosidasa de *Aspergillus fumigatus*, y de la forma más preferible la beta-glucosidasa de *Aspergillus fumigatus* de SEC ID n°: 70 o el polipéptido maduro de la misma. En otra forma de realización preferida, la beta-glucosidasa progenitora es codificada por la secuencia de nucleótidos contenida en el plásmido pEJG113 que está contenido en *Escherichia coli* NRRL B-30695, caracterizado por el hecho de que la secuencia de nucleótidos codifica un polipéptido que tiene actividad de beta-glucosidasa. En otra forma de realización preferida, la glucosidasa beta progenitora es codificada por la región de codificación del polipéptido maduro contenida en el plásmido pEJG113 que está contenido en *Escherichia* NRRL B-30695.

35

40

[0068] Se entenderá que para las especies mencionadas, la invención incluye los estados perfectos e imperfectos, y otros equivalentes taxonómicos, por ejemplo, anamorfos, independientemente del nombre de la especie por el cual son conocidos. Los expertos en la técnica fácilmente reconocerán la identidad de los equivalentes apropiados.

45

[0069] Las cepas de estas especies son fácilmente accesibles al público en una cantidad de colecciones de cultivos, tales como el 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).

50

55

[0070] Las beta-glucosidasas progenitoras también pueden ser identificadas y obtenidas a partir de otras fuentes incluidos los microorganismos aislados de la naturaleza (p. ej., tierra, abonos, agua, etc.) o muestras de ADN obtenidas directamente de materiales naturales (p. ej., tierra, abonos, agua, etc.) usando las sondas mencionadas anteriormente. Las técnicas para aislar microorganismos y ADN directamente de hábitats naturales son bien conocidas en la técnica. La secuencia de nucleótidos que codifica una beta-glucosidasa puede luego ser derivada seleccionando de manera similar una biblioteca genómica o de ADNc de otro microorganismo o muestra de ADN mezclado. Una vez que una secuencia de nucleótidos que codifica una beta-glucosidasa ha sido detectada con sonda/s adecuada/s como se describe en la presente, la secuencia se puede aislar o clonar utilizando técnicas conocidas por expertos en la materia (véase, por ejemplo, J. Sambrook, E.F. Fritsch, y T. Maniatus, 1989, Molecular Cloning, A Laboratory Manual, 2ª edición, Cold Spring Harbor, New York).

60

[0071] Tal y como se define en la presente, una beta-glucosidasa "aislada" es un polipéptido que está esencialmente libre de otros polipéptidos de glucosidasa no beta, por ejemplo, al menos aproximadamente 20% puro, preferiblemente al menos aproximadamente 40% puro, más preferiblemente aproximadamente 60% puro, incluso más preferiblemente aproximadamente 80% puro, de forma más preferible aproximadamente 90% puro, e incluso de forma más preferible aproximadamente 95% puro, según se determina por SDS-PAGE.

65

[0072] Las beta-glucosidasas progenitoras también pueden incluir polipéptidos fusionados o polipéptidos de fusión divisibles en los cuales otro polipéptido se fusiona en el N-término o el C-término del polipéptido o fragmento del mismo. Un polipéptido fusionado se produce por fusión de una secuencia de nucleótidos (o una parte de la misma) que codifica

otro polipéptido a una secuencia de nucleótidos (o una parte de la misma) de la presente invención. Las 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 éstas estén dentro del marco y que la expresión del polipéptido fusionado esté bajo control del/los mismo/s promotor/es y terminador. Las proteínas de fusión también pueden ser construidas usando tecnología de intein en la que las fusiones se crean post-traduccionalmente (Cooper et al., 1993, EMBO J. 12: 2575-2583; Dawson et al., 1994, Science 266: 776-779).

Variantes

[0073] En la presente invención, las variantes aisladas de una beta-glucosidasa progenitora comprenden una sustitución en una posición correspondiente a la posición 285 de los aminoácidos 20 a 861 de la SEC ID N°: 2 o correspondiente a la posición 285 de los aminoácidos 20 a 863 de la SEC ID N°: 70, en la que la variante tiene actividad de beta-glucosidasa y comprende las sustituciones Q202R + H285Q + D722G (o D724G); las sustituciones Q202R + H285Q, o las sustituciones H285Q + D722G (o D724G) comparado con la beta-glucosidasa progenitora. Preferiblemente, la variante, que tiene actividad de beta-glucosidasa, comprende una secuencia de aminoácidos que tiene un grado de identidad de al menos 70% aproximadamente, preferiblemente al menos 75% aproximadamente, más preferiblemente al menos 80% aproximadamente, más preferiblemente al menos 85% aproximadamente, incluso más preferiblemente al menos 90% aproximadamente, de forma más preferible al menos 95% aproximadamente, e incluso de forma más preferible al menos 97% aproximadamente a la secuencia de aminoácidos de la beta-glucosidasa progenitora. Para objetivos de la presente invención, el grado de identidad entre dos secuencias de aminoácidos se determina por el método de Clustal (Higgins, 1989, CABIOS 5: 151-153) usando el Software de MEGALIGN™ de LASERGENE™ (DNASTAR, Inc., Madison, WI) con una tabla de identidad y los siguientes parámetros de alineación múltiple: Penalización por gap de 10 y penalización por longitud de gap de 10. Los parámetros de alineación en parejas fueron Ktuple=1, penalizaciones por gap =3, ventanas =5 y diagonales=5.

[0074] Los aminoácidos esenciales en la beta-glucosidasa progenitora pueden identificarse según procedimientos conocidos en la técnica, tales como mutagénesis dirigida o mutagénesis de barrido de alanina (Cunningham y Wells, 1989, Science 244: 1081-1085). En esta última técnica, se introducen mutaciones de alanina únicas en cada residuo en la molécula, y las moléculas mutantes resultantes se evalúan para analizar su actividad biológica (es decir, actividad de beta-glucosidasa) para identificar residuos de aminoácidos que son críticos para la actividad de la molécula. Véase también, Hilton et al., 1996, J. Biol. Chem. 271: 4699-4708. El sitio activo de la enzima u otra interacción biológica también puede ser determinado por análisis físico de estructura, como se determina por técnicas tales como resonancia magnética nuclear, cristalografía, difracción electrónica o marcaje por fotoafinidad, conjuntamente con mutación de aminoácidos de sitio putativo de contacto. Véase, 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 también pueden ser inferidas del análisis de identidades con polipéptidos que se relacionan con un polipéptido según la invención.

[0075] Pueden realizarse y evaluarse sustituciones de aminoácidos únicas o múltiples usando métodos conocidos de mutagénesis, recombinación y/o bajarado, seguido de un procedimiento de selección pertinente, como los descritos por Reidhaar-Olson y Sauer, 1988, Science 241: 53-57; Bowie y Sauer, 1989, Proc. Natl. Acad. Sci. USA 86: 2152-2156; WO 95/17413; o WO 95/22625. Otros métodos que pueden ser usados incluyen PCR con tendencia al error, exposición en fago (p. ej., Lowman et al., 1991, Biochem. 30:10832-10837; patente de EE.UU. n°. 5.223.409; WO 92/06204) y mutagénesis dirigida a una región (Derbyshire et al., 1986, Gene 46:145; Ner et Al., 1988, DNA 7:127).

[0076] Los métodos de mutagénesis/bajarado se pueden combinar con métodos de selección automatizados de alto rendimiento para detectar actividad de polipéptidos mutagenizados clonados expresados por células huéspedes. La moléculas de ADN mutagenizadas que codifican polipéptidos activos se pueden recuperar de las células huéspedes y ser ordenadas rápidamente usando métodos estándar de la técnica. Estos métodos permiten la determinación rápida de la importancia de los residuos de aminoácidos individuales en un polipéptido de interés, y se pueden aplicar a polipéptidos de estructura desconocida.

[0077] En una forma de realización preferida, el número de sustituciones de aminoácidos en las variantes de la presente invención comprenden preferiblemente 4 sustituciones, más preferiblemente 3 sustituciones, incluso más preferiblemente 2 sustituciones, y de forma más preferible, 1 sustitución. En otra forma de realización preferida, el número de sustituciones de aminoácidos en las variantes de la presente invención consiste en preferiblemente 4 sustituciones, más preferiblemente 3 sustituciones, incluso más preferiblemente 2 y de forma más preferible, 1 sustitución.

[0078] En una forma de realización preferida, una variante de la presente invención consiste en 741 a 750, 751 a 760, 761 a 770, 771 a 780, 781 a 790, 791 a 800, 801 a 810, 811 a 820, 821 a 830, 831 a 840, 841 a 850, 851 a 860, 861 a 870, 871 a 880, 881 a 890, 891 a 900, 901 a 910, 911 a 920, 921 a 930, 931 a 940, o 941 a 950 amino ácidos.

[0079] Las variantes de la presente invención pueden comprender además una o más deleciones y/o inserciones de la secuencia de aminoácidos.

[0080] En otra forma de realización preferida, la variante comprende Gln como una sustitución en una posición correspondiente a la posición 285 de los aminoácidos 20 a 861 de la SEC ID N°: 2 o los aminoácidos 20 a 863 de la SEC ID N°: 70. En otra forma de realización preferida, la variante comprende la sustitución H285Q de los aminoácidos 20-861 de la SEC ID N°: 2 o los aminoácidos 20 a 863 de la SEC ID N°: 70.

5

[0081] En otra forma de realización más preferida, la variante comprende las sustituciones Q202R + H285Q de los aminoácidos 20 a 861 de la SEC ID N°: 2. En otra forma de realización preferida, la variante comprende las sustituciones H285Q + D722G de los aminoácidos 20 a 861 de la SEC ID N°: 2.

10

[0082] En otra forma de realización preferida, la variante comprende las sustituciones Q202R + H285Q de los aminoácidos 20 a 863 de la SEC ID N°: 70. En otra forma de realización más preferida, la variante comprende las sustituciones H285Q + D724G de los aminoácidos 20 a 863 de la SEC ID N°: 70.

15

[0083] En otra forma de realización más preferida, la variante comprende las sustituciones Q202R + H285Q + D722G de los aminoácidos 20 a 861 de la SEC ID N°: 2. En otra forma de realización preferida, la variante comprende las sustituciones Q202R + H285Q + D724G de los aminoácidos 20 a 863 de la SEC ID N°: 70. Tal y como se define en la presente, una "variante aislada" de una beta-glucosidasa progenitora es un polipéptido que está esencialmente libre de otros polipéptidos de glucosidasa no beta, por ejemplo, al menos aproximadamente 20% puro, preferiblemente al menos aproximadamente 40% puro, más preferiblemente aproximadamente 60% puro, incluso más preferiblemente aproximadamente 80% puro, de la forma más preferible aproximadamente 90% puro, y de la forma aún más preferible, aproximadamente 95% puro, según se determina por SDS-PAGE.

20

Plásmidos

25

[0084] El plásmido o los plásmidos usados para preparar variantes de beta-glucosidasa puede ser cualquier plásmido o vector que puede someterse a procedimientos de ADN recombinante. El plásmido que comprende una secuencia de nucleótidos que codifica una beta-glucosidasa puede ser preparado ligando la secuencia de nucleótidos en un plásmido adecuado, o por cualquier otro método adecuado. El plásmido contiene preferiblemente uno o más marcadores seleccionables descritos en la presente que permiten la fácil selección de células transformadas. La elección del plásmido frecuentemente dependerá de la célula huésped en la cual se va a introducir.

30

[0085] En la presente invención, el plásmido puede ser un plásmido que se replica de manera autónoma, es decir, un plásmido que existe como una entidad extracromosómica, cuya replicación es diferente de la replicación cromosómica.

35

[0086] El replicador de plásmido puede ser cualquier replicador de plásmido que medie la replicación autónoma que funciona en una célula. El término "replicador de plásmido" es definido en la presente como una secuencia que le permite a un plásmido o vector replicarse *in vivo*. Ejemplos de un replicador de plásmido útil en una célula de levadura son el origen de replicación de 2 micra, ARS1, ARS4, la combinación de ARS1 y CEN3 y la combinación de ARS4 y CEN6. Ejemplos de un replicador de plásmido útiles en una célula fúngica filamentosa 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). El aislamiento del gen AMA1 y la construcción de plásmidos o vectores que comprenden el gen se puede realizar según los métodos descritos en WO 00/24883.

40

[0087] La linealización del/los plásmido/s puede ser dirigida hacia cualquier sitio en el plásmido. El/los plásmido/s se puede/n linealizar por cualquier método adecuado conocido en la técnica, por ejemplo, digestión con una o más enzimas de restricción. Las extremidades linealizadas del plásmido pueden ser rellenadas con nucleótidos como se describe por Pompon et al., 1989, Gene 83: 15-24. No obstante, se prefiere no rellenar las extremidades linealizadas puesto que esto puede crear un desplazamiento del marco de lectura.

45

50

[0088] Para facilitar el proceso de selección, el plásmido es preferiblemente un vector de expresión en el que la secuencia de nucleótidos en cuestión está operativamente vinculada a segmentos adicionales requeridos para la transcripción del ADN. En general, el vector de expresión se deriva de un plásmido, un cósmido o un bacteriófago, o puede contener elementos de cualquiera de estos o todos estos. Para objetivos de la presente invención, los términos "plásmido" y "vector" se usan de forma intercambiable.

55

Fragmentos de ADN

[0089] La biblioteca de fragmentos de ADN para ser combinados de forma aleatoria (o "redistribuidos") con regiones homólogas en el/los plásmido/s linealizado/s por recombinación *in vivo* se puede preparar por cualquier método adecuado. Por ejemplo, el fragmento de ADN se puede preparar por amplificación por PCR (p. ej., PCR con tendencia al error) de un plásmido que comprende la secuencia de nucleótidos, usando cebadores específicos, por ejemplo, como se describe en la patente de EE.UU. n°. 4.683.202 o Saiki et al., 1988, Science 239: 487-491. El fragmento de ADN puede también ser aislado de un plásmido que comprende la secuencia de nucleótidos deseada por digestión con enzimas de restricción, seguida de aislamiento usando, por ejemplo, electroforesis.

60

65

5 [0090] El fragmento de ADN puede ser preparado de forma alternativa sintéticamente por métodos estándares establecidos, por ejemplo, el método de fosfoamidita descrito por Beaucage y Caruthers, 1981, Tetrahedron Letters 22: 1859-1869, o el método descrito por Matthes et al., 1984, EMBO Journal 3: 801-805. Según el método de fosfoamidita, los oligonucleótidos son sintetizados en un sintetizador de ADN automático, purificados, apareados, ligados y clonados en plásmidos adecuados.

10 [0091] El fragmento de ADN también puede ser de orígenes mezclados sintético y genómico, mezclado sintético y de ADNc o mezclado genómico y de ADNc preparados ligando fragmentos de origen de ADNc, genómico o sintético, los fragmentos correspondientes a varias partes de la secuencia de nucleótidos entera, conforme a técnicas estándares.

15 [0092] La biblioteca de fragmentos de ADN comprenden una o más mutaciones de la secuencia de nucleótidos, caracterizados por el hecho de que los fragmentos comprenden al menos dos regiones, una o más regiones que son homólogas a la región 5' o la región 3' del espacio en la secuencia de nucleótidos linealizada y/o la secuencia de plásmidos y una o más segundas regiones que son homólogas a la región 5' o la región 3' de los fragmentos de ADN de la biblioteca.

[0093] Las regiones del fragmento de ADN pueden ser cualquier secuencia que sea homóloga a la secuencia de nucleótidos y/o secuencia de plásmidos.

20 [0094] En una forma de realización preferida, las regiones del fragmento de ADN son una región 5' y/o una región 3' que flanquea un gen que codifica una beta-glucosidasa; o una región 5' y/o una región 3' de un gen que codifica una beta-glucosidasa.

25 [0095] En otra forma de realización preferida de la presente invención, el fragmento o los fragmentos de ADN son preparados bajo condiciones que conducen a una frecuencia de mutagénesis aleatoria baja, media o alta. Para obtener la frecuencia de mutagénesis baja la/s secuencia/s de nucleótidos (que comprenden el o los fragmentos de ADN) se pueden preparar por un método de amplificación por PCR estándar (US 4.683.202 o Saiki et al., 1988, Science 239: 487-491). Una frecuencia de mutagénesis media o alta se puede obtener mediante la realización de la amplificación por PCR bajo condiciones que reduzcan la fidelidad de la replicación por la polimerasa termoestable y aumenten la incorporación incorrecta de nucleótidos, por ejemplo como se describe por Dethler, 1992, GATA 9: 103-106; Leung et al., 1989, BioTechniques 1: 11-15.

35 [0096] La amplificación por PCR se puede combinar con una fase de mutagénesis usando un agente mutagenizante físico o químico adecuado, por ejemplo, uno que induzca transiciones, transversiones, inversiones, aleatorización, deleciones y/o inserciones.

40 [0097] En una forma de realización preferida, el o los fragmento/s de ADN que se debe/n redistribuir preferiblemente tiene/n una longitud de aproximadamente 15 bp a 8 kb, más preferiblemente aproximadamente 30 bp a 6 kb, incluso más preferiblemente aproximadamente 40 bp a 6 kb, incluso más preferiblemente aproximadamente 80 bp a 4 kb, y de forma más preferible aproximadamente 100 bp a 2 kb, para ser capaces de interactuar óptimamente con el plásmido linealizado.

Células fúngicas

45 [0098] La célula fúngica, en la cual la mezcla de secuencias de plásmidos/fragmentos de nucleótidos deben ser introducidas, puede ser cualquier célula fúngica útil en la presente invención. Una "célula fúngica de recombinación" es definida en la presente como una célula capaz de mediar la recombinación de varias secuencias de nucleótidos homólogas.

50 [0099] En una forma de realización preferida, la célula de recombinación fúngica es una célula de levadura. En una forma de realización más preferida, la célula de recombinación de levadura es una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces* o *Yarrowia*.

55 [0100] En una forma de realización más preferida, la célula de recombinación de levadura es una célula de *Kluyveromyces lactis*, *Saccharomyces carlsbergensis*, *Saccharomyces cerevisiae*, *Saccharomyces diastaticus*, *Saccharomyces douglasii*, *Saccharomyces kluyveri*, *Saccharomyces norbensis*, *Saccharomyces oviformis* o *Yarrowia lipolytica*.

60 [0101] En otra forma de realización preferida, la célula de recombinación fúngica es una célula micótica filamentosa. En una forma de realización más preferida, la recombinación fúngica filamentosa es una célula de *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolyopocladium* o *Trichoderma*.

65 [0102] En la forma de realización más preferida, la célula de recombinación fúngica filamentosa es una célula de *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* o *Aspergillus oryzae*. En otra forma de realización más preferida, la célula de recombinación fúngica filamentosa es una célula de *Fusarium bactridioides*, *Fusarium cerealis*, *Fusarium crookwellense*, *Fusarium culmorum*, *Fusarium graminearum*,

Fusarium gramineum, *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 otra forma de realización más preferida, la célula de recombinación fúngica filamentosa es una célula de *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* o *Trichoderma viride*.

[0103] En otra forma de realización más preferida, la célula de *Aspergillus* es una célula de *Aspergillus oryzae*.

[0104] En otra forma de realización más preferida, la célula de *Aspergillus* es una célula de *Aspergillus niger*.

[0105] En otra forma de realización más preferida, la célula de *Fusarium venenatum* es *Fusarium venenatum* A3/5, que fue originalmente depositada como *Fusarium graminearum* ATCC 20334 y recientemente reclasificada como *Fusarium venenatum* por Yoder y Christianson, 1998, Fungal Genetics and Biology 23: 62-80 y O'Donnell et al., 1998, Fungal Genetics and Biology 23: 57-67; al igual que equivalentes taxonómicos de *Fusarium venenatum* independientemente del nombre de especie por el cual son conocidos actualmente. En otra forma de realización más preferida, la célula de *Fusarium venenatum* es un mutante morfológico de *Fusarium venenatum* A3/5 o *Fusarium venenatum* ATCC 20334 como se describe en WO 97/26330.

[0106] Las células fúngicas se pueden transformar por un proceso que implica formación de protoplastos, transformación de los protoplastos y regeneración de la pared celular en un modo conocido *per se*. Los procedimientos adecuados para la transformación de las células huéspedes de *Aspergillus* y *Trichoderma* están descritos en EP 238 023 y Yelton et al, 1984, Proceedings of the National Academy of Sciences USA 81: 1470-1474. Métodos adecuados para transformar 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. y Simon, M.I., editores, Guide to Yeast Genetics and Molecular Biology, Methods in Enzymology, Volume 194, páginas 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.

30 **Recombinación *in vivo***

[0107] Un gran número de variantes o genes homólogos se pueden combinar en una transformación para crear eficazmente quimeras genéticas de los genes homólogos. La redistribución de estos genes, que codifican variantes mejoradas de genes de tipo salvaje, da como resultado quimeras que se pueden expresar y seguido de una selección para identificar aquellas quimeras con la combinación óptima de mutaciones provechosas. El proceso aumenta de manera múltiple el número de variantes mejoradas adicionales que se puede obtener en comparación con un proceso que sólo utiliza mutagénesis aleatoria (para obtener un resumen, véase Kuchner y Arnold, 1997, TIBTech 15:523-530). La mutagénesis aleatoria introduce mutaciones en una secuencia de nucleótidos objetivo, creando mutaciones deletéreas con mucha más frecuencia que las provechosas. En ciclos iterativos de esa mutagénesis, las mutaciones deletéreas se acumulan más rápidamente que las provechosas, enmascarando eficazmente la identificación de mutaciones provechosas durante la selección. La recombinación aleatoria entre dos o más secuencias de nucleótidos homólogos que contienen múltiples cambios de nucleótidos únicos. En sus secuencias de nucleótidos permite potencialmente que todos aquellos cambios de nucleótidos contenidos en una variante se separen uno del otro y sean combinados de forma aleatoria, en cambio, con cualquier mutación presente en otras variantes. Esta redistribución de mutaciones proporciona un medio por el cual las mutaciones de distintas secuencias progenitoras pueden ser combinadas entre sí de forma aleatoria para aumentar la probabilidad de combinación de los cambios de nucleótidos en una secuencia única de nucleótidos.

[0108] La recombinación eficaz de fragmentos de superposición múltiple usando el método de recombinación *in vivo* es un medio de generar quimeras a partir de variantes o genes homólogos. Una superposición tan pequeña como 15 bp es suficiente para la recombinación, y se puede utilizar para redistribución de dominios muy sencillo incluso de genes remotamente relacionados. En la redistribución de dominios, los bloques más grandes de ADN no homólogos son clasificados de forma aleatoria mediante extensiones de homología en sus terminales.

[0109] Se prefiere que al menos un ciclo de redistribución sea un ciclo de retrocruce con el fragmento o fragmentos de ADN inicialmente usados, que pueden ser el fragmento de ADN de tipo salvaje. Esto elimina las mutaciones no esenciales. Las mutaciones no esenciales también pueden ser eliminadas usando fragmentos de ADN de tipo salvaje como el material de ADN de entrada usado inicialmente.

[0110] Más de dos secuencias de nucleótidos pueden redistribuirse al mismo tiempo, y pueden ser ventajosas puesto que un gran número de variantes bastante diferentes pueden ser hechos rápidamente sin una abundancia de procedimientos iterativos. Al recombinar muchos fragmentos de la misma región, la superposición múltiple de los fragmentos aumentará la frecuencia de intercambio de ADN por sí sola, pero también es importante tener un número relativamente alto de cruces aleatorios en las regiones de superposición para recombinar las variantes/diferencias cercanamente localizadas.

[0111] Una superposición tan pequeña como 15 bp entre dos fragmentos es suficiente para obtener una recombinación eficaz. Por lo tanto, en la presente invención, es conveniente la superposición en el rango de 15 a 5000 bp, preferiblemente de 30 bp a 500 bp, especialmente 30 bp a 100 bp.

5 [0112] En la presente invención, preferiblemente 2 o más fragmentos superpuestos, más preferiblemente 2 a 50 fragmentos superpuestos, y de forma más preferible 2 a 10 fragmentos superpuestos, pueden ser usados ventajosamente como fragmentos de ADN en un ciclo de redistribución.

10 [0113] Además de permitir la creación de genes quiméricos, la utilización de fragmentos superpuestos es un método útil para la redistribución de dominios mediante la creación de superposiciones pequeñas entre fragmentos de ADN de distintos dominios y la selección para lograr la mejor combinación. Por ejemplo, en el caso de tres fragmentos de ADN, las regiones de superposición pueden ser de la siguiente manera: el primer extremo del primer fragmento se superpone al primer extremo del plásmido linealizado, el primer extremo del segundo fragmento se superpone al segundo extremo del primer fragmento, y el segundo extremo del segundo fragmento se superpone con el primer extremo del tercer fragmento, el primer extremo del tercer fragmento se superpone (como se indica más arriba) con el segundo extremo del segundo fragmento, y el segundo extremo del tercer fragmento se superpone con el segundo extremo del plásmido linealizado.

20 [0114] Se entiende que cuando se usan dos o más fragmentos de ADN como la materia prima, se prefiere tener superposiciones continuas entre las extremidades del plásmido y los fragmentos de ADN.

25 [0115] Aunque se prefiere redistribuir secuencias de nucleótidos homólogas en forma de fragmento/s de ADN y plásmido/s linealizado/s, también es posible redistribuir dos o más plásmidos linealizados que comprendan secuencias de nucleótidos homólogas que codifican polipéptidos. No obstante, en tal caso, es importante linearizar los plásmidos en diferentes sitios.

30 [0116] En la presente invención, dos o plásmidos linealizados y uno o más fragmentos de ADN homólogos pueden ser usados como la materia prima que se debe redistribuir. La proporción entre el/los plásmido/s linealizado/s y el/los fragmento/s de ADN homólogo preferiblemente están en el rango de 20:1 a 1:50, preferible de 2:1 a 1:10 (moles de plásmido:moles de fragmentos) con las concentraciones específicas que van de 1 pM a 10 M del ADN.

35 [0117] Los plásmidos linealizados se pueden escindir de manera que la superposición entre los fragmentos se delecte en el plásmido. La reparación del espacio en el plásmido luego requiere que los fragmentos se recombinen uno con el otro además de recombinarse con las extremidades del plásmido escindido para reconstruir un plásmido que se replica de manera autónoma y circular. En una forma de realización preferida, la linearización del plásmido o vector crea un espacio suficiente en la secuencia codificante de la secuencia de nucleótidos para forzar la recombinación homóloga de los fragmentos de ADN con las regiones correspondientes de la secuencia de nucleótidos, recreando un plásmido de replicación circular.

40 **Secuencias de nucleótidos**

[0118] La presente invención también se refiere a secuencias de nucleótidos aisladas que codifican variantes de beta-glucosidasas de la invención. Las secuencias de nucleótidos aisladas que codifican variantes de beta-glucosidasa de la presente invención pueden comprender además una o más deleciones y/o inserciones de la secuencia.

45 [0119] En una forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende Arg y Gln como sustituciones en posiciones correspondientes a las posiciones 202, y 285, respectivamente, de los aminoácidos 20 a 861 de la SEC ID N°: 2 o los aminoácidos 20 a 863 de la SEC ID N°: 70. En otra forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende Gln y Gly como sustituciones en posiciones correspondientes a las posiciones 285 y 722, respectivamente, de los aminoácidos 20 a 861 de la SEC ID N°: 2.

50 [0120] En otra forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende Gln y Gly como sustituciones en las posiciones correspondientes a las posiciones 285 y 724, respectivamente, de los aminoácidos 20 a 863 de la SEC ID N°: 70.

60 [0121] En otra forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende Arg, Gln y Gly como sustituciones en las posiciones correspondientes a las posiciones 202, 285 y 722, respectivamente, de los aminoácidos 20 a 861 de la SEC ID N°: 2.

[0122] En otra forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende Arg, Gln y Gly como sustituciones en las posiciones correspondientes a las posiciones 202, 285 y 724, respectivamente, de los aminoácidos 20 a 863 de la SEC ID N°: 70.

65 [0123] En otra forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende las sustituciones Q202R + H285Q de los aminoácidos 20 a 861 de la SEC ID N°: 2. En otra

forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende las sustituciones H285Q + D722G de los aminoácidos 20 a 861 de la SEC ID N°: 2.

5 [0124] En otra forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende las sustituciones Q202R + H285Q de los aminoácidos 20 a 861 de la SEC ID N°: 2. En otra forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende las sustituciones H285Q + D724G de los aminoácidos 20 a 863 de la SEC ID N°: 70.

10 [0125] En otra forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende las sustituciones Q202R + H285Q + D722G de los aminoácidos 20 a 861 de la SEC ID N°: 2. En otra forma de realización preferida, la secuencia de nucleótidos aislada codifica una variante de beta-glucosidasa que comprende las sustituciones Q202R + H285Q + D724G de los aminoácidos 20 a 863 de la SEC ID N°: 70.

15 [0126] El término "secuencia de nucleótidos aislada", según se utiliza en este caso, se refiere a una secuencia de nucleótidos que está esencialmente libre de otras secuencias de nucleótidos, por ejemplo, al menos aproximadamente 20% puro, preferiblemente al menos aproximadamente 40% puro, más preferiblemente al menos aproximadamente 60% puro, incluso más preferiblemente al menos aproximadamente 80% puro, y de la forma más preferible, al menos aproximadamente 90% puro, según se determina por electroforesis de agarosa. Por ejemplo, una secuencia de nucleótidos aislada se puede obtener por procedimientos estándares de clonación usados en ingeniería genética para
20 reemplazar la secuencia de nucleótidos desde su ubicación natural a un sitio diferente donde será reproducida. Los procedimientos de clonación pueden implicar escisión y aislamiento de un fragmento de nucleótidos deseado que comprende la secuencia de nucleótidos que codifica el polipéptido, inserción del fragmento en una molécula de vector, e incorporación del vector recombinante en una célula huésped donde se replicarán múltiples copias o clones de la secuencia de nucleótidos. La secuencia de nucleótidos puede ser de origen genómico, de ADNc, de ARN, semisintético,
25 sintético, o cualquier combinación de los mismos.

Constructos de ácidos nucleicos

30 [0127] La presente invención también se refiere a constructos de ácidos nucleicos que comprenden una secuencia de nucleótidos que codifica una variante de beta-glucosidasa de la presente invención operativamente ligada 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. Se entenderá que la expresión incluye cualquier fase implicada en la producción del polipéptido incluyendo, de modo enunciativo y no limitativo, transcripción, modificación postranscripcional, traducción, modificación postraduccional y secreción.

35 [0128] El "constructo de ácido nucleico" se define en la presente como una molécula de ácido nucleico, monocatenaria o bicatenaria, que se aísla de un gen de origen natural o que ha sido modificado para contener segmentos de ácido nucleico combinados y superpuestos en un modo que, de lo contrario, no existiría en la naturaleza. El término constructo de ácidos nucleicos es sinónimo del término cassette de expresión cuando el constructo de ácidos nucleicos contiene todas las secuencias de control requeridas para la expresión de una secuencia codificante de una variante de la presente invención. El término "secuencia codificante" es definido en la presente como una secuencia de nucleótidos que especifica directamente la secuencia de aminoácidos de su producto proteínico. Los bordes de una secuencia codificante genómica generalmente son determinados por el codón de iniciación ATG (eucariotas), o codones de inicio alternativo tales como GTG y TTG, localizados justo arriba del marco de lectura abierto en el extremo 5' del ARNm y una secuencia del terminador de transcripción localizado justo corriente abajo del marco de lectura abierto en el extremo 3' del ARNm. Una secuencia codificante puede incluir, de modo enunciativo y no limitativo, ADN, ADNc y secuencias de nucleótidos recombinantes.

50 [0129] Una secuencia de nucleótidos aislada que codifica una variante de beta-glucosidasa de la presente invención puede ser manipulada en una variedad de maneras para proporcionar la expresión de la variante. La manipulación de la secuencia de nucleótidos antes de su inserción en un vector puede ser conveniente o necesaria dependiendo del vector de expresión. Las técnicas para modificar las secuencias de nucleótidos utilizando métodos de ADN recombinante son bien conocidas en la técnica.

55 [0130] El término "secuencias de control" es definido en la presente para incluir todos los componentes que son necesarios o ventajosos para la expresión de una variante de beta-glucosidasa de la presente invención. Cada secuencia de control puede ser nativa o extranjera a la secuencia de nucleótidos que codifica la variante. Tales secuencias de control incluyen, de manera enunciativa y no limitativa, una secuencia líder, de poliadenilación, secuencia de propéptido, promotor, secuencia de péptido señal y de terminador de transcripción. Como mínimo, las secuencias de control incluyen un promotor y señales de parada transcripcional y traduccional. Las secuencias de control pueden ser provistas con enlaces con el fin de introducir sitios de restricción específicos que faciliten la ligación de las secuencias de control con la región de codificación de la secuencia de nucleótidos que codifica una variante de beta-glucosidasa de la presente invención. El término "operativamente vinculado" es definido en la presente como una configuración en la cual una secuencia de control es apropiadamente colocada en una posición relativa a la secuencia codificante de la
60 secuencia de nucleótidos de manera que la secuencia de control dirige la expresión de una variante de beta-glucosidasa.
65

[0131] La secuencia de control puede ser una secuencia promotora apropiada, que es reconocida por una célula huésped para la expresión de la secuencia de nucleótidos. La secuencia promotora contiene secuencias de control transcripcionales que median la expresión de la variante de beta-glucosidasa. El promotor puede ser cualquier secuencia de ácidos nucleicos que muestre actividad transcripcional en la célula huésped de elección incluidos promotores mutantes, truncados e híbridos, y pueden obtenerse a partir de genes que codifican polipéptidos extracelulares o intracelulares homólogos o heterólogos a la célula huésped.

[0132] Ejemplos de promotores adecuados para dirigir la transcripción de los constructos de nucleótidos de la presente invención en una célula huésped fúngica filamentosa son promotores obtenidos a partir 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*, triosafosfato isomerasa de *Aspergillus oryzae*, acetamidasa de *Aspergillus nidulans*, amiloglucosidasa de *Fusarium oxysporum*, proteasa de tipo tripsina de *Fusarium oxysporum* (WO 96/00787), beta-glucosidasa de *Trichoderma reesei*, celobiohidrolasa I 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 que el promotor de NA2-tpi (un híbrido de los promotores de los genes para alfa amilasa neutra de *Aspergillus niger* y triosa fosfato isomerasa de *Aspergillus oryzae*); equivalentes de los mismos; y promotores mutantes, truncados e híbridos de los mismos.

[0133] En un huésped de levadura, se obtienen promotores útiles de los genes para enolasa de *Saccharomyces cerevisiae* (ENO-1), galactokinasa de *Saccharomyces cerevisiae* (GAL1), alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa de *Saccharomyces cerevisiae* (ADH1,ADH2/GAP), triosa fosfato isomerasa de *Saccharomyces cerevisiae* (TPI), metalotionina de *Saccharomyces cerevisiae* (CUP1) y 3-fosfoglicerato kinasa de *Saccharomyces cerevisiae*. Otros promotores útiles para células huéspedes de levadura son descritas por Romanos et al., 1992, Yeast 8: 423-488.

[0134] La secuencia de control también puede ser una secuencia del terminador de transcripción adecuada, que se reconoce por una célula huésped para terminar la transcripción. La secuencia del terminador está operativamente ligada al término 3' de la secuencia de nucleótidos que codifica la variante de beta-glucosidasa. En la presente invención, se puede utilizar cualquier terminador que sea funcional en la célula huésped de elección.

[0135] Los terminadores preferidos para células huéspedes fúngicas filamentosas se obtienen a partir 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*.

[0136] Los terminadores preferidos para células huéspedes de levadura se obtienen a partir de los genes para enolasa de *Saccharomyces cerevisiae*, citocromo C (CYC1) de *Saccharomyces cerevisiae*, y gliceraldehído-3- fosfato deshidrogenasa de *Saccharomyces cerevisiae*. Otros terminadores útiles para células huéspedes de levadura son descritos por Romanos et al., 1992, supra.

[0137] La secuencia de control también puede ser una secuencia líder adecuada, una región no traducida de un ARNm que es importante para la traducción por la célula huésped. La secuencia líder está operativamente vinculada al término 5' de la secuencia de nucleótidos que codifica la variante de beta-glucosidasa. En la presente invención se puede utilizar cualquier secuencia líder que sea funcional en la célula huésped de elección.

[0138] Los líderes preferidos para las células fúngicas filamentosas huéspedes se obtienen de los genes para TAKA amilasa de *Aspergillus oryzae* y triosa fosfato isomerasa de *Aspergillus nidulans*.

[0139] Los líderes adecuados para células huéspedes de levadura se obtienen a partir de los genes para enolasa (ENO-1) de *Saccharomyces cerevisiae*, 3-fosfoglicerato kinasa de *Saccharomyces cerevisiae*, factor alfa de *Saccharomyces cerevisiae*, y alcohol deshidrogenasa/gliceraldehído-3-fosfato deshidrogenasa (ADH2/GAP) de *Saccharomyces cerevisiae*.

[0140] La secuencia de control puede también ser una secuencia de poliadenilación, una secuencia operativamente ligada al término 3' de la secuencia de codificación del polipéptido y que, al ser transcrita, es reconocida por la célula huésped como una señal para añadir residuos de poliadenosina al ARNm transcrito. En la presente invención, se puede utilizar cualquier secuencia de poliadenilación que sea funcional en la célula huésped de elección.

[0141] Las secuencias de poliadenilación preferidas para células fúngicas filamentosas huéspedes se obtienen a partir 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*.

[0142] Las secuencias de poliadenilación útiles para células huéspedes de levadura son descritas por Guo y Sherman, 1995, Molecular Cellular Biology 15: 5983-5990.

[0143] La secuencia de control puede también ser una región de codificación del péptido señal que codifica una secuencia de aminoácidos ligada al término amino de una beta-glucosidasa variante y dirige el 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 de codificación del péptido señal naturalmente vinculada en el marco de lectura de traducción con el segmento de la región de codificación que codifica la variante de beta-glucosidasa segregada. De forma alternativa, el extremo 5' de la secuencia codificante puede contener una región de codificación del péptido señal que sea extranjero a la secuencia codificante. La región de codificación del péptido señal extranjero puede ser requerida donde la secuencia codificante no contiene naturalmente una región de codificación del péptido señal. De forma alternativa, la región de codificación del péptido señal extranjero puede simplemente reemplazar la región de codificación del péptido señal natural para mejorar la secreción de la variante de beta-glucosidasa. No obstante, en la presente invención, se puede utilizar cualquier región de codificación del péptido señal que dirige el polipéptido expresado en la vía secretora de una célula huésped de elección.

[0144] Las regiones de codificación del péptido señal eficaces para células fúngicas filamentosas huéspedes son las regiones de codificación del péptido señal obtenidas a partir de los genes para TAKA amilasa de *Aspergillus oryzae*, amilasa neutra de *Aspergillus niger*, glucoamilasa de *Aspergillus niger*, proteinasa aspártica de *Rhizomucor miehei*, celulasa Cel45A de *Humicola insolens* y lipasa de *Humicola lanuginosa*.

[0145] Los péptidos señal útiles para células huéspedes de levadura se obtienen a partir de los genes para factor alfa de *Saccharomyces cerevisiae* e invertasa de *Saccharomyces cerevisiae*. Otras regiones de codificación del péptido señal útiles son descritas por Romanos et al., 1992, supra.

[0146] La secuencia de control también puede ser una región de codificación del propéptido que codifica una secuencia de aminoácidos situada en el término amino de una variante de beta-glucosidasa. El polipéptido resultante es conocido como una proenzima o un propolipéptido (o un zimógeno en algunos casos). Un propolipéptido generalmente es inactivo y se puede convertir en un polipéptido maduro activo por escisión catalítica o autocatalítica del propéptido del propolipéptido. La región de codificación del propéptido se puede obtener a partir de los genes para factor alfa de *Saccharomyces cerevisiae*, proteinasa aspártica de *Rhizomucor miehei* y lacasa de *Myceliophthora thermophila* (WO 95/33836).

[0147] En los casos en que las regiones de propéptido y péptido señal están presentes en el término amino de un polipéptido, la región del propéptido está situada junto al término amino de un polipéptido y la región del péptido señal está situada junto al término amino de la región del propéptido.

[0148] Puede también ser conveniente añadir secuencias reguladoras que permiten la regulación de la expresión de la beta-glucosidasa variante relativa al crecimiento de la célula huésped. Ejemplos de sistemas reguladores son los que causan que la expresión del gen se active o no en respuesta a un estímulo químico o físico, incluida la presencia de un compuesto regulador. En la levadura, se puede utilizar el sistema ADH2 o el sistema GAL1. En hongos filamentosos, el promotor de TAKA alfa-amilasa, promotor de glucoamilasa de *Aspergillus niger*, y el promotor de glucoamilasa de *Aspergillus oryzae* pueden ser utilizados como secuencias reguladoras. Otros ejemplos de secuencias reguladoras son los que permiten la amplificación génica. En sistemas eucarióticos, 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 codifican la beta-glucosidasa variante estaría operativamente vinculada con la secuencia reguladora.

Vectores de expresión

[0149] La presente invención también se refiere a vectores de expresión recombinante que comprende una secuencia de nucleótidos que codifica una beta-glucosidasa variante de la presente invención, un promotor y señales de parada transcripcional y traduccional. Las diferentes secuencias de control y de nucleótidos anteriormente descritas se pueden unir para producir un vector de expresión recombinante que puede incluir uno o más sitios de restricción convenientes para permitir la inserción o sustitución de la secuencia de nucleótidos que codifica la variante en tales sitios. De forma alternativa, la secuencia de nucleótidos se puede expresar por inserción de la secuencia de nucleótidos o una constructo de ácidos nucléicos que comprende 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 está operativamente vinculada con las secuencias de control apropiadas para la expresión.

[0150] El vector de expresión recombinante puede ser cualquier vector (p. ej., un plásmido o virus) que pueda ser convenientemente sometido a procedimientos de ADN recombinante y pueda provocar la expresión de la secuencia de nucleótidos. La elección del vector típicamente dependerá de la compatibilidad del vector con la célula huésped en la cual el vector será introducido. Los vectores pueden ser plásmidos circulares lineales o cerrados.

[0151] Los vectores de la presente invención contienen preferiblemente uno o más marcadores seleccionables que permiten la fácil selección de células transformadas. Un marcador seleccionable es un gen cuyo producto proporciona resistencia biocida o vírica, resistencia a metales pesados, prototrofia a auxótrofos y similares. Los marcadores

adecuados para las células huéspedes de levadura son ADE2, HIS3, LEU2, LYS2, MET3, TRP1 y URA3. Los marcadores seleccionables para el uso en una célula fúngica filamentosa huésped incluyen, de modo enunciativo y no limitativo, *amdS* (acetamidasa), *argB* (ornitina-carbamoiltransferasa), *bar* (fosfotricina acetiltransferasa), *hph* (higromicina fosfotransferasa), *niaD* (nitrato reductasa), *pyrG* (orotidina-5'-fosfato descarboxilasa), *sC* (sulfato adeniltransferasa) y *trpC* (antranilato sintasa), al igual que equivalentes de los mismos. Para el uso en una célula de *Aspergillus* se prefieren los genes *amdS* y *pyrG* de *Aspergillus nidulans* o *Aspergillus oryzae* y el gen *bar* de *Streptomyces hygrosopicus*.

[0152] El vector puede ser un vector de replicación autónoma, es decir, un vector que existe como una entidad extracromosómica, cuya replicación es diferente de la replicación cromosómica, por ejemplo, un plásmido, un elemento extracromosómico, un minicromosoma o un cromosoma artificial. El vector puede contener cualquier medio para asegurar la autorreplicación. De forma alternativa, el vector puede ser uno que, al ser introducido en la célula huésped, se integra en el genoma y se replica con el/los cromosoma/s en que ha sido integrado. Además, puede utilizarse un único vector o plásmido o dos o más vectores o plásmidos que juntos contengan el ADN total que se introducirá en el genoma de la célula huésped, o un transposón.

[0153] Los vectores de la presente invención preferiblemente contienen un elemento o elementos que permitan la integración del vector en el genoma de célula huésped o la replicación autónoma del vector en la célula independiente del genoma.

[0154] Para la integración en el genoma de la célula huésped, el vector puede depender de la secuencia de nucleótidos que codifica la variante o cualquier otro elemento del vector para la integración del vector en el genoma por recombinación homóloga o no homóloga. De forma alternativa, el vector puede contener secuencias de ácidos nucleicos adicionales para dirigir la integración por recombinación homóloga en el genoma de la célula huésped. Las secuencias de ácido nucleico adicionales permiten que el vector se integre en el genoma de la célula huésped en una ubicación o ubicaciones precisas en el/los cromosoma/s. Para aumentar la probabilidad de integración a una ubicación precisa, los elementos integracionales deberían contener preferiblemente un número suficiente de ácidos nucleicos, tal como 100 a 10.000 bp, preferiblemente 400 a 10.000 bp, y de forma más preferible 800 a 10.000 bp, que son altamente homólogos a la secuencia meta correspondiente para mejorar la probabilidad de recombinación homóloga. Los elementos integracionales pueden ser cualquier secuencia que sea homóloga a la secuencia meta en el genoma de la célula huésped. Además, los elementos integracionales pueden ser no codificantes o codificantes de secuencias de ácidos nucleicos. En cambio, el vector se puede integrar en el genoma de la célula huésped por recombinación no homóloga.

[0155] Para la replicación autónoma, el vector puede comprender además un origen de replicación que le permita al vector replicarse de manera autónoma en la célula huésped en cuestión. Ejemplos de orígenes de replicación para el uso en una célula huésped de levadura son el origen de replicación de 2 micra, ARS1, ARS4, la combinación de ARS1 y CEN3 y la combinación de ARS4 y CEN6. El origen de la replicación puede ser uno que tenga una mutación que haga que el funcionamiento sea termosensible en la célula huésped (véase, por ejemplo, Ehrlich, 1978, Proceedings of the National Academy of Sciences USA 75: 1433). Ejemplos de un replicador de plásmido útil en una célula fúngica filamentosa 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). El aislamiento del gen AMA1 y la construcción de plásmidos o vectores que comprenden el gen se puede realizar según los métodos descritos en WO 00/24883.

[0156] Se puede insertar más de una copia de una secuencia de nucleótidos de la presente invención en la célula huésped para aumentar la producción de una variante de beta-glucosidasa. Se puede obtener un aumento en el número de copias de la secuencia de nucleótidos integrando al menos una copia adicional de la secuencia en el genoma de la célula huésped o incluyendo un gen marcador seleccionable amplificable con la secuencia de nucleótidos donde las células que contienen copias amplificadas del gen marcador seleccionable, y por tanto las copias adicionales de la secuencia de nucleótidos, pueden seleccionarse cultivando las células en presencia del agente seleccionable apropiado.

[0157] Los procedimientos usados para ligar los elementos anteriormente descritos para construir los vectores de expresión recombinante de la presente invención son conocidos por los expertos en la técnica (véase, por ejemplo, Sambrook et al., 1989, supra).

55 Células huéspedes

[0158] La presente invención también se refiere a células huéspedes recombinantes, que comprenden una secuencia de nucleótidos que codifica una beta-glucosidasa variante, que es ventajosamente usada en la producción recombinante de la variante. Un vector que comprende una secuencia de nucleótidos de la presente invención se introduce en una célula huésped de modo que el vector es mantenido como un integrante cromosómico o como un vector extracromosómico que se auto-replica como se describe anteriormente. El término "célula huésped" comprende cualquier progenie de una célula madre que no es idéntica a la célula madre debido a mutaciones que ocurren durante la replicación. La elección de una célula huésped depende en gran parte del gen que codifica el polipéptido y su fuente.

[0159] La célula huésped puede ser cualquier eucariota, como un mamífero, un insecto, una planta o una célula fúngica.

[0160] La célula huésped puede ser cualquier célula fúngica. "Hongo" según se utiliza en este caso incluye el filo Ascomycota, Basidiomycota, Chytridiomycota y Zygomycota (según la definición de Hawksworth et al., en Ainsworth and Bisby's Dictionary of The Fungi, 8^o edición, 1995, CAB International, University Press, Cambridge, Reino Unido) al igual que el Oomycota (según se cita en Hawksworth et al., 1995, supra, página 171) y todos los hongos mitospóricos (Hawksworth et al., 1995, supra).

[0161] En una forma de realización preferida, la célula fúngica huésped es una célula de levadura. "Levadura", según se utiliza en este caso, incluye levadura ascosporigénea (Endomycetales), levadura basidiosporogénea y levadura de los Fungi Imperfecti (Blastomycetes). Ya que la clasificación de levadura puede cambiar en el futuro, para los objetivos de la presente invención, la levadura será definida como se describe en Biology and Activities of Yeast (Skinner, F.A., Passmore, S.M. y Davenport, R.R., eds, Soc. App. Bacteriol. Symposium Series No. 9, 1980).

[0162] En una forma de realización más preferida, la célula huésped de levadura es una célula de *Candida*, *Hansenula*, *Kluyveromyces*, *Pichia*, *Saccharomyces*, *Schizosaccharomyces* o *Yarrowia*.

[0163] En una forma de realización más preferida, 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 otra forma de realización más preferida, la célula huésped de levadura es una célula de *Kluyveromyces lactis*. En otra forma de realización más preferida, la célula huésped de levadura es una célula de *Yarrowia lipolytica*.

[0164] En otra forma de realización preferida, la célula huésped fúngica es una célula fúngica filamentosa. Los "hongos filamentosos" incluyen todas las formas filamentosas de la subdivisión Eumycota y Oomycota (según se define por Hawksworth et al., 1995, supra). Los hongos filamentosos generalmente son caracterizados por una pared micelial compuesta por quitina, celulosa, glucano, quitosano, manano y otros polisacáridos complejos. El crecimiento vegetativo es por elongación hifal y el catabolismo de carbono es estrictamente aeróbico. En cambio, el crecimiento vegetativo por levaduras tales como *Saccharomyces cerevisiae* es por injerto de un talo unicelular y el catabolismo de carbono puede ser fermentativo.

[0165] En una forma de realización más preferida, la célula fúngica filamentosa huésped es, de modo enunciativo y no limitativo, una célula de *Acremonium*, *Aspergillus*, *Fusarium*, *Humicola*, *Mucor*, *Myceliophthora*, *Neurospora*, *Penicillium*, *Thielavia*, *Tolyocladium*, o *Trichoderma*.

[0166] En una forma de realización más preferida, la célula fúngica filamentosa huésped es una célula de *Aspergillus awamori*, *Aspergillus foetidus*, *Aspergillus japonicus*, *Aspergillus nidulans*, *Aspergillus niger* o *Aspergillus oryzae*. En otra forma de realización más preferida, la célula fúngica filamentosa huésped 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 una forma de realización incluso más preferida, la célula fúngica filamentosa huésped es una célula de *Fusarium venenatum* (Nirenberg sp. nov.). En otra forma de realización más preferida, la célula fúngica filamentosa huésped es una célula de *Humicola insolens*, *Humicola lanuginosa*, *Mucor miehei*, *Myceliophthora thermophila*, *Neurospora crassa*, *Penicillium purpurogenum*, *Thielavia terrestris*, *Trichoderma harzianum*, *Trichoderma koningii*, *Trichoderma longibrachiatum*, *Trichoderma reesei* o *Trichoderma viride*. En otra forma de realización incluso más preferida, la célula fúngica filamentosa huésped es *Trichoderma reesei* RutC30.

[0167] Las células fúngicas se pueden transformar según los procedimientos descritos en la presente.

Métodos de producción

[0168] La presente invención también se refiere a métodos para producir una variante de beta-glucosidasa, que comprende:

- (a) cultivo de una célula huésped bajo condiciones adecuadas para la expresión de la variante, en donde la célula huésped comprende una secuencia de nucleótidos de la invención; y
- (b) recuperación de la variante del medio de cultivo.

[0169] En los métodos de producción de la presente invención, las células huéspedes se cultivan en un medio nutritivo adecuado para la producción de la beta-glucosidasa variante usando métodos conocidos en la técnica. Por ejemplo, la célula se puede cultivar por cultivo en matraz de agitación, o fermentación en pequeña escala o gran escala (incluyendo fermentaciones continuas, discontinuas, en flujo discontinuo, o estado sólido) en fermentadores de laboratorio o industriales realizados en un medio adecuado y bajo condiciones que permitan que el polipéptido sea expresado y/o aislado. El cultivo se desarrolla en un medio nutritivo adecuado que comprende fuentes de nitrógeno y carbono y sales inorgánicas, usando procedimientos conocidos en la técnica. Los 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 es segregado, se puede recuperar de lisatos celulares.

[0170] En una forma de realización alternativa, la variante de beta-glucosidasa no es recuperada, sino que se usa como una fuente de la variante una célula huésped de la presente invención que exprese una variante.

5 [0171] Las beta-glucosidasas variantes pueden ser detectadas usando métodos conocidos en la técnica que son específicos para los polipéptidos. Estos métodos de detección pueden incluir el uso de anticuerpos específicos, la formación de un producto enzimático, o la desaparición de un sustrato enzimático. Por ejemplo, se puede utilizar un ensayo enzimático para determinar la actividad del polipéptido como se describe en este caso en los ejemplos.

10 [0172] La beta-glucosidasa variante resultante se puede recuperar por métodos conocidos en la técnica. Por ejemplo, el polipéptido puede ser recuperado del medio nutritivo por procedimientos convencionales incluidos, de modo enunciativo y no limitativo, recogida, centrifugado, filtración, extracción, secado por pulverización, evaporación o precipitación.

15 [0173] Una beta-glucosidasa variante de la presente invención se puede purificar por una variedad de procedimientos conocidos en la técnica incluidos, de modo enunciativo y no limitativo, cromatografía (p. ej., de intercambio iónico, por afinidad, hidrofóbica, de cromatoenfoco y de exclusión por tamaño), procedimientos electroforéticos (p. ej., isoelectroenfoco preparatorio), solubilidad diferencial (p. ej., precipitación de sulfato amónico), SDS-PAGE o extracción (véase, por ejemplo, Protein Purification, J.-C. Janson y Lars Ryden, editores, VCH Publishers, New York, 1989).

20 **Degradación de biomasa a monosacáridos, disacáridos y polisacáridos**

[0174] Las variantes de beta-glucosidasa y las células huéspedes de la presente invención se pueden utilizar en la producción de monosacáridos, disacáridos y polisacáridos como materias primas químicas o de fermentación de biomasa para la producción de etanol, plásticos u otros productos o productos intermedios. Las variantes de beta-glucosidasa pueden estar en forma de un caldo de fermentación crudo con o sin las células eliminadas o en forma de una preparación enzimática semi-purificada o purificada. De forma alternativa, una célula huésped de la presente invención puede ser utilizada como una fuente de la variante en un proceso de fermentación con la biomasa.

30 [0175] La biomasa puede incluir, de modo enunciativo y no limitativo, recursos de madera, desperdicios municipales sólidos, papel usado y residuos de cosecha (véase, por ejemplo, Wiselogel et al., 1995, en Handbook on Bioethanol (Charles E. Wyman, editor), páginas 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 In Bioconversion of Lignocellulosics, in Advances in Biochemical Engineering/Biotechnology, T. Scheper, Director Editorial, Volumen 65, página 23-40, Springer- Verlag, New York).

35 [0176] El polisacárido predominante en la pared celular primaria de la biomasa es la celulosa, el segundo más abundante es la hemicelulosa, y el tercero es la pectina. La pared celular secundaria, producida después de que la célula ha dejado de crecer, también contiene polisacáridos y se refuerza a través de lignina polimérica reticulada de manera covalente a la hemicelulosa. La celulosa es un homopolímero de anhidrocelobiosa y, de ese modo, es un beta-(1-4)-D-glucano lineal, mientras que las hemicelulosas incluyen una variedad de compuestos, tales como xilanos, xiloglucanos, arabinoxilanos y mananos en estructuras ramificadas complejas con un espectro de sustituyentes. Aunque en general se encuentra celulosa polimorfa en el tejido vegetal principalmente como una matriz cristalina insoluble de cadenas paralelas de glucano. Las hemicelulosas normalmente tienen enlace de hidrógeno a celulosa, al igual que a otras hemicelulosas, lo cual ayuda a estabilizar la matriz de la pared celular.

45 [0177] Se utilizan tres clases importantes de glicohidrolasas para descomponer la biomasa celulósica:
 (1) Las "endo-1,4-beta-glucanasas" o 1,4-beta-D-glucano-4-glucanohidrolasas (EC 3.2.1.4), que actúan de forma aleatoria en sustratos de 1,4-beta-glucano solubles e insolubles.
 50 (2) Las "exo-1,4-beta-D-glucanasas" que incluyen las 1,4-beta-D-glucano glucohidrolasas (EC 3.2.1.74), que liberan D-glucosa de 1,4-beta-D-glucanos e hidrolizan D-celobiosa lentamente, y celobiohidrolasas (1,4-beta-D- glucano celobiohidrolasas, EC 3.2.1.91), que liberan D-celobiosa de 1,4-beta-glucanos.
 (3) Las "beta-D-glucosidasas" o beta-D-glucósido glucohidrolasas (EC 3.2.1.21), que actúan para liberar unidades de D-glucosa de celobiosa y celodextrinas solubles, al igual que un conjunto de glucósidos.

55 [0178] Estas tres clases de enzimas trabajan juntas sinérgicamente dando como resultado descomposición e hidrólisis eficaces de celulosa nativa de biomasa para producir azúcares reducidos.

60 [0179] Las variantes de beta-glucosidasa de la presente invención se pueden utilizar conjuntamente con las enzimas mencionadas anteriormente para degradar aún más el componente de celulosa del sustrato de biomasa, (véase, por ejemplo, Brigham et al., 1995, en Handbook on Bioethanol (Charles E. Wyman, editor), páginas 119-141, Taylor & Francis, Washington D.C.; Lee, 1997, Journal of Biotechnology 56: 1-24).

65 [0180] El etanol se puede producir por degradación enzimática de biomasa y conversión de los sacáridos liberados a etanol. Este especie de etanol frecuentemente es denominada Bioetanol o biocombustible. Esto puede ser usado como un aditivo o suplemento de combustible en mezclas de menos de 1% y hasta 100% (un sustituto de combustible).

Composiciones de detergentes

[0181] Las variantes de beta-glucosidasa de la presente invención pueden ser añadidas a la composición detergente y así convertirse en un componente de la misma.

[0182] La composición detergente de la presente invención puede, por ejemplo, ser formulada como una composición detergente de lavado a mano o a máquina incluida una composición de aditivo de lavado adecuada para el tratamiento previo de tejidos manchados y una composición de suavizante agregado al enjuague, o ser formulada como una composición detergente para el uso en operaciones de limpieza de superficies duras del hogar en general, o ser formulada para operaciones de lavado de vajilla a mano o máquina.

[0183] En un aspecto específico, la presente invención proporciona un aditivo detergente que comprende la enzima de la invención. El aditivo detergente al igual que la composición detergente puede comprender una o más enzimas adicionales tales como una proteasa, lipasa, cutinasa, una amilasa, carbohidrasa, celulasa, pectinasa, mananasa, arabinasa, galactanasa, xilanasas, oxidasas, por ejemplo, una lacasa y/o peroxidasa.

[0184] En general, las propiedades de la/s enzima/s elegida/s deberían ser compatibles con el detergente seleccionado, (es decir, pH óptimo, compatibilidad con otros ingredientes enzimáticos y no enzimáticos, etc.) y la/s enzima/s debería/n estar presente/s en cantidades eficaces.

[0185] Proteasas: Las proteasas adecuadas incluyen las de origen animal, vegetal o microbiano. Se prefiere el origen microbiano. Están incluidos mutantes creados genéticamente o modificados químicamente de proteínas. 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, subtilisina Novo, subtilisina Carisberg, subtilisina 309, subtilisina 147 y subtilisina 168 (descritas en WO 89/06279). Ejemplos de proteasas de tipo tripsina son tripsina (p. ej., de origen porcino o bovino) y la proteasa de *Fusarium* descrita en WO 89/06270 y WO 94/25583.

[0186] 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.

[0187] Las enzimas proteásicas preferidas disponibles comercialmente incluyen Alcalase™, Savinase™, Primase™, Duralase™, Esperase™, y Kannase™ (Novo Nordisk A/S), Maxatase™, Maxacal™, Maxapem™, Properase™, Purafect™, Purafect OxP™, FN2™ y FN3™ (Genencor International Inc.).

[0188] Lipasas: Las lipasas adecuadas incluyen las de origen bacteriano o fúngico. Están incluidos mutantes creados genéticamente o modificados químicamente de proteínas. Ejemplos de lipasas útiles incluyen lipasas de *Humicola* (sinónimo, *Thermomyces*), por ejemplo, de *H. lanuginosa* (*T. lanuginosus*) según se describe en EP 258 068 y EP 305 216 o de *H. insolens* según se describe en WO 96/13580, una lipasa de *Pseudomonas*, por ejemplo, de *P. alcaligenes* o *P. pseudoalcaligenes* (EP 218 272), *P. cepacia* (EP 331376); *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).

[0189] Otros ejemplos son variantes de lipasa tales como las descritas 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

[0190] Las enzimas de lipasa disponible comercialmente preferidas incluyen Lipolase™ y Lipolase Ultra™ (Novo Nordisk A/S).

[0191] Amilasas: Las amilasas adecuadas (α y/o β) incluyen las de origen bacteriano o fúngico. Están incluidos mutantes creados genéticamente o modificados químicamente de proteínas. Las amilasas incluyen, por ejemplo, α -amilasas obtenidas de *Bacillus*, por ejemplo, una cepa especial de *Bacillus licheniformis*, descrita en más detalle en GB 1.296.839.

[0192] 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.

[0193] Las amilasas disponibles comercialmente son Duramyl™, Termamyl™, Fungamyl™ y BAN™ (Novo Nordisk A/S), Rapidase™ y Purastar™ (de Genencor International Inc.).

[0194] Celulasas: Las celulasas adecuadas incluyen las de origen bacteriano o fúngico. Están incluidos mutantes creados genéticamente o modificados químicamente de proteínas. Las celulasas adecuadas incluyen celulasas de los

géneros *Bacillus*, *Pseudomonas*, *Humicola*, *Fusarium*, *Thielavia*, *Acremonium*, por ejemplo, las celulosas fúngicas producidas a partir de *Humicola Insolens*, *Myceliophthora thermophila* y *Fusarium oxysporum* descritas en US 4.435.307, 5.648.263, 5.691.178, 5.776.757 y WO 89/09259.

5 [0195] Las celulasas especialmente adecuadas son las celulasas alcalinas o neutras que tienen beneficios de cuidados del color. Ejemplos de tales celulasas son celulasas descritas en EP 0 495 257, EP 0 531 372, WO 96/11262, WO 96/29397, WO 98/08940. Otros ejemplos son variantes de celulasas tales como las descritas en WO 94/07998, EP 0 531315, 5,457,046, 5,686,593, 5,763,254, WO 95/24471, WO 98/12307 y PCT/DK98/00299.

10 [0196] Las celulasas disponibles comercialmente incluyen Celluzyme™, y Carezyme™ (Novo Nordisk A/S), Clazinase™, y Puradax HA™ (Genencor International Inc.) y KAC-500(B)™ (Kao Corporation).

15 [0197] Peroxidasas/oxidases: Las peroxidasas/oxidases adecuadas incluyen las de origen vegetal, bacteriano o fúngico. Están incluidos mutantes creados genéticamente o modificados químicamente de proteínas. Ejemplos de peroxidasas útiles incluyen peroxidasas de *Coprinus*, por ejemplo, de *C. cinereus* y variantes de las mismas como las descritas en WO 93/24618, WO 95/10602 y WO 98/15257.

[0198] Las peroxidasas disponibles comercialmente incluyen Guardzyme™ (Novo Nordisk A/S).

20 [0199] La/s enzima/s detergente/s se puede/n incluir en una composición detergente añadiendo aditivos separados que contengan una o más enzimas, o añadiendo un aditivo combinado que comprenda 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, un líquido, un compuesto acuoso, etc. Las formulaciones de aditivo detergente preferidas son granulados, en particular granulados no en polvo, líquidos, en particular líquidos estabilizados, o lodos.

25 [0200] Los granulados no en polvo 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 revestimiento ceroso son productos de poli(etileno óxido) (polietilenglicol; PEG) con pesos molares medios de 1000 a 20.000; nonilfenoles etoxilados que tienen 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 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 revestimiento que forman películas adecuadas para la aplicación por técnicas de lecho fluidificado se dan en GB 1483591. Las 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 los métodos establecidos. Las enzimas protegidas se pueden preparar según el método descrito en EP 238.216.

30 [0201] La composición detergente de la invención puede ser en cualquier forma conveniente, por ejemplo, una barra, un comprimido, un polvo, un gránulo, una pasta o un líquido. Un detergente líquido puede ser acuoso, típicamente con 70% de agua y 0-30% de solvente orgánico, o no acuoso.

40 [0202] La composición detergente comprende uno o más agentes tensioactivos, que pueden ser no iónicos incluidos semipolares y/o aniónicos y/o catiónicos y/o zwitteriónicos. Los agentes tensioactivos están típicamente presentes a un nivel de 0,1% a 60% en peso.

45 [0203] Cuando se incluye en el mismo, el detergente normalmente contiene de aproximadamente 1% a aproximadamente 40% de un tensioactivo aniónico tal como alquilbencenosulfonato lineal, alfa-olefinsulfonato, sulfato de alquilo (sulfato de alcohol graso), etoxisulfato alcohólico, alcanosulfonato secundario, éster metílico de ácido alfa-sulfo graso, ácido alquil o alquenilsuccínico, o jabón.

50 [0204] Cuando se incluye en el mismo, el detergente normalmente contiene de aproximadamente 0,2% a aproximadamente 40% de un tensioactivo no iónico como alcohol etoxilato, nonilfenol etoxilato, alquilpoliglicósido, óxido de alquildimetilamina, monoetanolamida de ácido graso etoxilado, monoetanolamida de ácido graso, polihidroxi alquil amida de ácido graso, o derivados N-acil N-alquil glucosamina ("glugamidas").

55 [0205] 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).

60 [0206] El detergente puede comprender uno o más polímeros. Ejemplos son carboximetilcelulosa, poli(vinilpirrolidona), poli(etilenglicol), poli(vinil) alcohol, poli(vinilpiridina-N-óxido), poli(vinilimidazol), policarboxilatos tales como poliacrilatos, copolímeros de ácido maleico/acrílico, y copolímeros de lauril metacrilato/ ácido acrílico.

65 [0207] 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

tetraacetililenodiamina o nonanoiloxibencenosulfonato. De forma alternativa, el sistema blanqueante puede comprender peroxiácidos de, por ejemplo, tipo amida, imida, o sulfona.

5 [0208] La/s enzima/s de la composición de detergente de la invención puede/n ser estabilizada/s 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 fenilo borónico tal como ácido 4-formilfenil borónico y la composición puede ser formulada como se describe en, por ejemplo, WO 92/19709 y WO 92/19708.

10 [0209] El detergente también puede 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 anti redeposición de suciedad, tintes, bactericidas, abrillantadores ópticos, hidrótopos, inhibidores de decoloración o perfumes.

15 [0210] En las composiciones detergentes cualquier enzima, en particular la enzima de la invención, se puede adicionar 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.

20 [0211] La enzima de la invención puede adicionalmente ser incorporada en las formulaciones detergentes descritas en WO 97/07202, que está incorporada al presente como referencia.

Plantas

25 [0212] La presente invención también se refiere a una planta transgénica, parte de planta, o célula vegetal que ha sido transformada con una secuencia de nucleótidos que codifica una beta-glucosidasa variante de la presente invención para expresar y producir la variante en cantidades recuperables. La variante se puede recuperar de la planta o parte de planta. De forma alternativa, la planta o parte de planta que contiene la variante recombinante puede ser utilizada como tal para mejorar la calidad de un alimento o pienso, por ejemplo, mejorando el valor nutritivo, de apetencia, y propiedades reológicas, o para destruir un factor antinutritivo.

30 [0213] La planta transgénica puede ser dicotiledónea o monocotiledónea. Ejemplos de plantas monocotiledóneas son hierbas, tales como poa de prados (poa pratense, Poa), hierba forrajera tal como Festuca, Lolium, césped templado, tal como Agrostis y cereales, por ejemplo, trigo, avena, centeno, cebada, arroz, sorgo, y maíz.

35 [0214] Ejemplos de plantas dicotiledóneas son tabaco, leguminosas, tales como altramuces, patata, remolacha azucarera, guisante, moldura y semilla de soja, y plantas crucíferas (familia de Brassicaceae), tales como coliflor, semilla de colza y el organismo modelo cercanamente relacionado *Arabidopsis thaliana*.

40 [0215] Ejemplos de partes de planta son tallo, callo, hojas, raíz, frutas, semillas y tubérculos, al igual que los tejidos individuales que comprenden estas partes, por ejemplo, epidermis, mesófilo, parénquima, tejidos vasculares, meristemas. Compartimentos específicos de la célula vegetal, como cloroplastos, apoplastos, mitocondria, vacuolas, peroxisomas y citoplasma también son considerados parte de una planta. Además, cualquier célula vegetal, cualquiera que sea el origen del tejido, se considera una parte de una planta. De igual modo, las partes de planta tales como tejidos específicos y células aisladas para facilitar la utilización de la invención también son consideradas partes de la planta, por ejemplo, embriones, endospermas, aleurona y revestimientos de semillas.

45 [0216] También está incluida dentro del campo de la presente invención la progenie de tales plantas, partes de la planta y células vegetales.

50 [0217] La planta transgénica o célula vegetal que exprese una variante de la presente invención se puede construir conforme a métodos conocidos en la técnica. En resumen, la planta o célula vegetal se construye incorporando uno o más constructos de expresión que codifica/n una variante de la presente invención en el genoma huésped de la planta y propaga la planta modificada o célula vegetal resultante en una planta transgénica o célula vegetal.

55 [0218] Convenientemente, el constructo de expresión es un constructo de ácidos nucleicos que comprende una secuencia de ácidos nucleicos que codifica una variante de la presente invención operativamente vinculada con secuencias reguladoras apropiadas requeridas para la expresión de la secuencia de ácidos nucleicos en la planta o parte de la planta de elección. Además, el constructo de expresión puede comprender un marcador seleccionable útil para identificar células huéspedes en las cuales el constructo de expresión ha sido integrada y las secuencias de ADN necesarias para la introducción del constructo en la planta en cuestión (esto depende del método de introducción de ADN que se utilizará).

60 [0219] La elección de secuencias reguladoras, tales como secuencias de terminador y promotor y opcionalmente secuencias de tránsito o señal, está determinada, por ejemplo, basándose en cuándo, dónde, y cómo se desea expresar la variante. Por ejemplo, la expresión del gen que codifica una variante de la presente invención puede ser

constitutiva o inducible, o puede ser desarrollable, específica de fase o tejido, y el producto genético puede ser dirigido a un tejido o parte de planta específica tal como semillas u hojas. Las secuencias reguladoras son, por ejemplo, descritas por Tague et al., 1988, *Plant Physiology* 86: 506.

5 [0220] Para la expresión constitutiva pueden utilizarse el 35S-CaMV, la ubiquitina 1 de maíz y el promotor de la actina 1 del arroz (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). Los promotores específicos de un órgano pueden ser, por ejemplo, un promotor de tejidos sumidero de almacenamiento tales 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 el promotor de glutelina, prolamina, globulina o 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 desconocido de *Vicia faba* (Conrad et al., 1998, *Journal of Plant Physiology* 152: 708-711), un promotor de una proteína de cuerpo de aceite de semilla (Chen et al., 1998, *Plant and Cell Physiology* 39: 935-941), el promotor *napA* de proteína de almacenamiento de *Brassica napus*, o cualquier otro promotor específico de semilla conocido en la técnica, por ejemplo, como se describe en WO 91/14772. Además, el promotor puede ser un promotor específico de hoja tal como el promotor de eritrocitos de arroz o tomate (Kyojuka et al., 1993, *Plant Physiology* 102: 991-1000, el promotor del gen de metiltransferasa de adenina de virus chlorella (Mittra y Higgins, 1994, *Plant Molecular Biology* 26: 85-93), o el promotor de gen *aldP* de arroz (Kagaya et al., 1995, *Molecular and General Genetics* 248: 668-674), o un promotor inducible por lesiones tal como el promotor *pin2* de la 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 salinidad o inducido por sustancias aplicadas exógenamente que activan el promotor, por ejemplo, etanol, estrógenos, hormonas de planta tales como etileno, ácido abscísico y ácido giberélico y metales pesados.

25 [0221] También puede utilizarse un elemento intensificador del promotor para conseguir mayor expresión de un polipéptido de la presente invención en la planta. Por ejemplo, el elemento intensificador del promotor 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*, describen el uso del primer intrón del gen de actina 1 de arroz para mejorar la expresión.

30 [0222] El gen marcador seleccionable y cualquier otra parte del constructo de expresión se pueden elegir de aquellos disponibles en la técnica.

35 [0223] El constructo de ácido nucleico se incorpora en el genoma de la planta según técnicas convencionales conocidas en la técnica, incluida la transformación mediada por *Agrobacterium*, transformación mediada por virus, microinyección, bombardeo de partículas, transformación biolística y electroporación (Gasser et al., 1990, *Science* 244: 1293; Potrykus, 1990, *Bio/Technology* 8: 535; Shimamoto et al., 1989, *Nature* 338: 274).

40 [0224] Actualmente, la transferencia de genes mediada por *Agrobacterium tumefaciens* es el método de elección para generar dicotiledóneas transgénicas (para obtener un resumen, véase Hooykas y Schilperoort, 1992, *Plant Molecular Biology* 19: 15-38) y también puede ser usada para transformar monocotiledóneas, aunque frecuentemente se utilizan otros métodos de transformación para estas plantas. Actualmente, el método de elección para generar monocotiledóneas transgénicas es el bombardeo de partículas (partículas de tungsteno u oro microscópico revestidas con el ADN transformante) 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 la transformación de monocotiledóneas se basa en la transformación de protoplasto como se describe por Omirulleh et al., 1993, *Plant Molecular Biology* 21: 415-428.

50 [0225] Después de la transformación, los transformantes que hayan incorporado el constructo de expresión se seleccionan y regeneran en plantas enteras según métodos bien conocidos en la técnica. Frecuentemente, el procedimiento de transformación se diseña para la eliminación selectiva de genes de selección durante la regeneración o en las siguientes generaciones usando, por ejemplo, cotransformación con dos constructos de T-ADN separados o escisión específica de sitio del gen de selección por una recombinasa específica.

55 [0226] La presente invención también se refiere a métodos para producir una variante de la presente invención que comprende (a) cultivo de una planta transgénica o una célula vegetal que comprende una secuencia de ácidos nucleicos que codifica una variante que tiene actividad de beta-glucosidasa de la presente invención bajo condiciones propicias para la producción de la variante; y (b) recuperación de la variante.

60 Otros Usos

65 [0227] Las variantes de beta-glucosidasa de la presente invención también pueden ser usadas en el tratamiento de tejidos como agentes de biopulido y para reducción de pelusa, frisado, modificación de textura y lavado a la piedra (N.K. Lange, en P. Suominen, T. Reinikainen (Eds.), *Trichoderma reesei Cellulases and Other Hydrolases*, Foundation for Biotechnical and Industrial Fermentation Research, Helsinki, 1993, páginas 263-272). Además, las variantes descritas también pueden ser usadas en el tratamiento de la madera para biorreducción a pasta o descortezado, fabricación de papel para modificación de fibra, blanqueamiento y reducción de refinación de costes de energía, tratamiento de aguas

blancas, importante para reciclaje de aguas residuales, reciclaje de fibra lignocelulósica tal como desentintado y tratamiento de fibras secundarias y utilización de residuos de madera (S.D, Mansfield y A.R. Esteghalian en S.D, Mansfield y J.N. Saddler (Eds.), Applications of Enzymes to Lignocellulosics, ACS Symposium Series 855, Washington, D.C., 2003, págs. 2- 29).

[0228] La presente invención es descrita adicionalmente por los siguientes ejemplos, los cuales no limitan el alcance de la invención.

Ejemplos

Materiales

[02294] Los productos químicos usados como tampones y sustratos fueron productos comerciales de al menos calidad reactiva.

Cepas

[0230] La cepa de levadura de *Saccharomyces cerevisiae* YNG318 (MAT α , ura3-52, leu-2 Δ 2, pep4 Δ 1, his4-539, cir⁺) fue usada para expresar las beta-glucosidasas de *Aspergillus oryzae* y *Aspergillus fumigatus* y sus variantes. Las cepas bacterianas usadas para generar plásmidos fueron células de *Epicurian coli* XL-10 Gold ultracompetente, células de *Epicurian coli* XL1-Blue competente para subclonación y células ultracompetentes de *Epicurian coli* SURE para electroporación (Stratagene, La Jolla, CA). La cepa de *Aspergillus oryzae* Jai250 (WO 99/61651) fue usada para la expresión de la beta-glucosidasa de *Aspergillus oryzae*. *Aspergillus fumigatus* PaHa34 fue usada como la fuente de beta-glucosidasa de Familia GH3A.

Medios y soluciones

[0231] El medio YPD fue compuesto por litro de 10 g de extracto de levadura, 20 g de bacto triptona y 40 ml de glucosa al 50%.

[0232] El medio de selección de levadura fue compuesto por litro de 6,7 g de base nitrogenada de levadura, 0,8 g de mezcla de suplemento completo (CSM, Qbiogene, Inc., Carlsbad, CA; sin uracilo y con 40 mg/ml de adenina), 5 g de ácidos de casamino (sin aminoácidos), 100 ml de 0,5 M de succinato pH 5,0, 40 ml de glucosa al 50%, 1 ml de 100 mM de CuSO₄, 50 mg de ampicilina y 25 mg de cloranfenicol.

[0233] El medio de placa de selección de levadura fue compuesto por litro de medio de selección de levadura suplementado con 20 g de bacto agar y 150 mg de 5-bromo-4-cloro-3-indolil-beta-D-glucopiranosido (X-Glc, INALCO SPA, Milán, Italia) pero sin ampicilina ni cloranfenicol.

[0234] El medio M400 está compuesto por litro de 50 g de maltodextrina, 2 g de MgSO₄·7H₂O, 2 g de KH₂PO₄, 4 g de ácido cítrico, 8 g de extracto de levadura, 2 g de urea, 0,5 ml de solución de metales traza AMG y 0,5 g de CaCl₂.

[0235] La solución de metales traza AMG fue compuesta por litro de 14,3 g de ZnSO₄·7H₂O, 2,5 g de CuSO₄·5H₂O, 0,5 g de NiCl₂·6H₂O, 13, 8 g de FeSO₄·7H₂O, 8,5 g de MnSO₄·H₂O y 3 g de ácido cítrico.

[0236] 1X Bs fue compuesto por litro de 2 g de MgSO₄·7H₂O, 2 g de K₂SO₄ y 10 g de KH₂PO₄.

[0237] Las placas de selección de *Aspergillus oryzae* de medio mínimo fueron compuestas por litro de 6 g de NaNO₃, 0,52 g de KCl, 1,52 g de KH₂PO₄, 1 ml de solución de elementos traza COVE, 20 g de agar Noble, 20 ml de glucosa al 50%, 2,5 ml de 20% de MgSO₄·7H₂O.

[0238] La solución de elementos traza COVE fue compuesta por litro de 0,04 g de NaB₄O₇·10H₂O, 0,4 g de CuSO₄·5H₂O, 1,2 g de FeSO₄·7H₂O, 0,7 g de MnSO₄·H₂O, 0,8 g de Na₂ MoO₂·2H₂O y 10 g de ZnSO₄·7H₂O.

[0239] El tampón para lisis de levadura fue compuesto por dodecil sulfato de sodio (SDS) al 1%, 10 mM de tris-HCl y 1 mM de EDTA, pH 8,0.

[0240] El medio de dextrosa de patata fue compuesto por litro de 39 gramos de dextrosa de patata (Difco).

[0241] Las placas de PDA fueron compuestas por litro de 39 gramos de agar de dextrosa de patata (Difco).

[0242] El medio MDU2BP fue compuesto por litro de 45 g de maltosa, 1 g de MgSO₄·7H₂O, 1g de NaCl, 2 g de K₂SO₄, 12 g de KH₂PO₄, 7 g de extracto de levadura, 2 g de urea y 0,5 ml de solución de metales traza AMG, pH a 5,0.

Preparación de Adn plásmido y secuenciación del ADN

[0243] El Adn plásmido de cepas de *E. coli* fue preparado usando un BioRobot 9600 (QIAGEN, Inc., Chatsworth, CA).

[0244] La secuenciación del ADN fue realizada en un ABI3700 (Applied Biosystems, Foster City, CA) usando química de terminador colorante (Giesecke et al., 1992, Journal of Virol. Methods 38: 47-60). Las secuencias fueron ensambladas con el uso de phred/phrap/consed (University of Washington, Seattle WA) con cebadores específicos de secuencia.

Ejemplo 1: Construcción de vector de expresión de *Saccharomyces cerevisiae* pSATE111

[0245] Un fragmento de ADN de 2.605 bp que comprende la región del codón de inicio ATG al codón de detención TAA de la secuencia codificante de beta-glucosidasa de *Aspergillus oryzae* (SEC ID n°: 1 para la secuencia de ADNc y SEC ID n°: 2 para la secuencia de aminoácidos deducida) fue amplificado por PCR de pJaL660 (WO 2002/095014) como molde con los cebadores 992127 (sentidos) y 992328 (antisentidos) mostrados más abajo.

992127: 5'-GCAGATCTACCATGAAGCTTGGTTGGATCGAG-3' (SEC ID n°: 3)

992328: 5'-GCCTCGAGTTACTGGCCTTAGGCAGCGAG-3' (SEC ID n°: 4)

El cebador 992127 tiene un sitio *Bg/II* corriente arriba y el cebador 992328 tiene un sitio *Xho I* corriente abajo.

[0246] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de PCR 1X con MgCl₂ (Roche Applied Science, Mannheim, Alemania), 0,25 mM de dNTPs, 50 µM de cebador 992127, 50 µM de cebador 992328, 80 ng de pJaL660, y 2,5 unidades de *Pwo* ADN-polimerasa (Roche Applied Science, Mannheim, Alemania). Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 (Eppendorf Scientific, Inc., Westbury, NY) programadas durante 1 ciclo a 94°C durante 5 minutos, seguidos de 25 ciclos cada uno a 94°C durante 60 segundos, 55°C durante 60 segundos, y 72°C durante 120 segundos (extensión final de 10 minutos). El producto de PCR luego fue subclonado en el vector PCR-Blunt II-TOPO usando el equipo de clonación PCR-Blunt II-TOPO (Invitrogen, Carlsbad, CA) siguiendo las instrucciones del fabricante para generar plásmido pSATE101 (Figura 1). El plásmido pSATE101 fue digerido con *Bg/II* y *Xho I* para liberar el gen de beta-glucosidasa. Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando 40 mM de Tris-acetato-1 mM de tampón de EDTA (TAE) donde una banda de producto de 2,6 kb fue cortada del gel y purificada usando un equipo de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA) según las instrucciones del fabricante.

[0247] El producto de PCR de 2,6 kb fue digerido y clonado en los sitios *Bam HI* y *Xho I* del vector de 2 µm de expresión de levadura pCu426 inducible del cobre (Labbe y Thiele, 1999, Methods Enzymol. 306: 145-53) para generar pSATE111 (Figura 2).

Ejemplo 2: Construcción del vector de expresión pMJ04

[0248] El vector de expresión pMJ04 fue construido por PCR amplificando el terminador del gen de exocelobiohidrolasa 1 (*cbh1*) de *Trichoderma reesei* de ADN genómico de *Trichoderma reesei* RutC30 (Montenecourt y Eveleigh, 1979, Adv. Chem. Ser. 181: 289 301) usando cebadores 993429 (antisentido) y 993428 (sentido) mostrados más abajo. El cebador de antisentido fue creado genéticamente para que tuviera un sitio *Pac I* en el extremo 5' y un sitio *Spe I* en el extremo 3' del cebador de sentido.

Cebador 993429 (antisentido): 5'-AACGTTAATTAAGGAATCGTTTTGTGTTT-3' (SEC ID n°: 5)

Cebador 993428 (sentido): 5'-AGTACTAGTAGCTCCGTGGCGAAAGCCTG-3' (SEC ID n°: 6)

[0249] El ADN genómico de *Trichoderma reesei* RutC30 fue aislado usando un equipo DNeasy Plant Maxi (Qiagen, Chatsworth, CA).

[0250] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de reacción 1X TernoPol (New England Biolabs, Beverly, MA), 0,3 mM de dNTPs, 100 ng de ADN genómico de *Trichoderma reesei* RutC30, 0,3 µM de cebador 993429, 0,3 µM de cebador 993428 y 2 unidades de Vent ADN polimerasa (New England Biolabs, Beverly, MA). Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 5 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 50°C, y 60 segundos a 72°C, seguidos de 25 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 65°C, y 120 segundos a 72°C (final extensión de 5 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 229 bp fue cortada del gel y purificada usando un equipo de extracción de gel QIAquick.

[0251] El fragmento de PCR resultante fue digerido con *Pac I* y *Spe I* y fue ligado en pAILo01 digerido con las mismas enzimas de restricción usando un equipo Rapid Ligation Kit (Roche, Indianapolis, IN), para generar pMJ04 (Figura 3).

Ejemplo 3: Construcción de vector de expresión pCaHj568

[0252] El plásmido de expresión pCaHj568 fue construido a partir de pCaHj170 (patente U.S. 5.763.254) y pMT2188. El plásmido pCaHj170 comprende la región de codificación de endoglucanasa V de *Humicola insolens* (EGV). El plásmido pMT2188 fue construido de la siguiente manera: El origen de replicación pUC19 fue amplificado por PCR de pCaHj483 (WO 98/00529) con los cebadores 142779 e 142780 mostrados más abajo El cebador 142780 introducido un sitio *Bbu I* en el fragmento de PCR.

142779: 5'-TTGAATTGAAAATAGATTGATTTAAAACCTTC-3' (SEC ID n°: 7)

142780: 5'-TTGCATGCGTAATCATGGTCATAGC-3' (SEC ID nº: 8)

[0253] El sistema Expand PCR (Roche Molecular Biochemicals, Basel, Suiza) fue usado para la amplificación siguiendo las instrucciones del fabricante y las posteriores amplificaciones por PCR. Los productos de PCR fueron separados en un gel de agarosa al 1% usando tampón TAE y un fragmento de 1160 bp fue aislado y purificado usando un equipo de centrifugación de extracción de gel Jetquick (Genomed, Wielandstr, Alemania).

[0254] El gen URA3 fue amplificado del vector de clonación de *Saccharomyces cerevisiae* pYES2 (Invitrogen, Carlsbad, CA) usando cebadores 140288 y 142778 a continuación. Cebador 140288 introducido en un sitio *Eco* RI en el fragmento de PCR.

140288: 5'-TTGAATTCATGGGTAATAACTGATAT-3' (SEC ID nº: 9)

142778: 5'-AAATCAATCTATTTTCAATTCATCATT-3' (SEC ID nº: 10)

[0255] Los productos de PCR fueron separados en un gel de agarosa al 1% usando tampón TAE y un fragmento de 1126 bp fue aislado y purificado usando un equipo de centrifugación de extracción de gel Jetquick.

[0256] Los dos fragmentos de PCR fueron fusionados mediante la mezcla y amplificación usando los cebadores 142780 e 140288 mostrados más arriba por empalme de método de superposición (Horton et al., 1989, Gene 77: 61-68). Los productos de PCR fueron separados en un gel de agarosa al 1% usando tampón TAE y un fragmento de 2263 bp fue aislado y purificado usando un equipo de centrifugación de extracción de gel Jetquick.

[0257] El fragmento resultante fue digerido con *Eco* RI y *Bbu*I y ligado al fragmento más grande de pCaHj483 digerido con las mismas enzimas. La mezcla de ligación fue usada para transformar la cepa de *pyrF*-negativa *E. coli* DB6507 (ATCC 35673) hecha competente por el método de Mandel y Hlga, 1970, J. Mol. Biol. 45: 154. Los transformantes fueron seleccionados en medio M9 sólido (Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, 2º edición, Cold Spring Harbor Laboratory Press) suplementados por litro con 1 g de casaminoácidos, 500 µg de tiamina y 10 mg de canamicina. Un plásmido de un transformante fue aislado y designado pCaHj527 (Figura 4).

[0258] El promotor de NA2-tpi presente en pCaHj527 fue sometido a mutagénesis dirigida a un sitio por un enfoque de PCR simple. Los nucleótidos 134-144 fueron convertidos de G TACTAAAACC a CCGT TAAATTT usando el cebador mutagénico 141223:

Cebador 141223: 5'-GATGCTGTTGACTCCGGAAATTTAACGGTTTGGTCTTGCATCCC-3' (SEC ID nº: 11)

Los nucleótidos 423-436 fueron convertidos de ATGCAATTTAAACT a CCGCAATTTAACGG usando el cebador mutagénico 141222:

Cebador 141222: 5'-GGTATTGCCTGCAGACGGCAATTTAACGGCTTCTGCGAATCGC-3' (SEC ID nº: 12)

[0259] El plásmido resultante fue designado pMT2188 (Figura 5).

[0260] La región de codificación de endoglucanasa V de *Humicola insolens* fue transferida de pCaHj170 como un fragmento *Bam* HI-*Sal* I en pMT2188 digerida con *Bam* HI y *Xho* I para generar pCaHj568 (Figura 6).

Ejemplo 4: Construcción del vector de expresión pMJ05

[0261] El vector de expresión pMJ05 fue construido por PCR amplificando la región de codificación de endoglucanasa V de *Humicola insolens* de 915 bp de pCaHj568 usando los cebadores HiEGV-F y HiEGV-R mostrados más abajo.

HiEGV-F (sentido): 5'-AAGCTTAAGCATGCGTTCCTCCCCCTCC-3' (SEC ID nº: 13)

HiEGV-R (antisentido): 5'-CTGCAGAATTCTACAGGCACTGATGGTACCAG-3' (SEC ID nº: 14)

[0262] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de reacción 1X ThermoPol, 0.3 mM de dNTPs, 10 ng de plásmido pCaHj568, 0.3 µM de cebador HiEGV-F, 0.3 µM de cebador HiEGV-R y 2 unidades de Vent ADN polimerasa. Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 5 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 50°C, y 60 segundos a 72°C, seguidos de 25 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 65°C, y 120 segundos a 72°C (extensión final de 5 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 937 bp fue cortado del gel y purificado usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

[0263] Este fragmento purificado de 937 bp fue usado como ADN molde para amplificaciones posteriores usando los siguientes cebadores:

HiEGV-R (antisentido): 5'-CTGCAGAATTCTACAGGCACTGATGGTACCAG-3' (SEC ID nº: 15)

HiEGV-F-superposición (sentido): 5'-ACCGCGGACTGCGCATCATGCGTTCCTCCCCCTCC-3' (SEC ID nº: 16)

[0264] Las secuencias del cebador en cursiva son homólogas a 17 bp del promotor *cbh1* de *Trichoderma reesei* y las secuencias de cebador subrayadas son homólogas a 29 bp de la región de codificación de endoglucanasa V de *Humicola insolens*. La superposición de 36 bp entre el promotor y la secuencia codificante permitió la fusión precisa del

fragmento de 994 bp que comprende el promotor *cbh1* de *Trichoderma reesei* al fragmento de 918 bp que comprende marco de lectura abierto de la endoglucanasa V de *Humicola insolens*.

5 [0265] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de reacción 1X ThermoPol, 0,3 mM de dNTPs, 1 ul de fragmento de PCR purificado de 937 bp, 0,3 µM de cebador de HiEGV-F-superposición, 0,3 µM de cebador HiEGV-R y 2 unidades de Vent ADN polimerasa. Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 5 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 50°C, y 60 segundos a 72°C, seguidos de 25 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 65°C, y 120 segundos a 72°C (final extensión de 5 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 945 bp fue cortado del gel y purificado usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

15 [0266] Se realizó una PCR separada para amplificar la secuencia del promotor *cbh1* de *Trichoderma reesei* que se extiende de 994 bp corriente arriba del codón de iniciación ATG del gen de ADN genómico RutC30 de *Trichoderma reesei* usando los siguientes cebadores (el cebador de sentido fue creado genéticamente para tener un sitio de restricción de *SaI* en el extremo 5'):

TrCBH1pro-F (sentido): 5'-AAACGTCGACCGAATGTAGGATTGTTATC-3' (SEC ID n°: 17)

TrCBH1pro-R (antisentido): 5'-GATGCGCAGTCCGCGGT-3' (SEC ID n°: 18)

20 [0267] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de reacción 1X ThermoPol, 0,3 mM de dNTPs, 100 ng de ADN genómico RutC30 de *Trichoderma reesei*, 0,3 µM de cebador TrCBH1pro-F, 0,3 µM de cebador TrCBH1pro-R, y 2 unidades de Vent ADN polimerasa. Las reacciones fueron incubadas 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 120 segundos a 72°C (extensión final de 5 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde un producto de banda de 998 bp fue cortado del gel y purificado usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

25 [0268] El fragmento de PCR purificado de 998 bp fue usado para como ADN molde para amplificaciones posteriores usando los siguientes cebadores:

30 TrCBH1pro-F: 5'-AAACGTCGACCGAATGTAGGATTGTTATC-3' (SEC ID n°: 19)

TrCBH1pro-R-superposición: 5'-GGAGGGGGGAGGAACGCATGATGCGCAGTCCGCGGT-3' (SEC ID n°: 20)

35 [0269] Las secuencias en cursiva son homólogas a 17 bp del promotor *cbh1* de *Trichoderma reesei* y las secuencias subrayadas son homólogas a 29 bp de la región de codificación de la endoglucanasa V de *Humicola insolens*. La superposición de 36 bp entre el promotor y la secuencia codificante permitió la fusión precisa del fragmento de 994 bp que comprende el promotor *cbh1* de *Trichoderma reesei* al fragmento de 918 bp que comprende el marco de lectura abierto de endoglucanasa V de *Humicola insolens*.

40 [0270] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de reacción 1X ThermoPol, 0,3 mM de dNTPs, 1 ul de fragmento de PCR purificado de 998 bp, 0,3 µM de cebador de TrCBH1pro-F, 0,3 µM de cebador TrCBH1pro-R-superposición y 2 unidades de Vent ADN polimerasa. Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 5 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 50°C, y 60 segundos a 72°C, seguidos de 25 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 65°C, y 120 segundos a 72°C (extensión final de 5 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 1017 bp fue cortado del gel y purificado usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

45 [0271] El fragmento de PCR del promotor *cbh1* de *Trichoderma reesei* de 1017 bp y los fragmentos de PCR de endoglucanasa V de *Humicola insolens* de 945 bp fueron usados como ADN molde para la amplificación posterior usando los siguientes cebadores para fusionar precisamente el promotor *cbh1* de *Trichoderma reesei* de 994 bp a la región de codificación de endoglucanasa V de *Humicola insolens* de 918 bp usando PCR de superposición.

50 TrCBH1pro-F: 5'-AAACGTCGACCGAATGTAGGATTGTTATC-3' (SEC ID n°: 21)

HiEGV-R: 5'-CTGCAGAATTCTACAGGCACTGATGGTACCAG-3' (SEC ID n°: 22)

55 [0272] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de reacción 1X ThermoPol, 0,3 mM de dNTPs, 0,3 µM de cebador TrCBH1pro-F, 0,3 µM de cebador HiEGV-R y 2 unidades de Vent ADN polimerasa. Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 5 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 50°C, y 60 segundos a 72°C, seguidos de 25 ciclos cada uno durante 30 segundos a 94°C, 30 segundos a 65°C, y 120 segundos a 72°C (extensión final de 5 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 1926 bp fue cortada del gel y purificada usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

60 [0273] El fragmento de 1926 bp resultante fue clonado en vector PCR-Blunt-II-TOPO (Invitrogen, Carlsbad, CA) usando el equipo de clonación ZeroBlunt TOPO PCR siguiendo el protocolo del fabricante. El plásmido resultante fue digerido con *NotI* y *Sall* y el fragmento de 1926 bp fue purificado y ligado en el vector de expresión pMJ04, que también fue digerido con las mismas dos enzimas de restricción, para generar pMJ05 (Figura 7).

Ejemplo 5: Construcción del vector de expresión pSMai130

[0274] Un fragmento de ADN de 2586 bp que se extiende del codón de iniciación ATG al codón de detención de TAA de la secuencia codificante de beta-glucosidasa de *Aspergillus oryzae* (SEC ID n°: 1 para la secuencia de ADNc y SEC ID n°: 2 para la secuencia de aminoácidos deducida; *E. coli* DSM 14240) fue amplificada por PCR de pJaL660 (WO 2002/095014) como molde con los cebadores 993467 (sentido) y 993456 (antisentido) que se muestran a continuación. Un sitio *Spe* I fue creado genéticamente en el extremo 5' del cebador antisentido para facilitar la ligación. Las secuencias de cebador en cursiva son homólogas al promotor *cbh1* de 24 bp de *Trichoderma reesei* y las secuencias subrayadas son homólogas a la región de codificación de beta-glucosidasa de *Aspergillus oryzae* de 22 bp.

Cebador 993467: 5'-ATAGTCAACCGCGGACTGCGCATCATGAAGCTTGGTTGGATCGAGG-3' (SEC ID n°: 23)

Cebador 993456: 5'-ACTAGTTTACTGGGCCTTAGGCAGCG-3' (SEC ID n°: 24)

[0275] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de amplificación *Pfx* (Invitrogen, Carlsbad, CA), 0,25 mM de dNTPs, 10 n de pJaL660, 6,4 µM de cebador 993467, 3,2 µM de cebador 993456, 1 mM de MgCl₂ y 2,5 unidades de polimerasa *Pfx* (Invitrogen, Carlsbad, California). Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 30 ciclos cada uno durante 60 segundos a 94°C, 60 segundos a 55°C, y 180 segundos a 72°C (final extensión de 15 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 2586 bp fue cortada del gel y purificada usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

[0276] Una PCR separada fue realizada para amplificar la secuencia del promotor *cbh1* de *Trichoderma reesei* que se extiende de 1000 bp corriente arriba del codón de iniciación ATG del gen, usando de cebador 993453 (sentido) y cebador 993463 (antisentido) mostrados más abajo para generar un fragmento de PCR de 1000 bp. Las secuencias de cebador en cursiva son homólogas del promotor *cbh1* de *Trichoderma reesei* de 24 bp y las secuencias de cebador subrayadas son homólogas de la región de codificación de beta-glucosidasa de *Aspergillus oryzae* de 22 bp. La superposición del promotor de 46 bp entre la secuencia codificante permitió la fusión precisa del fragmento de 1000 bp que comprende el promotor *cbh1* de *Trichoderma reesei* al fragmento de 2586 bp que comprende el marco de lectura abierto de beta-glucosidasa de *Aspergillus oryzae*.

Cebador 993453: 5'-GTCGACTCGAAGCCCGAATGTAGGAT-3' (SEC ID n°: 25)

Cebador 993463: 5'-CCTCGATCCAACCAAGCTTCATGATGCGGAGTCCGCGGTTGAGTA-3' (SEC ID n°: 26)

[0277] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de amplificación *Pfx*, 0,25 mM de dNTPs, 100 ng de ADN genómico *RutC30* de *Trichoderma reesei*, 6,4 µM de cebador 993453, 3,2 µM de cebador 993463, 1 mM de MgCl₂ y 2,5 unidades de polimerasa *Pfx*. Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 30 ciclos cada uno durante 60 segundos a 94°C, 60 segundos a 55°C, y 180 segundos a 72°C (extensión final de 15 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 1000 bp fue cortada del gel y purificada usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

[0278] Los fragmentos purificados fueron usados como ADN molde para amplificación posterior usando el cebador 993453 (sentido) y el cebador 993456 (antisentido) mostrado más arriba para fusionar precisamente el promotor *cbh1* de *Trichoderma reesei* de 1000 bp a la beta-glucosidasa de *Aspergillus oryzae* de 2586 bp por PCR de superposición.

[0279] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de amplificación *Pfx*, 0,25 mM de dNTPs, 6,4 µM de cebador 99353, 3,2 µM de cebador 993456, 1 mM de MgCl₂ y 2,5 unidades de polimerasa *Pfx*. Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 30 ciclos cada uno durante 60 segundos a 94°C, 60 segundos a 60°C, y 240 segundos a 72°C (extensión final de 15 minutos).

[0280] El fragmento de 3586 bp resultante fue digerido con *Sal* I y *Spe* I y ligado en pMJ04, digerido con las mismas dos enzimas de restricción, para generar pSMai130 (Figura 8).

Ejemplo 6: Construcción de pSMai135

[0281] La región de codificación de beta-glucosidasa de *Aspergillus oryzae* (menos la secuencia señal nativa putativa, véase la Figura 9) de Lys-20 al codón de detención TAA fue amplificada por PCR a partir de pJaL660 (WO 2002/095014) como molde con cebador de 993728 (sentido) y cebador 993727 (antisentido) que se muestran más abajo. Las secuencias en cursiva son homólogas de la secuencia señal de endoglucanasa V de *Humicola insolens* de 20 bp y las secuencias subrayadas son homólogas de la región de codificación de beta-glucosidasa de *Aspergillus oryzae* de 22 bp. Un sitio *Spe* I fue creado genéticamente en el extremo 5' del cebador antisentido.

Cebador 993728: 5'-TGCCGGTGTGGCCCTTGCCAAGGATGATCTCGGCTACTCCC-3' (SEC ID n°: 27)

Cebador 993727: 5'-GACTAGTCTTACTGGGCCTTAGGCAGCG-3' (SEC ID n°: 28)

[0282] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de amplificación *Pfx*, 0,25 mM de dNTPs, 10 ng de pJaL660, 6,4 µM de cebador 993728, 3,2 µM de cebador 993727 1 mM de MgCl₂ y 2,5 unidades de polimerasa *Pfx*. Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente

manera: 30 ciclos cada uno durante 60 segundos a 94°C, 60 segundos a 55°C, y 180 segundos a 72°C (extensión final de 15 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 2523 bp fue cortada del gel y purificada usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

[0283] Una amplificación por PCR separada fue realizada para amplificar el promotor *cbh1* de *Trichoderma reesei* de 1000 bp y secuencia señal de endoglucanasa V de *Humicola insolens* putativa de 63 bp (codón de iniciación ATG a Ala-21, Figura 10, SEC ID n°: 29), usando un cebador 993724 (sentido) y cebador 993729 (antisentido) que se muestran más abajo. Las secuencias de cebador en cursiva son homólogas de la secuencia señal de endoglucanasa V de *Humicola insolens* de 20 bp y las secuencias de cebador subrayadas son homólogas de la región de codificación de beta-glucosidasa de *Aspergillus oryzae* de 22 bp. El plásmido pMJ05, que comprende la región de codificación de endoglucanasa V de *Humicola insolens* bajo el control del promotor *cbh1* fue usado como un molde para generar un fragmento de 1063 bp que comprende el fragmento de la secuencia señal de endoglucanasa V de *Humicola insolens*/promotor *cbh1* de *Trichoderma reesei*. Una superposición de 42 bp fue compartida entre la secuencia señal de endoglucanasa V de *Humicola insolens*/promotor *cbh1* de *Trichoderma reesei* y la secuencia codificante de *Aspergillus oryzae* para proporcionar un enlace perfecto entre el promotor y el codón de iniciación ATG de beta-glucosidasa de *Aspergillus oryzae* de 2523 bp.

Cebador 993724: 5'-ACGCGTCGACCGAATGTAGGATTGTTATCC-3' (SEC ID n°: 30)

Cebador 993729: 5'-GGGAGTACGCGAGATCATCCTTGCCAAGGGCCAACACCGGCA-3' (SEC ID n°: 31)

[0284] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de amplificación *Pfx*, 0,25 mM de dNTPs, 10 ng/µl de pMJ05, 6,4 µM de cebador 993728, 3,2 µM de cebador 993727, 1 mM de MgCl₂ y 2,5 unidades de polimerasa *Pfx*. Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 30 ciclos cada uno durante 60 segundos a 94°C, 60 segundos a 60°C, y 240 segundos a 72°C (extensión final de 15 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 1063 bp fue cortada del gel y purificada usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

[0285] Los fragmentos de superposición purificada fueron usados como un molde para la amplificación usando cebador 993724 (sentido) y cebador 993727 (antisentido) anteriormente descritos para fusionar precisamente el fragmento de 1063 bp que comprende la secuencia señal de endoglucanasa V de *Humicola insolens*/ promotor *cbh1* de *Trichoderma reesei* al fragmento de 2523 bp que comprende el marco de lectura abierto de beta-glucosidasa de *Aspergillus oryzae* por PCR de superposición.

[0286] Las reacciones de amplificación (50 µl) fueron compuestas por tampón de amplificación *Pfx*, 0,25 mM de dNTPs, 6,4 µM de cebador 993724, 3,2 µM de cebador 993727, 1 mM de MgCl₂ y 2,5 unidades de polimerasa *Pfx*. Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado de la siguiente manera: 30 ciclos cada uno durante 60 segundos a 94°C, 60 segundos a 60°C, y 240 segundos a 72°C (extensión final de 15 minutos). Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 3591 bp fue cortado del gel y purificado usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

[0287] El fragmento de 3591 bp resultante fue digerido con *Sall* y *Spe I* y ligado en pMJ04 digerido con las mismas enzimas de restricción para generar pSMai135 (Figura 11).

Ejemplo 7: Construcción de vector de expresión de *Saccharomyces cerevisiae* pALFd1

[0288] El plásmido pALFd1 fue generado de pSATE111 para cambiar la señal de secreción de beta-glucosidasa de *Aspergillus oryzae* nativa con la señal de secreción de la endoglucanasa V de *Humicola insolens* para mejorar la producción y la secreción de beta-glucosidasa de *Aspergillus oryzae* en *Saccharomyces cerevisiae*. El plásmido pSATE111 fue digerido con *Xho I* y *Spe I* para liberar un fragmento de 2,6 kb (beta-glucosidasa de *Aspergillus oryzae*) y un fragmento de 6 kb (resto del vector). La digestión fue realizada en un gel de agarosa al 0,7% usando tampón TAE y el fragmento de 6 kb fue aislado por purificación de gel usando un equipo de extracción de gel QIAquick (QIAGEN Inc., Valencia, CA) siguiendo el protocolo del fabricante y ligado al fragmento de PCR de 2,6 kb, con la región de codificación de beta-glucosidasa de *Aspergillus oryzae* (menos la secuencia señal de secreción) y la secuencia señal de endoglucanasa V de *Humicola insolens*, que fue amplificada de pSMai135 usando cebadores 993950 y 993951 que se muestran más abajo. Los cebadores contienen sitios de restricción *Spe I* y *Xho I* en sus extremos para subclonación posterior en los sitios de restricción *Spe I* y *Xho I* de pSATE111.

Cebador 993950: 5'-AATCCGACTAGTGGATCTACCATGCGTTCCTCCCCCTCC-3' (SEC ID n°: 32)

Cebador 993951: 5'-GCGGGCCTCGAGTTACTGGCCTTAGGCAGCG-3' (SEC ID n°: 33)

[0289] Las reacciones de amplificación (100 µl) fueron compuestas por tampón ThermoPol de PCR, 0,20 mM de dNTPs, 0,14 µg de Adn plásmido de pSMai135, 50 µM de cebador 993950, 50 µM de cebador 993951 y 2 unidades de Vent ADN polimerasa. Las reacciones fueron incubadas en un Termociclador RoboCycler de gradiente 40 (Stratagene, la Jolla, CA) programado de la siguiente manera: un ciclo de 1 minuto a 95°C, y 25 ciclos cada uno durante 1 minuto a 95°C, 1 minuto a 60 o 64°C y 3 minutos a 72°C (extensión final de 10 minutos). Los productos reactivos fueron

visualizados en un gel de agarosa al 0,7% usando tampón TAE. Los fragmentos de 2.6 kb resultantes fueron purificados usando una purificación por PCR PCR MinElute (Qiagen, Chatsworth, CA) según las instrucciones del fabricante. Los fragmentos purificados fueron combinados y digeridos con *SpeI* y *Xho I* y ligados en pSATE111 digerido con las mismas dos enzimas de restricción para generar pALFd1 (Figura 12).

5

Ejemplo 8: Generación de bibliotecas primarias de beta-glucosidasa mutagenizada en *Saccharomyces cerevisiae*

10 [0290] En un esfuerzo por identificar regiones en la beta-glucosidasa de *Aspergillus oryzae* que son críticas para la termoestabilidad de la proteína, el gen entero de beta-glucosidasa de *Aspergillus oryzae* de tipo salvaje fue mutagenizado usando PCR con tendencia al error con secuencias homólogas al vector de expresión de levadura pSATE111, que puede sufrir recombinación *in vivo* entre dominios homólogos de fragmentos diferentes. Este proceso generó plásmidos de replicación circulares de una combinación de vector linealizado y productos de PCR.

15 [0291] El cebador 992328 (del Ejemplo 1) y el cebador AoJal660.2, que se muestra más abajo, fueron usados en la amplificación por PCR con tendencia al error del gen de beta-glucosidasa de pSATE101 para generar secuencias mutagenizadas que podrían ser clonadas en pSATE111 para la expresión de la enzima de beta-glucosidasa en levadura. Cebador AoJal660.2: 5'-AGGGTGAATGGGCGGAA-3' (SEC ID n°: 34)

20 [0292] Las amplificaciones de PCR con tendencia al error (50 µl) fueron compuestas por tampón 1X Taq con 1,5 mM de MgCl₂ (Promega Corporation, Madison, WI), 60 ng de pSATE101, 0,16 mM de DATP, 0,07 mM cada uno de dCTP, dGTP y dTTP, 50 µM de cebador AoJA1660.2, 50 µM de cebador 992328, 0,1 mM de MnCl₂ y 5 unidades de Taq ADN polimerasa (Promega Corporation, Madison, WI). Las reacciones de amplificación fueron incubadas como se describe en el ejemplo 1.

25

[0293] El plásmido pSATE111 fue escindido por digestión con *Eco RI* y *Sma I* y luego purificado en gel usando resina QiaexII (QIAGEN Inc., Valencia, CA). La digestión fue verificada por fraccionamiento de una alícuota de la digestión en un gel de agarosa al 0,7% usando tampón TAE y coloreando con bromuro de etidio donde los fragmentos previstos de 8,054 bp (vector escindido que contiene parte de la secuencia codificante de beta-glucosidasa) y 672 bp (del gen de beta-glucosidasa) fueron obtenidos. La digestión fue purificada usando resina QiaexII.

30

[0294] Tres partes alícuotas de µl de las reacciones de la PCR de más arriba fueron mezcladas con 0,5 µl del vector pSATE111 escindido para cotransformación en células competentes de *Saccharomyces cerevisiae* YNG318. Los fragmentos cotransformados contuvieron al menos 250 bp de secuencia de ADN homóloga en los extremos para facilitar la reparación de espacios del plásmido expresado. Las células competentes de *Saccharomyces cerevisiae* YNG 318 fueron preparadas antes de cada transformación siguiendo el Protocolo de transformación de levadura YEASTMAKER (CLONTECH Laboratories, Inc., Palo Alto, CA) con las siguientes modificaciones: (1) El volumen de cultivo de levadura usado para inocular la incubación durante toda la noche (16-20 horas) fue entre 100-1.000 µl; (2) la recuperación de células después de la transformación fue realizada en medio YPD durante 45 minutos a 30°C; y (3) la mezcla de transformación fue dividida en partes alícuotas para colocar en placas en medio de placa de selección de levadura mientras el resto de los transformantes fueron congelados a -80°C en un congelador de índice controlado (Nalge Nunc International, Rochester, NY).

35

40

45 [0295] Las placas fueron incubadas a 30°C durante 4 días aproximadamente. Las colonias que produjeron beta-glucosidasa activa se volvieron azules tras la incubación debido a la hidrólisis de beta-glucosidasa de X-Glc. La actividad de la biblioteca fue estimada por el porcentaje de colonias azules obtenido. La biblioteca fue 67% activa.

45

Ejemplo 9: Selección de bibliotecas de beta-glucosidasa

50 [0296] La biblioteca de beta-glucosidasa primaria fue extendida en Genetix QTray's (placas de Petri de 22 x 22 cm, Genetics Ltd., Hampshire, Reino Unido) e incubada durante 5 días a 30°C. Usando un Genetix QPix (Genetix Ltd., Hampshire, Reino Unido), las colonias activas fueron escogidas usando selección X-Glc en placas de 96 pocillos con medio de selección de levadura. Las placas fueron incubadas durante 7 días a 30°C. El tampón de selección (0,1 M de succinato pH 5) se añadió a las placas de crecimiento antes del inicio de la pantalla. Usar un robot ORCA (Beckman Coulter, Fullerton, CA), las placas de crecimiento fueron transportadas a un Multimek (Beckman Coulter, Fullerton, CA) y las muestras fueron tomadas de la placa de crecimiento y mezcladas en placas de 96 pocillos de policarbonato de fondo en V. Las muestras luego fueron tomadas de las placas de fondo en V y dispensadas en placas de fondo plano de 96 pocillos vacíos para un ensayo de placa inicial con p-nitrofenil-beta-D-glucopiranosidos como sustrato en 0,1 M de succinato pH 5 a temperatura ambiente. La placa de fondo en V fue transportada a un bloque de calentamiento de 96 pocillos adaptado e incubada a 65°C por un total de 10 minutos. Las muestras luego fueron tomadas de las placas con fondo en V sometidas al bloque de calentamiento y dispensadas en placas de fondo plano de 96 pocillos vacíos para un ensayo de placa final con p-nitrofenil-beta-D-glucopiranosidos como sustrato en 0,1 M de succinato pH 5 a temperatura ambiente. Las placas de ensayo iniciales y finales luego fueron transportadas a un Multidrop (Labsystems, Vantaa, Finlandia) donde se añadió el sustrato de p-nitrofenil-beta-D-glucopiranosido. Después de que el período de incubación de ensayo predeterminado haya terminado, entre 30 y 120 minutos, las placas de ensayo iniciales y finales

55

60

65

fueron templadas con tampón 2 M tris pH 8. Ambas placas fueron leídas usando un lector de placas Spectramax (Molecular Devices, Sunnyvale, CA) a 405 nm. La proporción de la lectura final a la lectura inicial fue calculada usando Microsoft Excel (Microsoft Corporation, Redmond, WA) para determinar la actividad residual del porcentaje (%AR). Según las mediciones de %AR, la selección de las bibliotecas construidas en el Ejemplo 8 generaron dos variantes: BG13 y BG14. La actividad residual producida por estas variantes fue 13% y 17% para BG13 y BG14, respectivamente.

Ejemplo 10: Biblioteca redistribuida de las variantes de beta-glucosidasa BG13 y BG14

[0297] Para redistribuir el ADN de las variantes de beta-glucosidasa BG13 y BG14 del Ejemplo 9, el Adn plásmido fue aislado de las variantes. Cada variante fue cultivada a 30°C durante toda la noche en 3 ml de medio de selección de levadura con 75 µg de cloranfenicol para prevenir la contaminación. Las muestras de 100 µl de los cultivos durante toda la noche fueron centrifugadas durante 3 minutos a 13.200 x g. El sobrenadante fue eliminado y el ADN fue aislado del granulado restante según el protocolo descrito por Kaiser y Auer, 1993, BioTechniques 14 (4): 552, salvo que se hayan utilizado 20 µl de tampón para lisis de levadura.

[0298] Los genes de beta-glucosidasa de *Aspergillus oryzae* de variantes BG13 y BG14 fueron amplificados por PCR usando los siguientes cebadores:

Cebador anidado pSATE1115': 5'-GACATTTTTGCTGTCAGTCA-3' (SEC ID n°: 35)

Cebador anidado pSATE1113': 5'-AATGTTACATGCGTACACGC-3' (SEC ID n°: 36)

[0299] Se llevaron a cabo tres reacciones de la PCR usando ADN rescatado de la variante BG13 como molde y 5 reacciones usando ADN rescatado de la variante BG14. Las reacciones de amplificación (100 µl) fueron compuestas por 0,5 µl de Adn plásmido de levadura, tampón AmpliTaq 1X sin MgCl₂, 125 µM cada uno de dATP, dCTP, dGTP y dTTP, 50 pmol de cada cebador, 1,5 mM de MgCl₂ y 1 unidad de ADN polimerasa de AmpliTaq (PE Applied Biosystems, Foster City, CA). Las reacciones fueron incubadas en un Eppendorf Mastercycler 5333 programado durante 1 ciclo a 95°C durante 5 minutos; 25 ciclos cada uno a 95°C durante 1 minuto, 50°C durante 1 minuto; y 72°C durante 3 minutos; y un ciclo de extensión final a 72°C durante 10 minutos. Las tres reacciones de la PCR para BG13 y cinco reacciones de la PCR para BG14 fueron combinadas y purificadas usando un equipo de purificación de PCR QIAquick (QIAGEN Inc., Valencia, CA). El ADN fue eluido en 30 µl de tampón EB (QIAGEN Inc., Valencia, CA). La concentración de los productos de PCR purificados obtenidos después de la amplificación de cada variante fue verificada por visualización en un gel de agarosa al 0,7% realizada en tampón TAE y coloreada con bromuro de etidio. Cada variante produjo 125 ng de ADN por microlitro.

[0300] Para redistribuir el ADN de las variantes BG13 y BG14, 3, 2 µl (400 ng) del producto de PCR BG13 y 3,2 µl (400 ng) del producto de PCR BG14 fueron combinados con 2 µl (400 ng) de pSATE111 escindido preparado como se describe en el Ejemplo 8, salvo que se haya utilizado *SpeI* y *XhoI* para escindir el vector y luego se hayan transformado en células de *Saccharomyces cerevisiae* YNG318 competentes recientemente realizadas como se describe en el ejemplo 8. La biblioteca generada fue 78% activa según el porcentaje de colonias azules generado.

[0301] En total, se escogieron 6.336 colonias activas de esta biblioteca y fueron seleccionadas como se describe en el Ejemplo 9, lo cual dio como resultado el aislamiento de 6 variantes termoestables mejoradas, cuatro de las cuales fueron designadas BG40; BG41; BG42 y BG43. Para la secuenciación del ADN de las variantes BG40; BG41; BG42 y BG43, el ADN fue aislado de una muestra de 100 µl de caldo de selección de levadura según el procedimiento de Kaiser y Auer, 1993, supra. El ADN aislado fue transformado en células competentes de electroporación de *E. coli* SURE según las instrucciones del fabricante. El Adn plásmido fue aislado como se describe anteriormente y luego fue secuenciado.

[0302] La región de codificación entera de cada gen de variante de beta-glucosidasa de *Aspergillus oryzae* fue ordenada usando 0,5 µl de Adn plásmido y 3,2 pmol de los siguientes cebadores:

AoJal660.1: posición 5'-GTTTCGGCTCAGGACTG-3': 2492 directo (SEC ID n°: 37)

AoJal660.1a: posición 5'-ACTTCCGCCATTACACC-3': 141 inverso (SEC ID n°: 38)

AoJal660.2: posición 5'-AGGGTGAATGGGCGGAA-3': 123 directo (SEC ID n°: 39)

AoJal660.2a: posición 5'-GGCGGAAATGCTCTTGT-3': 614 inverso (SEC ID n°: 40)

AoJal660.3: posición 5'-GGATGGCGGTAGAAACT-3': 469 directo (SEC ID n°: 41)

AoJal660.3a: posición 5'-GCGGTCCAATCACTCAT-3': 861 inverso (SEC ID n°: 42)

AoJal660.4: posición 5'-GCTACGGTTGCGAGAAT-3': 774 directo (SEC ID n°: 43)

AoJal660.4a: posición 5'-CTCAAGGGCAAGGCACC-3': 1232 inverso (SEC ID n°: 44)

AoJal660.5: posición 5'-GGTGCCCTTGCCCTTGAG-3': 1232 directo (SEC ID n°: 45)

AoJal660.5a: posición 5'-TTCGCTGCGGTCTTGAC-3': 1629 inverso (SEC ID n°: 46)

AoJal660.6: posición 5'-GTGGAAGAACGGCGACA-3': 1591 dir (SEC ID n°: 47)

AoJal660.6a: posición 5'-CCCAGCCGTAGTTAGAA-3': 2195 inv (SEC ID n°: 48)

AoJal660.7: posición 5'-CGTCCCAGTACTCCC-3': 2019 dir (SEC ID n°: 49)

AoJal660.7a: posición 5'-CCTGGAGCGGCAGTTTC-3': 2573 inv (SEC ID n°: 50)

AoJal660.8: posición 5'-GGTCGGTGTCCTTAACGG-3': 964 dir (SEC ID n°: 51)

AoJal.660.8a: posición 5'-ACTATCTGCAAACACAAGC-3': 292 inv (SEC ID n°: 52)

AoJa1660.9: posición 5'-CCTTTCACCTGGGGCA-3' 1802 dir (SEC ID n°: 53)

AoJal660.9a: posición 5'-GGAGTTACCAGACTCCTGGC-3' 1756 inv (SEC ID nº: 54)

AoJa1660.10a: posición 5'-ACCTTCCGAAACATGGTTAT-3' 1132 inv (SEC ID nº: 55)

5 [0303] La secuenciación indicó que había mutaciones consistentes en los mutantes aislados que condujeron a sustituciones de aminoácidos. La variante BG40 poseía 5 mutaciones, en las cuales 3 de ellas condujeron a sustituciones de aminoácidos. La primera mutación, que condujo a la sustitución G4S, fue encontrada en la secuencia del péptido señal predicha de beta-glucosidasa. Se encontraron tres mutaciones en el sitio activo de la proteína, pero sólo 1 de las mutaciones condujeron a una sustitución de aminoácido: H285Q. Se encontró una mutación final en el dominio X de la proteína: D384N. El dominio X es la región exterior a la región del sitio activo.

10 [0304] La variante BG41 obtuvo tres mutaciones con dos de ellas que condujeron a sustituciones de aminoácidos: G161S y H285Q. Todas las sustituciones fueron encontradas en el sitio activo de la proteína.

15 [0305] La variante BG42 obtuvo tres mutaciones con dos que fueron silenciosas. La única sustitución de aminoácido real fue H285Q.

[0306] La variante BG43 obtuvo tres mutaciones, una silenciosa. La sustitución de aminoácido E7V fue localizada en la secuencia señal predicha y la sustitución de aminoácido H285Q fue encontrada en el sitio activo de la proteína.

20 [0312] La única sustitución de aminoácidos encontrada en todas las variantes aisladas fue H285Q.

Ejemplo 11: Segunda ronda de redistribución de variantes mejoradas de beta-glucosidasa

25 [0308] Para generar una biblioteca redistribuida con varios de los mutantes previamente aislados, las regiones de codificación de beta-glucosidasa de las variantes BG41 y BG43 fueron amplificadas usando Adn plásmido aislado como se describe anteriormente. Una reacción PCR de 100 µl fue realizada como se describe en el Ejemplo 10 salvo que se haya utilizado 200 µM de cada dATP, dCTP, dGTP, y dTTP y tampón 1X AmpliTaq con 1,5 mM de MgCl₂. Los insertos de beta-glucosidasa amplificada (BG41 y BG43) fueron purificados como se describe en el Ejemplo 10 cada uno produciendo una concentración de ADN estimada de 125 ng por microlitro. Las regiones de codificación de las variantes BG13 y BG14 también fueron más amplificadas y purificadas para redistribuir usando 1 µl de Adn plásmido en una reacción PCR como se describe en el Ejemplo 10. El rendimiento de cada producto purificado fue 1,250 ng por microlitro para BG13 y 125 ng por microlitro para BG14.

35 [0309] Una quinta variante designada BG2 fue generada como se describe en el Ejemplo 8 usando los cebadores descritos en el Ejemplo 10, y aislada mediante selección como se describe en el Ejemplo 9. El ADN de la variante BG2 también fue adicionado a la redistribución con el ADN de BG13; BG14; BG41; BG43, pero el análisis posterior mostró que BG2 fue un falso positivo. El producto de PCR purificado para cada mutante fue combinado (125 ng cada uno, salvo BG14 donde se utilizaron 12,5 ng) con 400 ng de pSATE111 escindido con *Spe I/Xho I* para transformación por reparación de espacio en 50 µl de células competentes de *Saccharomyces cerevisiae* YNG318 como se describe en el Ejemplo 8. La biblioteca resultante fue 93,9% activo según el porcentaje de colonias azules.

45 [0310] La selección de la biblioteca fue realizada como se describe en el Ejemplo 9, salvo que las placas fueron incubadas a 68°C. En total, 15.439 clones activos fueron escogidos de la biblioteca y fueron seleccionados, lo cual dio como resultado el aislamiento de las variantes BG47; BG48 y BG 49, que tuvieron 53%, 87%, y 21% de actividad residual tras la incubación durante 10 minutos a 68°C usando p-nitrofenil-beta-D-glucopiranosida como sustrato según se describe en el Ejemplo 9.

50 [0311] Para la secuenciación del ADN de las variantes BG47; BG48 y BG 49, el ADN fue aislado como se describe en el ejemplo 10, excepto que el ADN fuera aislado de una muestra de 500 µl de caldo de selección de levadura y transformado células ultracompetentes en *E. coli* XL-10 Gold.

55 [0312] La secuenciación de estos mutantes reveló que la variante BG47 contenía una sustitución de aminoácido E7V en la secuencia señal, una mutación silenciosa y dos sustituciones de aminoácido en G161S y H285Q en el sitio activo. La variante BG48 contuvo las mismas mutaciones que las encontradas en la variante BG41 más sustituciones de aminoácido en A35T en el sitio activo y en D722G en el dominio X. La variante BG49 fue idéntica a BG43.

Ejemplo 12: Generación de una biblioteca aleatorizada específica del sitio en las posiciones G161 y H285

60 [0313] En un esfuerzo por identificar las mejores sustituciones en las posiciones 161 y 285, que estuvieron presentes en la variante BG41, los nucleótidos que codifican los aminoácidos en estas posiciones, G161S y H285Q, en BG41, fueron aleatorizados sustituyéndolos con NN(G/C). Por lo tanto, la región que codifica estos aminoácidos en el gen de beta-glucosidasa de *Aspergillus oryzae* fue amplificada por PCR y transformada con el pSATE111BG41 que fue digerido con *B/p I*, de modo que los fragmentos de la PCR amplificados que son homólogos al vector de expresión de levadura linealizada pSATE111BG41, con la excepción de los nucleótidos degenerados, puedan sufrir recombinación *in vivo* al transformarse en *Saccharomyces cerevisiae*. Este proceso generó plásmidos de replicación circulares de una combinación de vector linealizado y productos de PCR.

[0314] El cebador BG41SDMSuperior (que contiene nucleótidos degenerados en la posición que codifica el aminoácido 162, en negrita) y el cebador BG41SDMInferior (que contiene nucleótidos degenerados en la posición que codifica el aminoácido 285, en negrita), que se muestran más abajo, fueron usados en la amplificación por PCR del gen de beta-glucosidasa de pSATE111BG41 para generar una secuencia parcial del gen de beta-glucosidasa que contiene estas posiciones aleatorizadas y podría ser clonada en pSATE111BG41 para la expresión de la enzima de beta-glucosidasa en levadura.

Cebador BG41SDMSuperior:

5'-GGTAGAAACTGGGAANNSTTCTCACCGATCCAGCCCTC-3' (SEC ID nº: 56)

Cebador BG41InferiorInferior:

5'-GCCTACGCCGCTGTGNSAGCGGTCCAATCACT-3' (SEC ID nº: 57)

[0315] Las amplificaciones por PCR (100 µl) estaban compuestas de tampón de amplificación 1X *Pfx*, 150 ng de pSATE111BG41, 2 µl de 10 mM de mezcla de dATP, dTTP, dGTP y dCTP, 3 µl de 50 mM de MgSO₄, 50 µM de cebador BG41SDMSuperior, 50 µM de cebador BG41InferiorInferior, 2,5 unidades de ADN polimerasa Platinum *Pfx* (Invitrogen, Carlsbad, CA). Las condiciones de amplificación fueron un ciclo a 95°C durante 5 minutos; 25 ciclos cada uno a 95°C durante 1 minuto, 52, 55, 58 o 61°C durante 1 minuto y 72°C durante 30 segundos; y un ciclo final de 72°C durante 10 minutos. El bloque de calor luego fue a un ciclo de remojo de 4°C.

[0316] El plásmido pSATE111BG41 fue linealizado por digestión con *Bip* I y luego fue limpiado usando el equipo de eliminación de nucleótidos QIAquick (QIAGEN Inc., Valencia, CA) siguiendo las instrucciones del fabricante. Las cuatro reacciones de PCR mencionadas fueron combinadas y limpiadas usando un equipo de purificación de PCR QIAquick y eluidas en 10 µl de tampón EB siguiendo las instrucciones del fabricante. Una alícuota de 7 µl del producto de PCR limpiado (3,5 µg) fue combinada con 560 ng del vector pSATE111BG41 linealizado para cotransformación en células competentes YNG318 de *Saccharomyces cerevisiae* como se describe en el Ejemplo 8. La biblioteca fue 82,8% activa. La alta actividad de esta biblioteca sugirió que la aleatorización de la posición G161 y H285 no fue exitosa. No obstante, esta biblioteca fue seleccionada de todos modos como se describe en el Ejemplo 9.

Ejemplo 13: Tercera ronda de redistribución de variantes mejoradas de beta-glucosidasa

[0317] Para obtener otra variante mejorada, una biblioteca redistribuida con el ADN de las variantes previamente descritas BG2; BG13; BG14 y BG48 fue construida al igual que con la variante BG50, que surgió de la generación de una biblioteca aleatorizada específica de sitio como se describe en el Ejemplo 12 y fue aislada mediante selección como se describe en el Ejemplo 9. El ADN de esta variante fue parcialmente secuenciado y tuvo las mismas mutaciones como las presentes en su secuencia progenitora, BG41, pero con una mutación extra: Q202R. Por lo tanto, esta variante fue un producto de mutagénesis en la reacción de amplificación del vector pSATE111BG41. La generación de la biblioteca fue realizada como se describe en el Ejemplo 10. La selección de la biblioteca fue realizada como se describe en el Ejemplo 9, excepto que las placas fueron incubadas a 70°C. La biblioteca condujo al aislamiento de las variantes BG52; BG53 y BG54, que tuvo un %AR de 60%, 65%, 50%, respectivamente, a 70°C. La secuenciación del ADN de estas variantes fue realizada como se describe en el Ejemplo 10. La variante BG52 tuvo las mismas sustituciones que BG48 con una sustitución S15P en la secuencia señal predicha. La variante BG53 tuvo las mismas sustituciones que BG41 con la adición de una sustitución Q202R y la misma sustitución de aminoácido D722G presente en BG48. BG54 contuvo las siguientes sustituciones: S15P, G161S, Q202R y H285Q.

[0318] La tabla 1 resume las variantes obtenidas de la selección de las bibliotecas primarias o redistribuidas y sus sustituciones de aminoácidos. La tabla 2 muestra la termoestabilidad de las variantes a 65°C, 68°C y 70°C durante 10 minutos. Las mutaciones en la secuencia de ADN que condujeron a sustituciones de aminoácido están en negrita. Las mutaciones en la secuencia de ADN que no condujeron a sustituciones de aminoácido están en texto sencillo.

Tabla 1. Variantes de beta-glucosidasa termoestable

Mutante	Sustituciones de aminoácido	Origen de biblioteca
BG13	N/A	Primaria de WT BG
BG14	N/A	Primaria de WT BG
BG40	G4S ; V245V; H285Q ; P354P ; D384N	Redistribución BG13 y BG14
BG41	G161S ; V245V; H285Q	Redistribución BG13 y BG14
BG42	V245V; H285Q ; N533N	Redistribución BG13 y BG14
BG43	E7V ; V245V; H285Q	Redistribución BG13 y BG14
BG47	E7V ; C74C; G161S ; H285Q	Redistribución "BG2", BG13, BG14, BG43
BG48	A35T ; A151A; G161S ; V245V; H285Q ; D722G	Redistribución "BG2", BG13, BG14, BG43
BG49	E7V , V245V ; H285Q	Redistribución "BG2", BG13, BG14, BG43
BG50	Sec parcial G161S ; Q202R ; V245V; H285Q	Primario de BG41
BG52	S15P ; A35T ; G161S ; H285Q ; D722G	Redistribución "BG2", BG13, BG14, BG48, BG50
BG53	G161S ; Q202R ; H285Q ; D722G	Redistribución "BG2", BG13, BG14, BG48, BG50
BG54	S15P ; G161S ; Q202R ; H285Q	Redistribución "BG2", BG13, BG14, BG48, BG50

Tabla 2. Termoestabilidad de las variantes de beta-glucosidasa

Mutante	Sustituciones de aminoácido	%AR a 65°C	%AR a 68°C	%AR a 70°C
BG13	N/A	13%	N/A	N/A
BG14	N/A	17%	N/A	N/A
BG40	G4S; V245V; H285Q; P354P; D384N	54%	N/A	N/A
BG41	G161S; V245V; H285Q	79%	45%	N/A
BG42	V245V; H285Q; N533N	45%	N/A	N/A
BG43	E7V; V245V; H285Q	45%	31%	N/A
BG47	E7V; C74C; G161S; H285Q	85%	53%	N/A
BG48	A35T; A151A; G161S; V245V; H285Q; D722G	100%	87%	44%
BG49	E7V; V245V; H285Q	N/A	69% (a 69°C)	N/A
BG50	Sec. parcial G161S; Q202R; H285Q	N/A	21%	N/A
BG52	S15P; A35T; G161S; H285Q; D722G	N/A	N/A	60%
BG53	G161S; Q202R; H285Q; D722G	N/A	N/A	65%
BG54	S15P; G161S; Q202R; H285Q	N/A	N/A	50%

[0319] Las muestras de beta-glucosidasa fueron diluidas en el mismo tampón como el usado en el Ejemplo 9 a la misma actividad en relación de una con otra. La muestra fue dividida en dos tubos de ensayo de polipropileno, una parte de la muestra fue incubada sumergida en un baño María de temperatura controlada y la otra parte fue incubada a temperatura ambiente, ambas por un periodo de 42 horas como máximo. Al final del periodo de incubación, las muestras de ambos fueron colocadas en una placa de 96 pocillos. El sustrato de metilumbeliferil-beta-D-glucopiranosido (MUG) (200 µl de 0,5 µM de solución de MUG) se añadió a la placa de 96 pocillos de muestras y fue incubado a temperatura ambiente durante 15 minutos. La reacción fue detenida con la adición de 2M de tampón Tris pH 9,0, y la placa fue leída en un fluorómetro para obtener las Unidades Relativas Fluorescentes (URF) con excitación 365, emisión 454. La actividad residual porcentual fue determinada usando el mismo método que el descrito en el Ejemplo 9. Los resultados se muestran en la Tabla 3. En general, los resultados se correlacionaron con los resultados mostrados en la tabla 2.

Tabla 3. Termoestabilidad de las variantes de beta-glucosidasa a tiempos de incubación más prolongados

Mutaciones según sea apropiado	Otra ID según sea apropiado	%AR a 55°C durante 42 hr	%AR a 60°C durante 15 hr	%AR a 60°C durante 23,5 hr	%AR a 65°C durante 1 hr
A. niger WT	ANBG	12%	0%	0%	5%
A. oryzae WT	AOBG	69%	2%	0%	0%
H285Q	BG43	No realizada	9%	No realizada	1%
G161S; H285Q	BG41	88%	40%	35%	40%
A35T; G161S; H285Q; D722G	BG48	87%	41%	37%	48%

Ejemplo 14: Construcción del vector de expresión pAILo2

[0320] El vector de expresión pAILo1 fue construido modificando pBANE6 (patente U.S. 6.461.837), que comprende el promotor NA2-tpi, secuencia del terminador de amiloglucosidasa de *Aspergillus niger* (terminador AMG) y gen de acetamidasa de *Aspergillus nidulans* (*amdS*). La modificación de pBANE6 fue realizada eliminando primero tres sitios de restricción *Nco* I en las posiciones 2051, 2722 y 3397 bp del marcador de selección *amdS* por mutagénesis dirigida al sitio. Todos los cambios fueron diseñados para que fueran "silenciosos" dejando sin cambios la secuencia de la proteína real del producto genético *amdS*. La eliminación de estos tres sitios fue realizada simultáneamente con un equipo de mutagénesis dirigida GeneEditor (Promega, Madison, WI) según las instrucciones del fabricante usando los siguientes cebadores (el nucleótido subrayado representa la base cambiada):

AMDS3NcoMut (2050): 5'-GTGCCCCATGATACGCCTCCGG-3' (SEC ID n°: 58)

AMDS2NcoMut (2721): 5'-GAGTCGTATTTCCAAGGCTCCTGACC-3' (SEC ID n°: 59)

AMDS1NcoMut (3396): 5'-GGAGGCCATGAAGrGGACCAACGG-3' (SEC ID n°: 60)

[0321] Un plásmido que comprende los tres cambios de secuencia previstos fue luego sometido a mutagénesis dirigida, usando un equipo de mutagénesis 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) fueron usados para la mutagénesis:

Cebador superior para mutagenizar la secuencia del terminador de amiloglucosidasa (AMG) de *Aspergillus niger*:

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

Cebador inferior para mutagenizar la secuencia del terminador de amiloglucosidasa (AMG) de *Aspergillus niger*:

5'-CTGGTCTTCTACACGAAGGAAAGAGCATGGCTTACGGTGTCTG-3' (SEC ID n°: 62)

[0322] El último paso en la modificación de pBANE6 fue la adición de un nuevo sitio de restricción *Nco* I en el principio del poliligador usando un equipo de mutagénesis QuickChange y los siguientes cebadores (los nucleótidos subrayados representan las bases cambiadas) para producir pAILo1 (Figura 13).

Cebador superior para mutagenizar el promotor de amilasa de *Aspergillus niger* (NA2-tpi):
 5'-CTATATACACA ACTGGATTTACCATGGGCCCCGGCCGCAGATC-3' (SEC ID n°: 63)
 Cebador inferior para mutagenizar el promotor de amilasa de *Aspergillus niger* (NA2-tpi):
 5'-GATCTGCGGCCGCGGGCCCATGGTAAATCCAGTTGTGTATATAG-3' (SEC ID n°: 64)

5 [0323] El gen *amdS* de pAIlO1 fue cambiado con el gen *pyrG* de *Aspergillus nidulans*. El plásmido pBANE10 (Figura 14) fue usado como una fuente para el gen *pyrG* como un marcador de selección. El análisis de la secuencia de pBANE10 mostró que el marcador *pyrG* fue contenido dentro de un fragmento de restricción *Nsi* I y no contiene sitios de restricción *Nco* I ni *Pac* I. Ya que el *amdS* también es flanqueado por sitios de restricción *Nsi* I, la estrategia para alternar el
 10 marcador de selección fue un simple cambio de fragmentos de restricción *Nsi* I. El Adn plásmido de pAIlO1 y pBANE10 fue digerido con la enzima de restricción *Nsi* I y los productos fueron purificados por electroforesis en gel de agarosa. El fragmento de *Nsi* I de pBANE10 con el gen *pyrG* fue ligado a la estructura de pAIlO1 para reemplazar el fragmento de ADN *Nsi* I original con el gen *amdS*. Los clones recombinantes fueron analizados por digestión de restricción para determinar que tuvieran el inserto correcto y también su orientación. Se seleccionó un clon con el gen *pyrG* transcrito en
 15 el sentido contrario a las agujas del reloj. El plásmido nuevo fue designado pAIlO2 (Figura 15).

Ejemplo 15: Subclonación de la beta-glucosidasa termoestable mejorada BG41 y variantes BG48 en *Aspergillus oryzae* para caracterización de proteína

20 [0324] Las regiones de codificación de variantes de beta-glucosidasa BG41 y BG48 fueron subclonadas en el vector de *Aspergillus oryzae* pAIlO2 digerido con *Nco* I y *Pac* I para formar una unión perfecta con el ATG del gen y el promotor de amilasa de *Aspergillus niger* (NA2-tpi) y la secuencia del terminador de amiloglucosidasa de *Aspergillus niger*. Puesto que la región de codificación de beta-glucosidasa poseía dos sitios *Nco* I, la subclonación del gen de beta-
 25 glucosidasa en pAIlO2 fue realizada diseñando dos cebadores, mostrados más abajo, que expanden el gen de beta-glucosidasa en pSATE111 y también se aparean a pAIlO2 cerca de sitios *Nco* I y *Pac* I.

Aoryzaebeta-glucosidasaSuperior:

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

ACTGGATTTACCATG se aparea a pAIlO2 y AAGCTTGGTTGGATC se aparea a pSATE111.

Aoryzaebeta-glucosidasaInferior:

30 5'-AGTCACCTCTAGTTATTACTGGGCCTTAGG-3' (SEC ID n°: 66)

AGTCACCTCTAGTTA se aparea a pAIlO2 y TTAGCCTTAGG se aparea a pSATE111.

[0325] Para amplificar el ADN de variantes BG41 y BG48, 0,5 µl de cada molde de Adn plásmido fue usado en una
 35 reacción de 100 µl con tampón de reacción 1X ThermoPol con 2 mM de MgSO₄, 0,05 mM de cada dATP, dGTP, dCTP y dTTP, 50 pmol de cada cebador (cebadores Aoryzaebeta-glucosidasaSuperior y Aoryzaebeta-glucosidasaInferior) y 1 unidad de Vent ADN polimerasa. Dos reacciones de amplificación fueron realizadas donde la primera reacción fue sometida a 1 ciclo a 95°C durante 5 minutos; 25 ciclos cada uno a 95°C durante 1 minuto, 50°C durante 1 minuto, 72°C durante 3 minutos; y un ciclo de extensión final a 72°C durante 10 minutos, y la segunda reacción fue realizada bajo las
 40 mismas condiciones pero a una temperatura de apareamiento de 56°C. Una alícuota de cada producto de PCR fue realizada en un gel de agarosa al 0,7% usando tampón TAE, tal y como se describe anteriormente, generando bandas previstas de aproximadamente 3 kb. Ambas reacciones de la PCR fueron combinadas y fueron purificadas usando un equipo de purificación MinElute PCR y ADN y con elución del ADN en 10 µl de tampón EB. El rendimiento de cada producto de PCR purificado fue estimado en 250 ng por microlitro por visualización en un gel de agarosa al 0,7% usando tampón TAE.

45 [0326] El plásmido pAIlO2 digerido con *Pac* I y con extremo romo en el sitio *Nco* I fue concentrado por precipitación con 0,1 de volumen de 3 M de acetato sódico pH 5,0 y 2 volúmenes de 95% de etanol durante toda la noche a -20°C. El plásmido precipitado fue centrifugado a 13.200 x g usando una microcentrifugadora durante 15 minutos. El sobrenadante fue eliminado y el granulado fue lavado con 1 ml de etanol al 70%. El granulado precipitado fue centrifugado otra vez a 13.200 x g durante 15 minutos, seguido de eliminación del sobrenadante, secado del granulado al vacío y resuspensión en 20 µl de agua. El plásmido tuvo una concentración de 80 ng por microlitro. La concentración fue verificada por visualización en un gel de agarosa al 0,7% usando tampón TAE. La clonación del producto de PCR de beta-glucosidasa anteriormente descrita y el vector pAIlO2 digerido fue realizada usando un equipo de clonación de PCR BD In-Fusion (Stratagene, La Jolla, CA).

50 [0327] Los vectores de expresión resultantes con las regiones de codificación de la variante BG41 o BG48 fueron designados pALFd3BG41 (Figura 16) y pALFd3BG48 (Figura 17), respectivamente.

Ejemplo 16: Expresión de variantes de beta-glucosidasa BG41 y BG48 en *Aspergillus oryzae*

60 [0328] Aproximadamente 4,5 µg de pALFd3BG41 y 6,25 µg de Adn plásmido pALFd3BG48 fueron usados para transformar independientemente los protoplastos Jal250 de *Aspergillus oryzae*. Los protoplastos Jal250 de *Aspergillus oryzae* fueron preparados según el método de Christensen et al., Bio/Technology 6: 1419-1422.

65 [0329] La transformación de *Aspergillus oryzae* Jal250 con pALFd3BG41 liberó 5 transformantes independientes, mientras que la transformación con pALFd3BG48 liberó 15 transformantes independientes, donde 9 de estos fueron

además subcultivados. Cuatro días después, cada transformante independiente ha sido transferido a placas de selección medias mínimas de 100 mm, las esporas fueron transferidas de las placas de selección a placas de 24 pocillos con M400 de medio diluido 1 a 5 con 1X BS e incubadas a 34°C. Siete días después de la incubación, 10 µl de cada sobrenadante fueron analizados usando geles de SDS-PAGE al 8-16% (Invitrogen, Carlsbad, CA) según las instrucciones del fabricante. Los perfiles de SDS-PAGE de los cultivos de los cuatro transformantes pALFd3BG41 mostraron una banda mayoritaria de aproximadamente 120 kDa que correspondió al peso molecular de beta-glucosidasa de *Aspergillus oryzae*. Los perfiles de SDS-PAGE de los cultivos de los seis transformantes pALFd3BG4 también mostraron una banda mayoritaria de aproximadamente 120 kDa.

10 **Ejemplo 17: Determinaciones de termoestabilidad de las variantes de beta-glucosidasa BG41 y BG48 con celobiosa**

15 [0330] La termoestabilidad de las variantes de beta-glucosidasa BG41 y BG48 (caldos de fermentación no purificados) fue determinada incubando el caldo con 10 mM de celobiosa en 100 mM de tampón de citrato sódico con Tween-20 al 0,01% a pH 5,0 hasta 21 horas a 65°C.

20 [0331] Los resultados de la determinación de termoestabilidad de variantes de beta-glucosidasa BG41 y BG48, como se muestra en la Figura 18, demostraron que las variantes fueron significativamente más estables con el paso del tiempo que la beta-glucosidasa de *Aspergillus niger* (Novozym 188) o la beta-glucosidasa de *Aspergillus oryzae*.

20 **Ejemplo 18: Construcción de una variante de beta-glucosidasa con la sustitución de G161S**

25 [0332] La región de nucleótidos conteniendo la mutación de G a A para producir la sustitución de G161S fue subclonada en la secuencia codificante de pSATE111 para posterior caracterización de los efectos de esta sustitución de este aminoácido único. La mutación G a A fue localizada entre los sitios únicos *Spe I* y *Bpu 1102 I* del vector pSATE111. Los plásmidos pSATE111 (aproximadamente 50 µg) y pSATE111BG41 (aproximadamente 300 µg) fueron digeridos con *Spe I* y *Blp I*, un isoesquizómero de *Bpu 1102 I*. Las reacciones produjeron dos fragmentos: una con la mayor parte de pSATE111 (8146 bp) y un fragmento más pequeño de 580 bp de la secuencia codificante de la beta-glucosidasa con la mutación G a A. pSATE111 digerido fue tratado con fosfatasa alcalina de gamba para la desfosforilación de los productos de ADN digeridos añadiendo tampón de SAP 1X y 2 µl de SAP (Roche Applied Science, Manheim, Alemania) e incubando la reacción durante 10 minutos a 37°C seguido de incubación a 85°C durante 10 minutos para la inactivación enzimática. Ambas digestiones fueron realizadas en gel de agarosa al 0,7% y purificadas usando un equipo de purificación de gel QIAGEN según las instrucciones del fabricante.

35 [0333] pSATE111 digerido fue ligado al fragmento de 580 bp de la digestión de pATE111BG41 con la posición que codificaba el aminoácido 161 con la mutación de nucleótido que condujo a la sustitución del aminoácido G161S. La ligación fue realizada usando el Rapid DNA Ligation Kit (Applied Science, Manheim, Manheim, Alemania) siguiendo las instrucciones del fabricante.

40 [0334] La reacción de la ligación fue transformada en células competentes de subclonación XL1-Blue *E. coli* según las instrucciones del fabricante (Stratagene, La Jolla, CA). Después de la transformación, el Adn plásmido de una colonia aislada fue aislado como se describe en el Ejemplo 10 y secuenciadas, confirmando la presencia de la única mutación de G a A en la región de codificación entera del gen. Posteriormente, el Adn plásmido fue transformado en células competentes de levadura como se describe en el Ejemplo 8, que dio como resultado el aislamiento de una variante de beta-glucosidasa con la sustitución de G161S. Esta variante de beta-glucosidasa fue designada el mutante G161S. Puesto que BG43 contenía sólo la sustitución de H285Q y el mutante G161S contenía sólo la sustitución de G161S, los efectos de cada sustitución, G161S y H285Q, pudieron ser caracterizados individualmente.

50 **Ejemplo 19: Ensayo de los efectos de las mutaciones G161S y H285Q**

55 [0335] Una prueba de termoestabilidad fue realizada a 60°C durante 23 horas demostrando el efecto sinérgico de las mutaciones. Las muestras de las variantes de beta-glucosidasa fueron diluidas en el mismo tampón que el usado en el Ejemplo 9 con las mismas actividades enzimáticas en relación de una con la otra. Cada muestra fue dividida en dos tubos de ensayo de polipropileno, un tubo de las muestras fue incubado sumergido en un baño María a temperatura controlada a 60°C y la otra parte fue incubada a temperatura ambiente, ambas durante un periodo de 42 horas como máximo. Al final del periodo de incubación, las muestras de ambos fueron colocadas en una placa de 96 pocillos. El sustrato de metilumbeliferil-beta-D-glucopiranosido (MUG) (200 µl de 0.5 µM de solución de MUG) fue añadido a la placa de 96 pocillos de muestras e incubado a temperatura ambiente durante 15 minutos. La reacción fue detenida con la adición de tampón de 2 M tris, pH 9,0 y la placa fue leída en un fluorómetro para obtener las Unidades Fluorescentes Relativas (UFR) con excitación 365, emisión 454. La actividad residual porcentual fue determinada usando el mismo método descrito en el Ejemplo 9.

60 [0336] Los efectos de las mutaciones G161S y H285Q individualmente y combinadas en una molécula se muestran en la Figura 19. El reposo individual de estas mutaciones mostró que la combinación de estos en una molécula tiene un efecto superior en la termoestabilidad de la actividad de la beta-glucosidasa en tampón que cualquiera de ellos de

manera individual como se muestra en la Figura 19. La barra es el compuesto matemático de estas dos sustituciones de aminoácido.

Ejemplo 20: Identificación de un gen de la familia GH3A de glicosil hidrolasa en la secuencia genómica de *Aspergillus fumigatus*

[0337] Una búsqueda de TblastN (Altschul et al., 1997, Nucleic Acids Res. 25: 3389-3402) de la secuencia de genoma parcial de *Aspergillus fumigatus* (The Institute for Genomic Research, Rockville, MD) se efectuó usando como secuencia de búsqueda una secuencia de proteína de beta-glucosidasa de *Aspergillus aculeatus* (Nº de acceso P48825). Diferentes genes fueron identificados como homólogos putativos de familia GH3A basado en un alto grado de similitud a la secuencia de búsqueda en el nivel de aminoácidos. Una región genómica de aproximadamente 3000 bp con identidad mayor que 70% a la secuencia de búsqueda en el nivel de aminoácidos fue elegida para estudio posterior.

Ejemplo 21: Extracción de ADN genómico de *Aspergillus fumigatus*

[0338] *Aspergillus fumigatus* fue cultivado en 250 ml de medio de dextrosa de patata en un matraz oscilante disipado a 37°C y 240 r.p.m. Las micelas fueron recogidas por filtración, lavadas dos veces en tampón TE (10 mM de Tris-1 mM de EDTA), y congeladas bajo nitrógeno líquido. Los micelios congelados fueron molidos por mortero y mano de mortero a un polvo fino, que fue resuspendido en tampón de pH 8,0 con 10 mM de tris, 100 mM de EDTA, Tritón X-100 al 1%, 0,5 M de HCl de guanidina y 200 mM de NaCl. Se agregó RNasa libre de ADNsa a una concentración de 20 mg/litro y el lisado fue incubado a 37°C durante 30 minutos. El detrito celular fue eliminado por centrifugado, y el ADN fue aislado usando una columna Qiagen Maxi 500 (QIAGEN Inc., Chatsworth, CA). Las columnas fueron equilibradas en 10 ml de QBT lavadas con 30 ml de QC, y eluidas con 15 ml de QF (todos los tampones de QIAGEN Inc., Chatsworth, CA). El ADN fue precipitado en isopropanol, lavado en etanol al 70% y recuperado por centrifugado. El ADN fue resuspendido en tampón TE.

Ejemplo 22: Clonación del gen de beta-glucosidasa de la Familia GH3A y construcción de un vector de expresión de *Aspergillus oryzae*

[0339] Dos cebadores oligonucleótidos sintéticos mostrados más abajo fueron diseñados para amplificar por PCR un gen de *Aspergillus fumigatus* que codifica una beta-glucosidasa putativa de la familia GH3A del ADN genómico preparado en el Ejemplo 21. Un equipo de clonación InFusion (BD Biosciences, Palo Alto, CA) fue usado para clonar el fragmento directamente en el vector de expresión, pAIIo2, sin la necesidad de digeridos de restricción y ligación.

Cebador directo: 5'-ACTGGATTTACCATGAGATTTCGGTTGGCTCG-3' (SEC ID nº: 67)

Cebador inverso: 5'-AGTCACCTCTAGTTACTAGTAGACACGGGGC-3' (SEC ID nº: 68)

Las letras en negrita representan la secuencia codificante. La secuencia restante es homóloga a los sitios de inserción de pAIIo2.

[0340] Cincuenta picomoles de cada uno de los cebadores mencionados fueron usados en una reacción de 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 polimerasa de ADN Platinum Pfx, 1 µl de 50 mM de MgSO₄ y 2,5 µl de solución intensificadora 10X pCRx (Invitrogen, Carlsbad, CA) 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 conducido a un ciclo de remojo de 4°C.

[0341] Los productos reactivos fueron aislados en un gel de agarosa al 1,0% usando tampón TAE donde una banda de producto de 3 kb fue cortada del gel y purificada usando un equipo de extracción de gel QIAquick según las instrucciones del fabricante.

[0342] El fragmento luego fue clonado en el vector de expresión pAIIo2 usando un equipo de clonación Infusion. El vector fue digerido con Nco I y Pac I. El fragmento fue purificado por electroforesis de gel y purificación de gel QIAquick. El fragmento del gen y el vector digerido fueron ligados en una reacción dando como resultado el plásmido de expresión pEJG97 (Figura 20) en la cual la transcripción del gen de beta-glucosidasa de la familia GH3A estuvo bajo el control del promotor NA2-tpi. La reacción de ligación (50 µl) fue compuesta por tampón de InFusion 1X (BD Biosciences, Palo Alto, CA), 1X BSA (BD Biosciences, Palo Alto, CA), 1X BSA (BD Biosciences, Palo Alto, CA), 1 µl de enzima de InFusion (diluico 1:10) (BD Biosciences, Palo Alto, CA), 150 ng de pAIIo2 digerido con Nco I y Pac I, y 50 ng del producto de PCR purificado de beta-glucosidasa de *Aspergillus fumigatus*. La reacción fue incubada a temperatura ambiente durante 30 minutos. Un µl de la reacción fue usado para transformar las células de *E. coli* XL10 Solopac Gold (Stratagene, La Jolla, CA). Un transformante de *E. coli* conteniendo el plásmido pEJG97 fue detectado por digestión de restricción del Adn plásmido.

Ejemplo 23: Caracterización de la secuencia genómica de *Aspergillus fumigatus* que codifica una beta-glucosidasa de la familia GH3A

[0343] La secuenciación del ADN del gen de beta-glucosidasa de *Aspergillus fumigatus* de pEJG97 fue realizada como se describe anteriormente usando una estrategia de desplazamiento del cebador. Un modelo de gen para la secuencia

de *Aspergillus fumigatus* fue construida basado en la similitud con genes homólogos de *Aspergillus aculeatus*, *Aspergillus niger* y *Aspergillus kawachii*. La secuencia de nucleótidos (SEC ID n°: 69) y la secuencia de aminoácidos deducida (SEC ID n°: 70) del gen de beta-glucosidasa de *Aspergillus fumigatus* se muestran en la Figura 21. El fragmento genómico codificó un polipéptido de 863 aminoácidos, escindido por 8 intrones de 62, 55, 58, 63, 58, 58, 63 y 51 bp. El contenido %G+C del gen es 54,3%. Usando el programa de software SignalP (Nielsen et al., 1997, Protein Engineering 10: 1-6), se predijo un péptido señal de 19 residuos. La proteína madura predicha contiene 844 aminoácidos con una masa molecular de 91,7 kDa.

[0344] Una alineación comparativa de las secuencias de beta-glucosidasa fue determinada 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 alineación múltiple: Penalización de espacio de 10 y penalización de longitud de espacio de 10. Los parámetros de alineación en pareja fueron Ktuple=1, penalización de espacio=3, ventanas=5, y diagonales=5. La alineación mostró que la secuencia de aminoácidos deducida del gen de beta-glucosidasa de *Aspergillus fumigatus* compartió 78%, 76% y 76% de identidad de las secuencias de aminoácidos deducidas de beta-glucosidasa de *Aspergillus aculeatus* (número de registro P48825), *Aspergillus niger* (número de registro 000089) y *Aspergillus kawachii* (número de registro P87076).

Ejemplo 24: Expresión del gen de beta-glucosidasa de *Aspergillus fumigatus* de la familia GH3A en *Aspergillus oryzae* JAL250

[0345] Los protoplastos de *Aspergillus oryzae* JAL250 fueron preparados según el método de Christensen et al., 1988, Bio/Technology 6: 1419-1422. Cinco µg de pEJG97 (así como pALo2 como un control de vector) fue usado para transformar *Aspergillus oryzae* JAL250.

[0346] La transformación de *Aspergillus oryzae* JAL250 con pEJG97 produjo aproximadamente 100 transformantes. Diez transformantes fueron aislados en placas de PDA individuales.

[0347] La placas de PDA confluentes de cinco de los diez transformantes fueron lavadas con 5 ml de Tween 20 al 0,01% e inoculadas separadamente en 25 ml de medio MDU2BP en frascos de agitación de vidrio de 125 ml e incubadas a 34°C, 250 r.p.m. Cinco días después de la incubación, se analizaron 0,5 µl de sobrenadante de cada cultivo usando geles de SDS-PAGE de Tris-glicina al 8-16% (Invitrogen, Carlsbad, CA) según las instrucciones del fabricante. Los perfiles de SDS-PAGE de los cultivos mostraron que uno de los transformantes (designado transformante 1) tuvo una banda mayoritaria de aproximadamente 130 kDa.

Ejemplo 25: Extracción de ARN total de *Aspergillus oryzae*

[0348] El transformante de *Aspergillus oryzae* descrito en el Ejemplo 22 fue congelado con nitrógeno líquido y fue almacenado a -80°C. Posteriormente, el tejido congelado fue molido en un molinillo de café eléctrico con unos trozos de hielo seco adicionado para tener los micelios en polvo congelados. Luego, el material molido fue transferido con una espátula a un tubo cónico estéril de 50 ml que fue previamente llenado con 20 ml de Fenozol (Active Motif, Inc., Carlsbad, CA). Esta mezcla fue mezclada rápidamente para disolver el material congelado a una solución gruesa, y fue colocada en un baño María a 50°C durante 15 minutos. Cinco ml de cloroformo libre de RNasa se añadió a la mezcla y se removió enérgicamente. Luego, se dejó que la mezcla permaneciera a temperatura ambiente durante 10 minutos. Luego, la mezcla fue centrifugada a 2700 r.p.m. en un centrifugador Sorvall RT7 (Sorvall, Inc, Newtown, CT) a temperatura ambiente durante 20 minutos. La fase superior fue transferida a un tubo cónico nuevo y se añadió un volumen igual de fenol-cloroformo-isoamil alcohol (25:24:1). La mezcla fue removida y centrifugada durante 10 minutos. Este procedimiento fue repetido dos veces de modo que se realizaron tres extracciones de fenol-cloroformo isoamil alcohol. Luego, la fase superior fue transferida a un tubo nuevo y se añadió un volumen igual de cloroformo:isoamil alcohol (24:1). La mezcla fue removida otra vez y centrifugada durante 10 minutos. Después del centrifugado, la fase acuosa, aproximadamente 5 ml en este punto, fue transferida a un tubo Oak Ridge nuevo y se añadieron 0,5 ml de 3 M de acetato sódico pH 5,2 y 6,25 ml de isopropanol. La mezcla fue mezclada e incubada a temperatura ambiente durante 15 minutos. Posteriormente, la mezcla fue centrifugada a 12.000 x g durante 30 minutos, a 4°C en un Sorvall RC5B (Sorvall, Inc, Newtown, CT). Después del centrifugado, el sobrenadante fue eliminado y se añadieron cuidadosamente 18 ml de etanol al 70% al granulado. Se realizó otra fase de centrifugado durante 10 minutos a 4°C a 12.000 x g. El sobrenadante fue cuidadosamente eliminado y el granulado fue secado al aire. El granulado de ARN fue resuspendido en 500 µl de agua tratada con pirocarbonato de dietilo (DEPC). En este punto, el calentamiento a 65°C durante 10 minutos ayudó en la resuspensión. El ARN total fue almacenado a -80°C. Se realizó la cuantificación y la evaluación de la calidad del ARN en un Bioanalizador Agilent 2100 (Englewood; Co) usando chips de ARN. Todos los materiales y los reactivos usados en este protocolo estaban libres de ARNasa.

Ejemplo 26: Clonación de la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus*

[0349] El ARN total extraído del transformante de *Aspergillus oryzae* con pEJG97 fue usado para clonar la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus* (SEC ID n°: 71 para la secuencia de ADNc y SEC ID n°: 70 para la secuencia de aminoácidos deducida). El ARNm del ARN total fue purificado usando el equipo Poly(A) Purist Mag (Ambion, Inc., Austin, TX) siguiendo las instrucciones del fabricante. La secuencia de ADNc de beta-glucosidasa de

- Aspergillus fumigatus*, luego fue amplificada en dos fragmentos: un fragmento de 1.337 bp que se extiende del codón de iniciación ATG a la posición 1.332 (marcada como fragmento 5') y un segundo fragmento de ADN de 1.300 bp (marcado como fragmento 3') que se extiende de la posición 1.303 hasta el codón de detención usando el sistema ProStar UltraHF RT-PCR (Stratagene), siguiendo el protocolo del fabricante para una reacción de 50 µl usando 200 ng de ARNm poli-A, los cebadores Afuma (sentido) y Afumc (antisentido) para el fragmento 5' y cebadores Afumd (sentido) y Afumb (antisentido) para el fragmento 3' como se muestra a continuación:
- 5 Afuma: 5'- GGCTCATGAGATTCGGTTGGCTCGAGGTC -3' (SEC ID nº: 72)
 Afumc: 5'- GCCGTTATCACAGCCGCGGTCGGGGCAGCC-3' (SEC ID nº: 73)
 Afumd: 5'- GGCTGCCCGACCGCGGCTGTGATAACGGC-3' (SEC ID nº: 74)
 10 Afumb: 5'- GCTTAATTAATCTAGTAGACACGGGGCAGAGGCGC -3' (SEC ID nº: 75)
 El cebador Afuma tiene un sitio corriente arriba *Bsp* HI y el cebador Afumb tiene un sitio corriente abajo *Pacl*. Veintinueve nucleótidos en el extremo 3' del fragmento 1.337 se superpusieron con el extremo 5' del fragmento 1.303. En la región de superposición hubo un sitio *Sac* II único.
- 15 [0350] Ambos fragmentos fueron subclonados individualmente en el vector pCR4Blunt-TOPO (Invitrogen, Carlsbad, CA) usando el equipo de clonación de PCR Zero Blunt TOPO para secuenciación (Invitrogen, Carlsbad, CA), siguiendo el protocolo de fabricante, generando plásmidos de pCR4Blunt-TOPOAfcDNA5' y pCR4Blunt-TOPOAfcDNA3', con los fragmentos 5' y 3' respectivamente (Figuras 22 y 23).
- 20 [0351] La región de codificación entera de ambos fragmentos de beta-glucosidasa de *Aspergillus fumigatus* fue confirmada por secuenciación usando 0,5 µl de cada Adn plásmido y 3,2 pmol de los siguientes cebadores:
 BGLU1.dir.: 5'-ACACTGGCGGAGAAGG-3' (SEC ID nº: 76)
 BGLU2.dir.: 5'-GCCCAGGGATATGGTTAC-3' (SEC ID nº: 77)
 BGLU3.dir.: 5'-CGACTCTGGAGAGGGTTTC-3' (SEC ID nº: 78)
 25 BGLU4.inv.: 5'-GGACTGGGTCATCACAAG -3' (SEC ID nº: 79)
 BGLU5.inv.: 5'-GCGAGAGGTCATCAGCA -3' (SEC ID nº: 80)
 M13 directo: 5'-GTAAAACGACGGCCAGT-3' (SEC ID nº: 81)
 M13 inverso: 5'-CAGGAAACAGCTATGA-3' (SEC ID nº: 82)
- 30 [0352] Los resultados de la secuenciación indicaron la presencia de cambios de nucleótidos diferentes cuando se compara la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus* obtenida con la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus* deducida de los datos de genoma del Instituto para la Investigación Genómica (Rockville, MD). En la posición 500, T fue sustituido por C, de modo que la secuencia codificante GTT fue cambiada a GCT, de modo que la valina fue sustituida por alanina. En la posición 903, T fue sustituido por C, de modo que la secuencia codificante CCC fue cambiada a CCT, no obstante, este cambio fue silencioso. En la posición 2.191, G fue sustituido por C, de modo que la secuencia codificante CAG fue cambiada a GAG, de modo que el ácido glutámico fue sustituido por glutamina. Finalmente, en la posición 2.368, C fue sustituido por T, de modo que la secuencia codificante CTG fue cambiada a TTG, no obstante, este cambio también fue silencioso.
- 35 [0353] Una vez que los dos fragmentos han sido secuenciados, ambos clones con cada fragmento fueron digeridos con *Sac* II y *Pme* I usando aproximadamente 9 µg de cada Adn plásmido. La digestión de pCR4Blunt-TOPOAfcDNA5' con las enzimas mencionadas generaron un fragmento de 3.956 bp (con la mayor parte del vector) y un segundo fragmento de 1.339 bp (con el fragmento de 5' ADNc de beta-glucosidasa de *Aspergillus fumigatus*). La digestión del vector pCR4Blunt-TOPOAfcDNA3' con las mismas enzimas generaron un fragmento de 5.227 bp (con la mayor parte del vector pCR4Blunt-TOPO y el fragmento 3' de ADNc de beta-glucosidasa de *Aspergillus fumigatus*) y un segundo fragmento de 31 bp. El pCR4Blunt-TOPOAfcDNA3' digerido fue tratado con fosfatasa alcalina de gamba para la desfosforilación de los productos de ADN digeridos añadiendo tampón de SAP 1X y 1 µl de fosfatasa alcalina de gamba (Roche Applied Science, Mannheim, Alemania) e incubando la reacción durante 10 minutos a 37°C seguido de incubación a 85°C durante 10 minutos para la inactivación enzimática. Ambas digestiones realizadas en gel de agarosa al 0,7% usando tampón TAE y purificadas usando un equipo de purificación de gel QIAGEN según las instrucciones del fabricante.
- 40 [0354] La banda de 1.339 bp generada de la digestión de pCR4Blunt-TOPOAfcDNA5' y el fragmento de 5.227 bp generado de la digestión de pCR4Blunt-TOPOAfcDNA3' fueron ligados usando el Radip DNA Ligation kit siguiendo las instrucciones del fabricante. La reacción de ligación fue transformada en células competentes de subclonación de XL1-Blue de *E. coli* según las instrucciones del fabricante. Después de la transformación, el Adn plásmido de una colonia aislada fue secuenciado para confirmar que los fragmentos 5' y 3' del ADNc de beta-glucosidasa de *Aspergillus fumigatus* fueron subclonados en serie generando un vector pCR4Blunt-TOPOAfcDNA de 6.566 bp (Figura 24).
- 45 **Ejemplo 27: Construcción de los vectores de expresión de *Sacharomyces cerevisiae* pALFd6 y pALFd7**
- 50 [0355] El ADNc de beta-glucosidasa de *Aspergillus fumigatus* de longitud total fue amplificada por PCR usando los siguientes cebadores que tienen homología con el vector pCU426 y las secuencias 5' y 3' del ADNc de beta-glucosidasa de *Aspergillus fumigatus* como se indica:
 60 AfumigatusBGSuperior:

5'-CTTCTGTAGTGCAATATCATATAGAAGTCATCGACTAGTGATCTACCATGAGATTCGGTTGGCTCG-3' (SEC ID n°: 83)

ATGAGATTCGGTTGGCTCG tiene homología con el extremo 5' del cDNA de *Aspergillus fumigatus*

AfumigatusBGIinferior:

5'-GCGTGAATGTAAGCGTGACATAACTAATTACATGACTCGAGCTAGTAGACACGGGGCAGAG-3' (SEC ID n°: 84)

CTAGTAGACACGGGGCAGAG tiene homología con el extremo 3' del ADNc de *Aspergillus fumigatus*

[0356] La reacción de amplificación (100 µl) fue compuesta por 0,5 µl del plásmido pCR4Blunt-TOPOAfcDNA con la secuencia de ADNc de *Aspergillus fumigatus*, tampón de amplificación 1X Pfx, 50 µM cada uno de dATP, dCTP, dGTP y dTTP, 50 pmol de cada uno de los cebadores mencionados, 1,5 mM de MgSO₄ y 2,5 unidades de ADN polimerasa Pfx Platinum. Las reacciones fueron incubadas en un RoboCycler Gradient 40 programado durante 1 ciclo a 95°C durante 5 minutos; 25 ciclos cada uno a 95°C durante 1 minuto, 50°C durante 1 minuto y 72°C durante 3 minutos; y un ciclo de extensión final a 72°C durante 10 minutos. La reacción de PCR fue purificada usando un equipo de purificación de PCR QIAquick. El ADN fue eluido en 30 µl de tampón EB. El producto de PCR tuvo 37 bp de la secuencia de ADN homóloga que fue mezclado con 1 µl del vector pCU426 escindido con *Spe* I y *Xho* I para la cotransformación en células competentes de *Saccharomyces cerevisiae* YNG318 como se describe en el Ejemplo 8. Estas colonias no se volvieron azules como se preveía, sugiriendo algún error de secuenciación en la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus*. La secuenciación adicional de la secuencia de ADNc de *Aspergillus fumigatus* indicó una inserción de un nucleótido extra en la secuencia de ADNc, que interrumpió el marco de lectura abierto de la enzima.

[0357] Simultáneamente a la expresión del ADNc de beta-glucosidasa de *Aspergillus fumigatus* en *Saccharomyces cerevisiae*, la secuencia señal de endoglucanasa V de *Humicola insolens* fue cambiada con la secuencia señal nativa de la secuencia de ADNc de *Aspergillus fumigatus* también para la expresión en *Saccharomyces cerevisiae* para comparar la expresión de beta-glucosidasa de *Aspergillus fumigatus* con cada secuencia señal. La secuencia de ADNc de *Aspergillus fumigatus* fue amplificada por PCR con un cebador que tiene homología con la secuencia señal de endoglucanasa V de *Humicola insolens* en el vector pALFd1 y homología con el extremo 5' de la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus* madura. Los cebadores usados para la amplificación de la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus* son el cebador AfumigatusBGIinferior descrito anteriormente y el cebador HiEGVAfumigatus descrito a continuación:

HiEGVAfumigatus:

5'-CCGCTCCGCCGTTGTGGCCGCCCTGCCGGTGTGGCCCTTGCC**GAATTGGCTTTCTCT**-3' (SEC ID n°: 85)

GAATTGGCTTTCTCTCC tiene homología con el extremo 5' de la secuencia madura de *Aspergillus fumigatus*.

[0358] La reacción de amplificación (100 µl) fue compuesta por 0,5 µl de pCR4Blunt-TOPOAfcDNA, tampón de amplificación 1 X Pfx, 50 µM cada uno de dATP, dCTR, dGTP y dTTP, 50 pmol de cada cebador mencionado, 1,5 mM de MgSO₄ y 2, 5 unidades de ADN-polimerasa Pfx Platinum. Las reacciones fueron incubadas en un RoboCycler Gradient 40 programado durante 1 ciclo a 95°C durante 5 minutos; 25 ciclos cada uno a 95°C durante 1 minuto, 50°C durante 1 minuto y 72°C durante 3 minutos; y un ciclo de extensión final a 72°C durante 10 minutos. La reacción de PCR fue purificada usando un equipo de purificación de PCR QIAquick. El ADN fue eluido en 10 µl de tampón EB. Tres µl del producto de PCR purificado fue mezclado con 1,8 µl del vector pALFd1 escindido con *Eco* RI y *Xho* I para cotransformación en células competentes de *Saccharomyces cerevisiae* YNG318 como se describe en el Ejemplo 8. Estas colonias se volvieron de color azul claro. No obstante, una colonia se destacó en un azul oscuro. El rescate del ADN de esta colonia fue realizado como se describe en el Ejemplo 10 y el plásmido fue transformado en células competentes de electroporación de *E. coli* SURE (Stratagene, La Jolla, CA) para secuenciación. La secuenciación de longitud completa indicó que la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus* fue correcta. Este plásmido fue designado pALFd7 (Figura 25), que contuvo la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus* con secuencia señal de endoglucanasa V de *Humicola insolens* para la expresión de levadura.

[0359] Para producir un vector de expresión de levadura que contiene la secuencia de ADNc de *Aspergillus fumigatus* correcta con su secuencia señal nativa, la región que contiene la secuencia de nucleótidos correcta del vector de expresión de levadura que contiene la secuencia de ADNc de *Aspergillus fumigatus* con la secuencia señal de endoglucanasa V de *Humicola insolens* (pALFd7) fue amplificada por PCR usando el cebador BGLU.5inv y el siguiente cebador:

BGL7dir.: 5'-CTGGCGTTGGCGCTGTC-3' (SEC ID n°: 86)

[0360] La reacción de amplificación (100 µl) fue compuesta por 0,5 µl de pALFd7, tampón de amplificación 1X Pfx, 50 µM cada uno de dATP, dCTP, dGTP y dTTP, 50 pmol de cada cebador mencionado, 1,5 mM de MgSO₄ y 2,5 unidades de ADN-polimerasa Pfx Platinum. Las reacciones fueron incubadas en un RoboCycler Gradient 40 programado durante 1 ciclo a 95°C durante 5 minutos; 25 ciclos cada uno a 95°C durante 1 minuto, 50°C durante 1 minuto y 72°C durante 1 minuto; y un ciclo de extensión final a 72°C durante 10 minutos.

[0361] El fragmento de PCR de 701 bp fue purificado usando un equipo de purificación de PCR QIAquick. El ADN fue eluido en 10 µl de tampón EB. Tres µl del producto de PCR purificado fue mezclado con 3 µl del vector de expresión de levadura con secuencia de ADNc de *Aspergillus fumigatus* con la secuencia señal nativa y el nucleótido extra escindido con el vector *SacII* y *XmaI* para cotransformación en células competentes de *Saccharomyces cerevisiae* YNG318 como se describe como en el Ejemplo 8. Estas colonias se volvieron azules. El Adn plásmido fue rescatado de una colonia

azul escogida de forma aleatoria como se describe en el Ejemplo 10, y transformado en células competentes de electroporación de *E. coli* SURE (Stratagene, La Jolla, CA) para secuenciación. La secuenciación de longitud completa indicó que la secuencia de ADNc de beta-glucosidasa de *Aspergillus fumigatus* fue correcta. Este vector de expresión de levadura fue designado pALFd6 (Figura 26), el cual contuvo la secuencia de ADNc de *Aspergillus fumigatus* con su secuencia señal nativa.

Ejemplo 28: Construcción de una variante de beta-glucosidasa de *Aspergillus fumigatus* con sustitución de G161S

[0362] La región de nucleótidos que codificó el aminoácido G161 en la secuencia codificante de *Aspergillus fumigatus* pEJG97AfumFAM3A fue mutagenizada para producir una mutación de G a A que codificó la sustitución G161S presente en la variante BG53 de la secuencia de beta-glucosidasa de *Aspergillus oryzae*. La mutagénesis fue realizada usando el equipo de mutagénesis dirigida al sitio QuickChange (Stratagene, La Jolla, CA) y los siguientes cebadores (los nucleótidos subrayados representan el codón con la base cambiada que codificó la nueva sustitución de aminoácidos):

SDMG142SSuperior:
5'-GCGGCAGAATCTGGGAAAGCTTCTCTCCTG-3' (SEC ID n°: 87)

SDMG142SInferior:
5'-CAGGAGAGAAGCTTCCAGATTCTGCCGC-3' (SEC ID n°: 88)

La presencia de la mutación G a A fue confirmada por secuenciación. El vector nuevo fue designado pE7G97AfumFAM3AG142S (Figura 27). La mutación G a A fue localizada entre el fragmento de 231 bp obtenido por digestión del vector pEJG97AfumFAM3AG142S con *Eco* RI y *Bst* XI. Estos sitios fueron únicos en el vector pALFd7.

[0363] Tanto pEJG97AfumFAM3AG142S como pALFd7 (aproximadamente 6 µg cada uno) fueron digeridos con *Eco* RI y *Bst* XI. La reacción con pALFd7 produjo dos fragmentos, un con la mayor parte del vector (8504 bp) y un fragmento más pequeño de 231 bp de la secuencia codificante de beta-glucosidasa con la ubicación de tipo salvaje donde la mutación G a A fue creada en el vector pEJG97AfumFAM3AG142S. La reacción con pEJG97AfumFAM3AG142S produjo tres fragmentos, uno con la mayor parte del vector (7351 bp), un segundo fragmento más pequeño de 1254 bp y un fragmento más pequeño de 231 bp de la secuencia codificante de beta-glucosidasa con la mutación G a A. El pALFd7 digerido con fosfatasa alcalina de gamba para desfosforilación de los productos de ADN digeridos añadiendo tampón de SAP 1X y 1 ul de SAP de Roche (Roche Applied Science, Manheim, Alemania) e incubando la reacción durante 10 minutos a 37°C después de la incubación a 85°C durante 10 minutos para inactivación enzimática. Ambas digestiones fueron realizadas en gel de agarosa al 0,7% y purificadas usando un equipo de purificación de gel QIAGEN según las instrucciones del fabricante.

[0364] El pALFd7 digerido fue ligado al fragmento de 231 bp de la digestión de pEJG97AfumFAM3AG142S con la posición que codificó el aminoácido 161 con la mutación de nucleótidos que condujo a la sustitución del aminoácido G161S. La ligación fue realizada usando el Rapid DNA Ligation kit siguiendo las instrucciones del fabricante.

[0365] La reacción de ligación fue transformada en células competentes de subclonación XL1-Blue *E. coli* según las instrucciones del fabricante. Después de la transformación, el Adn plásmido de una colonia aislada fue secuenciado como se describe en el Ejemplo 10 y se confirmó la presencia de la única mutación de G a A en la región de codificación entera del gen. Posteriormente, el Adn plásmido fue transformado en células competentes de levadura como se describe en el Ejemplo 8, que dio como resultado el aislamiento de una variante de beta-glucosidasa de *Aspergillus fumigatus* con la sustitución de G161S. El vector de expresión de levadura con la variante de beta-glucosidasa de *Aspergillus fumigatus* con la sustitución de G161S fue designado pALFd7G142S (Figura 28).

Ejemplo 29: Construcción de una variante de beta-glucosidasa de *Aspergillus fumigatus* con sustitución de H266Q

[0366] La región de nucleótidos que codificó el aminoácido H285 en la secuencia codificante de *Aspergillus fumigatus* del vector pEJG97AfumFAM3A fue mutagenizada para producir una mutación C a A que codificó una sustitución de H285Q presente en la variante BG53 de beta-glucosidasa de *Aspergillus oryzae*. La mutagénesis fue realizada usando el equipo de mutagénesis dirigida QuickChange con los siguientes cebadores (los nucleótidos subrayados representan el codón con la base cambiada que codificó la nueva sustitución de aminoácidos):

SDMH266QSuperior:
5'-TGACTGAGCGCTCAACACAGCGGTGTCG-3' (SEC ID n°: 89)

SDMH266QInferior:
5'-CGACACCGCTGTGTTGAGCGCTCCAGTCA -3' (SEC ID n°: 90)

[0367] La presencia de la mutación C a A fue confirmada por secuenciación. El vector nuevo fue designado pEJG97AfumFAM3AH266Q (Figura 29). La mutación C a A fue localizada entre el fragmento de 270 bp obtenido por digestión de pEJG97 AfumFAM3AH266Q con *Blp* I, un isoesquizómero de *Bpu* 1102 I y *Xma* I. Estos sitios fueron únicos en vector el pALFd7. Tanto pEJG97AfumFAM3AH266Q como pALFd7 (aproximadamente 6 µg cada uno) fueron digeridos con *Blp* I y *Xma* I. La reacción con pALFd7 produjo dos fragmentos, uno con la mayor parte de vector (8465 bp) y un fragmento más pequeño de 270 bp de la secuencia codificante de beta-glucosidasa que contiene la ubicación de tipo salvaje donde la mutación C a A fue creada en el vector pEJG97AfumFAM3AH266Q. La reacción con

pEJG97AfumFAM3AH266Q produjo tres fragmentos, uno con la mayor parte de vector (6.331 bp), un segundo fragmento más pequeño de 2.235 bp, y un fragmento más pequeño de 270 bp de la secuencia codificante de beta-glucosidasa con la mutación C a A. El pALFd7 digerido fue tratado con fosfatasa alcalina de gamba para desfosforilación de los productos de ADN digeridos añadiendo tampón de SAP 1X y 2 ul de SAP de Roche (Roche Applied Science, Mannheim, Alemania) e incubando la reacción durante 10 minutos a 37°C seguido de incubación a 85°C durante 10 minutos para inactivación enzimática. Ambas digestiones fueron realizadas en gel de agarosa al 0,7% y purificadas usando un equipo de purificación de gel QIAGEN según las instrucciones del fabricante.

[0368] El pALFd7 digerido fue ligado al fragmento de 270 bp de la digestión de pEJG97AfumFAM3AH266Q que contiene la posición que codificó el aminoácido 161 con la mutación del nucleótido que condujo a la sustitución del aminoácido H285Q. La ligación fue realizada usando el Rapid DNA Ligation Kit siguiendo las instrucciones del fabricante.

[0369] La reacción de la ligación fue transformada en células competentes de subclonación XL1-Blue *E. coli* según las instrucciones del fabricante. Después de la transformación, el Adn plásmido de una colonia aislada fue secuenciado como se describe en el Ejemplo 10 y se confirmó la presencia de la única mutación de C a A en la región de codificación entera del gen. Posteriormente, el Adn plásmido fue transformado en células competentes de levadura como se describe en el Ejemplo 8, que dio como resultado el aislamiento de una variante de beta-glucosidasa de *Aspergillus fumigatus* con la sustitución de H285Q. El vector de expresión de levadura con la variante de beta-glucosidasa de *Aspergillus fumigatus* con la sustitución de G161S fue designado pALFd7H266Q.

Ejemplo 30: Construcción de una variante de beta-glucosidasa de *Aspergillus fumigatus* con sustituciones de G161S y H285Q

[0370] El plásmido pEJG97AfumPAM3AH266Q contenía la mutación C a A localizada en el fragmento de 270 bp obtenida por digestión de pE7G97AfumFAM3AH266Q con *Bsp* I y *Xma* I. El vector pALFd7G142S contenía la mutación G a A que produjo la sustitución de G161S en la secuencia de beta-glucosidasa de *Aspergillus fumigatus*. Tanto pEJG97AfumFAM3AH266Q como pALFd7G142S (aproximadamente 6 µg cada uno) fueron digeridos con *Bsp* I y *Xma* I. La reacción con pALFd7 produjo dos fragmentos, uno con la mayor parte de vector (8.465 bp) y un fragmento más pequeño de 270 bp de la secuencia codificante de beta-glucosidasa con la ubicación de tipo salvaje donde la mutación C a A fue creada en el vector pEIG97AfumFAM3AH266Q. La reacción con pEJG97AfumFAM3AH266Q produjo tres fragmentos, uno con la mayor parte de vector (6.331 bp), un segundo fragmento más pequeño de 2.235 bp, y un fragmento más pequeño de 270 bp de la secuencia codificante de beta-glucosidasa con la mutación C a A. pALFd7G142S digerido fue tratado con fosfatasa alcalina de gamba para desfosforilación de los productos de ADN digeridos añadiendo tampón de SAP 1X y 2 ul de SAP de Roche (Roche Applied Science, Mannheim, Alemania) e incubando la reacción durante 10 minutos a 37°C seguida de la incubación a 85°C durante 10 minutos para inactivación enzimática. Ambas digestiones fueron realizadas en gel de agarosa al 0,7% usando tampón TAE y purificadas usando un kit de purificación de gel QIAGEN según las instrucciones del fabricante.

[0371] El pALFd7G142S digerido fue ligado al fragmento de 270 bp de la digestión de pEJG97AfumFAM3AH266Q con la mutación de nucleótidos que condujo a la sustitución del aminoácido H285Q. La ligación fue realizada usando el equipo de ligación de ADN rápido siguiendo las instrucciones del fabricante.

[0372] La reacción de ligación fue transformada en células competentes de subclonación XL1-Blue *E. coli* según las instrucciones del fabricante. Después de la transformación, el Adn plásmido de una colonia aislada fue secuenciado como se describe en el Ejemplo 10 y se confirmó la presencia de la mutación G a A que codificó la sustitución de G161S y la mutación C a A que codificó la sustitución H285Q en la región de codificación entera del gen. Posteriormente, el Adn plásmido fue transformado en células competentes de levadura como se describe en el Ejemplo 8, que dio como resultado el aislamiento de una variante de beta-glucosidasa de *Aspergillus fumigatus* con las sustituciones de G161S y H285Q. El vector de expresión de levadura con la variante de beta-glucosidasa el *Aspergillus fumigatus* con las sustituciones de G161S y H285Q fue designado pALFd7G142SH266Q.

Ejemplo 31: Construcción de una variante de beta-glucosidasa de *Aspergillus fumigatus* con las sustituciones de H285Q y D724G

[0373] La región del nucleótido que codificó el aminoácido D724 en la secuencia codificante de *Aspergillus fumigatus* de pEJG97AfumFAM3A fue mutagenizada para producir una mutación A a G que codificó una sustitución homóloga de D722G presente en la variante BG53 de beta-glucosidasa de *Aspergillus oryzae*. La mutagénesis fue realizada usando el kit de mutagénesis dirigida al sitio QuickChange según las instrucciones del fabricante y los siguientes cebadores (los nucleótidos subrayados representan el codón con la base cambiada que codificó la nueva sustitución del aminoácido):

SDMD705GSuperior:

5'-GAGGATTCTTCTGGCGACCCGAACACTACGGC-3' (SEC ID n°: 91)

SDMD705GInferior:

5'-GCCGTAGTTCGGGTCGCCAGAAGAATCCTC-3' (SEC ID n°: 92)

[0374] La presencia de la mutación A a G fue confirmada por secuenciación. El vector nuevo fue designado pEJG97AfumFAM3AD705G. La mutación A a G fue localizada en el fragmento de 711 bp obtenido por digestión del vector pEJG97AfumFAM3AD705G con *Bst* EII que tiene 2 sitios en el pEJG97AfumFAM3AD705G. Para hacer el fragmento 711 bp con la mutación en el vector pALFd7H266Q, los vectores pEJG97AfumFAM3AD705G y pALFd7H266Q (aproximadamente 6 µg cada uno) fueron digeridos con *Bst* EII. La digestión con pALFd7H266Q produjo dos fragmentos, uno con la mayor parte del vector (8.024 bp) y un fragmento más pequeño de 711 bp de la secuencia codificante de beta-glucosidasa conteniendo la ubicación de tipo salvaje que codifica el aminoácido D724. La digestión con pEJG97AfumFAM3AD705G produjo dos fragmentos, uno con la mayor parte del vector (8.125 bp) y un fragmento más pequeño de 711 bp de la secuencia codificante de beta-glucosidasa con la mutación A a G. El pALFd7H266Q digerido fue tratado con fosfatasa alcalina de gamba para defosforilación de los productos de ADN digeridos añadiendo tampón de SAP 1X y 2 ul de SAP e incubando la reacción durante 10 minutos a 37°C seguido de la incubación a 85°C durante 10 minutos para la inactivación enzimática. Ambas digestiones fueron realizadas en gel de agarosa al 0,7% y purificadas usando un kit de purificación de gel QIAGEN según las instrucciones del fabricante.

[0375] El pALFd7H266Q digerido fue ligado al fragmento de 711 bp de la digestión pEJG97AfumFAM3AD705G conteniendo la mutación de nucleótidos que condujo a la sustitución del aminoácido D724G. La ligación fue realizada usando el kit Rapid DNA Ligation siguiendo las instrucciones del fabricante. La reacción de ligación fue transformada en células competentes de subclonación de XL1-Blue *E. coli* según las instrucciones del fabricante. Después de la transformación, el Adn plásmido de una colonia aislada fue secuenciado como se describe en el Ejemplo 10 y se confirmó la presencia de las mutaciones únicas de C a A y A a G que codifican las sustituciones de los aminoácidos G185S y H285Q, respectivamente, en la región de codificación entera del gen. Posteriormente, el Adn plásmido fue transformado en células competentes de levadura como se describe en el Ejemplo 8, que dio como resultado el aislamiento de una variante de beta-glucosidasa de *Aspergillus fumigatus* con las sustituciones de H285Q y D724G. El vector de expresión de levadura con la variante de beta-glucosidasa de *Aspergillus fumigatus* con las sustituciones de H285Q y D724G fue designado pALFd7H266QD705G.

Ejemplo 32: Ensayo de los efectos de las mutaciones G161S, H285Q, G161S y H285Q, y H285Q y D724G en beta-glucosidasa de *Aspergillus fumigatus*

[0376] Las mediciones de la actividad residual de beta-glucosidasas de *Aspergillus fumigatus* codificadas por los siguientes vectores se efectuó como se describe en el Ejemplo 9 a 80°C después 10 minutos: pALFd7, que contiene la secuencia nuclear de ADNc de beta-glucosidasa de *Aspergillus fumigatus* de tipo salvaje con la secuencia señal de endoglucanasa V de *Humicola insolens*, y los mismos constructos con las siguientes mutaciones, marcadas respectivamente, pALFd7G142S, con la sustitución del aminoácido G161S; pALFd7H266Q con la sustitución del aminoácido H285Q; pALFd7G142SH266Q, con las sustituciones de los aminoácidos G161S y H285Q y pALFd7H266QD705G, con las sustituciones de aminoácidos H285Q y D724G. La tabla 4 a continuación muestra la actividad residual relativa de las beta-glucosidasas a 80°C. Los resultados mostraron que las sustituciones de aminoácido mejoraron la actividad residual de la enzima de beta-glucosidasa de *Aspergillus fumigatus* a 80°C. La combinación de sustitución de los aminoácidos G161S y H285Q tuvo el máximo impacto en la mejora de la termoestabilidad de la enzima de beta-glucosidasa de *Aspergillus fumigatus*.

Tabla 4. Actividad residual relativa de variantes de beta-glucosidasa de *Aspergillus fumigatus*

Sustituciones de aminoácido	Actividad residual relativa a 80°C
A. fumigatus WT	1,00
A. fumigatus G161S	1,37
A. fumigatus H285Q	1.08
A. fumigatus G161S; H285Q	1,73
A. fumigatus H285Q; D724G	1,57

45 Depósito de material biológico

[0377] El siguiente material biológico ha sido 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 les ha dado los siguientes números de registro:

Depósito	Número de accesión	Fecha de depósito
<i>E. coli</i> TOP10 (pEJG113)	NRRL B-30695	17 de octubre de 2003
<i>E. coli</i> (pSATE111BG53)	NRRL B-30652	2 de mayo de 2003

[0378] Las cepas han sido depositadas bajo condiciones que aseguran que el acceso a los cultivos estará disponible durante la pendencia de la presente solicitud de patente a una designada por el Comisario de Patentes y Marcas Registradas que tendrá derecho a ello bajo 37 C.F.R. §1.14 y 35 U.S.C. §122. Los depósitos representan cultivos substancialmente puros de las cepas depositadas. Los depósitos están disponibles según lo establecido por las leyes de patentes extranjeras en los países donde se presentaron copias de la solicitud en cuestión, o su progenie. No

obstante, debería entenderse que la disponibilidad de un depósito no constituye una licencia para poner en práctica la invención en cuestión en derogación de los derechos de las patentes concedidas por acción gubernamental.

Listado de secuencias
5 [0386]
<110> NOVOZYMES INC.

<120> Variantes de beta-glucosidasas
10 <130> 10371.204-WO

<150> 60/467,767
<151> 2003-05-02
15 <150> 60/528,342
<151> 2003-12-09

<160> 92
20 <170> PatentIn versión 3.2

<210> 1
<211> 2586
<212> ADN
25 <213> Aspergillus oryzae

<220>
<221> CDS
<222> (1).. (2583)
30 <400> 1

ES 2 393 058 T3

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

ES 2 393 058 T3

ggt Gly	ttc Phe	tca Ser	cca Pro	gat Asp 165	cca Pro	gcc Ala	ctc Leu	acc Thr	ggt Gly 170	gta Val	ctt Leu	ttt Phe	gcg Ala	gag Glu 175	acg Thr	528
att Ile	aag Lys	ggt Gly	att Ile 180	caa Gln	gat Asp	gct Ala	ggt Gly	gtc Val 185	att Ile	gcg Ala	aca Thr	gct Ala	aag Lys 190	cat His	tat Tyr	576
atc Ile	atg Met	aac Asn 195	gaa Glu	caa Gln	gag Glu	cat His	ttc Phe 200	cgc Arg	caa Gln	caa Gln	ccc Pro	gag Glu 205	gct Ala	gcg Ala	ggt Gly	624
tac Tyr	gga Gly 210	ttc Phe	aac Asn	gta Val	agc Ser	gac Asp 215	agt Ser	ttg Leu	agt Ser	tcc Ser	aac Asn 220	gtt Val	gat Asp	gac Asp	aag Lys	672
act Thr 225	atg Met	cat His	gaa Glu	ttg Leu	tac Tyr 230	ctc Leu	tgg Trp	ccc Pro	ttc Phe	gcg Ala 235	gat Asp	gca Ala	gta Val	cgc Arg	gct Ala 240	720
gga Gly	gtc Val	ggt Gly	gct Ala	gtc Val 245	atg Met	tgc Cys	tct Ser	tac Tyr	aac Asn 250	caa Gln	atc Ile	aac Asn	aac Asn	agc Ser 255	tac Tyr	768
ggt Gly	tgc Cys	gag Glu	aat Asn 260	agc Ser	gaa Glu	act Thr	ctg Leu	aac Asn 265	aag Lys	ctt Leu	ttg Leu	aag Lys	gcg Ala 270	gag Glu	ctt Leu	816
ggt Gly	ttc Phe	caa Gln 275	ggc Gly	ttc Phe	gtc Val	atg Met	agt Ser 280	gat Asp	tgg Trp	acc Thr	gct Ala	cat His 285	cac His	agc Ser	ggc Gly	864
gta Val	ggc Gly 290	gct Ala	gct Ala	tta Leu	gca Ala	ggt Gly 295	ctg Leu	gat Asp	atg Met	tcg Ser	atg Met 300	ccc Pro	ggt Gly	gat Asp	gtt Val	912
acc Thr 305	ttc Phe	gat Asp	agt Ser	ggt Gly	acg Thr 310	tct Ser	ttc Phe	tgg Trp	ggt Gly	gca Ala 315	aac Asn	ttg Leu	acg Thr	gtc Val	ggt Gly 320	960
gtc Val	ctt Leu	aac Asn	ggt Gly	aca Thr 325	atc Ile	ccc Pro	caa Gln	tgg Trp	cgt Arg 330	gtt Val	gat Asp	gac Asp	atg Met	gct Ala 335	gtc Val	1008
cgt Arg	atc Ile	atg Met	gcc Ala 340	gct Ala	tat Tyr	tac Tyr	aag Lys	gtt Val 345	ggc Gly	cgc Arg	gac Asp	acc Thr	aaa Lys 350	tac Tyr	acc Thr	1056
cct Pro	ccc Pro	aac Asn 355	ttc Phe	agc Ser	tcg Ser	tgg Trp	acc Thr 360	agg Arg	gac Asp	gaa Glu	tat Tyr	ggt Gly 365	ttc Phe	gcg Ala	cat His	1104
aac Asn	cat His 370	gtt Val	tcg Ser	gaa Glu	ggt Gly	gct Ala 375	tac Tyr	gag Glu	agg Arg	gtc Val	aac Asn 380	gaa Glu	ttc Phe	gtg Val	gac Asp	1152
gtg Val 385	caa Gln	cgc Arg	gat Asp	cat His	gcc Ala 390	gac Asp	cta Leu	atc Ile	cgt Arg	cgc Arg 395	atc Ile	ggc Gly	gcg Ala	cag Gln	agc Ser 400	1200
act Thr	gtt Val	ctg Leu	ctg Leu	aag Lys 405	aac Asn	aag Lys	ggt Gly	gcc Ala	ttg Leu 410	ccc Pro	ttg Leu	agc Ser	cgc Arg	aag Lys 415	gaa Glu	1248
aag Lys	ctg Leu	gtc Val	gcc Ala 420	ctt Leu	ctg Leu	gga Gly	gag Glu	gat Asp 425	gcg Ala	ggt Gly	tcc Ser	aac Asn	tcg Ser 430	tgg Trp	ggc Gly	1296
gct	aac	ggc	tgt	gat	gac	cg	ggt	tgc	gat	aac	ggt	acc	ctt	gcc	atg	1344

ES 2 393 058 T3

Ala	Asn	Gly 435	Cys	Asp	Asp	Arg	Gly 440	Cys	Asp	Asn	Gly	Thr 445	Leu	Ala	Met		
gcc	tgg	ggt	agc	ggt	act	gcg	aat	ttc	cca	tac	ctc	gtg	aca	cca	gag	1392	
Ala	Trp 450	Gly	Ser	Gly	Thr	Ala 455	Asn	Phe	Pro	Tyr	Leu 460	Val	Thr	Pro	Glu		
cag	gcg	att	cag	aac	gaa	ggt	ctt	cag	ggc	cgt	ggt	aat	gtc	ttc	gcc	1440	
Gln 465	Ala	Ile	Gln	Asn	Glu 470	Val	Leu	Gln	Gly	Arg 475	Gly	Asn	Val	Phe	Ala 480		
gtg	acc	gac	agt	tgg	gcg	ctc	gac	aag	atc	gct	gcg	gct	gcc	cgc	cag	1488	
Val	Thr	Asp	Ser	Trp 485	Ala	Leu	Asp	Lys	Ile 490	Ala	Ala	Ala	Ala	Arg 495	Gln		
gcc	agc	gta	tct	ctc	gtg	ttc	gtc	aac	tcc	gac	tca	gga	gaa	ggc	tat	1536	
Ala	Ser	Val	Ser 500	Leu	Val	Phe	Val	Asn 505	Ser	Asp	Ser	Gly	Glu 510	Gly	Tyr		
ctt	agt	gtg	gat	gga	aat	gag	ggc	gat	cgt	aac	aac	atc	act	ctg	tgg	1584	
Leu	Ser	Val 515	Asp	Gly	Asn	Glu	Gly 520	Asp	Arg	Asn	Asn	Ile 525	Thr	Leu	Trp		
aag	aac	ggc	gac	aat	gtg	gtc	aag	acc	gca	gcg	aat	aac	tgt	aac	aac	1632	
Lys	Asn 530	Gly	Asp	Asn	Val	Val 535	Lys	Thr	Ala	Ala	Asn 540	Asn	Cys	Asn	Asn		
acc	ggt	gtc	atc	atc	cac	tcc	gtc	gga	cca	ggt	ttg	atc	gat	gaa	tgg	1680	
Thr 545	Val	Val	Ile	Ile	His 550	Ser	Val	Gly	Pro	Val 555	Leu	Ile	Asp	Glu	Trp 560		
tat	gac	cac	ccc	aat	gtc	act	ggt	att	ctc	tgg	gct	ggt	ctg	cca	ggc	1728	
Tyr	Asp	His	Pro	Asn 565	Val	Thr	Gly	Ile	Leu 570	Trp	Ala	Gly	Leu	Pro 575	Gly		
cag	gag	tct	ggt	aac	tcc	att	gcc	gat	gtg	ctg	tac	ggt	cgt	gtc	aac	1776	
Gln	Glu	Ser	Gly 580	Asn	Ser	Ile	Ala	Asp 585	Val	Leu	Tyr	Gly	Arg 590	Val	Asn		
cct	ggc	gcc	aag	tct	cct	ttc	act	tgg	ggc	aag	acc	cgg	gag	tcg	tat	1824	
Pro	Gly	Ala 595	Lys	Ser	Pro	Phe	Thr 600	Trp	Gly	Lys	Thr	Arg 605	Glu	Ser	Tyr		
ggt	tct	ccc	ttg	gtc	aag	gat	gcc	aac	aat	ggc	aac	gga	gcg	ccc	cag	1872	
Gly 610	Ser	Pro	Leu	Val	Lys	Asp 615	Ala	Asn	Asn	Gly	Asn 620	Gly	Ala	Pro	Gln		
tct	gat	ttc	acc	cag	ggt	ggt	ttc	atc	gat	tac	cgc	cat	ttc	gat	aag	1920	
Ser 625	Asp	Phe	Thr	Gln	Gly 630	Val	Phe	Ile	Asp	Tyr 635	Arg	His	Phe	Asp	Lys 640		
ttc	aat	gag	acc	cct	atc	tac	gag	ttt	ggc	tac	ggc	ttg	agc	tac	acc	1968	
Phe	Asn	Glu	Thr	Pro 645	Ile	Tyr	Glu	Phe	Gly 650	Tyr	Gly	Leu	Ser	Tyr 655	Thr		
acc	ttc	gag	ctc	tcc	gac	ctc	cat	ggt	cag	ccc	ctg	aac	gcg	tcc	cga	2016	
Thr	Phe	Glu	Leu 660	Ser	Asp	Leu	His	Val 665	Gln	Pro	Leu	Asn	Ala 670	Ser	Arg		
tac	act	ccc	acc	agt	ggc	atg	act	gaa	gct	gca	aag	aac	ttt	ggt	gaa	2064	
Tyr	Thr	Pro 675	Thr	Ser	Gly	Met	Thr 680	Glu	Ala	Ala	Lys	Asn 685	Phe	Gly	Glu		
att	ggc	gat	gcg	tcg	gag	tac	gtg	tat	ccg	gag	ggg	ctg	gaa	agg	atc	2112	
Ile	Gly 690	Asp	Ala	Ser	Glu	Tyr 695	Val	Tyr	Pro	Glu	Gly 700	Leu	Glu	Arg	Ile		
cat	gag	ttt	atc	tat	ccc	tgg	atc	aac	tct	acc	gac	ctg	aag	gca	tcg	2160	
His	Glu	Phe	Ile	Tyr	Pro	Trp	Ile	Asn	Ser	Thr	Asp	Leu	Lys	Ala	Ser		

ES 2 393 058 T3

705					710					715					720	
tct	gac	gat	tct	aac	tac	ggc	tgg	gaa	gac	tcc	aag	tat	att	ccc	gaa	2208
Ser	Asp	Asp	Ser	Asn	Tyr	Gly	Trp	Glu	Asp	Ser	Lys	Tyr	Ile	Pro	Glu	
				725					730					735		
ggc	gcc	acg	gat	ggg	tct	gcc	cag	ccc	cgt	ttg	ccc	gct	agt	ggg	ggg	2256
Gly	Ala	Thr	Asp	Gly	Ser	Ala	Gln	Pro	Arg	Leu	Pro	Ala	Ser	Gly	Gly	
			740					745					750			
gcc	gga	gga	aac	ccc	ggg	ctg	tac	gag	gat	ctt	ttc	cgc	gtc	tct	gtg	2304
Ala	Gly	Gly	Asn	Pro	Gly	Leu	Tyr	Glu	Asp	Leu	Phe	Arg	Val	Ser	Val	
		755					760					765				
aag	gtc	aag	aac	acg	ggc	aat	gtc	gcc	ggg	gat	gaa	ggt	cct	cag	ctg	2352
Lys	Val	Lys	Asn	Thr	Gly	Asn	Val	Ala	Gly	Asp	Glu	Val	Pro	Gln	Leu	
	770					775					780					
tac	ggt	tcc	cta	ggc	ggc	ccg	aat	gag	ccc	aag	gtg	gta	ctg	cgc	aag	2400
Tyr	Val	Ser	Leu	Gly	Gly	Pro	Asn	Glu	Pro	Lys	Val	Val	Leu	Arg	Lys	
	785				790					795					800	
ttt	gag	cgt	att	cac	ttg	gcc	cct	tcg	cag	gag	gcc	gtg	tgg	aca	acg	2448
Phe	Glu	Arg	Ile	His	Leu	Ala	Pro	Ser	Gln	Glu	Ala	Val	Trp	Thr	Thr	
				805					810					815		
acc	ctt	acc	cgt	cgt	gac	ctt	gca	aac	tgg	gac	ggt	tcg	gct	cag	gac	2496
Thr	Leu	Thr	Arg	Arg	Asp	Leu	Ala	Asn	Trp	Asp	Val	Ser	Ala	Gln	Asp	
			820					825					830			
tgg	acc	gtc	act	cct	tac	ccc	aag	acg	atc	tac	ggt	gga	aac	tcc	tca	2544
Trp	Thr	Val	Thr	Pro	Tyr	Pro	Lys	Thr	Ile	Tyr	Val	Gly	Asn	Ser	Ser	
		835					840					845				
cgg	aaa	ctg	ccg	ctc	cag	gcc	tcg	ctg	cct	aag	gcc	cag	taa			2586
Arg	Lys	Leu	Pro	Leu	Gln	Ala	Ser	Leu	Pro	Lys	Ala	Gln				
	850					855					860					

<210> 2
 <211> 861
 <212> PRT
 <213> Aspergillus oryzae

5

<400> 2

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

ES 2 393 058 T3

Gly Ile Arg Phe Ser Asp Tyr Asn Ser Ala Phe Pro Ala Gly Val Asn
100 105 110

Val Ala Ala Thr Trp Asp Lys Thr Leu Ala Tyr Leu Arg Gly Gln Ala
115 120 125

Met Gly Glu Glu Phe Ser Asp Lys Gly Ile Asp Val Gln Leu Gly Pro
130 135 140

Ala Ala Gly Pro Leu Gly Ala His Pro Asp Gly Gly Arg Asn Trp Glu
145 150 155 160

Gly Phe Ser Pro Asp Pro Ala 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 Met Asn Glu Gln Glu His Phe Arg Gln Gln Pro Glu Ala Ala Gly
195 200 205

Tyr Gly Phe Asn Val Ser Asp Ser Leu Ser Ser Asn Val Asp Asp Lys
210 215 220

Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg Ala
225 230 235 240

Gly Val Gly Ala Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr
245 250 255

Gly Cys Glu Asn Ser Glu Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu
260 265 270

Gly Phe Gln Gly Phe Val Met Ser Asp Trp Thr Ala His His Ser Gly
275 280 285

Val Gly Ala Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Val
290 295 300

Thr Phe Asp Ser Gly Thr Ser Phe Trp Gly Ala Asn Leu Thr Val Gly
305 310 315 320

Val Leu Asn Gly Thr Ile Pro Gln Trp Arg Val Asp Asp Met Ala Val
325 330 335

Arg Ile Met Ala Ala Tyr Tyr Lys Val Gly Arg Asp Thr Lys Tyr Thr
340 345 350

Pro Pro Asn Phe Ser Ser Trp Thr Arg Asp Glu Tyr Gly Phe Ala His
355 360 365

Asn His Val Ser Glu Gly Ala Tyr Glu Arg Val Asn Glu Phe Val Asp
 370 375 380
 Val Gln Arg Asp His Ala Asp Leu Ile Arg Arg Ile Gly Ala Gln Ser
 385 390 395 400
 Thr Val Leu Leu Lys Asn Lys Gly Ala Leu Pro Leu Ser Arg Lys Glu
 405 410 415
 Lys Leu Val Ala Leu Leu Gly Glu Asp Ala Gly Ser Asn Ser Trp Gly
 420 425 430
 Ala Asn Gly Cys Asp Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met
 435 440 445
 Ala Trp Gly Ser Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu
 450 455 460
 Gln Ala Ile Gln Asn Glu Val Leu Gln Gly Arg Gly Asn Val Phe Ala
 465 470 475 480
 Val Thr Asp Ser Trp Ala Leu Asp Lys Ile Ala Ala Ala Ala Arg Gln
 485 490 495
 Ala Ser Val Ser Leu Val Phe Val Asn Ser Asp Ser Gly Glu Gly Tyr
 500 505 510
 Leu Ser Val Asp Gly Asn Glu Gly Asp Arg Asn Asn Ile Thr Leu Trp
 515 520 525
 Lys Asn Gly Asp Asn Val Val Lys Thr Ala Ala Asn Asn Cys Asn Asn
 530 535 540
 Thr Val Val Ile Ile His Ser Val Gly Pro Val Leu Ile Asp Glu Trp
 545 550 555 560
 Tyr Asp His Pro Asn Val Thr Gly Ile Leu Trp Ala Gly Leu Pro Gly
 565 570 575
 Gln Glu Ser Gly Asn Ser Ile Ala Asp Val Leu Tyr Gly Arg Val Asn
 580 585 590
 Pro Gly Ala Lys Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr
 595 600 605
 Gly Ser Pro Leu Val Lys Asp Ala Asn Asn Gly Asn Gly Ala Pro Gln
 610 615 620
 Ser Asp Phe Thr Gln Gly Val Phe Ile Asp Tyr Arg His Phe Asp Lys
 625 630 635 640
 Phe Asn Glu Thr Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr Thr

				645						650						655
Thr	Phe	Glu	Leu	Ser	Asp	Leu	His	Val	Gln	Pro	Leu	Asn	Ala	Ser	Arg	
			660					665					670			
Tyr	Thr	Pro	Thr	Ser	Gly	Met	Thr	Glu	Ala	Ala	Lys	Asn	Phe	Gly	Glu	
		675					680					685				
Ile	Gly	Asp	Ala	Ser	Glu	Tyr	Val	Tyr	Pro	Glu	Gly	Leu	Glu	Arg	Ile	
	690					695					700					
His	Glu	Phe	Ile	Tyr	Pro	Trp	Ile	Asn	Ser	Thr	Asp	Leu	Lys	Ala	Ser	
705					710					715					720	
Ser	Asp	Asp	Ser	Asn	Tyr	Gly	Trp	Glu	Asp	Ser	Lys	Tyr	Ile	Pro	Glu	
				725					730					735		
Gly	Ala	Thr	Asp	Gly	Ser	Ala	Gln	Pro	Arg	Leu	Pro	Ala	Ser	Gly	Gly	
			740					745					750			
Ala	Gly	Gly	Asn	Pro	Gly	Leu	Tyr	Glu	Asp	Leu	Phe	Arg	Val	Ser	Val	
		755					760					765				
Lys	Val	Lys	Asn	Thr	Gly	Asn	Val	Ala	Gly	Asp	Glu	Val	Pro	Gln	Leu	
	770					775					780					
Tyr	Val	Ser	Leu	Gly	Gly	Pro	Asn	Glu	Pro	Lys	Val	Val	Leu	Arg	Lys	
785					790					795					800	
Phe	Glu	Arg	Ile	His	Leu	Ala	Pro	Ser	Gln	Glu	Ala	Val	Trp	Thr	Thr	
				805					810					815		
Thr	Leu	Thr	Arg	Arg	Asp	Leu	Ala	Asn	Trp	Asp	Val	Ser	Ala	Gln	Asp	
			820					825					830			
Trp	Thr	Val	Thr	Pro	Tyr	Pro	Lys	Thr	Ile	Tyr	Val	Gly	Asn	Ser	Ser	
		835					840					845				
Arg	Lys	Leu	Pro	Leu	Gln	Ala	Ser	Leu	Pro	Lys	Ala	Gln				
	850					855					860					

<210> 3
 <211> 32
 <212> ADN

<213> *Aspergillus oryzae*
 <400> 3
 5 gcagatctac catgaagctt ggttgatcg ag 32
 <210> 4
 <211> 30
 <212> ADN
 <213> *Aspergillus oryzae*
 10 <400> 4
 gcctcgagt actggcctt aggacgag 30
 <210> 5
 15 <211> 29
 <212> ADN
 <213> *Trichoderma reesei*
 <400> 5
 20 aacgtaatt aaggaatcgt tttgtgtt 29
 <210> 6
 <211> 29
 <212> ADN
 25 <213> *Trichoderma reesei*
 <400> 6
 30 agtactagta gctccgtggc gaaagcctg 29
 <210> 7
 <211> 31
 <212> ADN
 <213> *Escherichia coli*
 35 <400> 7
 ttgaattgaa aatagattga ttaaaact c 31
 <210> 8
 <211> 25
 40 <212> ADN
 <213> *Escherichia coli*
 <400> 8
 45 ttgcatgct aatcatggc atagc 25
 <210> 9
 <211> 26
 <212> ADN
 <213> *Saccharomyces cerevisiae*
 50 <400> 9
 ttgaattcat gggtaataac tgatat 26
 <210> 10
 55 <211> 32
 <212> ADN
 <213> *Saccharomyces cerevisiae*
 <400> 10
 60 aatcaatct atttcaatt caatcatca tt 32
 <210> 11
 <211> 45
 <212> ADN
 65 <213> *Aspergillus niger*

<400> 11
 ggatgctgtt gactccggaa atttaacggt ttggtcttgc atccc 45
 5 <210> 12
 <211> 44
 <212> ADN
 <213> Aspergillus niger
 <400> 12
 10 ggtattgtcc tgcagacggc aatttaacgg cttctgcgaa tcgc 44
 <210> 13
 <211> 29
 <212> ADN
 15 <213> Humicola insolens
 <400> 13
 aagcttaagc atgcgttctt cccccctcc 29
 20 <210> 14
 <211> 32
 <212> ADN
 <213> Humicola insolens
 25 <400> 14
 ctgcagaatt ctacaggcac tgatggtacc ag 32
 <210> 15
 <211> 32
 30 <212> ADN
 <213> Humicola insolens
 <400> 15
 35 ctgcagaatt ctacaggcac tgatggtacc ag 32
 <210> 16
 <211> 36
 <212> ADN
 <213> Humicola insolens
 40 <400> 16
 accgcggact gcgcatcatg cgttcctccc ccctcc 36
 <210> 17
 45 <211> 29
 <212> ADN
 <213> Trichoderma reesei
 <400> 17
 50 aaacgtcgac cgaatgtagg attgttatc 29
 <210> 18
 <211> 17
 <212> ADN
 55 <213> Trichoderma reesei
 <400> 18
 gatgcgcagt ccgcggt 17
 60 19
 <211> 29
 <212> ADN
 <213> Trichoderma reesei
 65 <400> 19
 aaacgtcgac cgaatgtagg attgttatc 29

<210> 20
 <211> 36
 <212> ADN
 5 <213> Trichoderma reesei

 <400> 20
 ggagggggga ggaacgcatg atgcgcagtc cgcggt 36

 10 <210> 21
 <211> 29
 <212> ADN
 <213> Trichoderma reesei

 15 <400> 21
 aaacgtcgac cgaatgtagg attgttat 29

 <210> 22
 <211> 32
 20 <212> ADN
 <213> Trichoderma reesei

 <400> 22
 25 ctgcagaatt ctacaggcac tgatgtacc ag 32

 <210> 23
 <211> 46
 <212> ADN
 <213> Aspergillus oryzae
 30
 <400> 23
 atagtcaacc gcggactgcg catcatgaag ctggttga tcgagg 46

 <210> 24
 <211> 26
 <212> ADN
 <213> Aspergillus oryzae

 <400> 24
 40 actagttac tggccttag gcagcg 26

 <210> 25
 <211> 26
 <212> ADN
 45 <213> Trichoderma reesei

 <400> 25
 gtcgactcga agcccgaatg taggat 26

 50 <210> 26
 <211> 45
 <212> ADN
 <213> Trichoderma reesei

 55 <400> 26
 cctgatcca accaagcttc atgatgcgca gtccgcggt gacta 45

 <210> 27
 <211> 42
 60 <212> ADN
 <213> Aspergillus oryzae

 <400> 27
 65 tgccggtgtt ggcccttgcc aaggatgatc tcgctactc cc 42

 <210> 28

<211> 28
 <212> ADN
 <213> Aspergillus oryzae

5 <400> 28
 gactagtctt actgggcctt aggcagcg 28

<210> 29
 <211> 63
 <212> ADN
 <213> Humicola insolens

<400> 29
atgggttccat ccccccctccat ccgctccgac gttgtggcag acctgcaggt gttggccctt 60
gcc 63

15 <210> 30
 <211> 30
 <212> ADN
 <213> Aspergillus oryzae

20 <400> 30
 acgcgtcgac cgaatgtagg attggtatcc 30

25 <210> 31
 <211> 42
 <212> ADN
 <213> Aspergillus oryzae

30 <400> 31
 gggagtacgc gagatcatcc ttggcaaggg ccaacaccgg ca 42

<210> 32
 <211> 40
 <212> ADN
 <213> Aspergillus oryzae

35 <400> 32
 aatccgacta gtgatctac catgcgttcc tccccctcc 40

40 <210> 33
 <211> 32
 <212> ADN
 <213> Aspergillus oryzae

45 <400> 33
 gcgggcctcg agttactggg ccttaggcag cg 32

<210> 34
 <211> 17
 <212> ADN
 <213> Aspergillus oryzae

50 <400> 34
 aggggtaatg ggcggaa 17

55 <210> 35
 <211> 20
 <212> ADN
 <213> Aspergillus oryzae

60 <400> 35
 gacattttg ctgtcagtca 20

<210> 36
 <211> 20
 <212> ADN
 <213> Aspergillus oryzae
 5
 <400> 36
 aatgttacat gcgtacacgc 20
 <210> 37
 <211> 17
 <212> ADN
 <213> Aspergillus oryzae
 10
 <400> 37
 gtttcggctc aggactg 17
 <210> 38
 <211> 17
 <212> ADN
 <213> Aspergillus oryzae
 15
 <400> 38
 acttccgccc attcacc 17
 <210> 39
 <211> 17
 <212> ADN
 <213> Aspergillus oryzae
 20
 <400> 39
 aggggtgaatg ggcggaa 17
 <210> 40
 <211> 17
 <212> ADN
 <213> Aspergillus oryzae
 25
 <400> 40
 ggcggaaatg ctcttgt 17
 <210> 41
 <211> 18
 <212> ADN
 <213> Aspergillus oryzae
 30
 <400> 41
 sggatggcgg tagaaact 18
 <210> 42
 <211> 17
 <212> ADN
 <213> Aspergillus oryzae
 35
 <400> 42
 gcggtccaat cactcat 17
 <210> 43
 <211> 17
 <212> ADN
 <213> Aspergillus oryzae
 40
 <400> 43
 gctacggttg cgagaat 17
 <210> 44
 <211> 17

<212> ADN
 <213> Aspergillus oryzae

 <400> 44
 5 ctcaaggcca aggcacc 17

 <210> 45
 <211> 17
 <212> ADN
 10 <213> Aspergillus oryzae

 <400> 45
 ggtgccttgc ccttgag 17

 15 <210> 46
 <211> 17
 <212> ADN
 <213> Aspergillus oryzae

 20 <400> 46
 ttcgctgcgg tcttgac 17

 <210> 47
 <211> 17
 25 <212> ADN
 <213> Aspergillus oryzae

 <400> 47
 30 gtggaagaac ggcgaca 17

 <210> 48
 <211> 17
 <212> ADN
 35 <213> Aspergillus oryzae

 <400> 48
 cccagccgta gtagaa 17

 <210> 49
 <211> 17
 <212> ADN
 40 <213> Aspergillus oryzae

 <400> 49
 45 cgtcccgata cactccc 17

 <210> 50
 <211> 17
 <212> ADN
 50 <213> Aspergillus oryzae

 <400> 50
 cctggagcgg cagtffc 17

 55 <210> 51
 <211> 18
 <212> ADN
 <213> Aspergillus oryzae

 60 <400> 51
 ggtcgggtgc ctaacgg 18

 <210> 52
 <211> 20
 65 <212> ADN
 <213> Aspergillus oryzae

<400> 52
 actatcctgc aaacacaagc 20

5 <210> 53
 <211> 16
 <212> ADN
 <213> Aspergillus oryzae

10 <400> 53
 cctttcactt ggggca 16

15 <210> 54
 <211> 20
 <212> ADN
 <213> Aspergillus oryzae

20 <400> 54
 ggagttacca gactcctggc 20

25 <210> 55
 <211> 20
 <212> ADN
 <213> Aspergillus oryzae

30 <400> 55
 acctccgaa acatggtat 20

35 <210> 56
 <211> 39
 <212> ADN
 <213> Aspergillus oryzae

40 <220>
 <221> misc_feature
 <222> (16)..(17)
 <223> n is a, c, g, o t

45 <400> 56
 ggtagaaact gggaannstt ctcaccagat ccagccctc 39

50 <210> 57
 <211> 33
 <212> ADN
 <213> Aspergillus oryzae

55 <220>
 <221> misc_feature
 <222> (16)..(17)
 <223> n es a, c, g, o t

60 <400> 57
 gcctagccg ctgtgnnsag cggccaatc act 33

65 <210> 58
 <211> 22
 <212> ADN
 <213> Aspergillus nidulens

70 <400> 58
 gtgcccattg atacgcctcc gg 22

75 <210> 59
 <211> 26
 <212> ADN
 <213> Aspergillus nidulens

<400> 59
 gagtctgatt tccaaggctc ctgacc 26

5 <210> 60
 <211> 23
 <212> ADN
 <213> Aspergillus nidulens

10 <400> 60
 ggaggccatg aagtggacca acg 23

15 <210> 61
 <211> 45
 <212> ADN
 <213> Aspergillus niger

20 <400> 61
 caccgtgaaa gccatgctct ttcctcgtg tagaagacca gacag 45

25 <210> 62
 <211> 45
 <212> ADN
 <213> Aspergillus niger

30 <400> 62
 ctggtctct acacgaagga aagagcatgg ctttcacggt gtctg 45

35 <210> 63
 <211> 44
 <212> ADN
 <213> Aspergillus niger

40 <400> 63
 ctatatacac aactggattt accatgggcc cgcggccgca gatc 44

45 <210> 64
 <211> 44
 <212> ADN
 <213> Aspergillus niger

50 <400> 64
 gatctgcggc cgcgggccca tggtaaacc agttgtgtat atag 44

55 <210> 65
 <211> 30
 <212> ADN
 <213> Aspergillus oryzae

60 <400> 65
 actggattta ccatgaagct tggttgatc 30

65 <210> 66
 <211> 30
 <212> ADN
 <213> Aspergillus oryzae

70 <400> 66
 agtcacctct agttattact gggccttagg 30

75 <210> 67
 <211> 31
 <212> ADN
 <213> Aspergillus fumigatus

80 <400> 67

ES 2 393 058 T3

actggattta ccatgagatt cggttggctc g 31

<210> 68
 <211> 31
 5 <212> ADN
 <213> Aspergillus fumigatus

<400> 68
 10 agtcacctct agttactagt agacacgggg c 31

<210> 69
 <211> 3060
 <212> ADN
 15 <213> Aspergillus fumigatus

<400> 69

atgagattcg gttggctcga ggtggccgct ctgacggccg cttctgtagc caatgcccag	60
gtttgtgatg ctttcccgtc attgtttcgg atatagttga caatagtcac ggaaataatc	120
aggaattggc tttctctcca ccattctacc cttcgccttg ggctgatggc cagggagagt	180
gggcagatgc ccatcgacgc gccgtcgaga tcgtttctca gatgacactg gcggagaagg	240
ttaaccttac aacgggtact ggggtgggttg cgactttttt gttgacagtg agctttcttc	300
actgaccatc tacacagatg ggaaatggac cgatgcgtcg gtcaaaccgg cagcgttccc	360
aggtaagctt gcaattctgc aacaacgtgc aagtgtagtt gctaaaacgc ggtggtgcag	420
acttggatc 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

ES 2 393 058 T3

gctggggcct gctgctggtc ctctcggcaa ataccggac ggcggcagaa tctgggaagg 720
cttctctcct gatccgggtc tctactggtgt acttttcgcc gaaactatca agggatcca 780
agacgcgggt gtgattgcta ctgccaagca ttacattctg aatgaacagg agcatttccg 840
acaggttggc gaggcccagg gatatggtta caacatcacg gagacgatca gctccaacgt 900
ggatgacaag accatgcacg agttgtacct ttggtgagta gttgacactg caaatgagga 960
ccttgattga tttgactgac ctggaatgca ggccctttgc agatgctgtg cgcggttaaga 1020
ttttccgtag acttgacctc gcgacgaaga aatcgctgac gaaccatcgt agctggcggt 1080
ggcgtgtca tgtgttccta caatcaaata aacaacagct acggttgta aacagtcaa 1140
actctcaaca agctcctcaa ggctgagctg ggcttccaag gcttcgtcat gagtgactgg 1200
agcgtcacc acagcgggtg cggcgtgcc ctcgctgggt tggatatgtc gatgcctgga 1260
gacatttctt tcgacgacgg actctccttc tggggcacga acctaactgt cagtgttctt 1320
aacggcaccg ttccagcctg gcgtgtcgat gacatggctg ttcgtatcat gaccgcgtac 1380
tacaaggttg gtcgtgaccg tcttcgtatt cccctaact tcagctcctg gaccgggat 1440
gagtacggct gggagcattc tgctgtctcc gagggagcct ggaccaaggt gaacgacttc 1500
gtcaatgtgc agcgcagtca ctctcagatc atccgtgaga ttggtgccgc tagtacagtg 1560
ctcttgaaga acacgggtgc tcttcctttg accggcaagg aggttaaagt ggggtttctc 1620
ggatgaagac ctggttccaa cccgtgggggt gctaacggct gccccgaccg cggctgtgat 1680
aacggcactc ttgctatggc ctggggtagt ggtactgcca acttccctta ccttgtcacc 1740
cccgagcagg ctatccagcg agaggtcatc agcaacggcg gcaatgtctt tgctgtgact 1800
gataacgggg ctctcagcca gatggcagat gttgcatctc aatccagggt agtgcgggct 1860
cttagaaaaa gaacgttctc tgaatgaagt tttttaacca ttgcgaacag cgtgtctttg 1920
gtgtttgtca acgccgactc tggagagggt ttcacagtg tcgacggcaa cgagggtgac 1980
cgcaaaaatc tactctgtg gaagaacggc gaggccgtca ttgacactgt tgcagccac 2040
tgcaacaaca cgatttggtt tattcacagt gttgggcccg tcttgatcga ccggtggtat 2100
gataaccca acgtcactgc catcatctgg gccggcttgc ccggtcagga gagtggcaac 2160
tccttggtcg acgtgctcta tggccgcgtc aaccagcg ccaagacccc gttcacctgg 2220
ggcaagactc gggagtctta cggggtccc ttgctcaccg agcctaaca tggcaatggt 2280
gctcccagc atgatttcaa cgagggcgtc ttactgact accgtcactt tgacaagcgc 2340
aatgagacc ccatttatga gttggccat ggcttgagct acaccactt tggttactct 2400
caccttcggg ttcaggccct caatagtctg agttcggcat atgtcccagc tagcggagag 2460
accaagcctg cgccaacctt tgggtgagatc ggtagtgccg ccgactacct gtatcccag 2520
ggtctcaaaa gaattacca gtttatttac ccttggctca actcgaccga cctcaggat 2580
tcttctgacg acccgaacta cggctgggag gactcggagt acattcccga aggcgctagg 2640
gatgggtctc ctcaaccctt cctgaaggct ggcggcgtc ctggtggtaa ccctaccctt 2700

ES 2 393 058 T3

tatcaggatc ttgttagggg gtcggccacc ataaccaaca ctggtaacgt cgccggttat 2760
gaagtcctc aattggtgag tgacccgcat gttccttgcg ttgcaatttg gctaactcgc 2820
ttctagtatg tttcactggg cggaccgaac gagcctcggg tcgttctgcg caagttcgac 2880
cgaatcttcc tggctcctgg ggagcaaaag gtttgaccca cgactcttaa ccgtcgtgat 2940
ctcgccaatt gggatgtgga ggctcaggac tgggtcatca caaagtaccc caagaaagtg 3000
cacgtcggca gctcctcgcg taagctgcct ctgagagcgc ctctgccccg tgtctactag 3060

<210> 70

<211> 863

5 <212> PRT

<213> *Aspergillus fumigatus*

<400> 70

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

ES 2 393 058 T3

195					200					205					
Tyr	Gly	Tyr	Asn	Ile	Thr	Glu	Thr	Ile	Ser	Ser	Asn	Val	Asp	Asp	Lys
	210					215					220				
Thr	Met	His	Glu	Leu	Tyr	Leu	Trp	Pro	Phe	Ala	Asp	Ala	Val	Arg	Ala
225					230					235					240
Gly	Val	Gly	Ala	Val	Met	Cys	Ser	Tyr	Asn	Gln	Ile	Asn	Asn	Ser	Tyr
				245					250					255	
Gly	Cys	Gln	Asn	Ser	Gln	Thr	Leu	Asn	Lys	Leu	Leu	Lys	Ala	Glu	Leu
			260					265					270		
Gly	Phe	Gln	Gly	Phe	Val	Met	Ser	Asp	Trp	Ser	Ala	His	His	Ser	Gly
		275					280					285			
Val	Gly	Ala	Ala	Leu	Ala	Gly	Leu	Asp	Met	Ser	Met	Pro	Gly	Asp	Ile
	290					295					300				
Ser	Phe	Asp	Asp	Gly	Leu	Ser	Phe	Trp	Gly	Thr	Asn	Leu	Thr	Val	Ser
305					310					315					320
Val	Leu	Asn	Gly	Thr	Val	Pro	Ala	Trp	Arg	Val	Asp	Asp	Met	Ala	Val
				325					330					335	
Arg	Ile	Met	Thr	Ala	Tyr	Tyr	Lys	Val	Gly	Arg	Asp	Arg	Leu	Arg	Ile
			340					345					350		
Pro	Pro	Asn	Phe	Ser	Ser	Trp	Thr	Arg	Asp	Glu	Tyr	Gly	Trp	Glu	His
		355					360					365			
Ser	Ala	Val	Ser	Glu	Gly	Ala	Trp	Thr	Lys	Val	Asn	Asp	Phe	Val	Asn
	370					375					380				
Val	Gln	Arg	Ser	His	Ser	Gln	Ile	Ile	Arg	Glu	Ile	Gly	Ala	Ala	Ser
385					390					395					400
Thr	Val	Leu	Leu	Lys	Asn	Thr	Gly	Ala	Leu	Pro	Leu	Thr	Gly	Lys	Glu
				405					410					415	
Val	Lys	Val	Gly	Val	Leu	Gly	Glu	Asp	Ala	Gly	Ser	Asn	Pro	Trp	Gly
			420					425					430		
Ala	Asn	Gly	Cys	Pro	Asp	Arg	Gly	Cys	Asp	Asn	Gly	Thr	Leu	Ala	Met
		435					440					445			
Ala	Trp	Gly	Ser	Gly	Thr	Ala	Asn	Phe	Pro	Tyr	Leu	Val	Thr	Pro	Glu
	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
 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
 Tyr Asp Asn Pro Asn Val Thr Ala Ile Ile Trp Ala Gly Leu Pro Gly
 565 570
 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
 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

<210> 71
 <211> 2589
 <212> ADN
 <213> Aspergillus fumigatus

5

<400> 71

ES 2 393 058 T3

atgagattcg	gttggctcga	ggtggccgct	ctgacggccg	cttctgtagc	caatgcccag	60
gaattggctt	tctctccacc	attctaccct	tcgccttggg	ctgatggcca	gggagagtgg	120
gcagatgcc	atcgacgcgc	cgtcgagatc	gtttctcaga	tgacactggc	ggagaaggtt	180
aaccttacia	cgggtactgg	atgggaaatg	gaccgatgcg	tcggtcaaac	cggcagcgtt	240
cccagacttg	gtatcaactg	gggtctttgt	ggccaggatt	cccctttggg	tatccgtttc	300
tctgacctca	actccgcctt	ccctgctggt	actaatgtcg	ccgcgacatg	ggacaagaca	360
ctcgcctacc	ttcgtggcaa	ggccatgggt	gaggaattca	acgacaaggg	cgtggacatt	420
ttgctggggc	ctgctgctgg	tcctctcggc	aaatacccg	acggcggcag	aatctgggaa	480
ggcttctctc	ctgatccggc	tctcactggt	gtacttttcg	ccgaaactat	caagggtatc	540
caagacgcgg	gtgtgattgc	tactgccaag	cattacattc	tgaatgaaca	ggagcatttc	600
cgacaggttg	gcgaggccca	gggatatggt	tacaacatca	cggagacgat	cagctccaac	660
gtggatgaca	agaccatgca	cgagttgtac	ctttggccct	ttgcagatgc	tgtgcgcgct	720
ggcgttggcg	ctgtcatgtg	ttcctacaat	caaatcaaca	acagctacgg	ttgtcaaaac	780
agtcaaactc	tcaacaagct	cctcaaggct	gagctgggct	tccaaggctt	cgtcatgagt	840
gactggagcg	ctcaccacag	cgggtgctggc	gctgccctcg	ctgggttggg	tatgtcgatg	900
cccggagaca	tttccttcga	cgacggactc	tccttctggg	gcacgaacct	aactgtcagt	960
gttcttaacg	gcaccgttcc	agcctggcgt	gtcgatgaca	tggctgttcg	tatcatgacc	1020

ES 2 393 058 T3

gcgtactaca aggttggtcg tgaccgtctt cgtattcccc ctaacttcag ctcttgacc 1080
 cgggatgagt acggctggga gcattctgct gtctccgagg gagcctggac caaggtgaac 1140
 gacttcgtca atgtgcagcg cagtcactct cagatcatcc gtgagattgg tgccgctagt 1200
 acagtgctct tgaagaacac gggtgctctt cctttgaccg gcaaggagg taaagtgggt 1260
 gttctcggtg aagacgctgg ttccaacccg tggggtgcta acggctgccc cgaccgcggc 1320
 tgtgataacg gcactcttgc tatggcctgg ggtagtggta ctgccaaactt cccttacctt 1380
 gtcacccccg agcaggctat ccagcgagag gtcacagca acggcggcaa tgtctttgct 1440
 gtgactgata acggggctct cagccagatg gcagatggtg catctcaatc cagcgtgtct 1500
 ttggtgtttg tcaacgccga ctctggagag ggtttcatca gtgtcgacgg caacgaggg 1560
 gaccgcaaaa atctcactct gtggaagaac ggcgaggccg tcattgacac tgttgtcagc 1620
 cactgcaaca acacgattgt ggttattcac agtgttgggc ccgtcttgat cgaccgggtg 1680
 tatgataacc ccaacgtcac tgccatcatc tgggccggct tgcccggca ggagagtggc 1740
 aactccctgg tcgacgtgct ctatggccgc gtcaacccca gcgccaagac cccgttcacc 1800
 tggggcaaga ctcgggagtc ttacggggct cccttgctca ccgagccta caatggcaat 1860
 ggtgctcccc aggatgattt caacgagggc gtcttcattg actaccgtca ctttgacaag 1920
 cgcaatgaga cccccattta tgagtttggc catggcttga gctacaccac ctttggttac 1980
 tctcaccttc gggttcaggc cctcaatagt tcgagttcgg catatgtccc gactagcgga 2040
 gagaccaagc ctgcccacac ctatggtgag atcggtagtg ccgccgacta cctgtatccc 2100
 gagggcttca aaagaattac caagtttatt tacccttggc tcaactcgac cgacctcgag 2160
 gattcttctg acgaccgaa ctacggctgg caggactcgg agtacattcc cgaaggcgct 2220
 agggatgggt ctctcaacc cctcctgaag gctggcggcg ctctggtgg taaccctacc 2280
 ctttatcagg atcttgtag ggtgtcggcc accataacca aactggtaa cgtcgcgggt 2340
 tatgaagtcc ctcaattgta tgtttcattg ggcggaccga acgagcctcg ggtcgttctg 2400
 cgcaagttcg accgaatctt cctggctcct ggggagcaaa aggtttggac cagactctt 2460
 aaccgtcgtg atctcgccaa ttgggatgtg gaggctcagg actgggtcat cacaaagtac 2520
 cccaagaaag tgcacgtcgg cagctcctcg cgtaagctgc ctctgagagc gcctctgccc 2580
 cgtgtctac 2589

<210> 72
 <211> 29
 <212> ADN
 <213> Aspergillus fumigatus

<400> 72
 ggctcatgag attcggtgg ctgaggtc 29

<210> 73
 <211> 30

<212> ADN
 <213> Aspergillus fumigatus

 <400> 73
 5 gccgttatca cagccgcggt cggggcagcc 30

 <210> 74
 <211> 30
 <212> ADN
 10 <213> Aspergillus fumigatus

 <400> 74
 ggctgccccg accgcgctg tgataacggc 30

 <210> 75
 <211> 35
 <212> ADN
 <213> Aspergillus fumigatus

 <400> 75
 20 gcttaattaa tctagtagac acggggcaga ggcgc 35

 <210> 76
 <211> 16
 <212> ADN
 25 <213> Aspergillus fumigatus

 <400> 76
 30 aactggcgg agaagg 16

 <210> 77
 <211> 18
 <212> ADN
 <213> Aspergillus fumigatus
 35
 <400> 77
 gccagggat atggttac 18

 <210> 78
 <211> 19
 <212> ADN
 <213> Aspergillus fumigatus
 40
 <400> 78
 45 cgactctgga gagggttc 19

 <210> 79
 <211> 19
 <212> ADN
 <213> Aspergillus fumigatus
 50
 <400> 79
 ggactgggtc atcacaag 19

 <210> 80
 <211> 17
 <212> ADN
 <213> Aspergillus fumigatus
 55
 <400> 80
 60 gcgagaggtc atcagca 17

 <210> 81
 <211> 17
 <212> ADN
 <213> Aspergillus fumigatus
 65

<400> 81
 gtaaacgac ggccagt 17
 5 <210> 82
 <211> 16
 <212> ADN
 <213> Aspergillus fumigatus
 10 <400> 82
 caggaaacag ctatga 16
 <210> 83
 <211> 69
 15 <212> ADN
 <213> Aspergillus fumigatus
 <400> 83
cttcttgta gtgcaatc atatagaagt catcgactag tggatctacc atgagattcg 60
gttggctcg 69
 20 <210> 84
 <211> 61
 <212> ADN
 <213> Aspergillus fumigatus
 25 <400> 84
gcgtgaatgt aagcgtgaca taactaatta catgactcga gctagtagac acggggcaga 60
g 61
 <210> 85
 <211> 60
 30 <212> ADN
 <213> Aspergillus fumigatus
 <400> 85
 ccgctccgcc gttgtggccg cctgcccgtt gttggccctt gccgaattgg ctttctctcc 60
 35 <210> 86
 <211> 17
 <212> ADN
 <213> Aspergillus fumigatus
 40 <400> 86
 ctggcgttg cgctgc 17
 <210> 87
 <211> 30
 45 <212> ADN
 <213> Aspergillus fumigatus
 <400> 87
 50 gcggcagaat ctggaaagc ttctctctg 30
 <210> 88
 <211> 30
 <212> ADN
 55 <213> Aspergillus fumigatus
 <400> 88
 caggagagaa gcttcccag attctgccg 30

ES 2 393 058 T3

<210> 89
<211> 29
<212> ADN
<213> Aspergillus fumigatus
5
<400> 89
tgactggagc gctcaacaca gcggtgtcg 29
10
<210> 90
<211> 29
<212> ADN
<213> Aspergillus fumigatus
15
<400> 90
cgacaccgct gtgttgagcg ctccagtca 29
20
<210> 91
<211> 30
<212> ADN
<213> Aspergillus fumigatus
<400> 91
gaggattctt ctggcgaccc gaactacggc 30
25
<210> 92
<211> 30
<212> ADN
<213> Aspergillus fumigatus
30
<400> 92
gccgtagtc gggtcgccag aagaatcctc 30

REIVINDICACIONES

- 5 1. Variante aislada de una beta-glucosidasa progenitora, que comprende una sustitución a una posición correspondiente a la posición 285 de los aminoácidos 20 a 861 de SEC ID n°: 2 o correspondiente a la posición 285 de los aminoácidos 20 a 863 de SEC ID n°: 70, donde la variante tiene actividad de beta-glucosidasa y comprende las sustituciones Q202R + H285Q + D722G (o D724G), las sustituciones Q202R + H285Q; o las sustituciones H285Q + D722G (o D724G) en comparación con la beta-glucosidasa progenitora.
- 10 2. Variante según la reivindicación 1, donde la beta-glucosidasa progenitora es (a) un polipéptido que comprende una secuencia de aminoácidos que tiene al menos 70% de identidad, preferiblemente al menos 75% de identidad, más preferiblemente al menos 80% de identidad, incluso más preferiblemente al menos 85% de identidad, de la forma más preferible al menos 90% de identidad, e de la forma incluso más preferible al menos 95% de identidad con los aminoácidos 20 a 861 de SEC ID n°: 2 o aminoácidos 20 a 863 de la SEC ID n°: 70 o (b) un polipéptido codificado por una secuencia de nucleótidos que se hibrida bajo condiciones de astringencia baja, preferiblemente condiciones de astringencia media, más preferiblemente condiciones de astringencia media alta, y de forma más preferible condiciones de astringencia alta con los nucleótidos 58 a 2583 de la SEC ID n°: 1 o los nucleótidos 58 a 2589 de la SEC ID n°: 71, o sus hebras complementarias.
- 20 3. Variante según la reivindicación 1, donde la beta-glucosidasa progenitora comprende la secuencia de aminoácidos de los aminoácidos 20 a 861 de SEC ID n°: 2 o aminoácidos 20 a 863 de la SEC ID n°: 70.
4. Variante según la reivindicación 1, donde la beta-glucosidasa progenitora consiste en la secuencia de aminoácidos de los aminoácidos 20 a 861 de SEC ID n°: 2 o aminoácidos 20 a 863 de la SEC ID n°: 70.
- 25 5. Variante según cualquiera de las reivindicaciones 1 a 4, que comprende una secuencia de aminoácidos con al menos 70% de identidad, preferiblemente al menos 75% de identidad, más preferiblemente al menos 80% de identidad, incluso más preferiblemente al menos 85% de identidad, de forma más preferible al menos 90% de identidad, e incluso de forma más preferible al menos 95% de identidad a la secuencia de aminoácidos de la beta-glucosidasa progenitora.
- 30 6. Variante según cualquiera de las reivindicaciones 1 a 5, que tiene una o más propiedades mejoradas en comparación con la beta-glucosidasa progenitora, donde las propiedades mejoradas son seleccionadas del grupo que consiste en actividad térmica, termoestabilidad, actividad de pH, estabilidad de pH, especificidad del sustrato, especificidad del producto y estabilidad química.
- 35 7. Variante según la reivindicación 6, donde la actividad térmica de la beta-glucosidasa variante es al menos 1,5 veces, preferiblemente al menos 2 veces, más preferiblemente al menos 5 veces, de forma más preferible al menos 7 veces, e incluso de forma más preferible al menos 20 veces más térmicamente activa que la beta-glucosidasa progenitora de la variante.
- 40 8. Secuencia de nucleótidos aislada que codifica la variante de cualquiera de las reivindicaciones 1 a 7.
9. Constructo de ácidos nucleicos que comprende la secuencia de nucleótidos según la reivindicación 8.
- 45 10. Vector de expresión que comprende el constructo de ácidos nucleicos según la reivindicación 9.
11. Célula huésped que comprende el constructo de ácidos nucleicos según la reivindicación 9.
- 50 12. Método para la producción de una variante de una beta-glucosidasa progenitora, que comprende:
 - (a) cultivo de la célula huésped según la reivindicación 11 bajo condiciones adecuadas para la expresión de la variante; y
 - (b) recuperación de la variante del medio de cultivo.
- 55 13. Método para degradar o convertir biomasa con contenido de celulosa y hemicelulosa, que comprende tratar la biomasa con una cantidad eficaz de una variante de beta-glucosidasa de cualquiera de las reivindicaciones 1 a 7 y recuperar la biomasa degradada.
- 60 14. Método según la reivindicación 13, que comprende además tratar la biomasa con una cantidad eficaz de endo-1,4-beta-glucanasa y exo-1,4-beta-D-glucanasa.
15. Método para degradar o convertir una biomasa con contenido de celulosa y hemicelulosa, que comprende tratar la biomasa con una célula huésped según la reivindicación 11 y recuperar la biomasa degradada.
- 65 16. Método según la reivindicación 15, que comprende además tratar la biomasa con una cantidad eficaz de endo-1,4-beta-glucanasa y exo-1,4-beta-D-glucanasa.

17. Planta que comprende la secuencia de nucleótidos según la reivindicación 8.

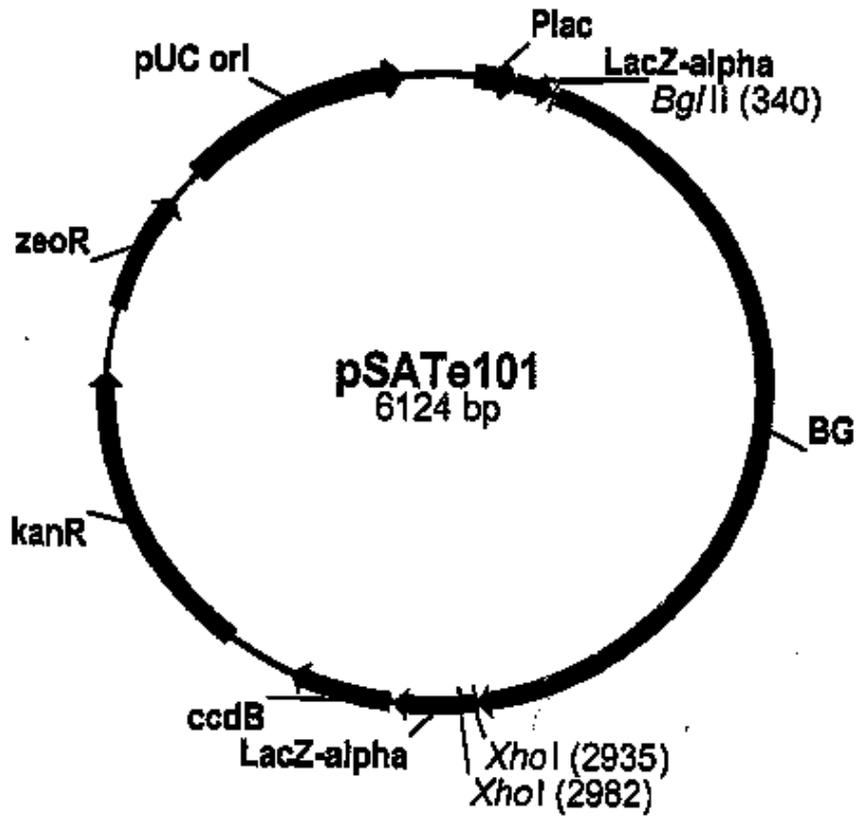


Fig. 1

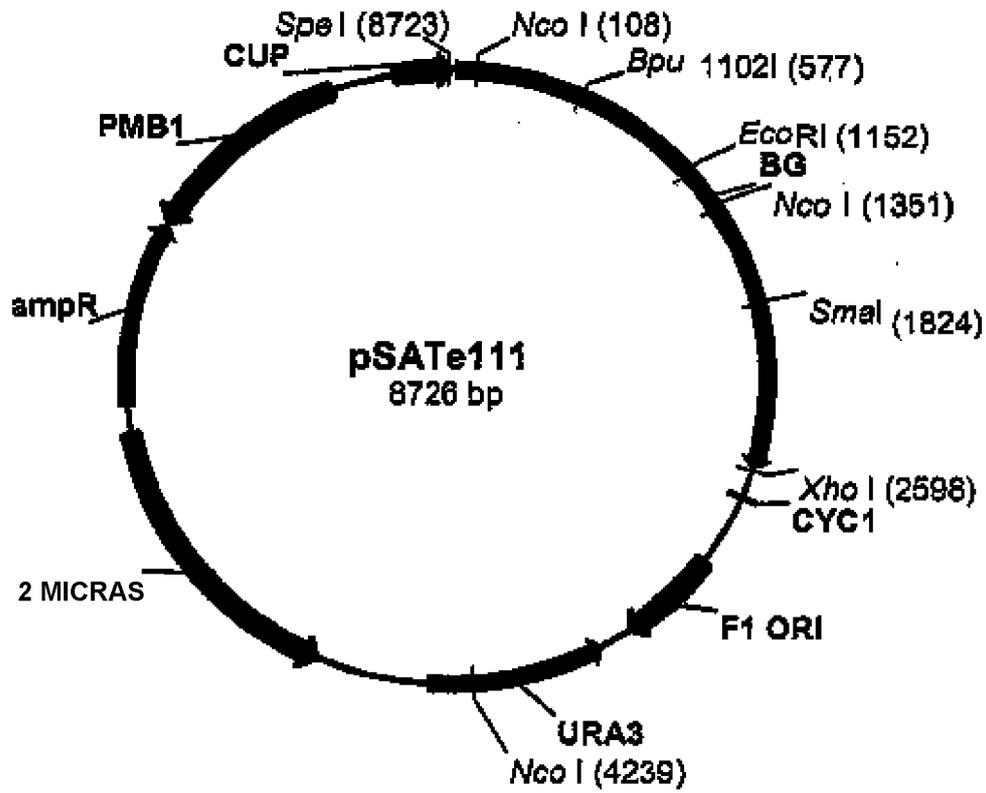


Fig. 2

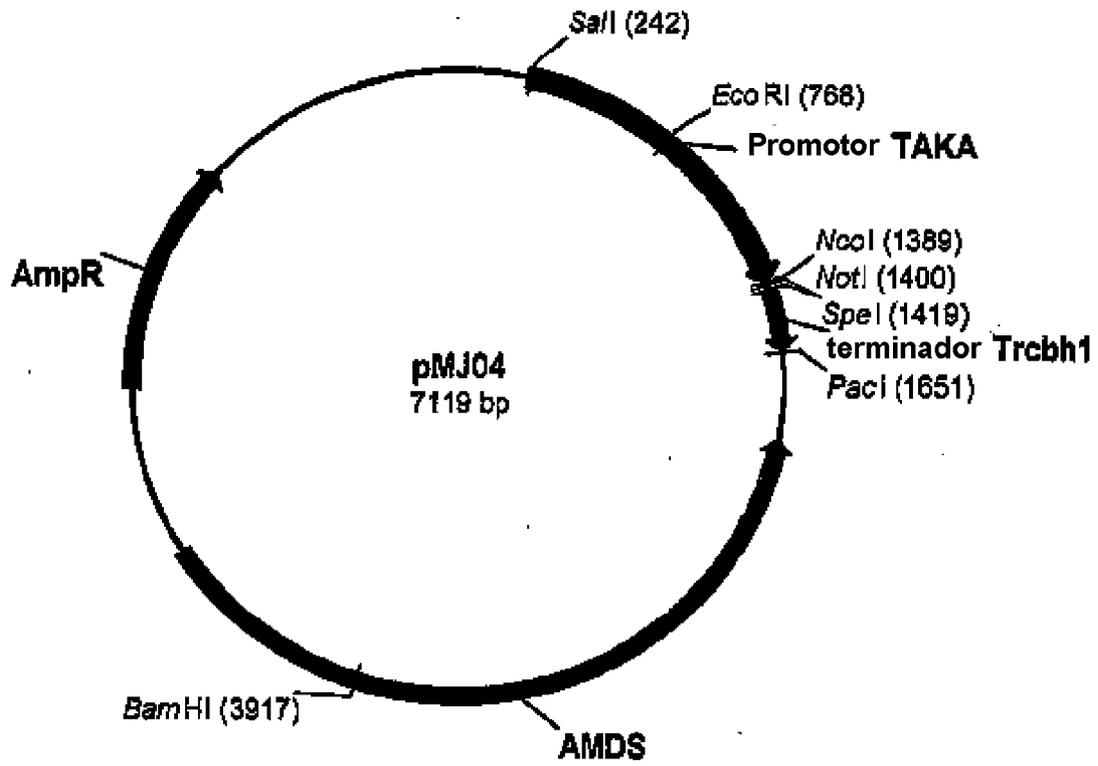


Fig. 3

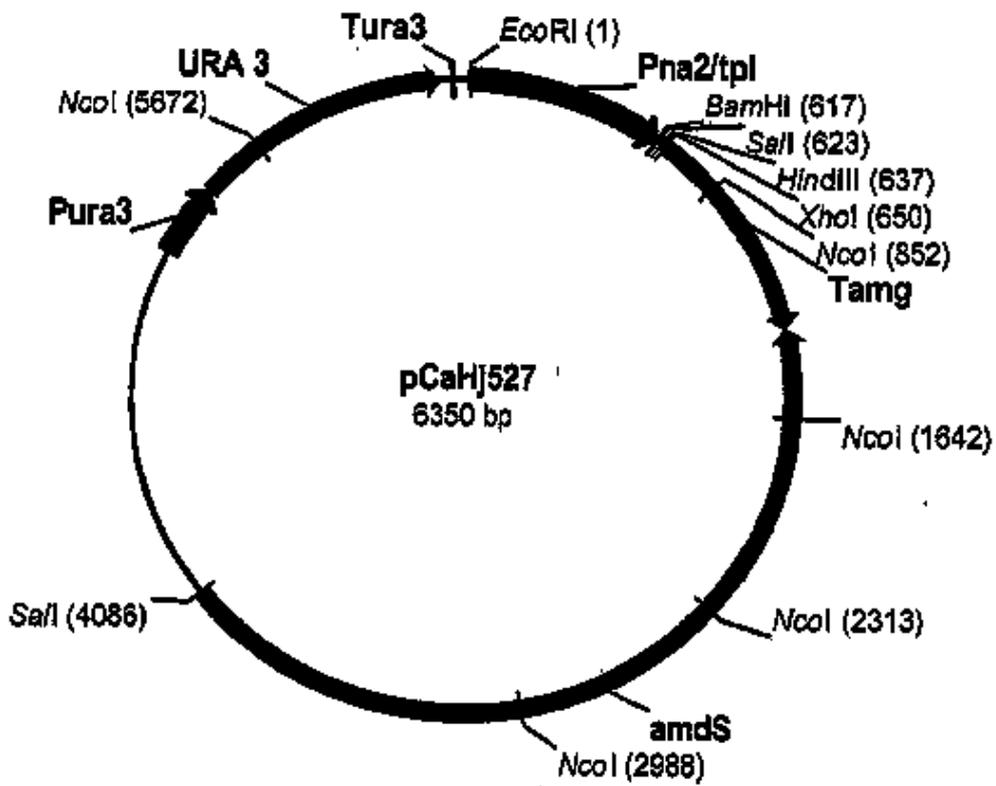


Fig. 4

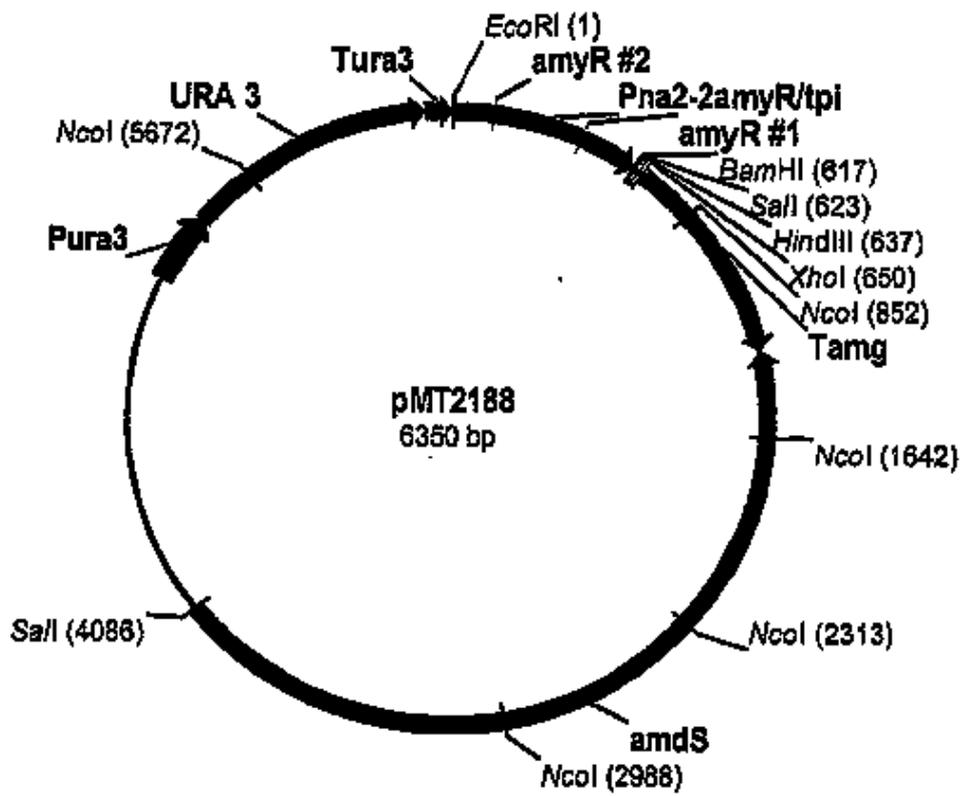


Fig. 5

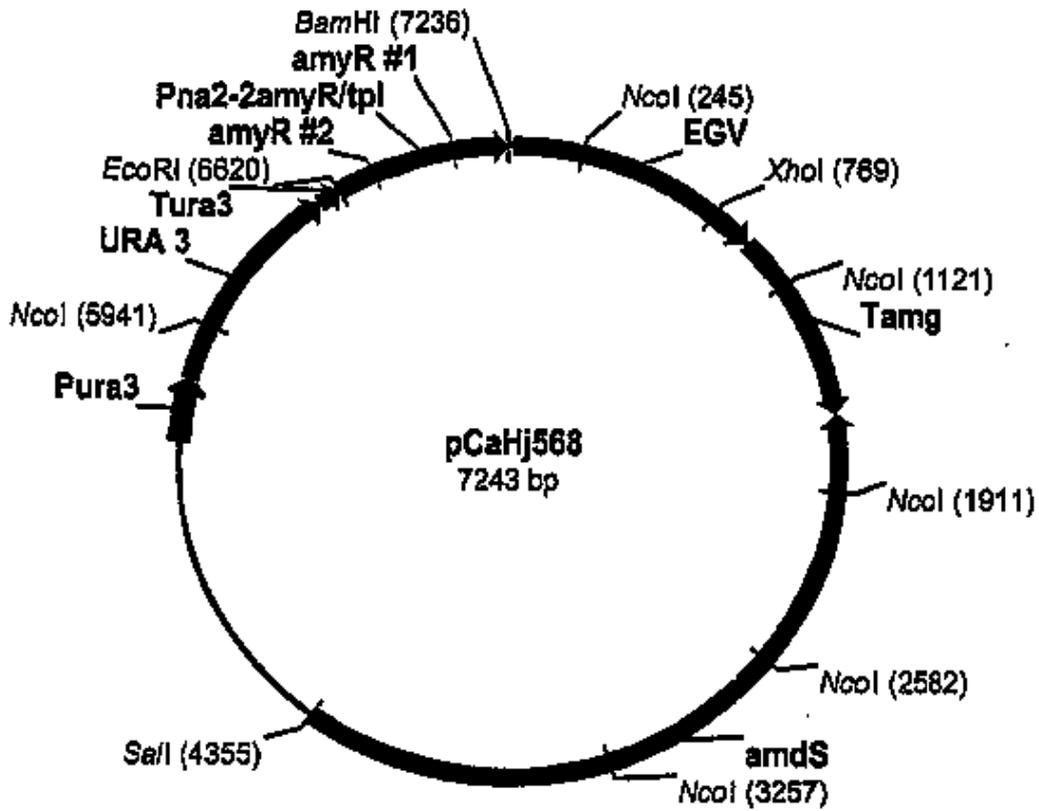


Fig. 6

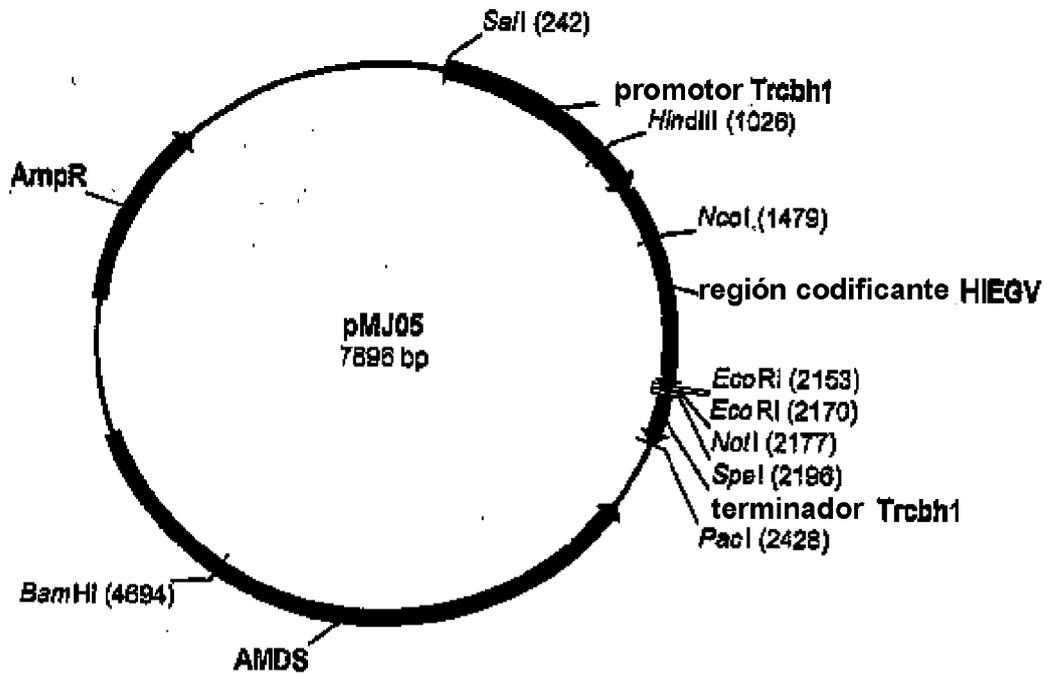


Fig. 7

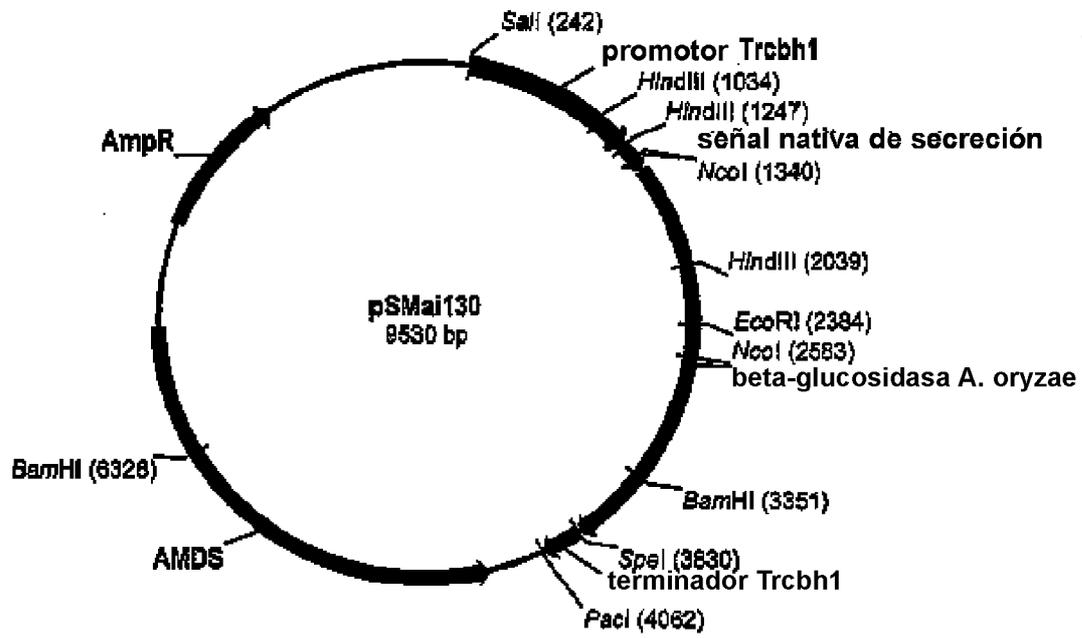


Fig. 8

atg aag ctt ggt tgg atc gag gtg gcc gca ttg gcg gct gcc tca gta 48
 Met Lys Leu Gly Trp Ile Glu Val Ala Ala Leu Ala Ala Ser Val 15
 1 5 10 15
 gtc agt gcc aag gat gat ctc gcg tac tcc cct cct ttc tac cct tcc 96
 Val Ser Ala Lys Asp Asp Leu Ala Tyr Ser Pro Pro Phe Tyr Pro Ser 30
 20 25 30
 cca tgg gca gat ggt cag ggt gaa tgg gcg gaa gta tac aaa cgc gct 144
 Pro Trp Ala Asp Gly Gln Glu Trp Ala Glu Val Tyr Lys Arg Ala 45
 35 40 45
 gta gac ata gtt tcc cag atg acg ttg aca gag aaa gtc aac tta acg 192
 Val Asp Ile Val Ser Gln Met Thr Leu Thr Glu Lys Val Asn Leu Thr 60
 50 55 60
 act gga aca gga tgg caa cta gag agg tgt gtt gga caa act ggc agt 240
 Thr Gly Thr Gly Trp Gln Leu Glu Arg Cys Val Gly Gln Thr Gly Ser 80
 65 70 75 80
 gtt ccc aga ctc aac atc ccc agc ttg tgt ttg cag gat agt cct ctt 288
 Val Pro Arg Leu Asn Ile Pro Ser Leu Cys Leu Gln Asp Ser Pro Leu 95
 85 90 95

Fig. 9A

ggt att cgt ttc tcg gac tac aat tca gct ttc cct gcg ggt gtt aat 336
 Gly Ile Arg Phe Ser Asp Tyr Asn Ser Ala Phe Pro Ala Gly Val Asn
 100 105 110
 gtc gct gcc acc tgg gac aag acg ctc gcc tac ctt cgt ggt cag gca 384
 Val Ala Ala Thr Trp Asp Lys Thr Leu Ala Tyr Leu Arg Gly Gln Ala
 115 120 125
 atg ggt gag gag ttc agt gat aag ggt att gac gtt cag ctg ggt cct 432
 Met Gly Glu Glu Phe Ser Asp Lys Gly Ile Asp Val Gln Leu Gly Pro
 130 135 140
 gct gct ggc cct ctc ggt gct cat ccg gat ggc ggt aga aac tgg gaa 480
 Ala Ala Gly Pro Leu Gly Ala His Pro Asp Gly Gly Arg Asn Trp Glu
 145 150 155 160
 ggt ttc tca cca gat cca gcc ctc acc ggt gta ctt ttt gcg gag acg 528
 Gly Phe Ser Pro Asp Pro Ala Leu Thr Gly Val Leu Phe Ala Glu Thr
 165 170 175
 att aag ggt att caa gat gct ggt gtc att gcg aca gct aag cat tat 576
 Ile Lys Gly Ile Gln Asp Ala Gly Val Ile Ala Thr Ala Lys His Tyr
 180 185 190

Fig. 9B

atc atg aac gaa caa gag cat ttc cgc caa caa ccc gag gct gcg ggt Ile Met Asn Glu Gln Glu His Phe Arg Gln Gln Pro Glu Ala Ala Gly 195 200 205	624
tac gga ttc aac gta agc gac agt ttg agt tcc aac gtt gat gac aag Tyr Gly Phe Asn Val Ser Asp Ser Leu Ser Ser Asn Val Asp Asp Lys 210 215 220	672
act atg cat gaa ttg tac ctc tgg ccc ttc gcg gat gca gta cgc gct Thr Met His Glu Leu Tyr Leu Trp Pro Phe Ala Asp Ala Val Arg Ala 225 230 235	720
gga gtc ggt gct gtc atg tgc tct tac aac caa atc aac aac agc tac Gly Val Gly Ala Val Met Cys Ser Tyr Asn Gln Ile Asn Asn Ser Tyr 245 250 255	768
ggt tgc gag aat agc gaa act ctg aac aag ctt ttg aag gcg gag ctt Gly Cys Glu Asn Ser Glu Thr Leu Asn Lys Leu Leu Lys Ala Glu Leu 260 265 270	816
ggt ttc caa ggc ttc gtc atg agt gat tgg acc gct cat cac agc ggc Gly Phe Gln Gly Phe Val Met Ser Asp Trp Thr Ala His His Ser Gly 275 280 285	864

Fig. 9C

912
 gta ggc gct gct tta gca ggt ctg gat atg tcg atg ccc ggt gat gtt
 Val Gly Ala Ala Leu Ala Gly Leu Asp Met Ser Met Pro Gly Asp Val
 290 295 300
 960
 acc ttc gat agt ggt acg tct ttc tgg ggt gca aac ttg acg gtc ggt
 Thr Phe Asp Ser Gly Thr Ser Phe Trp Gly Ala Asn Leu Thr Val Gly
 305 310 315 320
 1008
 gtc ctt aac ggt aca atc ccc caa tgg cgt gtt gat gac atg gct gtc
 Val Leu Asn Gly Thr Ile Pro Gln Trp Arg Val Asp Asp Met Ala Val
 325 330 335
 1056
 cgt atc atg gcc gct tat tac aag gtt ggc cgc gac acc aaa tac acc
 Arg Ile Met Ala Ala Tyr Tyr Lys Val Gly Arg Asp Thr Lys Tyr Thr
 340 345 350
 1104
 cct ccc aac ttc agc tcg tgg acc agg gac gaa tat ggt ttc gcg cat
 Pro Pro Asn Phe Ser Ser Trp Thr Arg Asp Glu Tyr Gly Phe Ala His
 355 360 365
 1152
 aac cat gtt tcg gaa ggt gct tac gag agg gtc aac gaa ttc gtg gac
 Asn His Val Ser Glu Gly Ala Tyr Glu Arg Val Asn Glu Phe Val Asp
 370 375 380

Fig. 9D

1200
 gtg caa cgc gat cat gcc gac cta atc cgt cgc atc ggc gcg cag agc
 Val Gln Arg Asp His Ala Asp Leu Ile Arg Arg Ile Gly Ala Gln Ser
 385 390 395 400
 act gtt ctg ctg aag aac aag ggt gcc ttg ccc ttg agc cgc aag gaa
 Thr Val Leu Leu Lys Asn Lys Lys Gly Ala Leu Pro Leu Ser Arg Lys Glu
 405 410 415
 1248
 aag ctg gtc gcc ctt ctg gga gag gat gcg ggt tcc aac tcg tgg ggc
 Lys Leu Val Ala Leu Leu Gly Glu Asp Ala Gly Ser Asn Ser Trp Gly
 420 425 430
 1296
 gct aac ggc tgt gat gac cgt ggt tgc gat aac ggt acc ctt gcc atg
 Ala Asn Gly Cys Asp Arg Gly Cys Asp Asn Gly Thr Leu Ala Met
 435 440 445
 1344
 gcc tgg ggt agc ggt act gcg aat ttc cca tac ctc gtg aca cca gag
 Ala Trp Gly Ser Gly Thr Ala Asn Phe Pro Tyr Leu Val Thr Pro Glu
 450 455 460
 1392
 cag gcg att cag aac gaa gtt ctt cag ggc cgt ggt aat gtc ttc gcc
 Gln Ala Ile Gln Asn Glu Val Leu Gln Gly Arg Gly Asn Val Phe Ala
 465 470 475 480
 1440

Fig. 9E

1488
 gtg acc gac agt tgg gcg ctc gac aag atc gct gcg gct gcc cgc cag
 Val Thr Asp Ser Trp Ala Leu Asp Lys Ile Ala Ala Ala Arg Gln
 485 490 495
 gcc agc gta tct ctc gtg ttc gtc aac tcc gac tca gga gaa ggc tat
 Ala Ser Val Ser Leu Val Phe Val Asn Ser Asp Ser Gly Glu Gly Tyr
 500 505 510
 1536
 ctt agt gtg gat gga aat gag ggc gat cgt aac aac atc act ctg tgg
 Leu Ser Val Asp Gly Asn Glu Gly Asp Arg Asn Asn Ile Thr Leu Trp
 515 520 525
 aag aac ggc gac aat gtg gtc aag acc gca gcg aat aac tgt aac aac
 Lys Asn Gly Asp Asn Val Val Lys Thr Ala Ala Asn Asn Cys Asn Asn
 530 535 540
 1632
 acc gtt gtc atc atc cac tcc gtc gga cca gtt ttg atc gat gaa tgg
 Thr Val Val Ile Ile His Ser Val Gly Pro Val Leu Ile Asp Glu Trp
 545 550 555 560
 1680
 tat gac cac ccc aat gtc act ggt att ctc tgg gct ggt ctg cca ggc
 Tyr Asp His Pro Asn Val Thr Gly Ile Leu Trp Ala Gly Leu Pro Gly
 565 570 575
 1728

Fig. 9F

1776
 cag gag tct ggt aac tcc att gcc gat gtg ctg tac ggt cgt gtc aac
 Gln Glu Ser Gly Asn Ser Ile Ala Asp Val Leu Tyr Gly Arg Val Asn
 580 585 590

 1824
 cct ggc gcc aag tct cct ttc act tgg ggc aag acc cgg gag tcg tat
 Pro Gly Ala Lys Ser Pro Phe Thr Trp Gly Lys Thr Arg Glu Ser Tyr
 595 600 605

 1872
 ggt tct ccc ttg gtc aag gat gcc aac aat ggc aac gga gcg ccc cag
 Gly Ser Pro Leu Val Lys Asp Ala Asn Asn Gly Asn Gly Ala Pro Gln
 610 615 620

 1920
 tct gat ttc acc cag ggt gtt ttc atc gat tac cgc cat ttc gat aag
 Ser Asp Phe Thr Gln Gly Val Phe Ile Asp Tyr Arg His Phe Asp Lys
 625 630 635 640

 1968
 ttc aat gag acc cct atc tac gag ttt ggc tac ggc ttg agc tac acc
 Phe Asn Glu Thr Pro Ile Tyr Glu Phe Gly Tyr Gly Leu Ser Tyr Thr
 645 650 655

 2016
 acc ttc gag ctc tcc gac ctc cat gtt cag ccc ctg aac gcg tcc cga
 Thr Phe Glu Leu Ser Asp Leu His Val Gln Pro Leu Asn Ala Ser Arg
 660 665 670

Fig. 9G

tac act ccc acc agt ggc atg act gaa gct gca aag aac ttt ggt gaa	2064
Tyr Thr Pro Thr Ser Gly Met Thr Glu Ala Ala Lys Asn Phe Gly Glu	
675 680 685	
att ggc gat gcg tcg gag tac gtg tat ccg gag ggg ctg gaa agg atc	2112
Ile Gly Asp Ala Ser Glu Tyr Val Tyr Pro Glu Gly Leu Glu Arg Ile	
690 695 700	
cat gag ttt atc tat ccc tgg atc aac tct acc gac ctg aag gca tcg	2160
His Glu Phe Ile Tyr Pro Trp Ile Asn Ser Thr Asp Leu Lys Ala Ser	
705 710 715 720	
tct gac gat tct aac tac ggc tgg gaa gac tcc aag tat att ccc gaa	2208
Ser Asp Asp Ser Asn Tyr Gly Trp Glu Asp Ser Lys Tyr Ile Pro Glu	
725 730 735	
ggc gcc acg gat ggg tct gcc cag ccc cgt ttg ccc gct agt ggt ggt	2256
Gly Ala Thr Asp Gly Ser Ala Gln Pro Arg Leu Pro Ala Ser Gly Gly	
740 745 750	
gcc gga gga aac ccc ggt ctg tac gag gat ctt ttc cgc gtc tct gtg	2304
Ala Gly Gly Asn Pro Gly Leu Tyr Glu Asp Leu Phe Arg Val Ser Val	
755 760 765	

Fig. 9H

aag gtc aag aac acg ggc aat gtc gcc ggt gat gaa gtt cct cag ctg Lys Val Lys Asn Thr Gly Asn Val Ala Gly Asp Glu Val Pro Gln Leu 770 775 780	2352
tac gtt tcc cta ggc ggc cct aat gag ccc aag gtg gta ctg cgc aag Tyr Val Ser Leu Gly Gly Pro Asn Glu Pro Lys Val Val Leu Arg Lys 785 790 795 800	2400
ttt gag cgt att cac ttg gcc cct tcg cag gag gcc gtg tgg aca acg Phe Glu Arg Ile His Leu Ala Pro Ser Gln Glu Ala Val Trp Thr Thr 805 810 815	2448
acc ctt acc cgt cgt gac ctt gca aac tgg gac gtt tcg gct cag gac Thr Leu Thr Arg Arg Asp Leu Ala Asn Trp Asp Val Ser Ala Gln Asp 820 825 830	2496
tgg acc gtc act cct tac ccc aag acg atc tac gtt gga aac tcc tca Trp Thr Val Thr Pro Tyr Pro Lys Thr Ile Tyr Val Gly Asn Ser Ser 835 840 845	2544
cgg aaa ctg ccg ctc cag gcc tcg ctg cct aag gcc cag taa Arg Lys Leu Pro Leu Gln Ala Ser Leu Pro Lys Ala Gln 850 855 860	2586

Fig. 91

ATG CGT TCC TCC CCC CTC CGC TCC GCC GTT GTG GCC GTG COG GTG TTG GCC CTT GCC
M R S S TCC P L L R S A V V A A L P V L A L A

Fig. 10

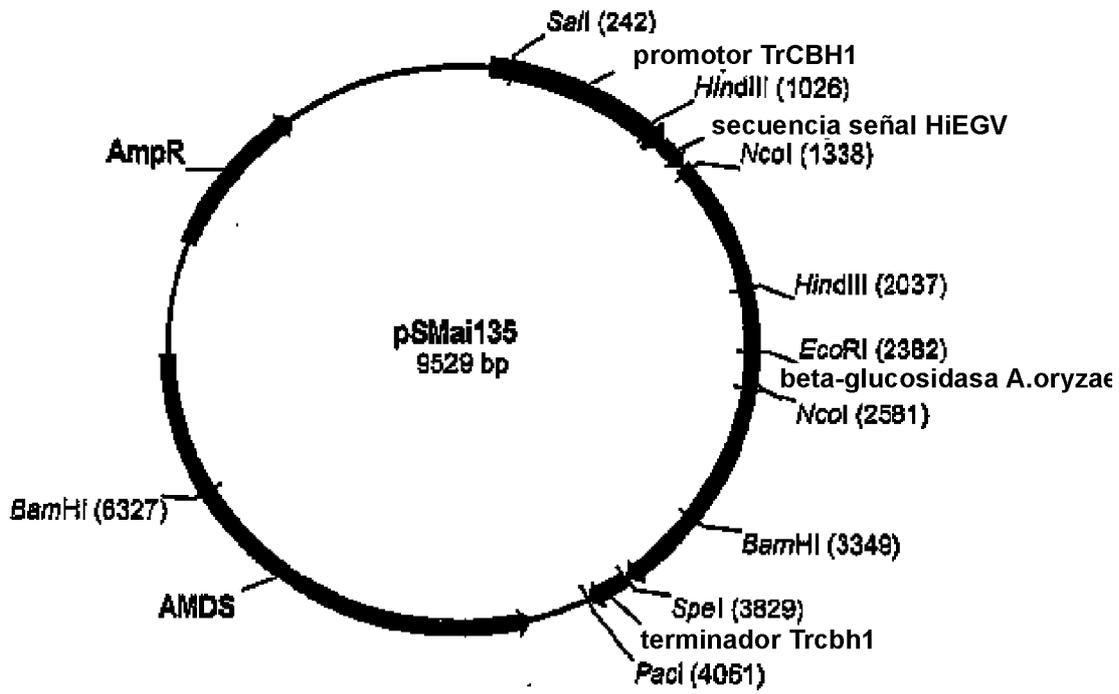


Fig. 11

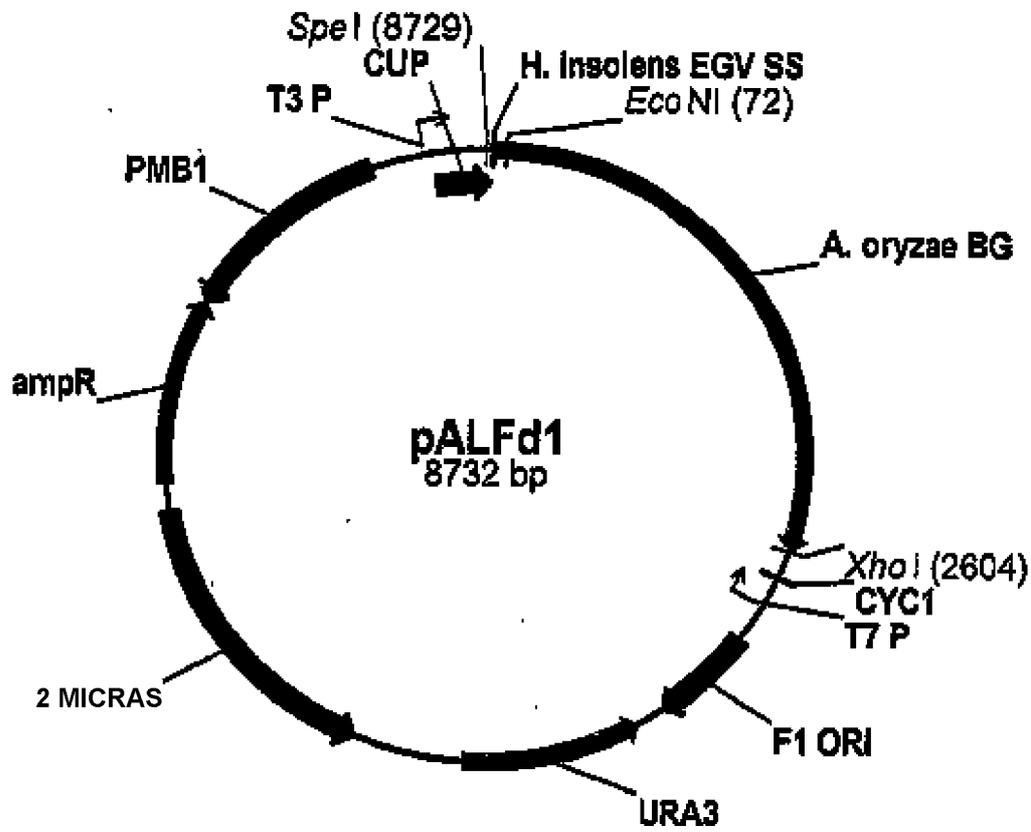


Fig. 12

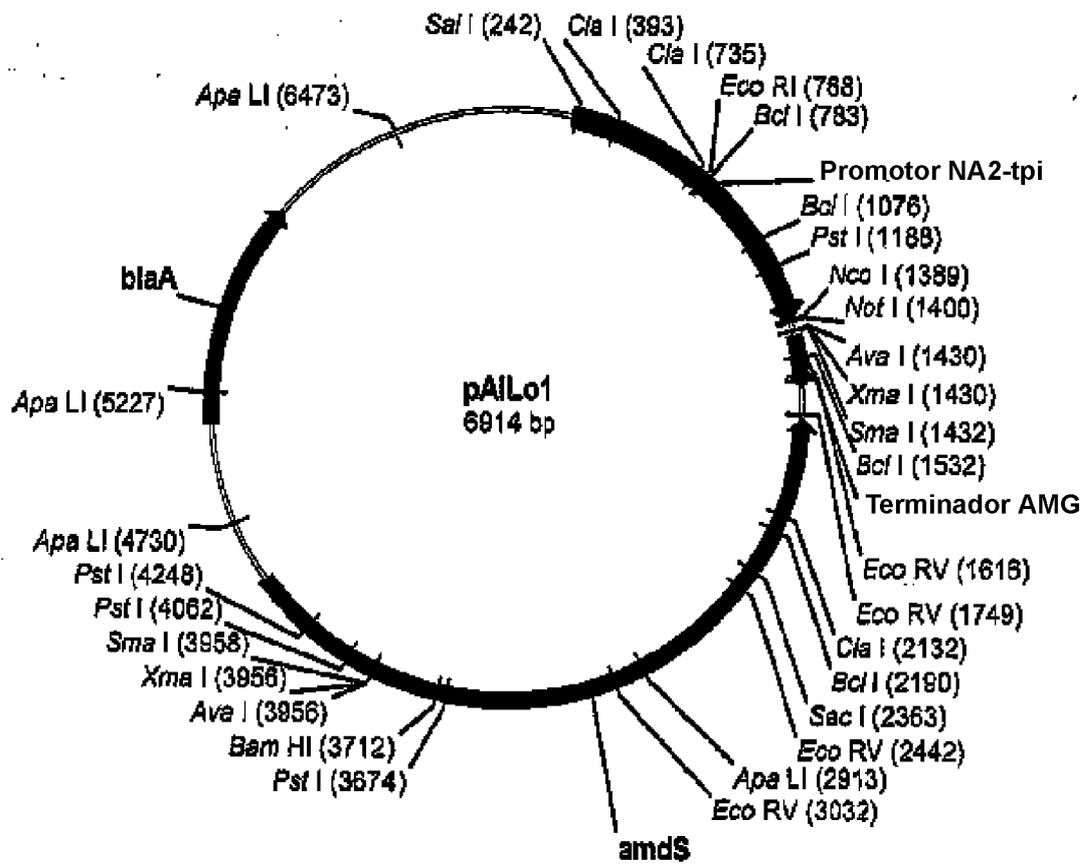


Fig. 13

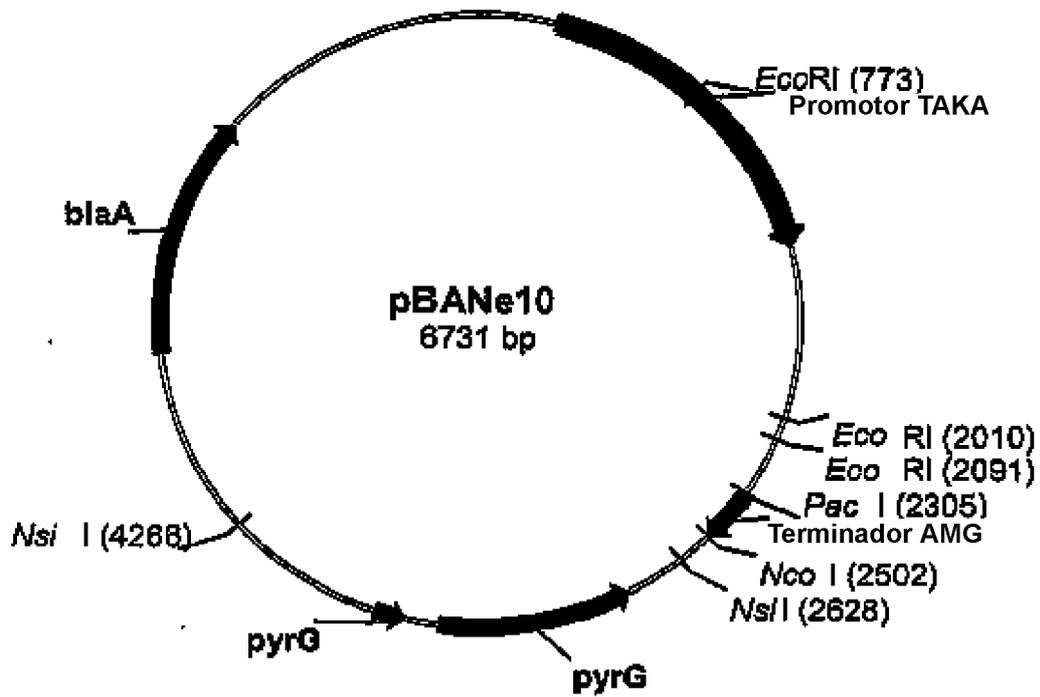


Fig. 14

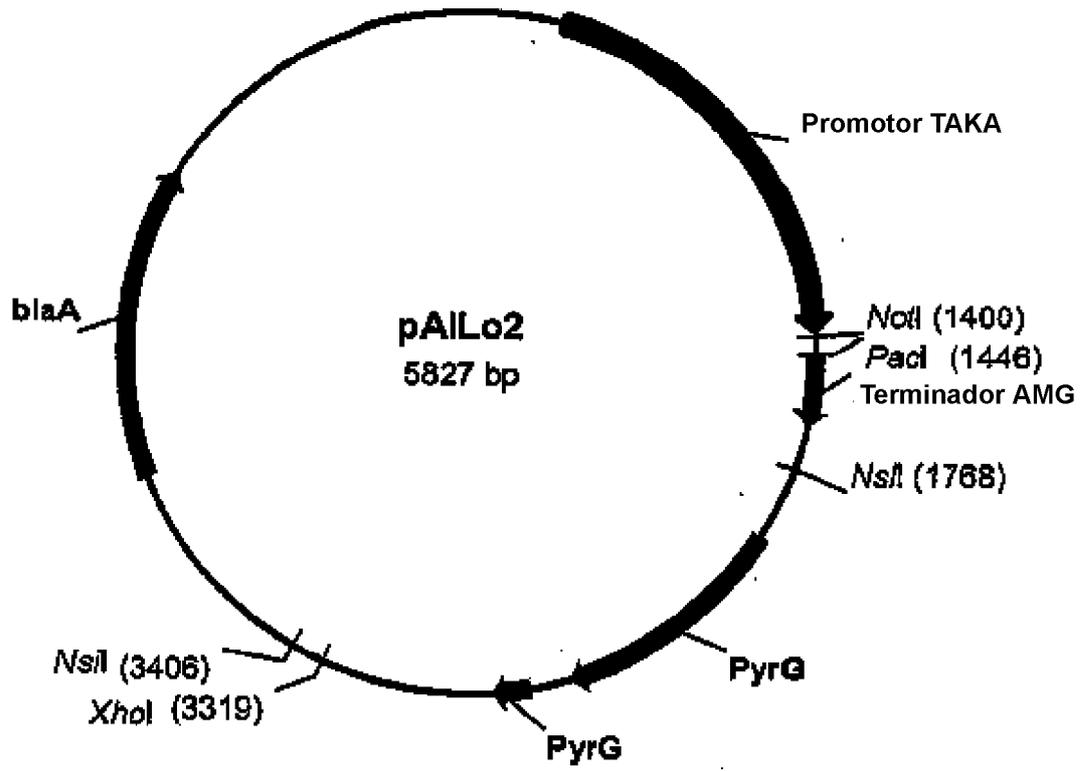


Fig. 15

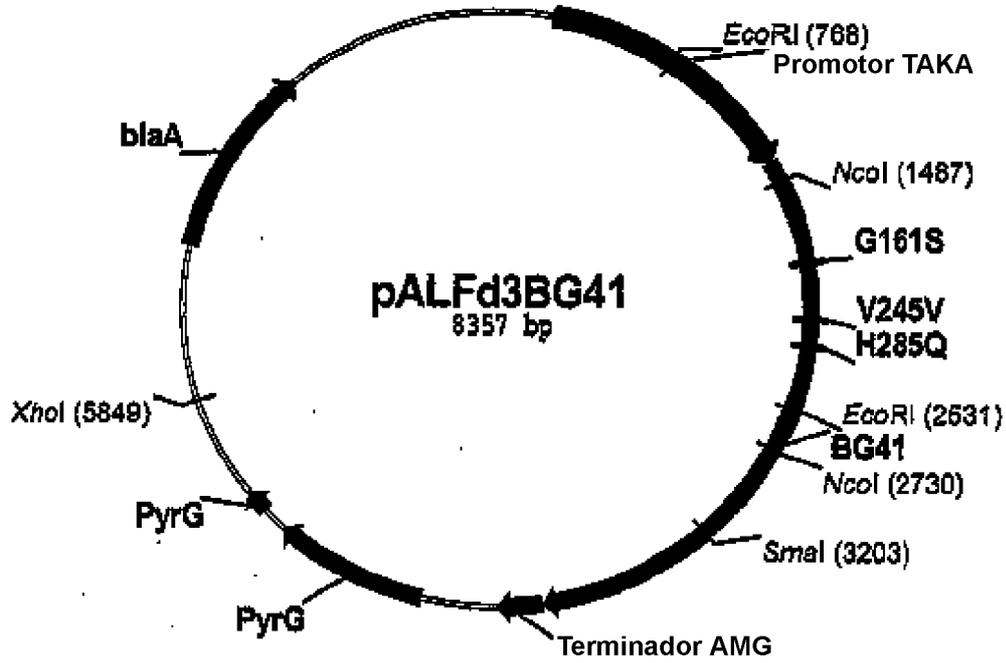


Fig. 16

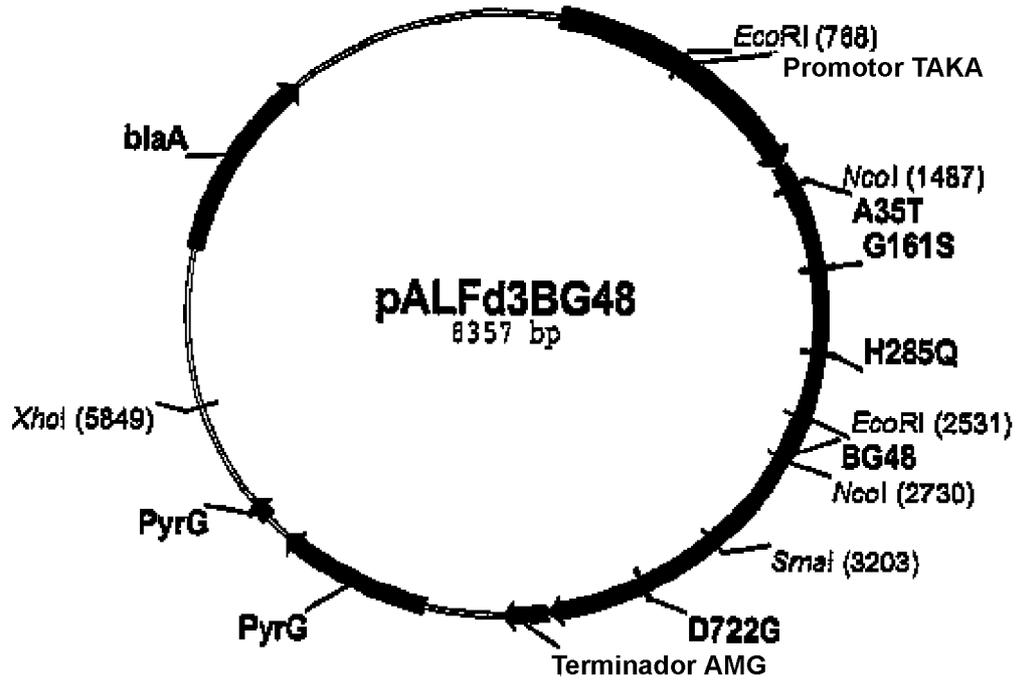


Fig. 17

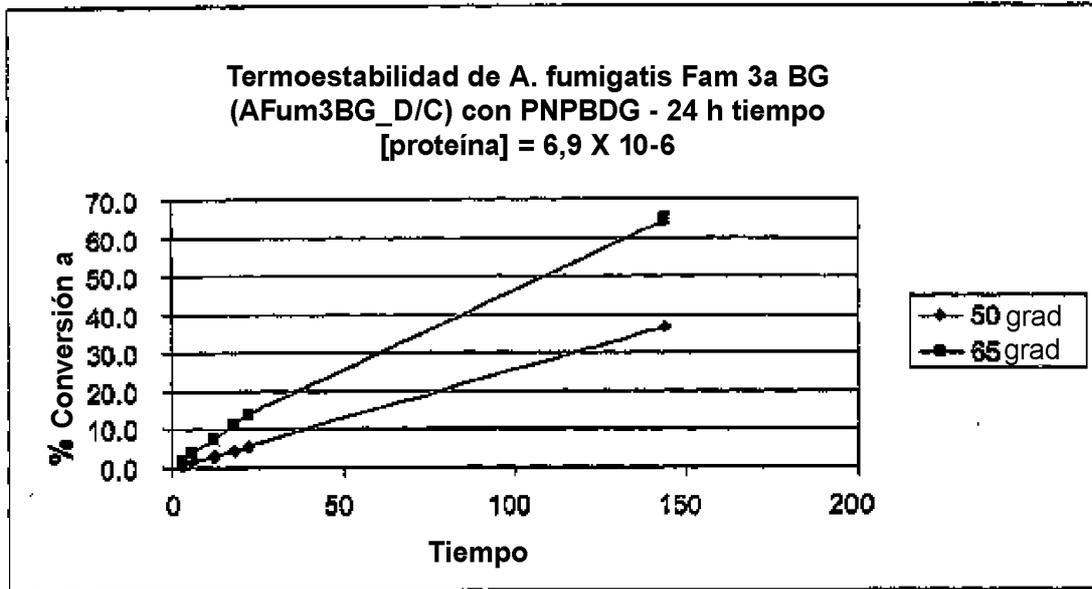


Fig. 18

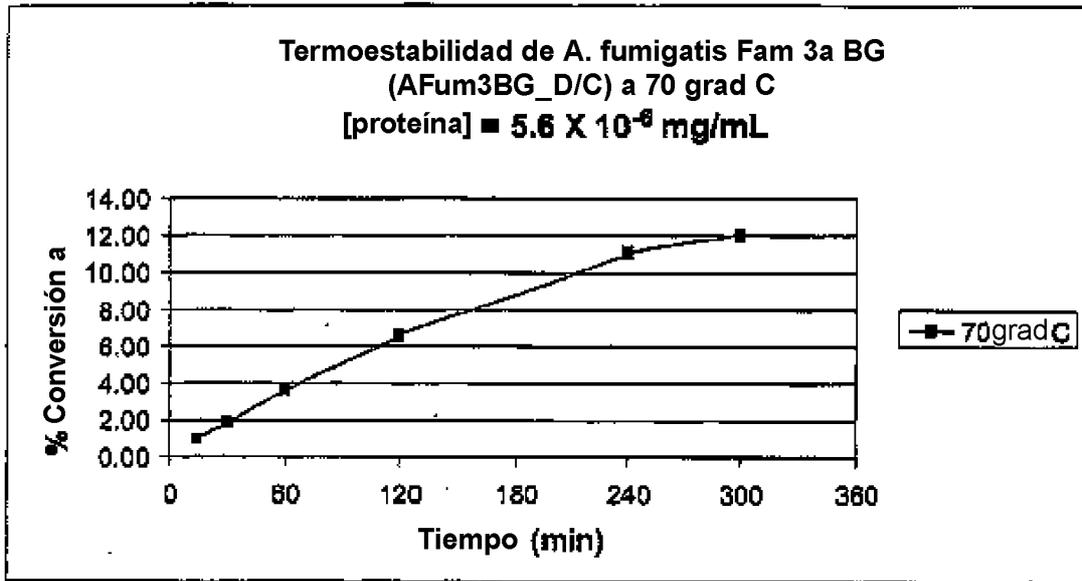


Fig. 19

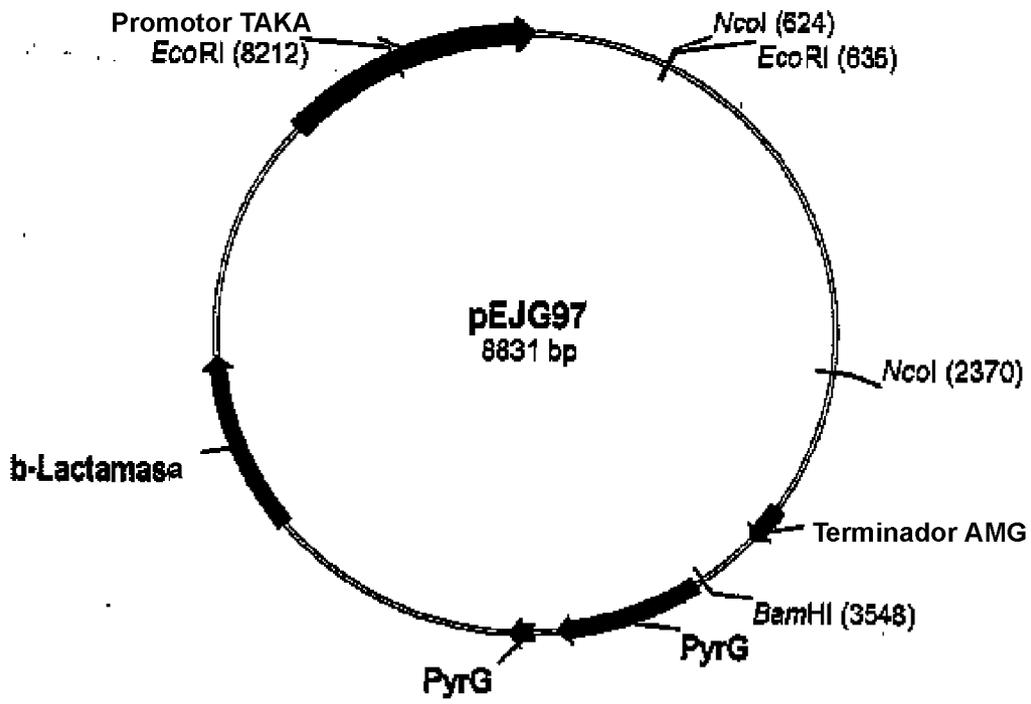


Fig. 20

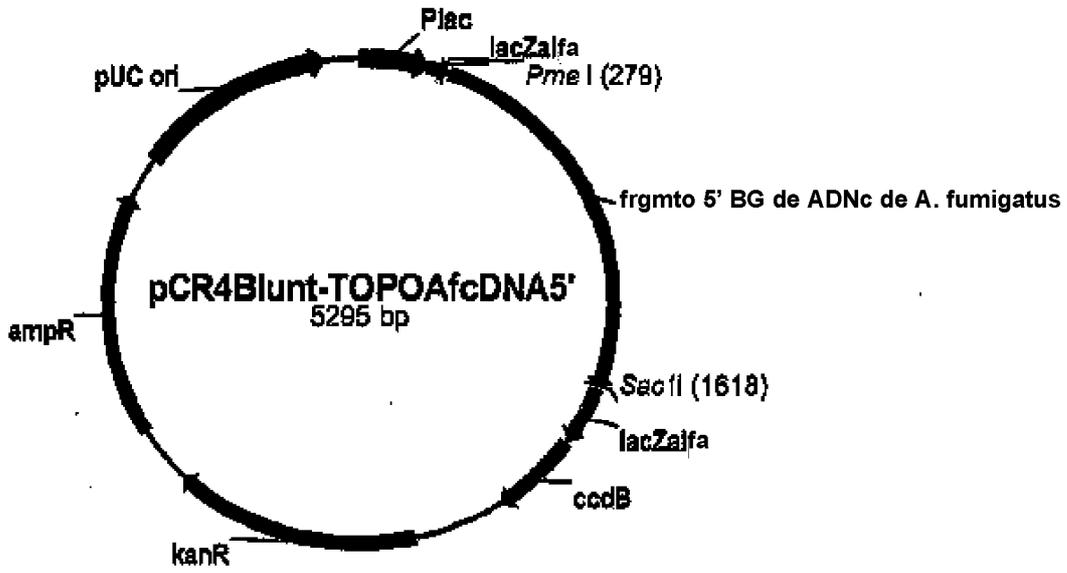


Fig. 22

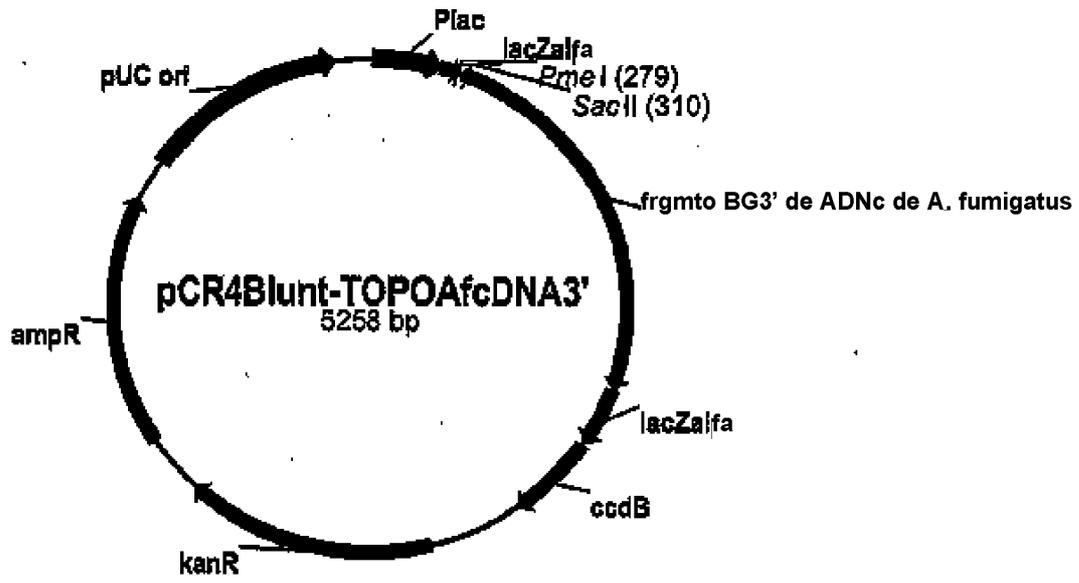


Fig. 23

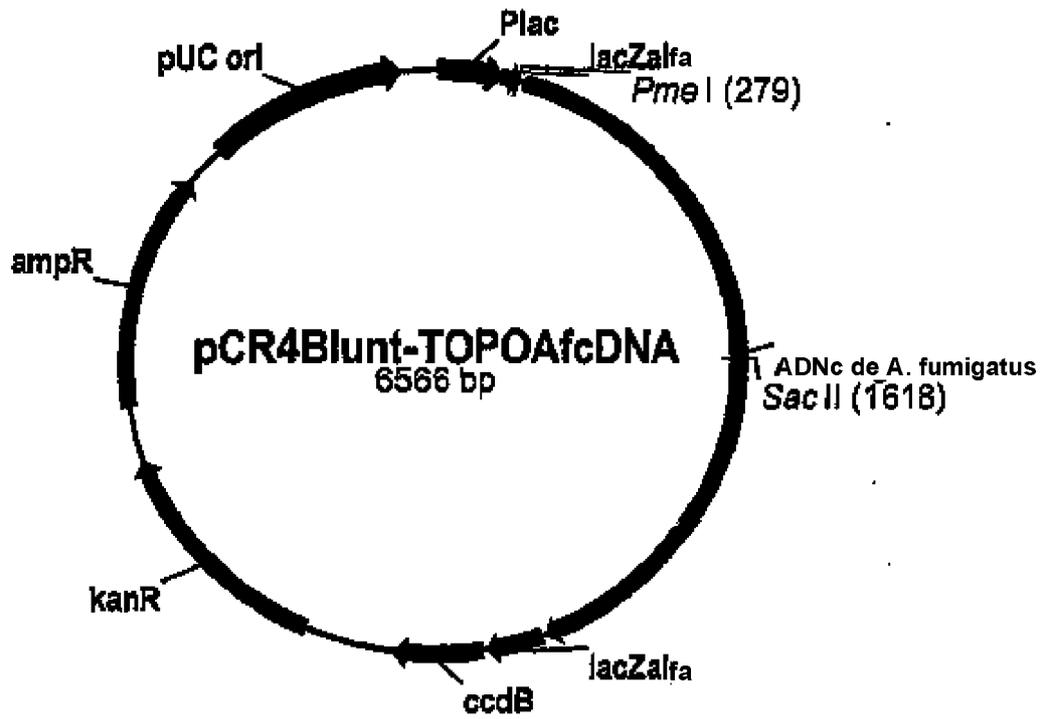


Fig. 24

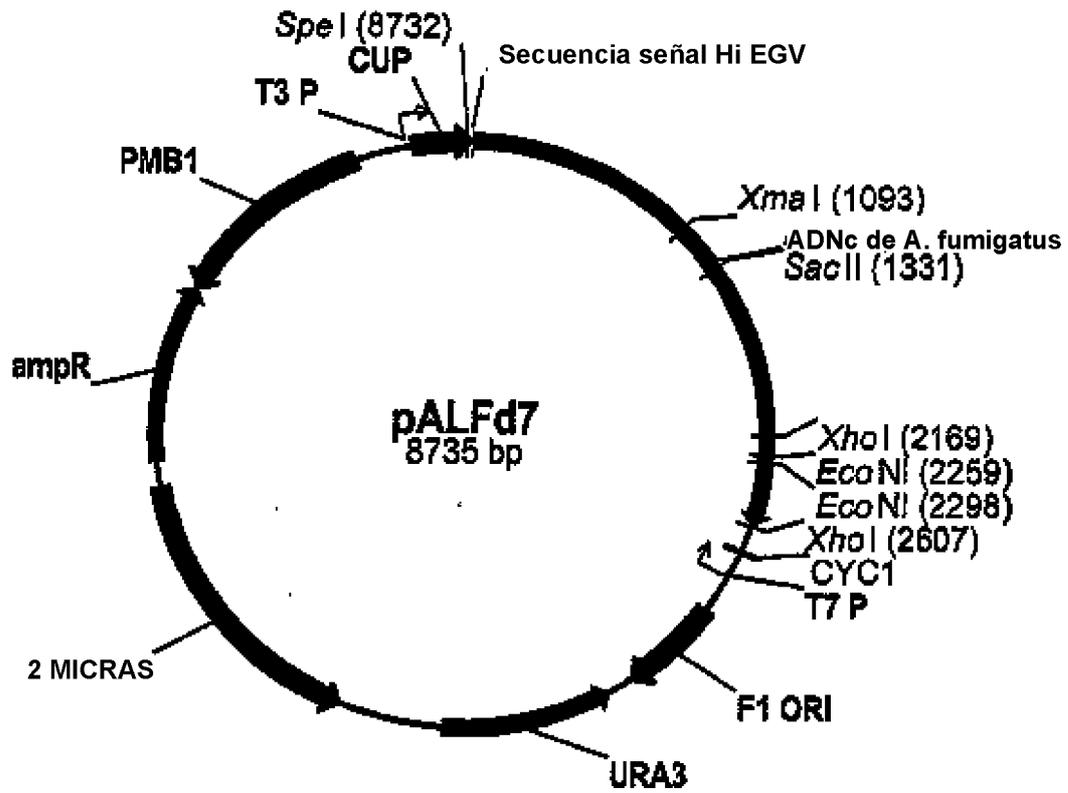


Fig. 25

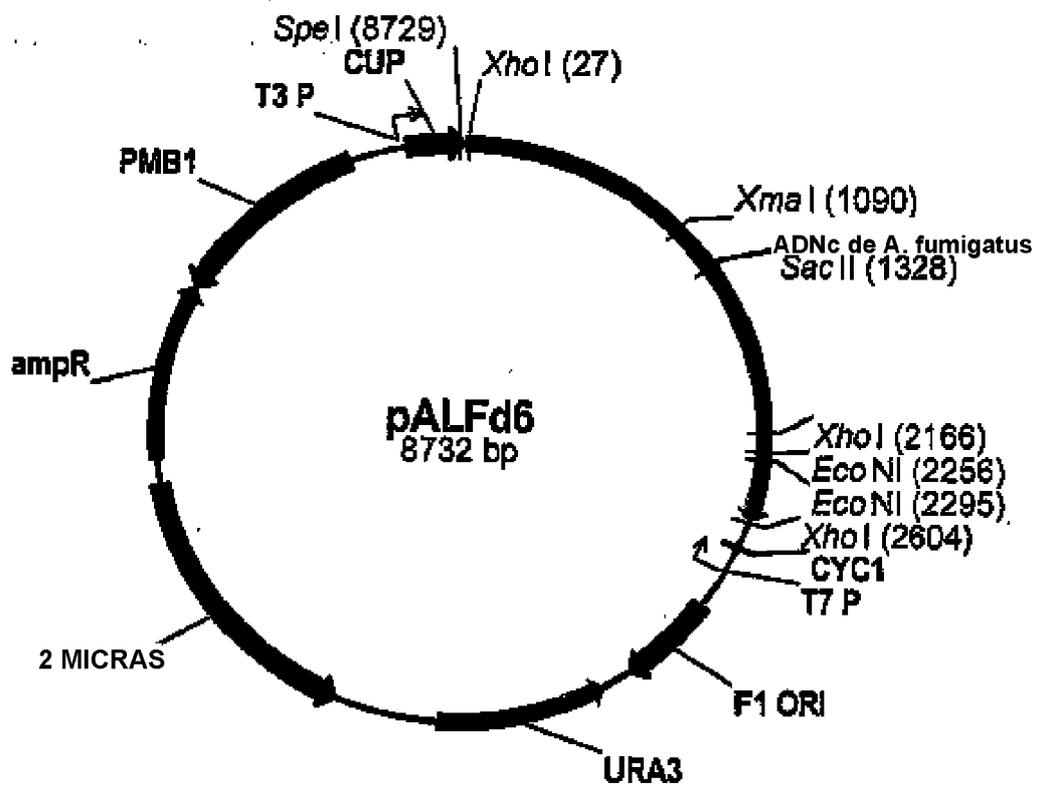


Fig. 26

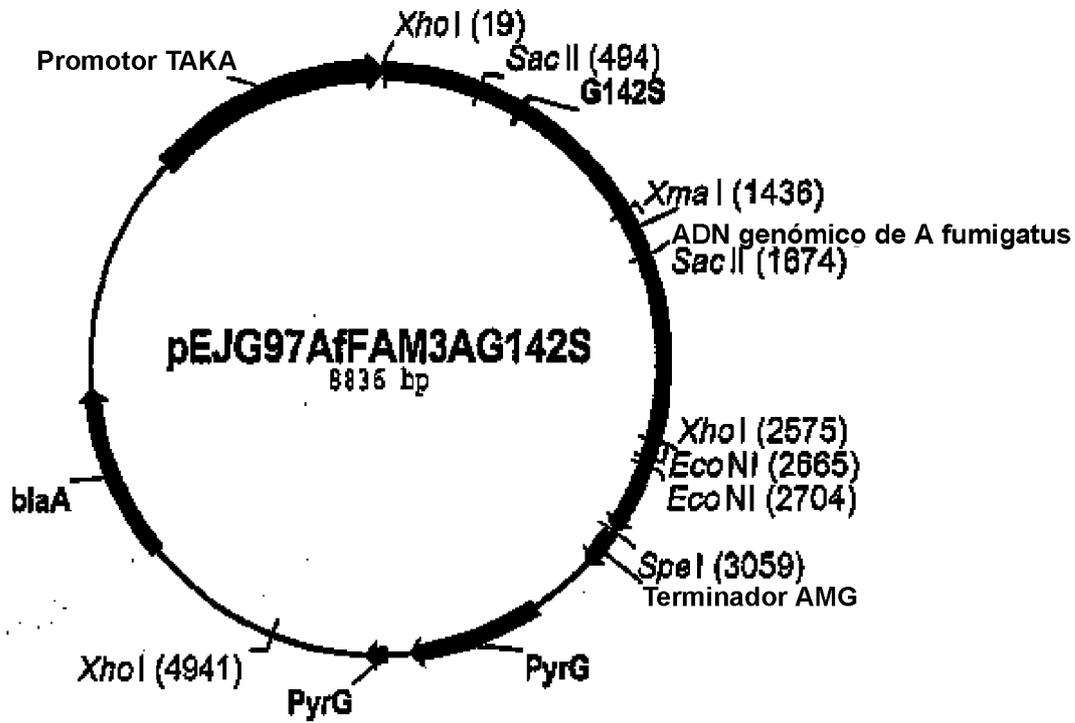


Fig. 27

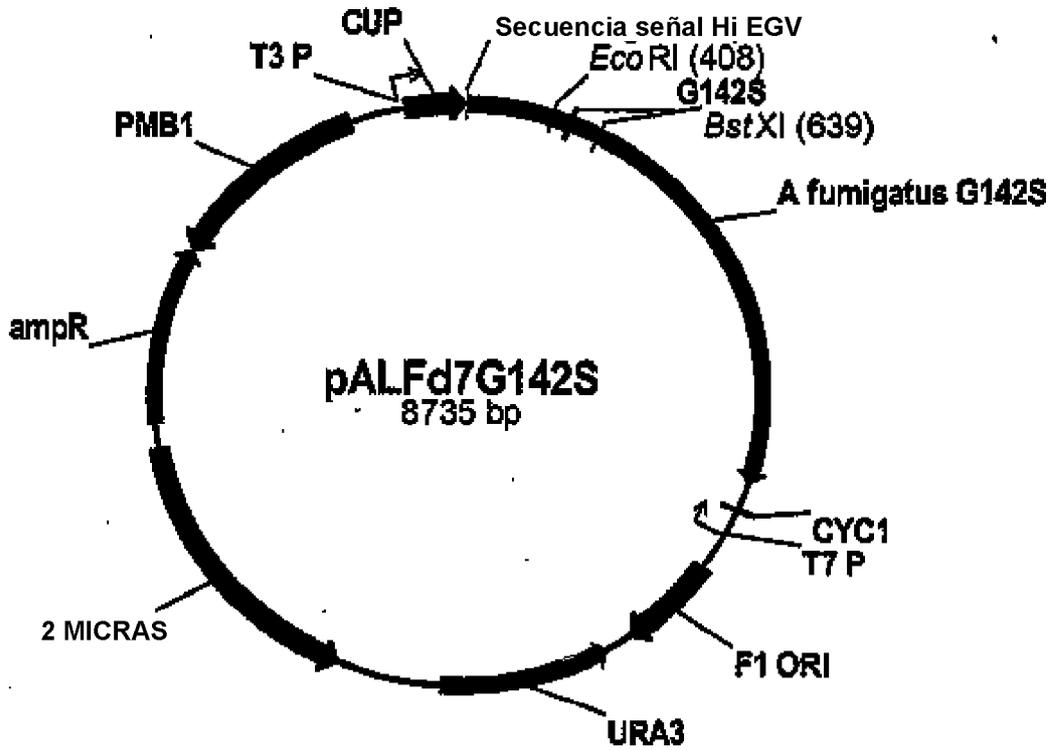


Fig. 28

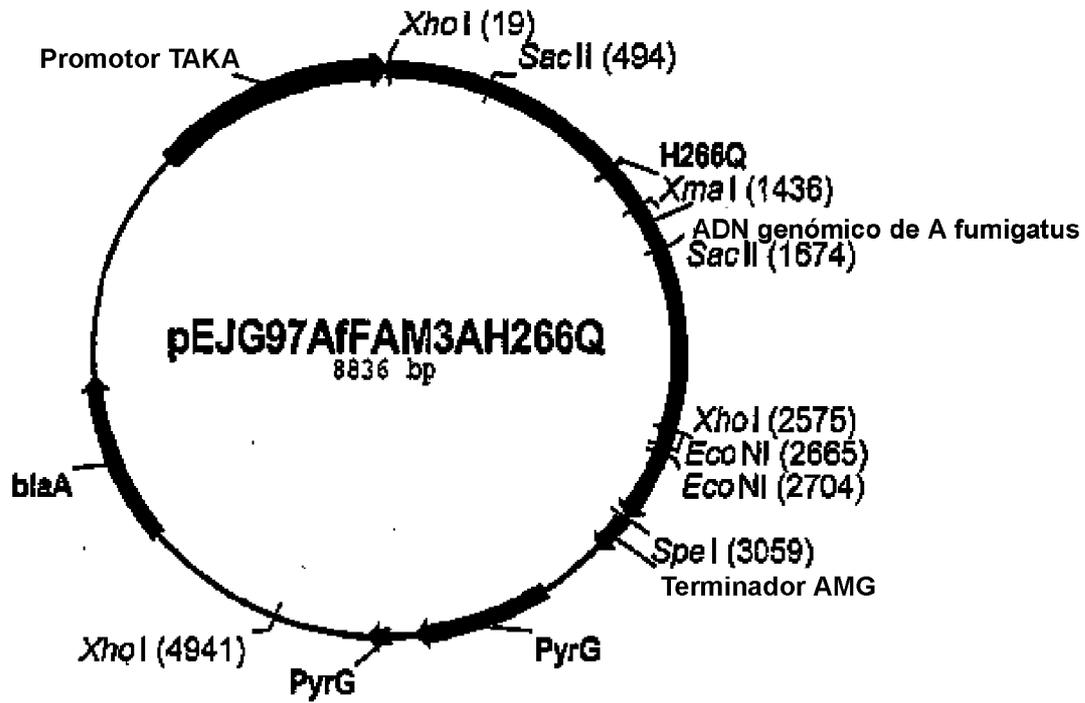


Fig. 29